CEPAL Review is prepared by the Secretariat of the Economic Commission for Latin America and the Caribbean. The views expressed in the signed articles, including the contributions of Secretariat staff members, however, represent the personal opinion of the authors and do not necessarily reflect the views of the Organization. The designations employed and the presentation of material in this publication do not imply the expression of any opinion whatsoever on the part of the Secretariat concerning the legal status of any country, territory, city or area or its authorities, or concerning the delimitation of its frontiers or boundaries.

CEPAL Review is published in Spanish and English versions three times a year.
Annual subscription costs for 2000 are US$ 30 for the Spanish version and US$ 35 for the English version. The price of single issues is US$ 15 in both cases.
The cost of a two-year subscription (2000-2001) is US$ 50 for the Spanish-language version and US$ 60 for English.
A subscription application form may be found just before the section "Recent ECLAC Publications".

UNITED NATIONS

ECLAC

ECONOMIC COMMISSION
FOR LATIN AMERICA AND
THE CARIBBEAN

JOSE ANTONIO OCAMPO
Executive Secretary

REYNALDO BAJRAJ
Deputy Executive Secretary

CEPAL

REVIEW

OSCAR ALTIMIR
Director of the Review

EUGENIO LAHERA
Technical Secretary

UNITED NATIONS

Notes and explanation of symbols
The following symbols are used in tables in the Review:

(…)	Three dots indicate that data are not available or are not separately reported.
(—)	A dash indicates that the amount is nil or negligible.
	A blank space in a table means that the item in question is not applicable.
(-)	A minus sign indicates a deficit or decrease, unless otherwise specified.
(.)	A point is used to indicate decimals.
(/)	A slash indicates a crop year or fiscal year, e.g., 1970/1971.
(-)	Use of a hyphen between years, e.g., 1971-1973, indicates reference to the complete number of calendar years involved, including the beginning and end years.

References to "tons" mean metric tons, and to "dollars", United States dollars, unless otherwise stated. Unless otherwise stated, references to annual rates of growth or variation signify compound annual rates. Individual figures and percentages in tables do not necessarily add up to the corresponding totals, because of rounding.

Guidelines for contributors
to CEPAL *Review*

The editorial board of the Review are always interested in encouraging the publication of articles which analyse the economic and social development of Latin America and the Caribbean. With this in mind, and in order to facilitate the presentation, consideration and publication of papers, they have prepared the following information and suggestions to serve as a guide to future contributors.

—The submission of an article assumes an undertaking by the author not to submit it simultaneously to other periodical publications.

—Papers should be submitted in Spanish, English, French or Portuguese. They will be translated into the appropriate language by ECLAC.

—Every article must be accompanied by a short summary (of about 300 words) giving a brief description of its subject matter and main conclusions. This summary will also be published on the ECLAC Home Page on the Internet.

—Papers should not be longer than 10 000 words, including the summary, notes and bibliography, if applicable, but shorter articles will also be considered.

—One copy of the original text should be submitted, accompanied by a copy on diskette (Word for Windows 95 format), to CEPAL Review, casilla 179-D, Santiago, Chile. In the absence of the copy on diskette, two printed or typed copies should be provided. Texts may also be sent by e-mail to: edesolminihac@eclac.cl.

—All contributions should be accompanied by a note clearly indicating the title of the paper, the name of the author, the institution he belongs to, his nationality, his fax and telephone numbers, and his e-mail address.

—**Footnotes should be kept to the minimum**, as should the number of tables and figures, which should not duplicate information given in the text.

—Special attention should be paid to the bibliography, which should not be excessively long. All the necessary information must be correctly stated in each case (name of the author or authors, complete title (including any subtitle), publisher, city, month and year of publication and, in the case of a series, the title and corresponding volume number or part, etc.).

—The editorial board of the *Review* reserve the right to make any necessary revision or editorial changes in the articles, including their titles.

—Authors will receive a one-year courtesy subscription to the *Review*, plus 30 offprints of their article, both in Spanish and in English, at the time of publication in each language.

CONTENTS

The role of the State
and the quality
of the public sector

Vito Tanzi

*Director,
Fiscal Affairs Department,
International Monetary Fund*
vtanzi@imf.org

The author of this article concludes that the quality of the public sector can be assessed only against the role of the State. In general, an efficient public sector should be able to achieve the State's objectives with the minimum degree of distortion of the market, with the lowest burden of taxation on taxpayers, with the smallest number of public employees, and with the lowest absorption of economic resources by the public sector. The public sector must be transparent in its processes and in its outcome. Corruption should have no part in the decisions made by bureaucrats and political leaders, and the resources in the hands of the public sector should be put to a use that maximizes their social rate of return. The quality of the public sector is also important for pursuing the objective of equity, which is now seen as one of the fundamental goals of the State. A high-quality public sector should make possible the pursuit of equity with the lowest costs in terms of efficiency. Finally, the so-called "first generation reforms" do not necessarily improve the quality of the public sector even though they may improve the quality of public policy. The pursuit of "first generation reforms" has, in fact, highlighted the need to improve the quality of the public sector, and for this to occur, "second generation reforms" are necessary.

I

The role of the State

*"What is the best government?
That which teaches us
to govern ourselves"*

Johann Wolfgang von Goethe

Human beings are social animals, and as a consequence they tend to aggregate in groups. Aggregation generates markets and the need for social institutions. Markets and economic activities are characterized by specialization on the part of the participants. Specialization leads to exchanges, and exchanges involve contracts, which may be implicit or explicit. Implicit contracts prevail in primitive societies and are often enforced by existing moral and social codes. As societies develop, they require more explicit or formal contracts. These, in turn, require institutions to register them, when registration is required, and to enforce them. In market economies individuals also accumulate assets and this accumulation leads to the need for property protection. Furthermore, they need personal protection against criminal elements within their groups or against outsiders. Thus, any aggregation of individuals requires the existence of at least rudimentary institutions that perform these functions and that can be considered as a minimal form of State. This is the essential role of the State in a primitive society or its minimal role in any society.[1]

As societies become more complex and as the groups that constitute them become larger and less homogeneous, the State must assume additional responsibilities if it wishes to promote to the full the welfare of the individuals that comprise it. For example, markets are not efficient when they are distorted by monopolies or when essential information is not available to the participants, or is too costly for them to obtain. This leads to the need for the State to regulate markets and to provide at least some essential information to the population. Thus, institutions must be created to satisfy these needs (World Bank, 1997).

There are some goods (defence, city streets, etc.) that only the State can provide because private individuals would not find it in their interest to provide them. One important reason for this is that once these goods are provided, the provider will not be able to exclude from their use individuals who do not wish to contribute to the cost of providing them. Another reason is that once they are produced, they can be consumed jointly by everyone, so that it would be inefficient for society as a whole to exclude some individuals from their use. Consequently, if these goods are essential, they must be provided (though not necessarily produced) by the State. These are the "public goods" of economics literature. Furthermore, the production or consumption of some goods or activities generates positive or negative externalities which are not captured by those who consume or produce them. If the State does not intervene, the market is likely to under-produce or over-produce these goods or activities. Once again, there may be a need for the public sector to intervene.[2]

All the above are examples of a general role of the State which, in public economics literature, goes under the name of resource allocation (see Musgrave, 1959).

[1] There is now an extensive literature that argues that even gangs or criminal associations of individuals require within themselves some organization that is a rudimentary form of government. See Skaperdas and Syropoulos (1995), Charap and Harm (forthcoming), and various papers in Fiorentini and Zamagni (eds.) (1999). For the concept of a minimal, or even of an ultraminimal, State see Nozick (1974).

[2] The traditional form of State intervention is to tax those who generate negative externalities and to subsidize those who generate positive externalities.

Although some problems of allocation can be dealt with by the use of regulations, often, if the State must perform this role, it will need resources. In principle, the State could directly appropriate these resources by forcing individuals to contribute their time or their wealth for the production of the required goods –as it does, for example, with compulsory military service or with some forms of expropriation. In medieval times roads were often built in this way. However, a more efficient alternative is to use resources raised through taxation. Thus, there is a need for a tax system and for institutions that collect taxes. And, of course, there is a need for institutions that spend the money and keep track of that spending. In all these activities, in a market economy the State is expected to minimize the cost that its activities impose on the market and on society at large. Thus, a public sector of high quality requires an efficient tax system and an expenditure system that minimizes inefficient and unproductive spending. It also requires a budget large enough to allow the State to perform its role in a satisfactory way (see Tanzi and Zee, 1997).

In addition to the essential or fundamental role of the State associated with the allocation of resources –a role that has been explicitly recognized and described by economists at least since Adam Smith's *Wealth of Nations*– two or even three additional economic roles have been assigned to the State in the present century. These are i) redistribution of income, ii) stabilization of economic activity, and iii) promotion of growth and employment. These latter tasks are less firmly based in economic theory than the allocation role, but they have been important in the practical implementation of policies in the second half of the twentieth century.[3]

Markets produce goods and services and provide incomes to those who participate in them. Depending on the initial situations of the individuals as regards ownership of real assets, of financial resources, and of human capital and talent in general, and also depending on their levels of effort, propensity to save, luck and risk taking, as well as on government policy, a given pattern of income (and wealth) distribution is determined. This distribution may or may not match the prevailing societal perception of the degree of inequality that is considered acceptable. Furthermore, because of physical or mental handicaps or other characteristics or situations, such as old age, unemployment, number of dependents in the family and so on, some individuals may not be able to generate an income sufficient to sustain themselves or their families. In these cases, in the twentieth century and in some countries, the State has been expected to provide them with transfers aimed at raising their consumption above a certain essential minimum level.[4]

The assignment of a redistributive role to the State has dramatically changed the character of its intervention in the economy because it has introduced into its actions purely political considerations or objectives. In fact, while the State's role in allocation may be defined (at least in principle) on the basis of technical and objective analysis, its proper redistributive role cannot be established objectively. For example, there is no objective way to determine the ideal Gini coefficient or the ideal ratio of the top quintile to the bottom quintile of the income distribution of a country. Thus, these decisions are unavoidably influenced by the biases of the individuals constituting the government in power.

Programmes aimed at redistributing income often require a lot of resources and the presence of institutions charged with their administration. For this reason, they tend to lead to large increases in public spending and in the level of taxation, as they have in the so-called "Welfare States".[5] At times, the redistribution objective has been pursued through progressive taxation, but more often than not it has required large bureaucracies and much public spending. In many countries a large proportion of the public employees are now engaged in the provision of public education and health services, and a large share of the total public spending of industrial countries goes for social programmes, including pensions. In some cases public employment has had the specific objective of redistributing income (see Alesina, Danninger and Rostagno, 1999).

Markets do not operate perfectly smoothly but are characterized by fluctuations that sometimes turn into recessions or even depressions that lead to growing unemployment and loss of output, and in the second part of the twentieth century they created another major justification for State intervention, namely, the maintenance of full or at least high employment and the stabilization of output. The promotion of this Keynesian objective is carried out by government officials and requires some corresponding institutions. These officials must have the

---

[3] These roles had not affected the behaviour of governments in market economies very much until the twentieth century.

[4] In previous centuries religious groups played a similar role in helping the poorest.

[5] The World Bank (1997) has argued that the process of taking over these new functions has distracted the State from its more fundamental role. Schansberg (1996) has argued that redistribution is often associated with poor policies that harm rather than help the poor.

capacity to change tax revenue and/or public spending in order to influence aggregate demand in desirable directions, and they must have the information and the technical knowledge and wisdom to make the right decisions.

Finally, in recent decades governments have also promoted policies aimed at raising the rate of growth or at creating employment. Import substitution and industrial policies are examples of attempts that have been made to raise the growth rate. Manipulation of the work week and public investments in particular activities or regions are examples of attempts at raising employment. Whatever the wisdom of these policies, special institutions must be responsible for executing them.

We can summarize the role of the State in a market economy in the following way:

i) it must establish and enforce formal rules in the economy, including the enforcement of contracts and the protection of property rights, as well as rules governing the extraction and use of public revenue;

ii) it must provide a legal and regulatory framework that reduces transaction costs –that is to say, the cost of dealing with other individuals in economic matters– and promotes market efficiency (through strategic market intervention in cases of market failure or through the production of some essential information);

iii) it must provide public goods and deal with obvious cases of externalities which cannot be resolved by negotiations between the private parties involved;

iv) it must promote macroeconomic stabilization; and

v) it must promote a distribution of income that is consistent with the prevailing view of society.

In order to perform these functions, the State needs some rules and institutions.

The role of the State has changed over time and across countries (see Van Creveld, 1999), partly because of the influence of changing or different views about that role and the evolution of technologies which influence what the State should do. Therefore, the quality of the public sector should be assessed in the light of the objectives and priorities of the government at any given point in time. These objectives may vary depending on the current needs of a particular country, and at times they may even conflict with one another. The State plays its role through the set of rules, laws, and institutions that make up the public sector. The higher the quality of that sector, the easier it will be for the State to play its role.

The quality of the public sector is defined here as the characteristic that allows the State to pursue its objectives in the most efficient way. Although it is difficult to separate the two concepts, the quality of the public sector is not necessarily the same thing as the quality of economic policy. Thus, a distinction must be made between the quality of the public sector and the quality of the policies that a given government may be pursuing at a given moment. A high- quality public sector is simply the instrument that facilitates the formulation and implementation of government policies. A good public sector makes it easier for the government to pursue good policies, but even a high-quality public sector cannot guarantee consistently good economic policies, because it cannot prevent policymakers from occasionally pursuing poor policies. However, one would hesitate to call a public sector "high-quality" if poor policies were frequently pursued and one would expect to find a close relation between the quality of the public sector, as defined here, and the quality of economic policy. In other words, over time a high-quality public sector is likely to promote good policies and a low-quality public sector is likely to promote poor policies.

The definition of a high-quality public sector used in this paper differs from that of "good government" used by other writers. For example, La Porta, López-de-Silanes, Shleifer and Vishny (1998) define good government as one that is "good for capitalist development". They do not distinguish between the quality of the public sector and the quality of public policies, yet a poor public sector will, of course, render the pursuit of good policies more difficult because it will not provide the government with the needed information and it will not guarantee that the policy decisions will not be distorted in the implementation stage (see, in this respect, Tanzi (forthcoming)).

II

The importance of rules

In the previous section the broad categories of State intervention in the economy were described. Each of these categories requires specific programmes, and each programme requires a legal mandate. In democracies with market economies the mandate given to the State for intervention in the economy is often specified, first in the Constitution and then in numerous laws and regulations that give a specific content to the normally general principles enunciated by the Constitution.[6] Thus, the Constitution and the laws and regulations establish the rules of the game that guide the actions of individuals and enterprises on the one hand and public institutions on the other. As North and Weingast (1994, p. 312) have pointed out, however, "[a] critical political factor is the degree to which the regime or sovereign is committed to or bound by these rules."[7]

1. The role of constitutions

The principles expressed in the constitution may be clear but are often not very specific. The constitution must be a living document that guides action, but it cannot address specific situations or anticipate activities or situations that did not exist when it was written. It has been reported that Napoleon rejected the draft of the constitution that the best legal minds of his time had prepared at his request because he felt that the document was too specific and thus potentially too restrictive and likely to become obsolete with the passing of time. He preferred a statement of general principles that could be interpreted with some flexibility over time. To a large extent the American constitution is a good example of such a statement. Only rarely has it required amendments, and the role of the Supreme Court in interpreting it has never been challenged.

Examples of constitutional principles that are not too clear abound. A well known example is Article 81 of the Italian Constitution, which seems to restrict deficit financing but, over the years, has been interpreted in ways that did not prevent the occurrence of large fiscal deficits (see Martino, 1989). Examples of excessively restrictive constitutional limits are also common. An important example of this is the 1988 Brazilian Constitution, which, in recent years, has prevented the national government from making important and badly needed reforms in fiscal federalism and in the pension system. Other examples are the Indian Constitution, which has prevented the introduction of a national sales tax, or the Pakistani Constitution, which has limited the scope of any sales tax imposed by the national government to goods only and has prevented the taxation of agricultural incomes (see also Dethier and Shapiro, 1998).

A problem with constitutions is that they tend to reflect the preoccupations and the political forces of the time when they are written. A good example of this characteristic is the Italian Constitution which came into force in 1948. Its first article sets the tone by stating that "Italy is a democratic republic founded on labour." It subsequently states, in Article 4, that all citizens have a right to work. In its third chapter (Title III), which specifically deals with "economic relations", it lists many rights of workers but only *limitations* on the property rights of individuals, specifying that the State: i) can put limits on the use of property; ii) can expropriate it for the national interest, and iii) can impose various limits on the use of property in order to "establish equitable social relations". An Italian author has contrasted this "Republic of the workers" with the old "Republic of the property owners" (see Rodotá, 1995, p. 352).

These limitations on property rights must be compared with Edmund Burke's opinion that "a law against property is a law against industry" or with Adam Smith's opinion that "[the] acquisition of valuable and extensive property ... necessarily requires the establishment of civil government. Where there is no property ... civil government is not so necessary".[8] If the proper economic role of the State in a market economy requires the protection of the property rights of individuals, as much recent literature has argued,[9] the Italian Constitution –at least in

---

[6] For a comprehensive treatment of the role of legal institutions see Davis and Trebilcock (1999).

[7] Under authoritarian governments or dictatorships the Constitution may play a marginal role. Stalin's Constitution for the Soviet Union was a very good one, but it did not prevent him from doing whatever he wanted. He was obviously not bound by constitutional limitations.

[8] Both cited in Landes, 1999, on pages 32 and 33 respectively.

[9] See, for example, Landes (1999), North and Thomas (1973), North and Weingast (1994),and Davis and Trebilcock (1999).

its formal declarations– seems reluctant to assign that role to the State. It should therefore not come as a surprise that economic policies and institutions in Italy have developed in line with the Italian Constitution and have, at times, allowed policies (rent controls, expropriation of land with very low compensation, etc.) that are not consistent with the principle of protection of property. This may also explain why Italy has one of the lowest scores, in terms of "economic liberty", among the countries assessed by the experts of the Economic Freedom Network.

In conclusion, the role that constitutions play in determining the quality of the public sector in democratic market economies cannot be exaggerated.[10] It is the Constitution that encourages or allows certain actions on the part of the governing body and the citizens. Thus, at least in principle, the laws and the regulations that govern a country must be consistent with the constitutional principles as interpreted by the Supreme Court or by an equivalent body. In recent years, there has been a trend in many countries to try to modernize their Constitutions. However, the results have not always been very satisfactory, and at times the revised constitutions have been poorly drafted, complex and confusing.

Some economists, including James Buchanan and Francesco Forte, have argued that the economic or fiscal role of the constitution should be one of establishing limits on governmental action, as is done for example in the Swiss Constitution. According to these authors, the constitution should stress what the State *cannot* do rather than what it should do (see Forte (ed.), 1998). These authors tend to favour limitations on tax rates, on levels of public spending, or on fiscal deficits, the objective obviously being to provide the greatest scope for the working of the market. The Maastricht Stability and Growth Pact might be seen as an example of such a limit in a constitution.

In some circumstances –and again Brazil comes to mind– a change in the constitution to remove obstacles to desirable and important reforms may be a necessary condition for raising the quality of the public sector and of public policy.

2. Laws

While the constitution sets, or at least should set, the general principles that guide a country's policies, the latter are permitted and directed by specific laws. It could be argued in general terms that the quality of the public sector is enhanced when the laws are relatively few, are clearly written (and thus not subject to conflicting interpretations), are comprehensive, and do not conflict with each other. Difficulties have arisen when there are too many laws, they are not clear, and they do not cover all relevant areas of economic activities, or provide conflicting signals.

It has been reported that while some European countries have only a few thousand active laws, others have tens of thousands (see Ferro, Lo Faso and Salvemini, 1999). In such circumstances it may be difficult to find one's way in this legal jungle. The legal system may reflect a problem conceptually similar to that associated with the Y2K difficulty. It occurs because every time a new law is enacted, all the existing laws should be scrutinized and, if necessary, revised to make sure that all the elements in them are consistent with the new law. This, of course, does not happen. When the existing laws are very numerous, and especially when they are not clear, it is almost impossible to ascertain this. Thus, at some point elements imbedded in laws which were enacted in the past but are still on the books may be seen to conflict with the new laws. When this happens, the directives to citizens, as well as to the institutions charged with implementing the programmes contemplated by the laws, become confused. This problem could be termed one of legal inconsistency. It often characterizes the relations between national and subnational levels of government, or between, say, pension and health programmes on one hand and annual budget laws on the other.[11] Zoning laws and laws dealing with the environment have also suffered from this problem. At times, one law and a given institution authorize a certain land use but another law and another institution prohibit this. These conflicts may lead to costly mistakes and to uncertainty about property rights, and could negatively affect market decisions.

Thus, to repeat, the quality of the public sector is enhanced when the laws are written clearly and cover all the necessary areas, when they do not lead either the public or public officials to make conflicting interpreta-

---

[10] This of course assumes that the rule of law is well established in the country, so that the constitution is taken seriously. It also assumes that effective Supreme Courts exist. In Italy, the "constitutional court" started operating ten years after the approval of the Constitution (see Rodotá, 1995, p. 353).

[11] For example, the budget law may assign a given budgetary allocation for the health sector but the laws that determine the performance of the health sector may call for a higher level of spending. See Reviglio (1999) for a specific example from Italy.

tions, when their number is as small as feasible, and when they do not conflict with each other. Over the years there have been attempts on the part of some countries (France, Italy, New Zealand, Russia) either to codify the existing laws or to simplify them. However, as these actions would give benefits only over the long run, they do not get much political support and the attempts are often abandoned before they produce the necessary results.[12]

3. Regulations

Laws are often accompanied by specific regulations.[13] These can be classified in three groups: economic regulations, safety regulations, and information regulations. Regulations explain procedures or elaborate on the content of the laws or simply impose rules on individuals and enterprises. In some cases the laws are so complex that they require a very large body of regulations. For example, the regulations covering the U.S. Internal Revenue Code are reported to cover 18,000 pages, while the Internal Revenue Code itself is some 2,300 pages long. Regulations issued by an executive authority or by a regulatory agency can themselves be very complex, may not be easily accessible by the public, and may overlap with other regulations. In some countries they may not even be published.[14] Yet, as a recent OECD report states: "Regulation is perhaps the most pervasive form of State intervention in economic activity" (OECD, 1999, p. 179).

Because many forms of regulations do not require budgetary appropriations or formal approval by the legislature, they tend to be less scrutinized. Thus, there is often an oversupply of regulations and their interpretation is all too often left to the bureaucrats who administer them. Therefore, the possibility of confusion or even abuse is very high. Regulations have been identified by the literature on this subject as one of the major causes of corruption, because the bureaucrats in charge may abuse them for their personal gain (see Tanzi, 1998b). They have also been shown to impose very large welfare costs on the economy.[15]

Because of the dynamic character of economies, and because of fast technological change, it is common to find countries with too many anachronistic, useless or even damaging regulations and too few necessary regulations relating to new economic activities.[16] This is an area where excess and scarcity often coexist. Many countries are now struggling to create needed regulations for the financial and banking sectors, for the use of the information superhighway (Internet), for genetic research, for the use of some drugs, and for several other important new areas.

To sum up, a high-quality public sector must have enough clear rules to guide economic (and other) activities, but not rules that are so numerous or so vague as to give excessive power to bureaucrats or to create uncertainty among those who make economic decisions. In general, the rules should specify what is *not* allowed rather than authorize what *is* allowed. Discretion by bureaucrats should whenever possible be kept to a minimum. It has been reported that there are countries where some routine activities such as a request for a tax incentive or an application to open a small enterprise may need 30 or 40 authorizations signed by public employees in as many agencies or offices (see, *inter alia,* De Soto, 1987). This cannot be consistent with a high-quality public sector or with an efficient market.

An important step would be to make a periodic inventory (say, every ten years) of all the existing regulations, so that a "regulatory budget" could be established and could be pruned of redundant and anachronistic regulations while complementing it with new required ones and clarifying the confusing ones. Such an attempt has been made, with mixed results, by a few countries including Hungary and Argentina. Such a process, while costly, would raise the overall quality of the public sector. Another useful step would be the creation of a one-stop, or "single window" centre where individuals can obtain all the permits and authorizations they need for their activities. Such centres have been created in a few places and are reported to have led to the elimination of many existing regulations and a reduction in corruption.[17]

---

[12] See, for example, Guy (1996), Braibant (1996), Mattarella (1994), Smith and Richardson (1999) and Tan and Tower (1992).

[13] Regulations may be legislated and thus be laws themselves, or they may simply be issued by public agencies.

[14] In a country in which I worked many years ago tax incentives regulations were not publicly available in any form. Thus, the bureaucrats making the decisions had all the knowledge and their decisions could not be challenged by the taxpayers. These cases of asymmetric information between the State and citizens are not uncommon.

[15] See United States Congress, Senate Committee on the Budget (1984); Dixit (1996); Laffont and Tirole (1993); FIEL (1988); OECD (1999), and Jacobs (1999).

[16] A couple of years ago, it was reported in the American press that in California, bread had to be sold only in pieces of an exact, well specified weight. This regulation, which was still on the books, had been introduced at the beginning of the century when most buyers did not have scales, so that they could be easily cheated by unscrupulous bakers. In Italy and Germany discounts on sales by shops must be approved by the municipality and must be limited in time.

[17] Centres of this type now exist in Bologna (Italy) and in Salvador de Bahía (Brazil). In Bologna an attempt is being made to allow access to this one-stop centre through the Internet.

In addition to the formal constitutional rules and the rules specified in the laws and regulations, the quality of the public sector may be affected by informal norms or arrangements that influence the economic behaviour of individuals and the behaviour of the public sector. Such norms i) influence the choice of presidents, ministers, and other high-level officials; ii) set the pattern of appointments in the civil service, and iii) influence contacts between the State and the private sector.[18] These informal norms may be based on religious, social or political considerations. Since they are of an informal nature and are based on cultural characteristics, they are difficult to change. Even so, the application of arms-length relationships and the rule of law in all aspects of public sector behaviour should be a goal in the search for a higher-quality public sector (see Tanzi, 1995).

III

Political and procedural rules

The constitution, the laws and the regulations establish the broad legal powers of the public sector; or, to put it differently, they set the rules of the game that should determine the behaviour of the public sector and the regulation of the market. The importance of these rules cannot be overestimated. Several authors, including Buchanan, Alesina, von Hagen, Poterba, Tabellini and Persson, have argued that political arrangements such as fiscal federalism and fiscal decentralization, proportional or non-proportional representation in parliaments, the frequency of elections, the choice of presidential versus non-presidential types of governments, the role of the ministry of finance as a super-ministry, the rules that apply to the budgetary process (for example, whether it starts with a macroeconomic constraint which reflects a collective view on priorities or whether it allows pressures for spending to be determined through the political influence of each minister), whether parliament can modify the content of budgetary proposals or must vote on the whole budget, whether the central bank is independent, and so on have a significant impact on fiscal and macroeconomic outcomes.[19] These situations have been modelled using strong assumptions and sophisticated game theories and have subsequently been subjected to empirical testing (see, among others, Tabellini and Persson (forthcoming) and Poterba and von Hagen (eds.), 1999).

I will not review this literature, which is still evolving and which, at times, has come to conflicting conclusions. While recognizing its importance and its potential contribution to explaining economic policy, the focus in the present paper is on other aspects and, specifically, on the *quality* of the public institutions. Political and procedural rules are more likely to affect policies than the quality of the public sector or of public institutions. It is the institutions that confront the citizens and implement the policies. But, of course, by changing the incentives under which policy-makers and institutions operate, the political and procedural rules may affect the behaviour and thus the quality of the institutions, and vice versa. The rules are just a set of instructions. They are not yet the policies. Until the game is played, these rules remain just pieces of paper, and the game is played by the institutions charged with carrying out these instructions,[20] which may or may not carry out the instructions in a faithful and efficient way.[21] The public sector is composed of many institutions, some more important than the others. It is the performance and efficiency of these institutions which, to a large extent, determine the quality of the public sector.

[18] In some countries high public positions are almost inherited. Some jobs in particular institutions are passed from one member of a family to another. In others, party affiliation facilitates access to government jobs and a change in government creates a large number of vacancies because many jobs are considered as political appointments.

[19] Ricardo Haussmann has argued in support of a politically independent Fiscal Council that would restrict annual public debt accumulation to agreed levels. The analogy with the idea of an independent central bank is obvious. Others have argued in favour of a fiscal policy that to some extent is insulated from political pressures (see, for example, Blinder (1997) and Kopits and Symansky (1998)).

[20] It can be argued that there are two sets of games taking place. The first is the one that sets the rules and the second is the one that implements them.

[21] In a recent paper I argued that public policy is often distorted by the existence of principal-agent problems (see Tanzi (forthcoming)).

IV

The quality of public institutions

The quality of the public sector may be affected by the absence of some essential institutions or by the poor performance of the existing institutions. For example, in many countries there are no institutions responsible for enforcing competition, for forcing full disclosure on the part of financial institutions or for obliging enterprises whose shares are traded on the stock market to present proper accounts. As a consequence, the market may function less well because of cronyism and monopoly powers or because of lack of essential information. The performance of public institutions depends on many factors including i) tradition and reputation; ii) the resources they have available and the discretion over their use; iii) the clarity of their mandate; iv) their organization; v) the incentives they are given; vi) the quality of their leadership and staff, and vii) the freedom they have in terms of reorganization.

Let us take one of the most fundamental institutions as an example: namely, the tax administration. Its performance will depend in part on its tradition and reputation. A tax administration which has been efficient, honest, and proud of its work in the past is likely to continue to be so in the future unless it suffers truly fundamental shocks. By the same token, it is very difficult in the short run to change a corrupt and inefficient administration.[22] Its performance will also depend on the resources that it has available for hiring capable employees and paying good salaries, investing in new computer technology, carrying out necessary audits and so on. The clarity of its mandate –for example, to enforce the tax laws fairly and objectively– is also important, while its day-to-day independence from political pressures is essential. When the mandate becomes unclear, either because the laws are not transparent or because the institution is subjected to political interference that forces it to accommodate to the special circumstances of some taxpayers, problems develop. This has been the case, for example, in some transition economies and some developing countries,[23] where political interference has reduced the quality of the tax administration. The organization of the tax ad-

ministration is also important, as is the set of incentives that it is given. If an institution is poorly organized, or if good or bad performances are equally rewarded, the contribution of that institution to the quality of the public sector will be low.[24]

Attempts are currently being made to strengthen the incentives for good tax administration by making such administrations politically independent, like the central banks,[25] and by negotiating explicit contracts between the government and the tax administration that require quantitatively specified levels of performance. In Australia, for example, the government guarantees the tax administration a given level of resources over a three-year-period, and in return the tax administration undertakes to deliver certain quantifiable results and outputs.[26]

This brings us to two other important and related aspects of public institutions, namely i) the synergy among public institutions, and ii) the enforcement mechanisms. These are treated here as two separate aspects, although to a large extent they are two faces of the same coin.

1. Synergy

Like different elements of an ecological system, public institutions work together and support one another, so that it may not be possible to have, say, a first-class tax administration in an environment where other institutions, such as the treasury or other important ministries, the judiciary, or even the post office do not function well.[27] Often the same weaknesses affect different institutions, so that attempting to improve just one institution, when the others need equal attention, is not likely to generate the desired results in the long run. This has

---

[22] Occasionally countries have opted to close down unsatisfactory offices or departments and open new ones, thus starting again from scratch.

[23] See for example Tanzi, 1998a.

[24] The rules on hiring, promotion, firing and other matters relating to public servants obviously also play an important role in determining the quality of the public sector.

[25] Thus, the tax administration is given greater discretion in the use of resources.

[26] The cost of collection may be of importance in this context. See, for example, Highfield (1999).

[27] For example, in countries where tax evaders are not punished by the courts, it will be more difficult for a tax administration to fight tax evasion. In countries where the mail does not operate well, it will be difficult for the tax administration to rely on it to contact taxpayers and vice versa.

been the experience in transition economies where, for example, the establishment of a good treasury system has not improved the quality of public expenditure management much because the budgetary process has continued to generate budgets that are so unrealistic that no treasury could finance or manage them. In some of these cases the result has been the accumulation of arrears on the part of the government, matched by a similar accumulation of arrears by taxpayers (see Potter and Diamond, 1999).

Inter-institutional externalities (either positive or negative) are very important and must be recognized and dealt with in any attempt at improving the quality of the public sector, although unfortunately, as far as I know this aspect has not been addressed in the literature. For example, when the judiciary does not work well, many other institutions suffer. The same could be said of the educational system. A holistic approach that simultaneously addresses problems in different institutions is likely to be necessary. However, such an approach, which is inevitably difficult to follow, must be guided by a clear strategy and by the proper sequencing of the changes required and made. If this approach requires more time to implement than the political horizon of the government that introduces it, it is less likely that it will be fully implemented. This is the reason why the quality of the public sector changes only slowly over time.

2. Enforcement mechanisms

The quality of the public sector will depend to a considerable extent on the existence of controls and enforcement mechanisms. Some of these mechanisms must operate within the institutions themselves. For example, efficient internal auditors' offices can improve the functioning of the institutions and provide some guarantees that the latter will not stray away from their basic mandate. But these mechanisms may not be sufficient. In other cases the enforcement mechanisms must cut across institutions, as occurs when top-level audit institutions specializing in controls and enforcement have responsibility for controlling spending and revenue collection. These institutions often draw their mandate from the constitution itself and act as independent bodies.

Examples of such top-level audit institutions are the General Accounting Office (GAO) in the United States, the Cour des Comptes in France, the Corte dei Conti in Italy, the Comptroller-General's Office in many Latin American countries, and so on. Historically, these audit institutions have focused too much on whether the institutions have complied with the *letter* of the law rather

than its spirit. Thus, in many cases they have paid less attention to the performance of the institutions in terms of outputs and outcomes than to the question of whether they have complied with legalistic requirements. This kind of auditing is of limited value, except, perhaps, for ensuring accountability, because it does not promote the quality of the public sector in its fundamental objective of serving the public. It does not guarantee that the public is getting value for the money spent by the public sector.

In recent years there has been a movement to focus on performance and output, rather than on formalities and on input. This movement tries to assess public spending in terms of its economy, efficiency and effectiveness. It requires quantitative indicators of public performance and of the cost of public sector activities. It has had its strongest expression in New Zealand and Australia, but it has also been spreading to other countries in modified forms.[28] It has brought with it many changes in contractual arrangements and in the organization of public institutions. For example, in the countries that have adopted it, civil service jobs no longer carry lifetime tenure and many constraints on the actions of those who run departments have been removed. The government, as the principal, now makes contracts with a public agency on what the latter must deliver, and the head of that agency becomes personally responsible for the outcome. If the agreed goals are not met, the contracts of the heads of the agencies are not renewed or their salaries are reduced. Thus strong economic incentives are being linked to performance. The final impact of these changes in countries whose cultures are very different from those of Australia and New Zealand remains to be assessed.[29] It would be hard to implement these changes, for example, in countries where it is difficult to fire workers, almost regardless of their performance. It is easy to see the conflict between this approach and the spirit of the Italian Constitution.

---

[28] While the Australian system initially focused on the outcomes of public spending, the New Zealand system has always concentrated on the *measurable* output. It should be clear that outputs and outcomes are not the same thing. Outcomes may be more difficult to measure and may be affected by public spending, but with long lags. For example, educational spending may be assessed in terms of students in school (an output) or in terms of growth in human capital (an outcome). Health spending may be assessed in terms of operations performed or in terms of impact on the length and the quality of life. Obviously, the goal of the government should be to influence outcomes, but these are more difficult to measure than outputs.

[29] For a skeptical assessment of these changes, see Schick (1998).

This discussion of enforcement mechanisms and controls would not be complete without a reference to two fundamental points.[30] The first is the weakness of the cash accounting approach, which has traditionally been used for the measurement of government operations, as a means of securing good efficiency controls. The second is the measurement of consumer reactions and preferences with respect to the services rendered by public agencies, which has only recently begun to be studied. Before addressing these two points, it is important to mention that any concept of the quality of government services must be related to the cost of those services and to the resources available for generating them. The greater the resources available, the better one should expect the public services to be. For a given level of available resources, however, the higher the efficiency of the public agency, the higher will be the quality of the services provided.

The first question, then, is how to measure the level of resources used. Traditionally, the fiscal accounts (and hence the budget) have relied on cash accounting or cash transactions. Thus, the cost of, say, agency X is measured by the cash transfer that agency X receives from the budget. However, accounts based on cash transactions have many weaknesses which are now more fully recognized by accountants and economists. The need for accrual accounting is now well established, at least, in principle, although practical difficulties in its use are likely to delay for some time the transition from cash accounting.[31] Accrual accounting measures much more accurately the true costs of public sector activities, and it would thus make possible audits that are more meaningful in comparing outcomes with true costs. For example, cash accounting ignores the opportunity cost of using public sector assets if this use does not result in cash transfers (see Tanzi and Prakash (forthcoming)). Thus, an activity that receives only small cash transfers but uses very valuable land or buildings is now assessed as having low costs, which is obviously wrong.

The second question relates to the role of citizens, as taxpayers and consumers of public services, in evaluating the latter. Most of the reforms to date –particularly those agency-based reforms that distinguish between principal (the government) and agent (the public agencies)– represent actual or implicit contracts between two

producer interests. An example from the private sector might be a contract between a regulatory body for the private sector and the monopoly industry it regulates. The regulatory body is supposed to represent the consumer in some sense, but its direct role is basically to avoid excessive profits for the producer rather than to specifically protect consumer interests. What is missing in the administrative arrangements in Australia and New Zealand is the voice of the consumer. Unlike the private sector, where consumers can state their preferences directly (providing there is an adequate degree of competition within the market), this possibility does not exist in public services. Therefore, some reforms in the 1990s have been aimed at trying to find surrogate indicators of consumer preferences. At least three approaches are worth mentioning in this respect:

i) Some evaluation studies have canvassed the views of particular consumers, not in relation to an organized consumer survey, but as part of a wider exercise aimed at finding out, for example, parents' views on the education given to their children. Such exercises have tended to be partial and confined to individual sectors.

ii) Another approach has been based on formal consumer surveys that may cover a number of different services. The World Bank and individual countries or institutions have carried out some initiatives in this area.[32] While this does represent an advance on the past, the limitations of this approach are also significant. First, it is essentially opinion-related rather than measuring some actions or providing specific responses by the consumer to quality concerns. Second, it is subject to the usual problem that a well-designed survey can produce more positive views than might really be justified of services provided by the public sector.

iii) These weaknesses in both evaluations and consumer survey-type approaches led the United Kingdom in the early 1990s to seek an alternative approach. From this emerged the "Citizens' Charter", which tries to involve the citizens in the setting of standards. This has been done in a number of ways. Citizens' representatives or other consumer interest groups are also involved in the setting of standards and performance contracts between the government and agencies, not just the ministry of finance. Moreover, having set standards, consumers are also given the right and means to complain, through the provision of telephone hot lines, etc. Consumer response is taken into account when assessing the

---

[30] I am grateful to Barry Potter for calling my attention to these two points.
[31] The Department of Statistics of the IMF is in the process of producing a new manual of government finance statistics based on accrual accounting.

[32] For example, users of public hospitals have been asked to assess the quality of the services provided, or taxpayers have been asked their opinions on some aspects of tax administration.

overall performance of an agency manager. The compulsory use of name badges for officials –so that consumers can identify them and send in responses that reward good service and punish bad attention– is a small but significant element. This is still another example of how incentives are slowly being introduced in the public sector. Furthermore, in certain cases consumers are entitled to some form of compensation, if the public service provider fails to meet the required standards. The simplest example is that season ticket holders on certain rail services are entitled to refunds if delays in the running of the trains exceed a certain limit.

None of these approaches is wholly satisfactory, but at all events the inclusion of consumers' views in setting standards and commenting upon their compliance or otherwise, and thus influencing the future allocation of public sector resources, is likely to become a major theme over the next few years. It is also likely that incentives and penalties will play an increasing role in promoting a higher quality public sector.

In addition to the internal controls and those performed by auditing institutions such as the GAO or the Cour des Comptes, there are some other institutions whose work and efficiency are an essential ingredient in the quality of the whole public sector. Among these the system of justice is of the greatest importance. The role of the justice system in all its manifestations in enforcing contracts, in protecting property rights, in ensuring the safety of individuals, in keeping corruption under control, and in improving the efficiency of other institutions is fundamental. Thus, it is no exaggeration to say that the quality of the public sector of a country and the functioning of its market depend significantly on the performance of its justice system. It is for this reason that the system of justice is receiving a lot of attention in many countries, including European countries such as France, Italy and Portugal and the majority of Latin American countries.

In many countries the system of justice has been in crisis in recent years. Individual and property rights are not protected, contracts are not enforced, processes take years or even decades to be concluded, and so on. In many cases individuals who break the law are not punished, are punished only lightly, or are not punished until much later, when the deterrent effect of the punishment is lost. In some countries, the slow-moving or even corrupt system of justice has provided an implicit incentive for tax evasion, corruption and other illegal activities, because those caught committing crimes can count on the inefficiency or corruptibility of the justice system in order to escape punishment. In some countries it takes ten years for the government to determine whether someone accused of tax evasion should really pay the taxes claimed or not. Furthermore, often the penalties imposed are insignificant. This is a good example of the cross-institutional externalities mentioned earlier. The same may occur in the case of those who do not comply with the terms of contracts. For example, the proliferation of bad loans and of financial crises is in part a direct consequence of the excessively low and much-delayed penalties on those who do not comply with the terms of their contracts.[33] In some countries apartments or houses are left empty, rather than being rented out, because of the difficulty that the owners would have in reclaiming them at the end of the contract. The failure of the justice system clearly encourages tenants to ignore the terms of their contracts, and imposes efficiency costs on society as a whole.

When justice is corrupt or inefficient, it also becomes unjust, because some people are more adroit at taking advantage of its weak points. Law-abiding citizens are the ones who end up paying a higher price, and economic activity suffers. Equal access to justice –and to justice that is administered in a timely fashion– must remain one of the fundamental goals of the State. It is also one of the main requirements for an efficient market. If this goal is not met the quality of the public sector will remain poor (see Guigou, 1999).

---

[33] In many countries bankruptcy has become almost a trivial matter, of little consequence. The establishment of suitable bankruptcy laws has therefore become an important issue in such cases.

V

Measuring the general quality of the public sector

Although knowledgeable individuals may have an *a priori* idea of the general quality of the public sector of a given country, it would be difficult or even impossible to make objective measurements of that quality. In principle, one could conduct surveys of perceptions of such quality, using the same techniques adopted, for example, in surveys of corruption. In order to obtain acceptable results, however, the informational requirements on the part of the respondents would be extremely high, so that the quality of the responses would tend to be low. It might be easier to evaluate the quality of each of the major institutions that make up the public sector and somehow weigh their importance for the general quality of the public sector. However, given the number of such institutions and the knowledge required to assess them, such an exercise would also be very costly and would not necessarily achieve the desired results. An alternative way would be to simply measure the economic and social performance of a country, focussing on the output or, better, on the outcome and attributing the results obtained to the quality of the public sector. However, this approach would also have its limitations.

In recent years some institutions and scholars have begun to focus on particular features that capture some significant aspects of the quality of public sectors. The IMF, for example, has been focussing on types of statistics that countries could produce and make available to the public. Some of these data relate to the public sector. The assumption would be that countries that are willing and able to generate good public sector data and make them available on a timely basis to the public have a higher quality public sector.[34] The Fund has also started to assess the *transparency* of fiscal policy and fiscal institutions in the light of a set of general principles of fiscal transparency. The assumption is that lack of transparency is an indication of lower quality of the public sector and that this lack of transparency promotes inefficiency, poor policies, and various problems of governance. In time, transparency reports could be available for most countries. If they were comparable and com-

prehensive, these reports would serve as proxies for informal assessments of the quality of the public sector.

In recent years, a lot of attention has been paid to issues of governance and corruption in public institutions. These issues also bear on the quality of the public sector. It has been recognized that lack of transparency in the way institutions operate and lack of controls promote inefficiency and corruption. There is now a large literature on transparency and on corruption. Australia and New Zealand have once again led the way in promoting techniques aimed at increasing transparency. In the new "architecture" of the international financial system, it is recommended that countries should become more transparent in their policy-making. More transparency would probably mean less corruption and generally a higher-quality and better-performing public sector. However, much needs to be learned about these relationships. It must also be recognized that tests of transparency may be either superficial or deep. Superficial tests would not be very useful for understanding what is going on in a public sector, while deep tests might require a large amount of highly specialized and costly resources.

Corruption is certainly a sign of a low-quality public sector. Various groups have been preparing corruption indexes for a large number of countries (see Tanzi, 1998b). These indexes claim to measure "perceptions" of corruption. It is likely that to some extent, and assuming that the perceptions reflect reality, these indexes of corruption can also be taken as proxies for the quality of the public sector. However, caution is needed because, quite apart from the quality of the corruption indexes, a country could have a totally honest bureaucracy or political leadership but very inefficient policies and institutions. Although important, corruption is only one aspect of poor public sector quality.

Some authors, such as Rauch and Evans (forthcoming), have provided estimates of bureaucratic performance for many developing countries, thus providing measures of yet another variable that has a bearing on the quality of the public sector. Keefer and Knack (1997), for their part, after defining institutional quality as "objective evaluations ... of the institutions that protect property and contractual rights", attempt a measurement based

---

[34] The data themselves might tell more about the quality of public policy than the quality of the public sector.

on several indicators, but it is not clear to what extent these measure the quality of public sector institutions.

There are also other variables that may provide information on the quality of the public sector. Some relate to its efficiency, some to the policies followed. For example, the relationship between spending in a given category –say health and education– and the outcome of that spending –such as lives saved, successful operations, reduction in the incidence of certain diseases, educational achievements– would be an indication of efficiency (see Gupta, Honjo and Verhoeven, 1997). On the other hand, measurements of fiscal sustainability might be indications of poor policy rather than of poor quality of the public sector.

VI

Conclusions

The quality of the public sector can only be assessed against the background of the role of the State. If the public sector allows the State to promote its goals in an efficient and successful way, it can be argued that the public sector is of high quality. However, the goals must be realistic because even a very efficient public sector will be unable to fulfill unrealistic goals. Thus, in general the quality of the public sector cannot be measured by the quality of the policy outcome, although the two are obviously closely linked, especially over the long run.

In general, an efficient public sector should be able to achieve the State's objectives with the minimum degree of distortion of the market, with the lowest burden of taxation on taxpayers, with the smallest number of public employees, and with the lowest absorption of economic resources by the public sector. The public sector must be transparent in its processes and in its outcome. Corruption should have no part in the decisions made by bureaucrats and political leaders, while the resources in the hands of the public sector should be used in a way that maximizes their social rate of return. Such a public sector would be clearly "market augmenting", to use a now fashionable expression, and it would pay particular attention to the protection of property rights and the enforcement of contracts.

The quality of the public sector is also important in the pursuit of equity, which is now seen as one of the fundamental goals of the State. A public sector which facilitates the pursuit of equity in all its aspects must, *ceteris paribus*, be deemed of higher quality than one that does not do so. However, as argued earlier, it is difficult to determine the optimal role of the State in this area, and it is easy to see how some policies that redistribute income can have disincentive effects. These effects are more likely to arise when the public sector is of low quality and thus makes it easier for policies to be distorted during the implementation stage (see, for example, Alesina, Danninger and Rostagno, 1999, and Schansberg, 1996). A high quality public sector should make possible the pursuit of equity at the lowest cost in terms of efficiency.[35]

Finally, the discussion in this paper should have made it clear that the so-called "first generation reforms" which have been so popular in many countries in the 1980s and 1990s do not necessarily improve the quality of the public sector, even though they may improve the quality of public policy. The pursuit of "first generation reforms" has, in fact, highlighted the need to improve the quality of the public sector, and for this to occur, "second generation reforms" are necessary.

(Original: English)

---

[35] An important role in this area has been assigned to the State by Sen (1999), who identifies development with freedom and who assigns to the State the function of providing access to education and health for all. A high-quality public sector would facilitate the achievement of Sen's objective.

Bibliography

Alesina, A., S. Danninger and M. Rostagno (1999): *Redistribution Through Public Employment: The Case of Italy*, IMF working paper, No. 177, Washington, D.C., International Monetary Fund (IMF), December.

Blinder, A. (1997): Is government too political?, *Foreign Affairs*, vol. 76, N° 6, New York, Council on Foreign Relations, Inc.

Braibant, G. (1996): Utilité et difficultés de la codification, *Droits*, N° 24, Paris, Presses Universitaires de France, December.

Charap, J. and C. Harm (forthcoming): Institutionalized corruption and the kleptocratic State, in C. Menard (ed.), *Institutions, Contracts and Organizations: Perspectives from New Institutional Economics*, Northampton, Massachusetts, Edward Elgar.

Davis, K. and M. J. Trebilcock (1999): *What role do legal institutions play in development?*, paper presented at the Conference on Second Generation Reforms, Washington, D.C., IMF, 8-9 November.

De Soto, H. (1989): *The Other Path,* New York, Harper and Row.

Dethier, J. J. and T. Shapiro (1998): Constitutional rights and the reform of social entitlements, in L. Bokros and J. J. Dethier (eds.), *Public Finance Reform During the Transition: The Experience of Hungary*, Washington, D.C., World Bank.

Dixit, A. K. (1996): *The Making of Economic Policy*, Cambridge, Massachusetts, MIT Press.

Ferro, P., S. Lo Faso and G. Salvemini (1999): L´azione della pubblica amministrazione per la competitività internazionale in presenza di vincoli di finanza pubblica, in M. Bordignon and D. Da Empoli (eds.), *Concorrenza fiscale in un'economia internazionale integrata*, Rome**,** Franco Angeli.

FIEL (Foundation for Latin American Economic Research) (1988): *Los costos del Estado regulador*, Buenos Aires.

Fiorentini, G. and S. Zamagni (eds.) (1999): *The Economics of Corruption and Illegal Markets*, Northampton, Massachusetts, Edward Elgar.

Forte, F. (ed.) (1998): Le regole della costituzione fiscale, *Politeia*, vol. 14, N° 49, Milan, Politeia.

Guigou, E. (1999): Justice: Du Ministère des Affaires au Ministère du Droit, *Le Monde*, Paris, 1 September.

Gupta, S., K. Honjo and M. Verhoeven (1997): *The Efficiency of Government Expenditure: Experiences in Africa*, IMF working paper, N° 153, Washington, D.C., IMF.

Guy, S. (1996): Une utopie: la codification, *Revue française de droit constitutionnel*, N° 26, Paris, Presses Universitaires de France.

Highfield, R. (1999): Tax administration: Understanding and using the cost of collection ratio, Washington, D.C.**,** November, *mimeo.*

Jacobs, S. H. (1999): *The second generation of regulatory reforms*, paper presented at the Conference on Second Generation Reforms, Washington, D.C., IMF, 8-9 November.

Keefer, P. and S. Knack (1997): Why don't poor countries catch up? A cross-national test of an institutional expla-

nation, *Economic Inquiry*, vol. XXXV, Huntington Beach, California, Western Economic Association, July.

Kopits, G. and S. Symansky (1998): *Fiscal Policy Rules*, Occasional Paper N°162, Washington, D.C., IMF.

Laffont, J. J. and J. Tirole (1993): *A Theory of Incentives in Procurement and Regulation*, Cambridge, Massachusetts, MIT Press.

Landes, D. S. (1999): *The Wealth and Poverty of Nations: Why Some are So Rich and Some So Poor*, New York, W.W. Norton & Company.

La Porta, R., F. Lopez-de-Silanes, A. Shleifer and R. Vishny (1998): *The quality of government*, Cambridge, Massachusetts, August, *mimeo*.

Martino, A. (1989): Fisco e costituzione, in V. Uckmar (ed.), *Esperienze straniere e prospettive per l´ordinamento tributario italiano*, Padua, Italy, CEDAM.

Mattarella, B. G. (1994): La codification du droit: réflexions sur l´expérience française contemporaine, *Revue française de droit administrative*, vol. 10, N° 4, Paris, Sirey, July-August.

Musgrave, R. A. (1959): *The Theory of Public Finance*, New York, McGraw-Hill.

North, D. C. and R. Thomas (1973): *The Rise of the Western World: A New Economic History*, Cambridge, U.K., Cambridge University Press.

North, D. C. and B. Weingast (1994): Constitutions and commitment: The evaluation of institutions governing public choice in seventeenth century England, in T. Persson and G. Tabellini (eds.), *Monetary and Fiscal Policy*, vol. 1, Cambridge, Massachusetts, MIT Press.

Nozick, R. (1974): *Anarchy, State and Utopia*, New York, Basic Books.

OECD (Organization for Economic Cooperation and Development) (1999): *OECD Economic Outlook*, Paris, December.

Poterba, J. M. and J. Von Hagen (eds.) (1999): *Fiscal Institutions and Fiscal Performance*, Chicago, Illinois, National Bureau of Economic Research, Inc. (NBER)/University of Chicago Press.

Potter, B. and J. Diamond (1999): *Guidelines for Public Expenditure Management*, Washington, D.C., IMF.

Rauch, J. E. and P. B. Evans (forthcoming): Bureaucratic structure and bureaucratic performance in less developed countries, *Journal of Public Economics*, Amsterdam, The Netherlands, Elsevier Science B.V.

Reviglio, F. (1999): *Budget transparency for public expenditure control: The case of Italy*, Washington, D.C., December, *mimeo*.

Rodotá, S. (1995): Le libertá e i diritti, in R. Romanelli (ed.), *Storia dello Stato italiano*, Rome, Donzelli Editore.

Schansberg, D. E. (1996): *How Poor Government Policy Harms the Poor*, Boulder, Colorado, Westview Press.

Sen, A. (1999): *Development as Freedom*, New York, Alfred A. Knopf.

Schick, A. (1998): Why most developing countries should not try New Zealand's reforms, *The World Bank Research Observer*, vol. 13, N° 1, Washington, D.C., World Bank, February.

Skaperdas, S. and C. Syropoulos (1995): Gangs as primitive states, in G. Fiorentini and S. Peltzman (eds.), *The Economics of Organised Crime*, Cambridge, U.K., Cambridge University Press.

Smith, D. and E. Richardson (1999): The readability of Australia's taxation laws and supplementary material: An empirical investigation, *Fiscal Studies*, vol. 20, N° 3, London, Institute for Fiscal Studies, September.

Tabellini, G. and T. Persson (forthcoming): *Fiscal Policy in Representative Democracies*, Cambridge, Massachusetts, MIT Press.

Tan, L. M. and G. Tower (1992): Readability of tax laws: An empirical study in New Zealand, *Australian Tax Forum*, vol. 9, N° 3, Clayton, Victoria, Australia, Monash University.

Tanzi, V. (1995): Corruption: arm's-length relationships and markets, in G. Fiorentini and S. Peltzman (eds.), *The Economics of Organised Crime*, Cambridge, U.K., Cambridge University Press.

————(1998a): *Essential Fiscal Institutions in Selected Economies in Transition*, Discussion paper, No. 53, Washington, D.C., Collegium Budapest, Institute for Advanced Study, November.

————(1998b): Corruption Around the World, *Staff Papers*, vol. 45, No. 4, Washington, D.C., IMF, December.

————(forthcoming): Rationalizing the government budget: Or why fiscal policy is so difficult, in A. O. Krueger, *Economic Policy Reform: The Second Stage*, Chicago, Illinois, University of Chicago Press.

Tanzi, V. and T. Prakash (forthcoming): The cost of government and the use of public assets, Washington, D.C., *mimeo.*

Tanzi, V. and H. Zee (1997): Fiscal Policy and Long-Run Growth, *Staff Papers*, vol. 44, N° 2, Washington, D.C., IMF, June.

United States Congress, Senate Committee on the Budget (1984): President's private sector survey on cost control (Grace Commission): Hearing before the Committee on the Budget, *United States Senate, Ninety-Eighth Congress*, First session, Washington, D.C., 7 November.

Van Creveld, M. (1999): *The Rise and Decline of the State*, Cambridge, U.K., Cambridge University Press.

World Bank (1997): The State in a changing world, *World Development Report,* Washington, D.C.

The effects of growth
and economic reform
on income distribution
in Latin America

Samuel A. Morley

*International Food Policy
Research Institute*
s.morley@cgiar.org

The purpose of this article is to investigate the determinants
of the distribution of income in Latin America, focusing in
particular on two questions: one is the relationship between
the distribution and income, while the other is the impact of
the package of structural economic reforms that have been
adopted in Latin America in recent years. Two main conclu-
sions are drawn from the econometric evidence. There ap-
pears to be a robust and significant relationship between the
distribution and income. It has the inverted U-shape that
Kuznets predicted, but this relationship has been shifting in
a regressive direction over time. Growth is now a good deal
less progressive than it used to be. In the aggregate that means
that further growth in Latin America is unlikely to improve
the distribution much, if at all, so supplementary measures
will have to be taken. Among those suggested by the regres-
sions are the maintenance of low inflation rates and invest-
ment in education. Generally speaking, the structural reforms
appear to have a regressive effect on distribution, but that
effect is small and not very robust statistically. Reforms in
different areas have differing effects on equity. Trade re-
form is regressive in all of our specifications, but it is insig-
nificant in all but the nationwide sample. Tax reform is un-
ambiguously regressive, and opening up the capital account
is unambiguously progressive. The results for trade and tax
reform and capital account liberalization are the most ro-
bust and significant. For the other two reforms –privatization
and financial reform– the available data were not good
enough to give a clear answer.

I

Introduction

There have been many previous efforts to econometrically estimate models of the relationship between the level or growth rate of income and its distribution. Most have been estimations of the Kuznets relationship, using cross-country distribution and income data[1] (Ahluwalia, 1976; Anand and Kanbur, 1993; Bruno, Ravallion and Squire, 1996; Clarke, 1995; Deininger and Squire, 1996; De Janvry and Sadoulet (forthcoming); Fields,1994; Ravallion and Chen, 1997). All but the last of these studies use a world-wide sample of countries. The difficulty with that approach, as Fields points out, is that since Latin America is a middle-income region and has the highest inequality in the world, one can get an apparent inverse U-shaped Kuznets curve simply because of the choice of the sample. Fields found that if he put in a dummy for the Latin American observations, the supposed relationship between income and inequality disappeared. Deininger and Squire (1996) found exactly the same thing. Bruno, Ravallion and Squire (1996), using data from 63 surveys covering 44 countries, tested the Kuznets hypothesis for both levels and changes over time. In no case could they find evidence of an inverted U shape, and in no case was the relationship between the distribution and income significantly different from zero. Ravallion and Chen (1997) regressed changes in the Gini against changes in mean real consumption over 64 periods in 67 countries and found a negative and significant relationship between the two in the full sample. However, when they excluded the observations from Eastern Europe and Central Asia from the sample, the relationship disappeared.

In this study, we will econometrically estimate a distribution function using a pooled cross-section time series of observations from 16 countries in Latin America. Clearly there are serious econometric difficulties in attempting to use pooled cross section time series evidence to capture a time series relationship for a single country. Essentially one is assuming that the relationship between distribution and income in today's high income countries is similar to what the less developed countries can expect when they reach the income level of more advanced countries. In other words unknown country-specific effects do not affect the relationship between income and the distribution. Obviously there are many factors that vary across countries that might be expected to have an impact on the distribution or on its responsiveness to changes in income, and these need to be included in the model. In addition we use a fixed-effects model with country-specific constants to capture any unknown country-specific factors affecting the distribution relationship.

The data base used in our econometric estimation is much larger than those typically used in Kuznets curve estimation in the Latin American region, and it owes a large debt to the pioneering work in data collection by ECLAC, the IDB and the World Bank. The sample consists of 261 observations in 16 countries from 1960 to 1997. No country was included in the survey unless it had at least four separate observations.

Figure 1 gives a simple plot of the Ginis and income levels for all the observations in our sample. There are obviously serious problems of comparability between many of the surveys and therefore also between the distribution estimates that come from those surveys. Some of the surveys are based on the distribution of family income, others on family income per capita. Most of the surveys are based on income, but there are some that use expenditure data instead because it is measured with so much less error and is a better indicator of permanent income.

[1] Simon Kuznets (1955) found an inverted "U" shaped relationship between income and income distribution, using historical data for England and the United States, and hypothesized that it was mainly explainable by the movement of population from the low-income rural economy to the higher-income urban economy.

FIGURE 1

Gini Coefficients
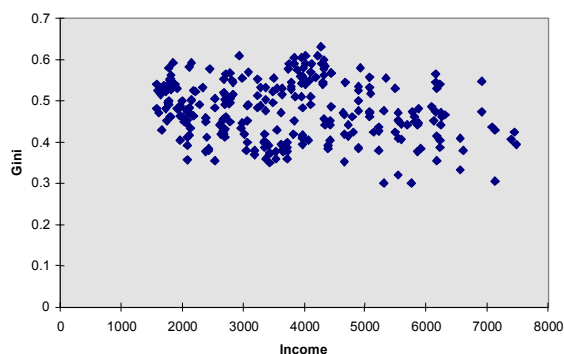
Another important distinction is between urban and national surveys. Both Fields (1994) and Deininger and Squire (1996) have argued strongly that one should only use national surveys in any analysis of the distribution of income. But in many countries (Bolivia, Ecuador and Paraguay, for example) until very recently urban surveys were all that were available, and in another two (Argentina and Uruguay) they are still the only surveys available. One therefore has the choice of excluding these countries from any analytical work or of attempting to control for systematic differences between urban and national surveys by the use of dummy variables. We have chosen the latter strategy. Not only does this expand substantially the size of our sample but it also permits us to see whether there are any systematic differences of reaction in income or income growth between the urban sector and the national surveys. To check on the sensitivity of the estimates to this aggregation, we will display results for the urban and national samples separately.

II

The model for the determinants

of income distribution

We can write the general regression model for the distribution as follows:

$$Gini_{it} = A_i + B_i Y_{it} + C_i 1/Y_{it} + DZ_{it} + ER_{it} + FS_i + GT_t Y_{it} + HT + \text{error} \quad [1]$$

where i denotes countries and t denotes year.

The Gini coefficient will be our measure of income distribution. A is a regression constant which may vary across countries but, in our model, not across years. Y is income. Z is a vector of variables such as inflation, land distribution and education which we hypothesize may have an effect on the distribution. R is an index of economic reform and S is a vector of dummies which reflect various sample characteristics such as whether the sample is urban and whether it is based on household income or income per capita or on expenditure rather than income. T is a trend variable.

The first two income terms represent the Kuznets relationship. Our hypothesis is that this relationship has an inverted U shape in which inequality rises with income growth at low levels of income, but falls with growth above a given income level. This hypothesis trans-

lates into the expectation that both B and C will be negative and significant.

With regard to trend, we introduced two separate trend terms in equation [1] to test the hypothesis that there are significant shifts in the K-curve over time. If H is negative, the Kuznets curve shifts down over time (i.e., the distribution becomes more progressive). But we also hypothesize that there may be a systematic change in the relationship between income and income distribution. Our hypothesis is that this change is regressive: i.e., that G is positive. Perhaps for technological reasons, growth now is more regressive than it used to be. If G is positive, the slope of the K-curve changes over time. To the left of the inflection point, where the curve itself is upward-sloping, the slope gradually gets steeper. To the right of the inflection point, where the slope itself is negative, the trend makes the slope gradually flatter. Furthermore, the interaction term makes the inflection point itself shift gradually to the right over time, extending the range over which growth is regressive. Thus the trend terms tell two opposing stories. The trend term on the intercept is progressive and shifts the K-curve down, but the interaction term is regressive.

We have argued that the distribution of assets should have an effect on the distribution of income. We will include two measures of asset distribution here, one for the distribution of land, the other for the distribution of education. The first is a dummy variable which equals one for those countries with an unequal distribution of land.[2] We have used two variants of this variable, one which takes a value of one for all observations in the countries with unequal land distribution, while the other has a value of one for the national but not the urban observations. This means, for example, that in the second variant all the Paraguay observations have a zero for this variable even though Paraguay has a very unequal distribution of land. This is because all the observations for that country are urban.

With respect to education, our hypothesis is that the relative supply of more and less educated labour will have a significant effect on relative wages and the distribution of income. We have used a number of separate indicators of the supply of educated and less educated labour. NOSCHOOL is the percentage of the adult population with no schooling, PRIMARY is the percentage with no more than primary schooling, and HIGH is the percentage with more than secondary education. We also attempted to use measures of the variance of education levels across the adult population, but the problem with

this variable is that educational improvements which increase the supply of high school and university graduates will in many case increase the measured variance instead of reducing it.

Inflation is another important variable which could be expected to have a powerful effect on the distribution. Labour markets react fairly quickly to moderate but not to extreme rates of inflation. When the inflation rate is low, nominal wages adjust and there may be little change in wage structure due to price changes. This does not happen in episodes of hyperinflation, when wage adjustments (particularly in the minimum wage) may lag behind the rate of inflation. Furthermore, even if nominal wages are raised by the full amount of inflation, it is still true that the average level of the real wage over the adjustment period is a negative function of inflation. This factor is not particularly important when the rate of inflation is low, but it becomes exceedingly important when the rate is high, which is one of the reasons why the interval between adjustments gets shorter in periods of hyperinflation. The implication of all of this is that high rates of inflation may have an impact on income distribution, but the relationship is highly non-linear. To test for this we have included an inflation dummy which takes a value of one for any year in which the annual inflation exceeds 1,000% but is zero elsewhere.

III

The structural economic reforms

Given our interest in the impact of the economic reforms, it was essential to have some sort of quantifiable index with which to compare the extent of reforms between countries or the progress of reforms over time in a single country. Our attempt to do this is described more fully in Morley, Machado and Pettinato (1999) and is an extension of the work initiated by Eduardo Lora at the IDB (see Lora, 1998).

Our index is a simple average of reform indexes in five areas: trade, finance, tax, privatization and capital account. In each area we tried to choose indicators such as tariff or tax rates which reflect government policy, rather than proxies for those policies such as openness

to trade or the government deficit. Each index is normalized to come between zero and 100, with the latter being assigned to the country and year in which the sector was the most reformed or free from distortion or government intervention, and zero, to the country and year with the greatest degree of intervention.[3] We do not mean to imply by this procedure that a high value for an index is necessarily better than a low one, but only that the sector is closer to a pure market solution without government intervention.

The trade reform index is the average of two sub-components: the average level and the dispersion of tar-

---

[2] We were forced to use a dummy for this variable rather than a numerical estimate because estimates are not available for some countries, while in others they appear to be based on different measures.

[3] Formally, each sub-index is defined as $I_{it} = (IR_{it} - Min)/(Max-Min)$, where IR is the raw value of the index in country I, year t, and Max and Min are the maximum and minimum values of the raw index for all countries over the period 1970-1995.

iffs. We were unable to obtain a satisfactory measure of non-tariff restrictions, and this represents a weakness of the index because in some cases such as Brazil such restrictions significantly affected the timing of trade reform. Domestic financial reform is the average of three sub-indexes: the control of bank borrowing and lending rates and the reserves-to-deposits ratio. Tax reform has four subcomponents: the maximum marginal tax rate on corporate and personal incomes, the value added tax (VAT) rate, and the efficiency of the VAT. Our index for privatization is equal to 100 minus the percentage of value added in State-owned enterprises in the non-agricultural GDP. Capital account reform is the average of four sub-components reflecting the extent of government control of foreign investment, limits on repatriation of profits and interest, controls on external borrowing, and capital outflows. Unlike the other indexes, this one is based on a subjective interpretation of the descriptions in the IMF's annual Balance of Payments Arrangements publication.

In this definition of reform we make a clear distinction between what we call structural reform and macroeconomic policy reforms such as government deficit reduction, inflation control and exchange rate management, which could be called stabilization reforms. In many countries the two types of reforms were adopted together to deal with balance of payments or hyperinflation crises. Here, we are explicitly studying only the structural reforms. The impact of stabilization reforms will be captured through their effect on inflation and the level of income.

To some extent, the choice of what to include in our measurement of structural reform was arbitrary. Argu-

ably the two most important excluded reforms were those in the labour market and financial market regulation. We excluded the former because in Lora's study the labour reform index by countries changed very little between 1985 and 1995, and because his measure would not have captured the important changes that occurred in the Southern Cone countries in the 1970s. We excluded financial market regulation because of the lack of an adequate measure for it. Elimination of price subsidies is another reform which was not included, but which probably had important effects in some countries.

Figure 2 is a graphic presentation of the region-wide averages for each of our reform indexes. It gives a quick picture of what has been reformed most and when the process occurred. Figure 3 shows the progression of reforms in each of the countries in our regression sample. Note that these indexes are non-weighted, simple averages of the values of the indexes for each of the countries.

FIGURE 2

REFORM INDEXES, 1970-1995

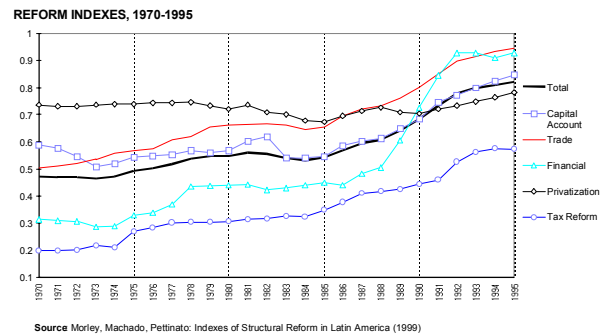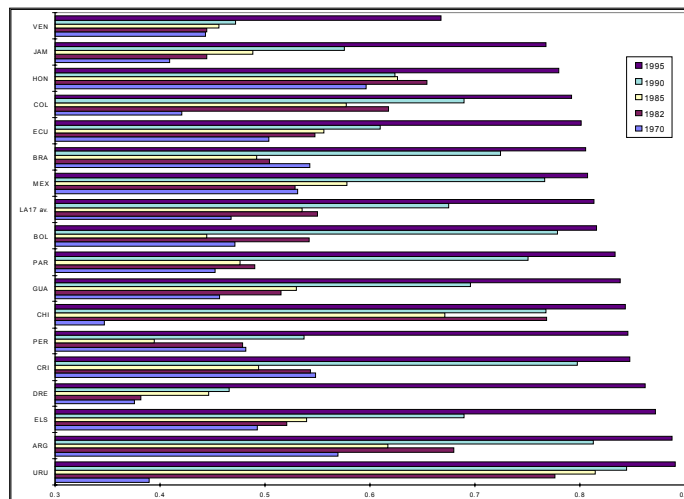**Source** Morley, Machado, Pettinato: Indexes of Structural Reform in Latin America (1999)

FIGURE 3

Latin America and the Caribbean (17 countries): Reform indexes by country, 1970-1995

Source: Morley, Machado and Pettinato (1999).

IV

The impact of the reforms

What effect might these reforms be expected to have on equity? When one analyses the effects of the reform package as a whole, the broad thrust is to remove any sort of insulation from the market's determination of the allocation of resources. Trade reforms remove tariff protection from domestic production, while financial reforms and privatization reduce government influence over the allocation of resources. Balance of payment reforms integrate foreign and domestic capital markets and reduce the capacity of government to control capital movements. Similarly, labour market reform increases labour flexibility or, to put it another way, reduces labour's ability to defend itself against either market-driven fluctuations in demand, or alternatively wage reductions. Altogether, this adds up to quite a big leap into a new world whose transition costs are justified by expected increases in efficiency, income and growth. Whether or not the reforms have had these positive expected benefits, it is quite clear that very little attention has been paid to the winners and losers in the process or to its distributional implications.

There have been several recent attempts to examine the relationship between reforms and income distribution. Albert Berry recently published a set of case studies on nine countries of the region (see Berry, ed., 1998) in which he finds that in every case but Costa Rica, and possibly Colombia, the period of reforms coincided with a very sharp increase in inequality. The expectation that trade reform would lead to a narrowing in wage differentials has not been borne out in practice, indicating either that Latin America's comparative advantage does not lie in unskilled-labour-intensive products or that the opening has forced a shift in technology in favour of more capital- and skill-intensive production. The data for his study end fairly early in the decade of the reforms, so it is unclear whether the rises in inequality that he observed are part of a short-run adjustment or an unfavourable long-run trend.

Victor Bulmer-Thomas's study (Bulmer-Thomas, 1996) comes to equally pessimistic conclusions, though for somewhat different reasons. A theoretical analysis of each of the different reforms leads the contributors to his volume to the conclusion that, taking all the reforms together, real wages will fall, unemployment will go up, real interest rates will rise, there will be a rise in

informalization and there will be an increase in the concentration of wealth, all of which are regressive. But the evidence to decide whether or not these predictions are reasonable was weak, since his study only extended up to 1992. Basically, his conclusion was that the problem with the new economic model was not so much connected with equity as with whether or not the new dependence on markets and the private sector would be capable of producing adequate, steady and sustainable growth rates of per capita income.

Londoño and Szekely (1998) of the IDB come to quite a different conclusion. Using cross-country regressions as opposed to country case studies, they find that equity is positively related to both growth and investment. These in turn are positively related to the structural reforms of the new economic model, leading to the conclusion that the reforms are progressive. This is confirmed by a direct correlation of income shares of different quintiles of the population with indexes of the different reforms. While there was no significant relationship between income shares and most of the indexes, trade liberalization was positively related to the income share of the bottom quintile and negatively related to the share of the top quintile. In the view of these authors, unlike many other researchers, trade reform helped the poor and unskilled.

There is a growing literature on the effects of trade reform on wage inequality (Robbins, 1995 and 1996; Wood, 1994 and 1997; Edwards, 1997; Ocampo and Taylor, 1998). The general conclusion of all these studies is that wage inequality has risen in those countries which opened their internal markets to external competition. While an increase in wage inequality does not necessarily translate into an increase in inequality of total income, these results suggest caution in accepting the Hecksher-Ohlin assertion that trade should help countries with large supplies of unskilled labour. Wood (1994) argues that the experience of East Asia in the 1960s and 1970s supports the theory that greater trade openness tends to narrow the wage gap between skilled and unskilled workers in developing countries. In Latin American, however, since the mid-1980s increased openness has widened wage differentials. Wood (1997) thinks that this conflict of evidence is probably not the result of differences between East Asia and Latin America but rather

the result of differences between the 1960s and the 1980s: specifically, the entry of China into the world market and perhaps the advent of new technologies biased against unskilled workers.

Spilimbergo, Londoño and Szekely (1997) point out that what really matters is each country's factor endowments, including land, relative to the average world effective supply of each factor. They too find that trade openness is associated with higher inequality, for constant factor endowments, but the effect depends on the relative abundance of each type of factor. Inequality increases in countries that are relatively well endowed with skills, but it declines in countries which are well endowed with physical capital and land. Since, in their sample, the factor endowments for Latin America are relatively close to world averages, the effect of trade openness on inequality is modest: a rise of 10% in the Latin American countries' openness index only raises the average Gini coefficient by 0.63 of a point.

One should not ignore the demand side in considering this question. The purpose of trade reform is to switch the production of tradeables away from inefficient import substitutes to exportables in which countries have a comparative advantage. The connection with income distribution comes from the differences in factor demands between these two types of products. It is thus a question of relative factor-intensity. But there is a demand side to consider as well. The success of the old import-substitution, inward-looking development strategy depended to a large extent on a growing internal market. If there is going to be satisfactory growth under that sort of strategy, there has to be a growing middle class with growing purchasing power. Growing real wages are an integral part of that strategy. The mature capitalist economies long ago discovered that both the owners of capital and their workers could profit from a strategy in which rising wages increased both costs and profits at the same time, thanks to increases in the size of the internal market induced by rising wage payments.

The export-led growth strategy is quite a different matter. Its success depends on controlling costs, and the internal market is irrelevant. In the export model, rising real wages are a clear threat to growth. They do not have the positive indirect effect through demand that they have in the inward-looking growth strategy. Countries embarking on the outward-looking growth path are making their wage levels hostage to wage levels and labour costs in other countries. It may well be that the advantages of greater efficiency in export production compared with import substitution outweigh the disadvantages of this wage competition, so that workers are better off. But the

grounds for that presumption certainly are not immediately obvious, particularly in the large economies.

What is the likely effect of liberalizing the capital account? What this reform does is to integrate more closely the local and international capital markets, thus bringing local interest and profit rates, adjusted for risk, closer to rates in the rest of the world. Whether or not this is progressive depends on the reactions of foreign and domestic owners of capital. If foreign investors have been deterred from a country because of controls on repatriation of capital and profits, the reforms should induce a foreign capital inflow. The distributional effect of this is ambiguous. Wage/profit ratios should fall because of the rise in the capital/labour ratio, which is a progressive development, but at the same time, if capital and skilled labour are complementary, the skill differential will rise, which is regressive. A similar ambiguity results from the actions of domestic owners of capital. One of the reasons for liberalizing the capital account was to lift restrictions on capital outflows by domestic savers and investors, and if there was previously excess demand for foreign exchange under capital controls, the reforms should cause a capital outflow, with results just the reverse of those described for foreign capital inflows.

Aside from the effect of these reforms on factor supply and factor demand, removing barriers to capital movements increases the bargaining power of capital in its negotiations with both labour and the government. That is likely to be regressive, for if investors are free to move from one country to another, governments will find it far more difficult to tax capital or to pass regulations that force businesses to shoulder more of the cost of infrastructure or labour regulation. Indeed, in a world of perfect capital mobility, countries will be forced to compete in offering generous tax holidays, subsidized credits and other costly assistance as a way of attracting foreign capital. But it is not only foreign capital that is affected. The same argument is valid for domestic capital. Both government and labour will be forced to accept arrangements that are sufficiently generous to ensure that domestic entrepreneurs and holders of wealth are content to leave their money invested in their home country. In this way, opening up the capital account shifts the balance of power in favor of the holders of capital. This is one of the reasons why there has been a shift away from the taxation of corporate profits and a big reduction in the top marginal income tax rate in most Latin American countries in recent years.

Financial reforms eliminated controls on interest rates, reduced compulsory reserve requirements of banks and reduced the use of directed or subsidized credit. The

direct effect of this on income distribution is probably small, but to the extent that these reforms increased private saving and investment, they should be progressive.

The fourth component of the reform project is tax reform. Two major measures have been widely adopted in this respect. The first was the value added tax. Reformers favoured this tax because they argued that while all taxes have distorting effects on private decisions, these are less with an across-the-board value added tax than with either tariffs or high marginal income tax rates. In addition, of course, there should be less tax evasion with VAT than with an income tax based system. VAT was introduced in the 1970s in nine of the 17 countries for which we have data. In the 1980s it was adopted in all the remaining countries in the region, and in addition there was an increase in the coverage or efficiency of this tax in most countries.

A second element of tax reform was the reduction in marginal tax rates on corporate and personal income, which significantly reduced the progressiveness of income tax. Every country in the region has reduced its top marginal tax rate since 1970. Not all have gone as far as Uruguay, which eliminated personal income tax altogether, but overall the average marginal rate on personal income has fallen from around 50% in 1970 to about 25% in 1995, while the corporate rate has fallen from 37% in 1970 to 29% in 1995. Almost all these changes have taken place since 1985.

From the distributional standpoint, the effect of these changes in the tax system was to shift the burden of the tax system away from the wealthy and toward the middle and lower classes. The introduction and later expansion of the value added tax was a shift away from the taxation of income toward the taxation of consumption. Since the poor consume a greater fraction of their income than the rich, this change must have been regressive, except in certain countries which exempted basic necessities from the tax.

The changes in income tax amplified the trend toward greater regressiveness. Top marginal tax rates on personal income were lowered and the corporate tax rate was cut by over 20%. In addition to increased non-neutrality, the impact of the tax reforms should also depend on the fraction of national income being taxed. This aspect has not been included in our index of reform. While a full analysis of the incidence of all these changes is beyond the scope of this paper, it is almost certain that they were regressive, although it should be noted that if

tax reform was part of a programme of deficit reduction and inflation control, its overall impact may well have been progressive.

Another important component of the reforms made in the region was privatization. State enterprises were a key component of the old development model, which has been dramatically redesigned by the reforms we are analyzing. The impact of privatization on income distribution depends on three elements. First, whether or not the sales price of the assets of the State-owned enterprises reflects their true market value. If it is less, buyers have received a gift from taxpayers. Second, for public utilities like electricity, telephone and water companies, the impact depends on what happens to the price of the services they provide to the public. In many cases publicly-owned utilities subsidized their customers by selling below cost. Transferring that sort of company to the private sector and eliminating the subsidy could be either progressive or regressive, depending on who their customers were. One might expect that to be regressive, but a recent study of gasoline and electricity pricing in Venezuela and Peru came to the opposite conclusion, considering that those wealthy enough to have electric appliances and cars came from the top, not the bottom of the distribution (Márquez and others, 1993). In fact, most of this sort of subsidy probably benefited the middle class.

Neither of the effects we have been discussing so far will be reflected in our distribution data, because the latter are based on earnings and not on expenditure or wealth. A result of privatization which *is* reflected in the earnings data, however, is its effect on labour demand and employment. Labour productivity in the typical State-owned enterprises (SOEs) was low. For political reasons many governments seemed more interested in using these enterprises to create jobs than to provide good service at the lowest possible cost, but when the enterprises were sold, all this had to change. Privatization operations in places like Chile and Argentina were blamed for a good deal of the job destruction and rising unemployment that accompanied reform. The distributional impact of this depends on who the displaced employees were. There is no good study of this question, but judging by the labour force profile of the typical State-owned enterprise, these jobs were largely in the middle of the earnings distribution scale. Thus, privatization is likely to have mainly hurt the middle class, both because they were the main users of subsidized SOE services and also the main employees of State-owned firms.

V

Econometric results

Tables 1-4 give our best estimates of the determinants of the distribution. Table 1 shows the results for the overall average reform index, using all 262 observations, both national and urban. Table 2 separates the urban and national samples as a check on the robustness of our results to alternative aggregations. Tables 3 and 4 show the effect of each of the five different areas of reform, first for the entire sample and then for the national and urban samples considered separately.

In table 1 we show four alternative regressions, three with fixed effects and one with a common constant. The first two regressions (columns 1 and 2) use the same variables to show the difference between using cross-

section weights or pooled least squares. Cross-section weights are used in all the remaining reported results. The third regression shows the effect of adding a trend to the constant terms. The fourth gives some idea of what explains differences between the constants across countries.

Perhaps the most important result is that the general regression model fits the data well, explaining between 85% and 97% of the total variance in the Gini coefficient over time and across countries. In addition the coefficient estimates and significance appear to be robust and consistent across the alternative fixed effects regressions. Of the four specifications, those with fixed effects

TABLE 1

Combined sample results

			Fixed effects				Single intercept	
			Cross-sectional weights					
	Pooled least squares —1—		—2—		—3—		—4—	
Variable	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
Income	-0.0001	-7.0504	-0.0001	-7.317398	-0.0001	-7.9707	-0.000102	-5.2978
1/Income	-260.3067	-2.9407	-208.7411	-2.671	-336.4263	-4.0021	-69.31056	-1.1264
Urban	-0.0336	-5.3983	-0.0333	-6.9739	-0.0320	-6.7043	-0.0278	-3.7291
ECLAC	-0.0522	-8.3100	-0.0510	-9.2131	-0.0530	-9.4478	-0.0568	-6.8944
Expenditure	-0.0874	-4.6603	0.0875	-2.9325	-0.0838	-2.8620	-0.1109	-9.4870
Inflation	0.0114	1.2774	0.0112	1.3838	0.0138	1.7027	0.0510	3.8580
Household	-0.0089	-1.5645	-0.0118	-2.3858	-0.0154	-3.0176	-0.0144	-1.8140
High	-0.0065	-2.8047	-0.0082	-3.2362	-0.0039	-1.3811	-0.0080	-5.4828
Primary	0.2482	3.1068	0.1965	2.7995	0.2311	3.2581	0.0167	0.3300
Trend*Income	0.0000	6.0880	0.0000	6.2408	0.0000	7.1460	0.0000	4.6505
Reform	0.0261	1.2573	0.0303	1.6333	0.0633	2.8999	-0.0095	-0.3703
Trend					-0.0030	-3.1760	-0.0015	-1.4856
Landdist							0.0364	4.5922
Constant							0.7387	8.2049
R-squared	0.8620		0.9756		0.9755		0.9341	
Adjusted-R-squared	0.8468		0.9729		0.9726		0.9306	
Standard error of regression	0.0274		0.0273		0.0268		0.0467	
Log likelihood	1789.3480		1791.4680		1792.1660		1577.077	
Durbin-Watson	1.5275		1.6061		1.6065		0.6725	
Mean dependent variable	0.4758		0.5256		0.5214		0.5387	
Standard deviation of dependent variable	0.0701		0.1659		0.1623		0.1772	
Sum of squared residuals	0.1768		0.1751		0.1687		0.5402	
F-statistic	146.8103		940.4532		845.5984		270.2431	
df (degree of freedom)	261		261		261		261	

and cross-section weights (nos. 2 and 3) have the best fit, and we will refer to their coefficients in the discussion that follows. A Wald test on the sum of squared residuals of the fixed and common constant regressions (compare regressions 3 and 4) decisively rejects the hypothesis that there is a common Kuznets curve across the different countries. Unspecified country-specific factors significantly affect the level of inequality for a given level of income.

What do our results tell us about the existence and/or shape of the Kuznets curve?

i) *The coefficients of income and the inverse of income.* Both are negative, and both are highly significant in all the fixed effects specifications (table 1). This result means that one cannot reject the hypothesis that there is a stable and identifiable relationship between income and inequality in the region (a Kuznets curve), and that this relationship has the same inverted U shape that Kuznets found for Britain and the United States. Inequality rises at low levels of income, but at some income level there is an inflection point after which inequality begins to decline as income increases. This is an important result. But it leaves open the question of whether there really is a single Kuznets curve for all the countries. To test that, we reran the model, permitting coefficients B and C of equation [1] to differ across countries (results not shown). Doing that significantly improves the fit of the regression, which means that there are differences between countries in the way that inequality reacts to changes in income. Even in this estimation, however, 12 out of 16 of the B_i coefficients and 10 out of 16 of the C_i coefficients are negative, and only two countries (Bolivia and Paraguay) have a positive and significant B_i. It thus seems fair to conclude that while there are significant differences across countries, the average values of B and C shown in table 1 are quite representative of the typical or average relationship between income and inequality in the region. It also leaves open the question of whether there is a single curve representing both the urban and national samples. We will look more closely at that question when we discuss the urban and national results below.

ii) *Education.* Education is an important qualifier of our discussion of the Kuznets curve. We included three education variables in the model: the percentage of the adult population with no schooling (NOSCHOOL), the percentage with no more than primary schooling (PRIMARY), and the percentage with university education (HIGH). High percentages of poorly educated workers have quite a large and regressive effect on income distribution. In Argentina, for example, the share of adults

with no more than primary school education has fallen from 81% in 1974 to 64% in 1996.[4] According to regression 2 in table 1 that improvement alone should have lowered the Gini coefficient by about three percentage points (197*.17). At the same time the negative coefficient for HIGH tells us that increasing the share of university graduates in the adult population is progressive (shifts the Kuznets curve down). Note that the absolute size of the effect of expanding the university graduate proportion is much smaller than the effect of reducing the share of the poorly educated, suggesting that one gets a bigger distributional impact by spending money to reduce the size of the "primary schooling or less" group than by expanding the coverage of high school education and universities. That is exactly the same message that one gets from the comparison of education profiles between East Asia and Latin America (see Morley (forthcoming), chap. 3).

iii) *Urban vs. national.* In table 1 the urban dummy is negative, significant and robust. On average the Gini coefficient can be expected to be about 3 percentage points lower in the urban surveys than in the national ones. If one compares the Kuznets curves implied by the regressions reported in table 1 with those for the urban and national samples reported separately in table 2, one sees that while the general form is the same (all three have an inverted U shape), the coefficients and significance for the various income and trend terms are different enough to make it worth displaying the curves from all three estimations (figure 4). As expected, the urban curve in the figure is lower than the national curve over the relevant income range. It also peaks earlier and is slightly flatter than the national curve. That is consistent with the rationale that Kuznets described for the relationship. According to him, the distribution changes with development because people move from the low-income rural sector to the higher-income urban sector. At low levels of aggregate income, the urban sector is small, so this change in structure increases inequality (i.e., the national curve is upward sloping). After a certain point, however, when the urban sector has got big enough, continued rural-urban migration reduces inequality because it shrinks the size of the group that is poor (i.e., after the inflection point the curve turns down). But *within* the urban sector there is much less reason to expect income growth to have these effects. The urban sector is more homogeneous, so the gains from income growth ought to spread more evenly through it. This implies that the

---

[4] This finding is consistent with results from other cross-section studies. See the discussion in Bruno, Ravallion and Squire (1996).

urban Kuznets curve should be quite flat –flatter than the national curve- which indeed it is.

iv) *Inflation.* As we assumed in our hypothesis, episodes of high inflation (more than 1,000% per year in our model) are regressive. On average these episodes add about one percentage point to the Gini coefficient. This effect is robust to alternative estimation methods, but it is not significant in any regressions with an interaction term between the trend and income.

v) *Sample characteristics.* All of the sample characteristics had a significant effect on the level of the Kuznets curve. Distribution measures based on expenditure rather than income had Gini's about nine percentage points lower than those based on income. They also had significantly flatter slopes, which we found by putting an interaction term (not shown) in the regression. This is what one would expect. To the extent that variations in income are temporary and not permanent, expenditures should reflect the latter more than the former. That will tend to imply that expenditures tend to change less than measured income across different levels of income, sig-

FIGURE 4
Kuznets curves
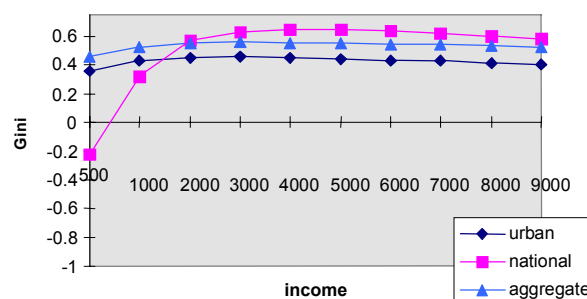
nifying that the distribution of expenditures is more equal than the distribution of measured income. Another sample characteristic is whether or not the distribution is based on family income or per capita family income. Surveys based on family income have Gini's which are about one percentage point below those based on per capita family income, and this difference is significant. It reflects the fact that poor families tend to systematically have more family members. Finally, the ECLAC distributions are

TABLE 2

Results for aggregate reform index

	Urban sample				National sample	
Variable	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
	—1—		—2—		—3—	
Constant	0.638532	7.192282	0.63856	7.220107	0.909964	14.38141
Income	-0.0000655	-3.460007	-0.0000649	-3.458332	-0.0000788	-6.047558
1/Income	-77.0826	-0.834087	-78.90591	-0.858713	-557.4071	-6.043466
ECLAC	-0.058326	-6.515791	-0.058633	-6.613158		
Inflation	0.050199	3.27413	0.049649	3.271231	0.02444	1.198074
Reform	0.012198	0.320243			0.136146	3.457715
Noschool	0.001908	3.415206	0.001871	3.437436	0.002566	5.538963
Trend	-0.003353	-1.493729	-0.003089	-1.485051	-0.003769	-2.996997
Trend*Income	0.00000136	2.952585	0.00000135	2.946409	0.00000124	3.677888
High					-0.010499	-7.236491
Per household					-0.063041	-6.952319
Expenditure					-0.076538	-5.187983
R-squared		0.499424		0.498985		0.705668
Adjusted R-squared		0.465197		0.469264		0.682121
Standard error						
of regression		0.040767		0.040611		0.041473
Sum of squared residues		0.194446		0.194616		0.215004
Log likelihood		229.0684		229.0132		245.6077
Durbin-Watson statistic		0.721265		0.716844		1.062467
Mean dependent variable		0.449939		0.449939		0.49977
Standard deviation of						
dependent variable		0.055745		0.055745		0.073559
F-statistic		14.59135		16.78887		29.96904
Prob(F-statistic)		0		0		0
df (degree of freedom)		125		125		136

systematically 5-6 percentage points more equal than the others because of their treatment of home consumption and other sources of under-reporting.

vi) *Reforms*. Here, we look first at the effects of the average reform index, deferring consideration of each of the sub-indices for later. As the reader can see from tables 1 to 3, overall the reforms have a regressive effect on the distribution. The coefficient is positive in all three samples, and is significant in some of them. The effect, however, is relatively small. According to the reform coefficient for regressions 2 and 3 in table 1, increasing the average reform index by 10% can be expected to increase the Gini coefficient by between 1/3 and 2/3 of a percentage point. While this effect is not large, the sign does confirm the assertions of Berry (ed.) (1998) and Bulmer-Thomas (ed.) (1996). They used historical evidence up to about 1994 for a smaller cross-section of countries to show that inequality had widened after the imposition of the neoliberal reform package. The evidence here comes from a much larger cross-section of countries and a far longer time period, but it points to the same conclusion.

Two notes of caution are called for here: first, it should be remembered that when we talk here about the effect of these structural reforms, we mean their direct impact, and not whatever effect the reforms may have had through inflation or income. If the reforms increased the growth rate or led to lower inflation, as they seem to have done in some countries, the positive effect of those two factors may outweigh the direct regressive effect on inequality of the reforms themselves.[5] Second, as we will see below, different reforms appear to have quite dramatically different and offsetting effects on the distribution. One will get quite different conclusions if the pattern of reform differs from the across-the-board average change being considered here.

vii) *Land distribution*. In regression 4 of table 1 we re-estimated the model with a single constant and added a measure of land distribution to see whether this may be one of the reasons why country-specific intercept terms differ. As the reader can see, the land distribution variable is highly significant and positive, adding about

---

[5] In a recent paper Escaith and Morley (forthcoming) have estimated the effect of this same package of reforms on economic growth. Their results show that the average reform index did not have a significant effect on the growth rate, because the various components of the reform package had different and sometimes opposing effects on growth.

TABLE 3

Effects of subindexes of reform on level of inequality

Variable	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
Urban	-0.0305	-6.3250	-0.0314	-6.6106	-0.0307	-6.4613
ECLAC	-0.0523	-9.5654	-0.0513	-9.5852	-0.0516	-9.6664
Expenditure	-0.0813	-2.7260	-0.0851	-2.8787	-0.0876	-2.9229
Inflation	0.0165	1.9793	0.0179	2.1863	0.0186	2.2736
Household	-0.0132	-2.7286	-0.0133	-2.7530	-0.0127	-2.6238
High	-0.0093	-3.5635	-0.0088	-3.4433	-0.0084	-3.2990
Primary	0.1441	2.0320	0.1519	2.1561	0.1467	2.0859
Trend*Income	0.0000	5.7801	0.0000	5.7123	0.0000	5.9264
Income	-0.0001	-7.2460	-0.0001	-7.1979	-0.0001	-7.3759
1/Income	-251.9838	-3.3943	-250.8982	-3.3851	-234.4582	-3.2187
Privatization	0.0604	3.0108	0.0567	2.8810	0.0587	2.9490
Tax reform	0.0514	2.7174	0.0473	2.5774	0.0521	2.9200
Financial reform	-0.0228	-2.3371	-0.0245	-2.5559	-0.0177	-2.0436
Trade reform	0.0282	1.6202	0.0247	1.4514		
Capital account	-0.0156	-0.9176				
R-squared	0.9785		0.9782		0.9778	
Adjusted R-squared	0.9757		0.9755		0.9751	
Standard error of regression	0.0269		0.0269		0.0269	
Mean dependent var	0.5375		0.5371		0.5359	
S.D. dependent variable	0.1726		0.1717		0.1706	
Sum of squared residues	0.1670		0.1674		0.1686	
F-statistic	751.4683		802.2768		855.2841	
Prob (F-statistic)	0.0000		0.0000		0.0000	
df (degree of fredom)	261		261		261	

3.6 percentage points to the average Gini, which is a large effect. The change also affects the significance and/or size of some of the other variables as well. Inflation, for example, becomes a larger and more significant factor. So do differences in the university education variable. All this suggests that differences in inflation, in educational profile and in the distribution of land are among the reasons why income distributions differ across the countries in the region. But they are not the only differences. If they were, we would not have been able to significantly improve the fit of the model by using country-specific constants.

viii) *Trend.* Our results indicate that there are important shifts in the K-curve over time. We attempted to capture this by introducing two trend terms, one in the constant and the other in the K-curve itself. Regression 3 in table 1 puts the trend in the intercept term. It is negative and significant, suggesting a gradual improvement in inequality over time, other things being equal. But the interaction term tells a different story. It is positive and significant in all of the regressions, including the one with the common constant, and in the regressions for the urban and national samples considered separately. This means that the slope of the K-curve changes over time. Since the coefficient is positive, it means that to the left of the inflection point, where the curve itself is upward sloping, the slope is gradually getting steeper, while to the right of the inflection point, where the slope itself is negative, the trend is making the slope gradually flatter. Furthermore, the interaction term makes the inflection point itself shift gradually to the right over time, extending the range over which growth is regressive. Thus the

trend terms tell two opposing stories. The trend term on the intercept is progressive, shifting the K-curve down. But the interaction term is regressive, so that growth has become systematically less progressive than it used to be. To illustrate all this we show the K-curve for Brazil for 1970, 1980 and 1996 (figure 5). For purposes of comparison, all the relevant variables other than trend are set at their 1996 values. We used the coefficients from regression 3 of table 1 for this calculation. Series one is 1970, two is 1980 and three is 1996. As the reader can see, the progressive shift downward in the intercept is increasingly dominated by the outward shift in the curve and its change of slope. Both these changes make growth less progressive than it would otherwise be, for not only is the interaction reducing the slope of the curve but it also means that the country is moving from one curve to another, thus making the improvement in inequality per unit of growth less than it would be if the country was moving down a stationary K-curve.

FIGURE 5

Kuznets Curves for Brazil
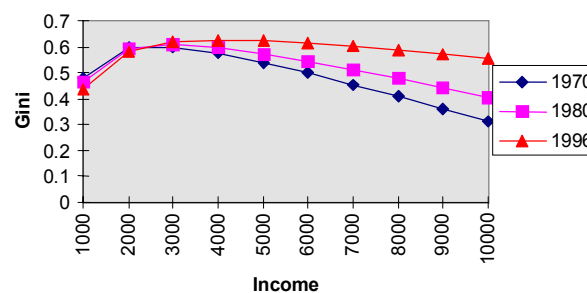
VI

Sub-indexes of reform

We will now look at tables 3 (on page 34) and 4, which show the results for each of the five areas of reform. Table 3 uses the entire sample, while table 4 shows the results for the urban and national samples separately. Table 5 summarizes the results. It is quite clear from table 5 that the various reforms have different and offsetting effects on equity. In all three samples some of the reforms have a significant regressive and some a significant progressive effect. This explains why the overall average reform indexes seem to have little effect on inequality.

Comparing the combined regressions with the separate urban and national regressions, the results for trade, capital and tax reform are a good deal more robust than those for the other two reforms. Trade reform has been

regressive, more so in the national than in the urban regressions. This suggests that the negative effect on agriculture of the loss of protection and price subsidies was more significant that the loss of protection in the manufacturing sector. The theoretical case for trade reform rested on the idea that increased openness should favour Latin America's most abundant factor, assumed to be unskilled labour. That should have improved the distribution. But our econometric evidence says that it has not worked out that way in practice. If anything, the effect has been the opposite. That is consistent with the findings of Donald Robbins (1996), who has presented evidence that trade liberalization has led to the widening of skill differentials. These results are somewhat stronger

Table 4

Results for the subindexes of reform

Variable	National sample		Urban sample	
	Coefficient	t-statistic	Coefficient	t-statistic
Constant	0.617703	9.312883	0.64272	7.573593
Income	-0.0000291	-2.266974	-0.0000613	-3.50241
1/Income	-194.1939	-2.145082	57.34615	0.633523
Trend	-0.001197	-1.103537	-0.003263	-1.587047
Per household	-0.035718	-4.341743		
Expenditure	-0.054946	-4.229492		
Trend*Income	0.000000602	2.038764	0.00000156	3.64106
High	-0.009492	-7.60405		
Noschool	0.002715	6.58017	0.001456	2.318449
Inflation			0.034843	2.458605
ECLAC			-0.067288	-8.141843
Trade	0.081856	3.098284	0.013151	0.435331
Finance	0.017831	0.954765	0.041569	2.327227
Tax	0.108429	3.669748	0.030098	1.255619
Capital	-0.12167	-6.162735	-0.136489	-4.821436
Privatization	0.025497	0.998892	-0.043319	-1.70211
R-squared		0.811166		0.610982
Adjusted R-squared		0.791045		0.56967
Standard error of regression		0.033625		0.036569
Sum of squared residues		0.137939		0.151112
Log likelihood		275.7889		244.9528
Durbin-Watson statistic		1.380068		1.07592
Mean dependent variable		0.49977		0.449939
Standard deviation of dependent variable		0.073559		0.055745
F-statistic		40.31313		14.78958
Prob(F-statistic)		0		0
df (degree of fredom)		136		125

TABLE 5
Effect of Reforms on the Kuznets Curve

	Combined	Urban	National
Privatization	regressive*	progressive	regressive
Financial	progressive*	regressive*	regressive
Tax	regressive*	regressive	regressive*
Trade	regressive	regressive	regressive*
Cap Acct	progressive	progressive*	progressive*

* significant at 1% level

than those of Spilimbergo, Londoño and Szekely (1997), who found that "Trade openness also has a negligible effect over income distribution in Latin America", mainly because relative factor endowments in Latin America are very close to world averages weighted by population and openness (*ibid.*, p. 30). Our conclusions are not consistent with the work of Londoño and Szekely (1998), who found a significant positive relationship between trade reforms in the 1985-1995 period and the income share of the bottom quintile for a panel of 13 countries in the region. However, their regressions did not include urban observations, such as those for Argentina or Bolivia. Nor

did they include any variables other than the reform indexes in the regressions. Thus the effects that they assign to the reforms may well actually come from other policies or variables.

In contrast to trade reform, opening the capital account has been progressive. Reducing barriers to capital mobility has attracted a great deal of foreign capital to Latin America, and theoretically this should have reduced profit rates and increased the demand for labour, all of which should be progressive. The results of the present study indicate that this has indeed been the case. The tax reforms, however, have shifted the Kuznets curve up towards greater inequality. There are clear theoretical arguments explaining why this is so. Switching from progressive income taxes to a flatter tax structure and substituting VAT or consumption taxes for income taxes and tariffs shifts the tax burden away from the rich. As for the other two reforms, the variations in the signs and significance of the coefficients on privatization and financial reforms suggest that our data are not good enough to give an unambiguous answer regarding the effect these two reforms have had.

VII

The effect of economic growth on income distribution: an application of the estimated Kuznets curve

One of the central questions facing anyone analysing trends in income distribution is the effect of growth on future inequality. Supposing that there are no policy changes other than growth: will Latin America become more equitable or not? Is high inequality nothing more than a phase which will be overcome by growth?

Our regressions shed a good deal of light on this question. First of all, since the K-curve has an inverted U shape, we know that some countries are undoubtedly to the left of the inflection point. For them growth is going to be inequitable. For the remaining countries growth should improve things. However, there is the complicating factor of the interaction between trend and income, which is making the growth-equity relationship steadily less progressive. This is partly because it moves the inflection point to the right and partly because it makes the curve itself flatter on the downward portion and steeper on the upward portion.

In 1996 the inflection point of the Kuznets curve lay at just under US$ 4,000 per capita. That means that all seven of the high-income countries in the region were beyond the inflection point, so that for them growth was equalizing. The remaining nine were on the rising portion of the K-curve, and for them the Gini coefficient rises with growth. For the region as a whole, a simple average of the individual country elasticities was .0224, indicating that if all countries grew by the same amount, there would be a slight rise in the average Gini coefficient. Overall, growth in 1996 was not equalizing because the weight of those nine countries on the rising portion of the K-curve was greater than the weight of those on the equalizing part of the curve. If one were to weight the elasticities by either population or income, however, that conclusion would be reversed, because all the big and relatively prosperous countries (Argentina, Brazil, Colombia, Mexico and Venezuela) were on the falling part of their curves.

At the same time, one must remember that because of the trend term the high-income country K-curves are tending to get flatter over time, while the low-income curves are getting steeper. This means that growth has become less and less progressive. To show this, we re-calculated the elasticities using the parameters of the 1970 K-curve. With those parameters, the aggregate distribution elasticity would have been -0.18, which means that a growth rate that raised average inequality in 1996 would have lowered it in 1970. Some feature of the economy is making economic growth significantly less progressive than it used to be. We suspect that the culprit is skill-intensive growth, but we cannot prove that. Whatever the cause is, however, the implication is that if nothing else changes, the impact of growth on inequality in the future is likely to be more regressive than it is today.

VIII

Conclusions

Two main conclusions may be drawn from the econo-metric evidence presented above:

1. There appears to be a robust and significant relation-ship between distribution and income. It has the inverted U-shape that Kuznets predicted, but the relationship has been shifting in a regressive direction over time, so that growth is now a good deal less progressive than it used to be. In the aggregate that means that further growth in Latin America is unlikely to improve the distribution much, if at all. Supplementary measures will therefore have to be taken. Among those suggested by the regres-sions are maintaining low inflation rates and investing in education. Giving new entrants to the labour force more education at any level is progressive, but countries will get a much bigger reduction in inequality if they start at the bottom, universalizing the coverage of pri-mary education and then broadening the coverage of secondary and university education.

2. On aggregate, the structural reforms appear to have a regressive effect on the distribution, but this effect is both small and not very robust statistically. We refer here only to the structural economic reforms, not the macroeco-nomic stabilization measures which were often adopted at the same time. It should also be noted that we are re-ferring only to the direct effect of the reforms, apart from whatever impact they may have had through their effect on growth or inflation. The reason the direct effect is small or insignificant seems to be that reforms in differ-ent areas have offsetting effects on equity. Trade reform is regressive in all of our specifications, but it is insig-nificant in all but the national sample. Tax reform is un-ambiguously regressive, and opening up the capital ac-count is unambiguously progressive. Our results for trade and tax reform and capital account liberalization are the most robust and significant; for the other two reforms, our data were not good enough to give us a clear answer.

(Original: English)

Bibliography

Ahluwalia, M. (1976): Inequality, poverty and development, *Journal of Development Economics*, vol. 3, N° 4, Amsterdam, The Netherlands, North-Holland Publishing Company.

Altimir, O. and L. Beccaria (1999): *Distribución del ingreso en la Argentina*, "Reformas económicas" series, N° 40, Santiago, Chile, Economic Commission for Latin America and the Caribbean (ECLAC).

Anand, S. and S. M. Kanbur (1993): The Kuznets process and the inequality-development relationship, *Journal of Development Economics,* vol. 40, N° 1, Amsterdam, The Netherlands, North-Holland Publishing Company.

Berry, A. (ed.) (1998): *Poverty, Economic Reform and Income Distribution in Latin America*, London, Lynne Rienner.

Birdsall, N., C. Graham and R. H. Sabot (eds.) (1998): *Beyond Trade-Offs: Market Reform and Equitable Growth in Latin America*, Washington, D.C., Inter-American Development Bank (IDB).

Birdsall, N. and J. L. Londoño (1997): *Asset Inequality Does Matter: Lessons from Latin America*, Working paper series, N° 344, Washington, D.C., IDB.

Birdsall, N., D. Ross and R. Sabot (1995): Inequality and growth reconsidered: Lessons from East Asia, *The World Bank Economic Review*, vol. 9, N° 3, Washington, D.C., World Bank.

Bruno, M., M. Ravallion and L. Squire (1996): *Equity and Growth in Developing Countries: Old and New Perspectives on the Policy Issues*, Policy research working paper, N° 1563, Washington, D.C., World Bank.

Bulmer-Thomas, V. (ed.) (1996): *The New Economic Model in Latin America and its Impact on Income Distribution and Poverty*, London, Macmillan.

Burki, S. J. and G. E. Perry (1997): *The Long March: A Reform Agenda for Latin America and the Caribbean in the Next Decade*, Washington, D.C., World Bank.

Cárdenas, M. and R. Bernal (1999): *Changes in the Distribution of Income and the New Economic Model in Colombia*, "Reformas económicas" series, N° 36, Santiago, Chile, ECLAC.

Clarke, G. R. (1995): More evidence on income distribution and growth, *Journal of Development Economics*, vol. 47, N° 2, Amsterdam, The Netherlands, North-Holland Publishing Company.

Deininger, K. and L. Squire (1996): A new data set measuring income inequality, *The World Bank Economic Review*, vol. 10, No. 3, Washington, D.C., World Bank.

De Janvry, A. and E. Sadoulet (in the press): *Growth, Poverty and Inequality in Latin America: A Causal Analysis, 1970-94*, Berkeley, California, University of California.

ECLAC (1997): *The Equity Gap. Latin America, the Caribbean and the Social Summit*, LC/G.1954/Rev.1-P, Santiago, Chile, United Nations publication, Sales N° E.97.II.G.11.

ECLAC (several years): *Social Panorama of Latin America,* Santiago, Chile.

Edwards, S. (1995): *Crisis and Reform in Latin America: From Despair to Hope*, Washington, D.C., World Bank.

————(1997): Trade policy, growth and income distribution, *The American Economic Review,* vol. 87, N° 2, Washington, D.C., American Economic Association.

Escaith, H. and S. Morley (forthcoming): *The Impact of Structural Reforms on Growth in Latin America and the Caribbean: An Empirical Estimation*, Working paper, Santiago, Chile, ECLAC.

Ffrench-Davis, R. and H. Reisen (eds.) (1998): *Capital Flows and Investment Performance: Lessons from Latin America*, Paris, Organization for Economic Co-operation and Development (OECD).

Fields, G. (1994): Data for measuring poverty and inequality changes in the developing countries, *Journal of Development Economics*, vol. 44, N° 1, Amsterdam, The Netherlands, North-Holland Publishing Company.

Gindling, T. and A. Berry (1992): The performance of the labour market during recession and structural adjustment: Costa Rica in the 1980s, *World Development*, vol. 20, N° 11, Oxford, U.K., Pergamon Press Ltd.

IDB (Inter-American Development Bank) (1997): *Economic and Social Progress in Latin America. 1997 Report*, Washington, D.C.

————(1998): *Economic and Social Progress in Latin America. 1998-1999 Report*, Washington, D.C.

Jacome, L., C. Larrea and R. Vos (1998): Políticas macroeconómicas, distribución y pobreza en el Ecuador, in E. Lanuza, L. Taylor and S. Morley (eds.), *Política macroeconómica y pobreza en América Latina y el Caribe*, New York, United Nations Development Programme (UNDP).

Kuznets, S. (1955): Economic growth and income inequality, *The American Economic Review*, vol. 45, N° 1, Washington, D.C., American Economic Association.

Lanuza, E., L. Taylor and S. Morley (eds.) (1998): *Política macroeconómica y pobreza en América Latina y el Caribe*, New York, UNDP.

Larrañaga, O. J. (1999): *Distribución de ingresos y crecimiento económico en Chile*, "Reformas económicas" series, N° 35, Santiago, Chile, ECLAC.

Londoño, J. L. and M. Szekely (1997): *Persistent Poverty and Excess Inequality: Latin America, 1970-1995*, Working paper series, No. 357, Washington, D.C., IDB.

————(1998): Sorpresas distributivas después de una década de reformas, *Pensamiento iberoamericano*, special issue, Madrid, Fundación Centro Español de Estudios de América Latina.

Lora, E. (1998): Una década de reformas estructurales en América Latina: qué se ha reformado y cómo medirlo, *Pensamiento iberoamericano*, special issue, Madrid, Fundación Centro Español de Estudios de América Latina.

Lora, E. and F. Barrera (1998): El crecimiento económico en América Latina después de una década de reformas estructurales, *Pensamiento iberoamericano*, special issue, Madrid, Fundación Centro Español de Estudios de América Latina.

Márquez, G. and others (1993): Fiscal policy and income distribution in Venezuela, in R. Hausmann and R. Rigobon (eds.), *Government Spending and Income Distribution in Latin America*, Washington, D.C., IDB.

Morley, S. A. (1995): *Poverty and Inequality in Latin America*, Baltimore, Maryland, Johns Hopkins.

————(1998): La pobreza en tiempos de recuperación económica y reforma en América Latina: 1985-1995, in

E. Lanuza, L. Taylor and S. Morley (eds.), *Política macroeconómica y pobreza en América Latina y el Caribe*, New York, UNDP.

———(forthcoming): *The Distribution Problem in Latin America*, Santiago, Chile, ECLAC.

Morley, S. A., R. Machado and S. Pettinato (1999): *Indexes of Structural Reform in Latin America*, "Reformas económicas" series, N° 12, Santiago, Chile, ECLAC.

Neri, M. and J. M. Camargo (1999): *Structural Reforms, Macroeconomic Fluctuations and Income Distribution in Brazil*, "Reformas económicas" series, N° 39, Santiago, Chile, ECLAC.

Ocampo, J. A., M. J. Pérez, C. Tovar and F. J. Lasso (1998): Macroeconomía, ajuste estructural y equidad en Colombia: 1978-1996, in E. Lanuza, L. Taylor and S. Morley (eds.), *Política macroeconómica y pobreza en América Latina y el Caribe*, New York, UNDP.

Ocampo, J. A. and L. Taylor (1998): Trade liberalisation in developing economies: Modest benefits but problems with productivity growth, macro prices and income distribution, *The Economic Journal*, vol. 108, N° 450, Oxford, U.K., The Royal Economic Society.

Pereira, R. and W. Jiménez (1998): Políticas macroeconómicas, pobreza y equidad en Bolivia, in E. Lanuza, L. Taylor and S. Morley (eds.), *Política macroeconómica y pobreza en América Latina y el Caribe*, New York, UNDP.

Psacharopoulos, G., S. Morley, A. Fizbein, H. Lee and B. Wood (1997): *Poverty and Income Distribution in Latin America: The Story of the 1980s*, Technical paper, N° 351, Washington, D.C., World Bank.

Ravallion, M. and S. Chen (1997): What can new survey data tell us about recent changes in distribution and poverty?, *The World Bank Economic Review*, vol. 11, N° 2, Washington, D.C., World Bank.

Robbins, D. (1995): *Trade, Trade Liberalization and Inequality in Latin America and East Asia: Synthesis of Seven Country Studies*, Cambridge, Massachusetts, Harvard Institute for International Development.

———(1996): *HOS Hits Facts: Facts Win. Evidence on Trade and Wages in the Developing World*, Cambridge, Massachusetts, Harvard Institute for International Development.

Saavedra, J. and J. J. Díaz (1999): *Desigualdad del ingreso y del gasto en el Perú antes y después de las reformas estructurales*, "Reformas económicas" series, N° 34, Santiago, Chile, ECLAC.

Sheahan, J. and E. V. Iglesias (1998): Kinds and causes of inequality in Latin America, in N. Birdsall, C. Graham and R. H. Sabot (eds.), *Beyond Tradeoffs: Market Reform and Equitable Growth in Latin America*, Washington, D.C., IDB.

Spilimbergo, A., J. L. Londoño and M. Székely (1997): *Income Distribution, Factor Endowments and Trade Openness*, Working paper series, N° 356, Washington, D.C., IDB.

Székely, M. (1998): *The Economics of Poverty, Inequality and Wealth Accumulation in Mexico*, New York, St. Martin's Press.

Trejos, J. D. (1999): *Reformas económicas y distribución del ingreso en Costa Rica*, "Reformas económicas" series, N° 37, Santiago, Chile, ECLAC.

Wood, A. (1994): *North-South Trade, Employment and Inequality: Changing Fortunes in a Skill-Driven World*, Oxford, U.K., Clarendon Press.

———(1997): Openness and wage inequality in developing countries: The Latin American challenge to East Asian conventional wisdom, *World Bank Economic Review*, vol. 11, N° 1, Washington, D.C., World Bank.

Education and *income distribution* in urban Brazil, *1976-1996*

Francisco H.G. Ferreira

Professor, Department of Economics,
Catholic University of
Rio de Janeiro
fferreira@econ.puc-rio.br

Ricardo Paes de Barros
Institute of Applied
Economic Research (IPEA),
Rio de Janeiro
rbarros@ipea.gov.br

Despite tremendous macroeconomic instability, Brazil's urban income distributions in 1976 and 1996 appear, at first glance, deceptively similar. Mean household income per capita was stagnant, with a minute accumulated growth of 4.3% over the two decades. The Gini coefficient hovered just above 0.59 in both years, and the incidence of poverty (with respect to a poverty line of R$60/month at 1996 prices) was effectively unchanged at 22%. Yet, behind this apparent stability, a powerful combination of labour market, demographic and educational dynamics were at work, one effect of which was to generate a substantial increase in extreme urban poverty. Using a micro-simulation-based decomposition methodology which endogenizes labour incomes, individual occupational choices and education decisions, we show that the distribution of incomes was being affected, on the one hand, by a decline in average returns to both education and experience and by impoverishing changes in the structure of occupations and labour force participation (all of which tended to increase poverty), and on the other hand by an increase in educational endowments across the distribution and a progressive reduction in dependency ratios (both of which tended to reduce poverty).

I

Introduction

Both by the standards of its own previous growth record, during the 'Brazilian miracle' years of 1968-1973, and by those of other leading developing countries thereafter, notably in Asia, the two decades between 1974 and 1994 (i.e., between the first oil shock and the return of stability with the Plano Real) were dismal for Brazil. First and foremost, they were characterized by persistent macroeconomic disequilibrium, the main symptoms of which were stubbornly high and accelerating inflation and a GDP time-series marked by unusual volatility and a very low positive trend.

The macroeconomic upheavals during this period involved three price and wage freezes (during the Cruzado Plan of 1986, the Bresser Plan of 1987 and the Verão Plan of 1989), all of which were followed by higher inflation rates; one temporary financial asset freeze (with the Collor Plan of 1990); and finally a successful currency reform followed by the adoption of a nominal anchor in 1994 (the Plano Real). The national currency changed its name four times.[1] Throughout the period, macroeconomic policy was almost without exception characterized by relative fiscal laxity and growing monetary stringency.

In addition, substantial structural changes were taking place. Brazil's population grew by 46.6% between 1976 and 1996 and also became more urban (the urbanization rate rose from 68% to 77%). The average level of education of the population over 10 years of age rose from 3.2 to 5.3 effective years of schooling.[2] Open unemployment grew steadily more prevalent. The sectoral composition of the labour force changed away from agriculture and manufacturing, and towards services. The degree of formalization of the labour force declined substantially: the proportion of formal workers (wage workers with formal documentation) was almost halved, from just under 60% to just over 30% of all workers. And yet, despite the macroeconomic turmoil and continuing structural changes, a casual glance at the indicators of inequality and headline poverty given in table 1 might suggest that little had changed in Brazilian urban income distribution between 1976 and 1996.

As is often the case, however, casual glances may turn out to be misleading. This apparent distributional stability belies a number of powerful, and often countervailing, changes in four areas: the returns to education in the labour markets; the distribution of educational achievements over the population; the pattern of occupational choices; and the demographic structure resulting from household fertility choices. In this paper, we note two 'puzzles' about the evolution of Brazil's urban income distribution in the 1976-1996 period, and suggest explanations for them.

The first puzzle is posed by the combination of (slow) growth in mean incomes and stable or slightly declining inequality, on the one hand, and rising extreme poverty on the other. We argue that this can only be explained by the growth in the size of a group of very poor households, who appear to be effectively excluded both from the labour markets and from the system of formal safety nets. This group is trapped in indigence at the very bottom of the urban Brazilian income distribution, and contributes to the rises in poverty measures which are particularly sensitive to the depth –P(1)– and severity –P(2)– of poverty, particularly when poverty is defined with respect to a low poverty line.[3] This is starkly cap-

[1] The changes were from Cruzeiro to Cruzado in 1986; from Cruzado to Novo Cruzado in 1989; from Novo Cruzado back to Cruzeiro in 1990, and from Cruzeiro to Real in 1994.

[2] 'Effective' years of schooling are based on the last grade completed, and are thus net of repetition. All the figures are from Ferreira and Paes de Barros, 1999.

[3] All poverty measures reported in this paper are the P(α) class of decomposable measures developed by Foster, Greer and Thorbecke (1984). An increase in α implies an increase in the weight given to the distance between people's incomes and the poverty line.

TABLE 1

Brazil: General economic indicators, 1976, 1981, 1985 and 1996

	1976	1981	1985	1996
Annual per capita GNP (in constant 1996 Reals)	4,040	4,442	4,540	4,945
Annual inflation rate[a]	42%	84%	190%	9%
Open unemployment[b]	1.82%	4.26%	3.38%	6.95%
Average years of schooling[c,d]	3.23	4.01	4.36	5.32
Rate of urbanization[d]	67.8%	77.3%	77.3%	77.0%
Self-employed as share of labour force[d]	27.03%	26.20%	26.19%	27.21%
Share of formal employment[d,e]	57.76%	37.97%	36.41%	31.51%
Mean (urban) household per capita income[d]	265.10	239.08	243.15	276.46
Inequality (Gini)[d]	0.595	0.561	0.576	0.591
Inequality (Theil - T)[d]	0.760	0.610	0.657	0.694
Poverty incidence (R$ 30/month)[f]	0.0681	0.0727	0.0758	0.0922
Poverty incidence (R$ 60/month)[f]	0.2209	0.2149	0.2274	0.2176

[a] Percent, from January to December. Based on the IGP-DI for 1976, and on the INPC-R for all other years.

[b] Based on the IBGE Metropolitan Unemployment Index.

[c] For all individuals 10 years of age or older, in urban areas.

[d] Calculated by the authors from the urban national household survey samples.

[e] Defined as the number of employees 'com carteira' as a fraction of the sum of all wage employees and self-employed workers.

[f] Urban only, monthly and spatially deflated. Expressed in constant 1996 Reals.

tured by figure 1 in section II below, which plots the observed (truncated) Pen parades for the four years being studied.[4] The main endogenous channel through which the marginalization of this group is captured in our model is a shift in their occupational 'decisions' away from either wage or self-employment towards unemployment or exit from the labour force.

The evidence which we will examine in section IV below reveals downward shifts in the earnings-education profile, controlling for age and gender, in both the wage-earning and self-employment sectors.[5] Although the profile is slightly convex, the magnitude of the shift implies a decline in the (average) rate of return to education for all relevant educational levels. Similarly, average returns to experience also fell unambiguously for 0-50 years of experience (see figure 5). The combined effect of these changes –the 'price effect'– was an increase in simulated poverty for all measures and for both lines. Simulated inequality also rose, albeit much more mildly. Both effects were exacerbated when the changes (up to 1996) in the determinants of labour force participation decisions were also taken into account.

The second puzzle, then, is what forces counterbalance these price and occupational choice effects, so as to explain the observed stability in inequality and 'headline' poverty.[6] We find that they were fundamentally the combination of increased education endowments, which move workers up along the flattening earnings-education slope, together with an increase in the correlation between family income and family size, caused by a more than proportional reduction in dependency ratios and family sizes for the poor. This demographic factor had both direct effects on per capita income –through a reduction in the denominator– and indirect effects, through participation decisions and higher incomes.

We addressed these issues by means of a micro-simulation-based decomposition of distributional changes, developed by Bourguignon, Ferreira and Lustig (1998), which itself builds upon the work of Almeida dos Reis and Paes de Barros (1991) and Juhn, Murphy and Pierce (1993). The approach has two distinguishing features. First, unlike other dynamic inequality decompositions, such as that proposed by Mookherjee and Shorrocks (1982), it decomposes the effects of changes on an entire distribution, rather than on a scalar summary statistic (such as the mean log deviation). This allows for much greater versatility: within the same framework, a wide range of simulations can be performed to investigate the effects of changes in specific parameters on any number of inequality or poverty measures (and

---

[4] The idea of "parades" of income recipients was developed by Jan Pen. Pen parades are the mathematical inverse of distribution functions; that is: they plot the incomes earned by each person (or group of persons), when these are ranked by income.

[5] This shift is from 1976 to 1996, and takes place after upward shifts in the 1980s. See Figure 4.

[6] By 'headline' poverty, we mean the incidence of poverty computed with respect to the R$60/month poverty line.

then for any number of poverty lines or assumptions about equivalence scales). Second, the evolving distribution which it decomposes is a distribution of household incomes per capita (with the recipient unit generally being the individual). Therefore, moving beyond pure labour market studies, the effect of household composition on living standards and participation decisions is explicitly taken into account. As it turns out, these are of great importance for a fuller understanding of the dynamics involved.

The remainder of the paper is organized as follows. Section II briefly reviews the main findings of the litera-

ture on income distribution in Brazil over the period studied and presents summary statistics and dominance comparisons for the four observed distributions we analyze: 1976, 1981, 1985 and 1996. Section III outlines how the basic model of Bourguignon, Ferreira and Lustig (1998) was adapted to the case of Brazil. Section IV presents the results of the estimation stage and discusses some of their implications. Section V presents the main results of the simulation stage and decomposes the observed changes in poverty and inequality. Finally, section VI presents the conclusions and draws some policy implications.

II

Income distribution in Brazil from 1976 to 1996:
a brief review of the literature and of our data set

There is little disagreement in the existing literature about the broad trends in Brazilian inequality since reasonably reliable data first became available in the 1960s. The Gini coefficient rose substantially during that decade, from around 0.500 in 1960 to 0.565 in 1970 (see Bonelli and Sedlacek, 1989). There was a debate over the causes of this increase, spearheaded by Albert Fishlow on the one hand, and Carlos Langoni on the other, but there was general agreement that the 1960s had seen substantially increased dispersion in the Brazilian income distribution.[7]

The 1970s displayed a more complex evolution. Income inequality rose between 1970 and 1976, reached a peak in that year, but then fell – both for the distribution of total individual incomes in the economically active population (EAP) and for the complete distribution of household per capita incomes – from 1977 to 1981 (see Bonelli and Sedlacek (1989); Hoffman (1989) and Ramos (1993)). The recession year of 1981 was a local minimum in the inequality series, whether measured by the Gini or the Theil-T index. After 1981, inequality rose during the recession years of 1982 and 1983. Some authors report small declines in some indices in 1984, but the increase was resumed in 1985. 1986, the year of the Cruzado Plan, saw a break in the series, caused both by

a sudden (if short-lived) decline in inflation and a large increase in reported household incomes. Stability and economic growth led to a decline in measured inequality, according to all authors. Thereafter, however, with the failure of the Cruzado Plan stabilization attempt and the return to stagflation, inequality resumed its upward trend, with the Gini finishing the decade at 0.606.

The general trends identified in the existing literature are mirrored in the statistics for the years with which we concern ourselves in this paper, namely 1976, 1981, 1985 and 1996. The distributions for each of these years come from the Pesquisa Nacional por Amostra de Domicilios (National Household Survey –PNAD), run by the Brazil–ian Geographical and Statistical Institute (IBGE). Except where otherwise explicitly specified, we deal with distributions for urban areas only, where the welfare concept is total household income per capita (in constant 1996 Reals, spatially deflated to adjust for regional differences in average cost of living) and the unit of analysis is the individual. Details of the PNAD sampling coverage and methodology, sample sizes, the definition of key income variables, spatial and temporal deflation issues, and adjustments with respect to the National Accounts baseline are discussed in Ferreira and Paes de Barros (1999).

Table 2 below presents a number of summary statistics for these distributions, in addition to the mean, which was given in table 1 above. The four inequality indices used throughout this paper are the Gini coefficient and three members of the Generalized Entropy

---

[7] The Fishlow-Langoni debate concerned the importance of education vis-a-vis repressive labour market policies in determining the high level of Brazilian inequality. See, for example, Fishlow (1972), Langoni (1973) and Bacha and Taylor (1980).

TABLE 2

**Brazil: Basic distributional statistics for different
degrees of household economies of scale**

	1976	1981	1985	1996
Median (1996 R$)[a]	127.98	124.04	120.83	132.94
Inequality				
Gini - $\theta = 1.0$	0.595	0.561	0.576	0.591
Gini - $\theta = 0.5$	0.566	0.529	0.548	0.567
E(0) - $\theta = 1.0$	0.648	0.542	0.588	0.586
E(0) - $\theta = 0.5$	0.569	0.472	0.524	0.534
E(1) - $\theta = 1.0$	0.760	0.610	0.657	0.694
E(1) - $\theta = 0.5$	0.687	0.527	0.580	0.622
E(2) - $\theta = 1.0$	2.657	1.191	1.435	1.523
E(2) - $\theta = 0.5$	2.254	0.918	1.134	1.242
Poverty - R$30/ month				
P(0) - $\theta = 1.0$	0.0681	0.0727	0.0758	0.0922
P(0) - $\theta = 0.5$	0.0713	0.0707	0.0721	0.0847
P(1) - $\theta = 1.0$	0.0211	0.0337	0.0326	0.0520
P(1) - $\theta = 0.5$	0.0235	0.0315	0.0303	0.0442
P(2) - $\theta = 1.0$	0.0105	0.0246	0.0224	0.0434
P(2) - $\theta = 0.5$	0.0132	0.0226	0.0204	0.0357
Poverty - R$60/ month				
P(0) - $\theta = 1.0$	0.2209	0.2149	0.2274	0.2176
P(0) - $\theta = 0.5$	0.2407	0.2229	0.2382	0.2179
P(1) - $\theta = 1.0$	0.0830	0.0879	0.0920	0.1029
P(1) - $\theta = 0.5$	0.0901	0.0875	0.0927	0.0960
P(2) - $\theta = 1.0$	0.0428	0.0525	0.0534	0.0703
P(2) - $\theta = 0.5$	0.0471	0.0508	0.0521	0.0625

[a] For urban areas only, spatially deflated.

Class of inequality indices, $E(\phi)$. Specifically, we have chosen $E(0)$, also known as the mean log deviation or the Theil–L index; $E(1)$, more commonly known as the Theil–T index, and $E(2)$, which is half of the square of the coefficient of variation. These provide a useful range of sensitivities to different parts of the distribution. $E(0)$ is more sensitive to the bottom of the distribution, while $E(2)$ is more sensitive to higher incomes. $E(1)$ is somewhere in between, whereas the Gini places greater weight around the mean.

We also present three poverty indices from the Foster-Greer-Thorbecke additively decomposable class $P(\alpha)$. $P(0)$, also known as the headcount index, measures poverty incidence, $P(1)$ is the normalized poverty deficit, and $P(2)$ is an average of squared normalized deficits, thus placing greater weight on incomes furthest from the poverty line. We calculate each of these with respect to two poverty lines, representing R$1 and R$2 per day at 1996 prices.[8]

Each of these poverty and inequality indices is presented both for the (individual) distribution of total household incomes per capita and for an equivalized distribution using the Buhmann, Rainwater, Schmaus and Smeeding (1988) parametric class of equivalence scales (with $\theta = 0.5$). This provides a rough test that the trends we describe are robust to different assumptions about the degree of economies of scale in consumption within households. As usual, per capita incomes generate an upper limit for inequality measures, whereas allowing for some proportion of local public goods within households raises the income of (predominantly poor) very large households and lowers inequality. In the case of the poverty measures, the poverty lines were adjusted as follows: $z^* = z[\mu(n)]^{1-\theta}$, where $\mu(n)$ is the mean household size in the distribution (see Deaton and Paxson, 1997).

Table 2 also confirms that the evolution of inequality over the period is marked by a decline from 1976 to 1981 but a subsequent deterioration over the remaining two sub-periods. This trend is robust to the choice of equivalence scale, proxied here by two different values for θ, although the inequality levels are always lower when we allow for economies of scale within households.

---

[8] At 1996 market exchange rates, this was roughly equal to U$1and U$2. In real terms, this would be slightly less than the conventional poverty lines of PPP U$ 1 and 2 valued at 1985 prices which the World Bank often uses for international comparisons, due to US inflation in the intervening decade.

TABLE 3
Brazil: Stochastic dominance results

	1976	1981	1985	1996
1976			F	
1981				L
1985				L
1996				

It is also robust to the choice of inequality measure, at least as regards the inequality increases from 1981 to 1996 and from 1985 to 1996, as the Lorenz dominance results identified in table 3 indicate.

The results for poverty are more ambiguous. With respect to the higher poverty line, the incidence is effectively unchanged throughout the period (and even displays a slight decline for the equivalized distribution). *P(1)* and *P(2)*, however, show increases over the period, and these become both more pronounced and more robust with respect to θ as the concavity of the poverty measure increases. This suggests that the depth and severity of poverty, affected mostly by falling incomes at the very bottom of the distribution, were on the rise.

These results are reflected in table 3, where a letter in cell (*i, j*) indicates that the distribution for year *i* second order or Lorenz stochastically dominates (L) or first-or-

der stochastically dominates (F) over that for year *j*. 1981 and 1985 both display Lorenz dominance over 1996, as suggested above. There is only one case of first-order welfare dominance throughout the period: social welfare measured in terms of money was unambiguously higher in 1976 than in 1985. Indeed, all poverty measures reported for both of our lines (and for $\theta = 1.0$) are higher in 1985 than in 1976.[9] This is conspicuously not the case for a comparison between 1976 and 1996. Whereas poverty measures very sensitive to the poorest are higher for 1996, poverty incidence for the 'higher' lines (R$ 60 per month) falls from 1976 to 1996, suggesting a crossing of the distribution functions. Figure 1 shows this crossing, by plotting the Pen parades ($F^{-1}(y)$) –truncated at the 60th percentile– for all four years analyzed. Note that whereas 1976 lies in all cases above 1985, all other pairs cross. In particular, 1976 and 1996 cross somewhere near the 17th percentile.

Before we turn to the model used to decompose changes in the distribution of household incomes, which will shed some light on all of these changes, it will prove

---

[9] Note that this first-order welfare dominance is not robust to a change in θ to 0.5.
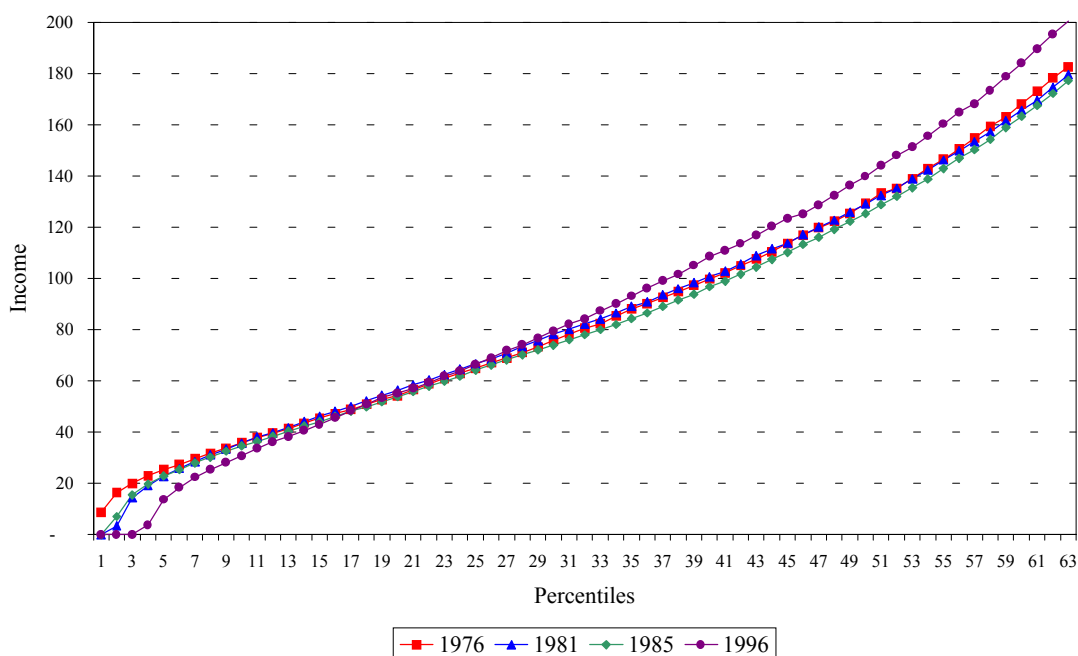
FIGURE 1
Brazil: Truncated Pen parades, 1976-1996

TABLE 4

Brazil: Educational and labour force participation statistics, by gender and race[a]

	1976	1981	1985	1996
Average years of schooling (males)	3.32	4.04	4.36	5.20
Average years of schooling (females)	3.14	3.99	4.37	5.43
Average years of schooling (blacks and mixed race)	–	–	–	4.20
Average years of schooling (whites)	–	–	–	6.16
Average years of schooling (Asians)	–	–	–	8.13
Labour force participation (males)	73.36%	74.63%	76.04%	71.31%
Labour force participation (females)	28.62%	32.87%	36.87%	42.00%
Labour force participation (blacks and mixed race)	–	–	–	55.92%
Labour force participation (whites)	–	–	–	56.41%
Labour force participation (Asians)	–	–	–	54.88%

[a] Average 'effective' years of schooling for persons ten years of age or older, in urban areas.
Labour force participation refers to urban areas only.

helpful to gather some evidence on the evolution of educational attainment (as measured by average effective years of schooling) and labour force participation, for different groups in the Brazilian population, partitioned by gender and ethnicity. Table 4 presents these statistics.

As may be seen, there was some progress in average educational attainment in urban Brazil over this period. Average effective years of schooling for all individuals ten years or older, as reported in table 1, rose from 3.2 to 5.3. In fact, this piece of good news will prove of vital importance for having prevented a more pronounced increase in poverty. Table 4 reveals that the male-female educational gap has now been eliminated, with females older than ten being on average slightly more educated than males. Clearly, this must imply a large disparity in

favour of girls in recent cohorts. While a cohort analysis of educational trends is beyond the scope of this paper,[10] such a rapid reversal may in fact warrant a shift in public policy towards programmes aimed at keeping boys in school, without in any way discouraging the growth in female schooling. Finally, the remarkable disparity in educational attainment across ethnic groups should be noted, with Asians substantially above average and blacks and those of mixed race below it.

As for labour force participation, the persistent and substantial increase in female participation –from 29% to 42% over the two decades– was partly mitigated by a decline in male participation rates. Those trends notwithstanding, the male-female participation gap remains high, at around 30 percentage points. There is little evidence of differential labour force participation across ethnic groups.

III

The model and the decomposition methodology

Let us now turn to the Brazilian version of the general semi-reduced-form model for household income and labour supply developed by Bourguignon, Ferreira and Lustig (1998). It is used here to investigate the evolution of the distribution of household incomes per capita over the two decades from the mid-1970s to the mid-1990s. Specifically, we analyze the distributions for 1976, 1981, 1985 and 1996 and simulate changes between them. As noted above, the paper covers only Brazil's urban areas

(which account for some three-quarters of its population). The general model therefore collapses to two occupational sectors: wage earners and self-employed in urban areas.[11]

---

[10] See Duryea and Szekely (1998) for such an educational cohort analysis of Brazil and other Latin American countries.
[11] In Brazil, wage earners include employees with or without formal documentation ('com ou sem carteira'). The self-employed are termed own-account workers ('conta propria').

Total household income is given by:

$$Y_h = \sum_{i=1}^{n} w_i L_i^w + \sum_{i=1}^{n} \pi_i L_i^{se} + Y_{0h} \qquad [1]$$

where w_i are the total wage earnings of individual i, L^w is a dummy variable that takes a value of 1 if individual i is a wage earner (and zero otherwise); π_i are the self-employment profits of individual i; L^{se} is a dummy that takes the value of 1 if individual i is self-employed (and zero otherwise); and Y_0 is income from any other sources, such as transfers or capital incomes. Equation [1] is not estimated econometrically. It aggregates information on the first right-hand-side term from equations [2] and [4], on the second term from equations [3] and [4] and on the third term directly from the household data set.

The wage-earnings equation is given by:

$$Log w_i = X_i^P \beta^w + \varepsilon_i^w \qquad [2]$$

where $X_i^P = (ed, ed^2, exp, exp^2, D_g)$. *Ed* denotes completed effective years of schooling. Experience (*exp*) is defined simply as: age - education − 6, though a more desirable definition would require the age when a person first entered employment, a variable which is not available for 1976.[12] D_g is a gender dummy, which takes the value *1* for females and zero for males. w_i are the monthly earnings of individual i. e_i is a residual term which captures any other determinant of earnings, including any unobserved individual characteristics, such as innate talent. This extremely simple specification was chosen so as to make the simulation stage of the decomposition feasible, as described below. Analogously, the self-employed earnings equation is given by:

$$Log \pi_i = X_i^P \beta^{se} + \varepsilon_i^{se} \qquad [3]$$

Equations [2] and [3] are estimated by simple ordinary least squares (OLS). Equation [2] is estimated for all employees, whether or not heads of household and whether or not they have formal sector documentation ('com or sem carteira'). Equation [3] is estimated for all self-employed individuals (whether or not heads of households). Because the errors ε are unlikely to be independent of the exogenous variables, a sample selec-

tion bias correction procedure might be used. However, the standard Heckman procedure for sample selection bias correction requires equally strong assumptions about the orthogonality between the error terms ε and ξ (from the occupational choice multinomial *logit* below). The assumptions required to validate OLS estimation of [2] and [3] are no more demanding than those required to validate the results of the Heckman procedure. We assume, therefore, that all errors are independently distributed, and do not correct for sample selection bias in the earnings regressions.

We now turn to the labour force participation model. Because we have a two-sector labour market (segmented into the wage employment and self-employment sectors), labour force participation and the choice of sector (occupational choice) could be treated in two different ways. One could assume that the choices are sequential, with a participation decision independent of the occupational choice, and the latter conditional on the former. This approach, which would be compatible with a sequential *probit* estimation, was deemed less satisfactory than one in which individuals face a single three-way choice, between staying out of the labour force, working as employees, or engaging in self-employment. Such a choice can be estimated by a multinomial logit model. According to that specification, the probability of being in state s ($= 0, w, se$) is given by:

$$P_i^s = \frac{e^{Z_i^o \gamma_s}}{e^{Z_i^o \gamma_s} + \sum_{j \neq s} e^{Z_i^o \gamma_j}} \quad \text{where } s, j = (0, w, se) \qquad [4]$$

where the explanatory variables differ for household heads and other household members, by assumption, as follows.

For household heads:

$$Z_1^h = \left(\begin{array}{l} X_1^P; n_{0-13}, n_{14-65}, n_{>65}, \frac{1}{n_{14-65}} \sum_{-1} D_{14-65} ed, \left[\frac{1}{n_{14-65}} \sum_{-1} D_{14-65} ed \right]^2 \\ , \frac{1}{n_{14-65}} \sum_{-1} D_{14-65} age, \left[\frac{1}{n_{14-65}} \sum_{-1} D_{14-65} age \right]^2, \frac{1}{n_{14-65}} \sum_{-1} D_{14-65} Gd, D \end{array} \right)$$

Note that this is essentially a reduced-form model of the labour supply, where own earnings are replaced by the variables that determine them, according to [2] or [3].

For other members of the household:

$$Z_i^h = \left(\begin{array}{l} X_i^P; n_{0-13}, n_{14-65}, n_{>65}, \frac{1}{n_{14-65}} \sum_{-i} D_{14-65} ed, \left[\frac{1}{n_{14-65}} \sum_{-i} D_{14-65} ed \right]^2, \\ \frac{1}{n_{14-65}} \sum_{-i} D_{14-65} age, \left[\frac{1}{n_{14-65}} \sum_{-i} D_{14-65} age \right]^2, \frac{1}{n_{14-65}} \sum_{-i} D_{14-65} Gd, D_1^{se}, L_1^w w_1, D \end{array} \right)$$

---

[12] Since education is given by the last grade completed, and is thus net of repetition, this definition will overestimate the experience of those who repeated grades at school and hence bias the experience coefficient downwards. The numbers involved are not so large as to alter any conclusions on trends, however.

Where n_{k-m} is the number of persons in the household whose age falls between k and m; D_{14-65} is a dummy that takes the value 1 for individuals whose age is between 14 and 65; D^{se} is a dummy for a self-employed head, and the penultimate term is the earnings of a wage-earning head. These last two variables establish a direct conduit for the effect of the head's occupational choice (and possibly income) on the participation decisions of other members. D is a dummy variable that takes the value of 1 if there are no individuals aged 14-65 in the household. The sums defined over $\{-j\}$ are sums over $\{\forall \iota \in h / j\}$.

The multinomial logit model in [4] corresponds to the following discrete choice process:

$$s = \underset{j}{Argmax} \{U_j = Z_i^h \gamma_j + \xi_j, j = (0, w, se)\} \quad [5]$$

where Z is given above, separately for household heads and other members; the expressions ξ_j are random variables with a double exponential density function, and U_j may be interpreted as the utility of alternative j. Once the vector γ_j is estimated by [4] and a random term ξ is drawn, each individual chooses an occupation j so as to maximize the above utility function.

Once equations [2], [3] and [4] have been estimated, we have two vectors of parameters for each of the four years in our sample ($t \in \{1976, 1981, 1985, 1996\}$): β_t from the earnings equations for both wage earners and the self-employed (including constant terms α_t), and γ_t from the participation equation. In addition, from equation [1] we have Y_{0ht} and Y_{ht}. Let $X_{ht} = \{X_i^P, Z_i^h \mid \forall i \in h\}$ and $\Omega_{ht} = \{\varepsilon_i^w, \varepsilon_i^{se}, \xi_i^j \mid i \in h\}$. We can then write the total income of household h at time t as follows:

$$Y_{ht} = H(X_{ht}, Y_{oht}, \Omega_{ht}; \beta_t, \gamma_t) \quad h = 1, \dots, m \quad [6]$$

Based on this representation, the distribution of household incomes:

$$D_t = \{Y_{1t}, Y_{2t}, \dots, Y_{mt}\} \quad [7]$$

can be rewritten as:

$$D_t = D[\{Y_{ht}, Y_{oht}, \Omega_{ht}\}, \beta_t, \gamma_t] \quad [8]$$

Where $\{.\}$ refers to the joint distribution of the corresponding variables over the whole population.

We are interested in understanding the evolution of D_t over time, or possibly that of a set of alternative summary poverty or inequality measures defined on the basis of it. Based on the representation of a distribution given by [8], changes in the distribution of incomes can

be decomposed into price effects (β), occupational choice effects (γ), endowment effects (X, Y_0) and residual effects (Ω), as outlined in Ferreira and Paes de Barros (1999). The simplest decomposition applies to those arguments which are exogenous to the household: that is, βs, γs, and the variance of the various residual terms. Changing the values of βs amounts to assuming a change in the rate of return on human capital variables in equation [2] and [3]. We refer to this as a "price effect".

Calculating the price and occupational choice effects is reasonably straightforward, once the relevant exogenous parameters have been estimated. Estimating individual endowment effects requires a further step, since elements of the X and Y vectors are jointly distributed, and a change in the value of any one variable must be understood as being conditional on all the other observable characteristics.

Specifically, if we are interested in the effect of a change in the distribution of a single specific variable X_k on the distribution of household incomes between times t and t', it is first necessary to identify the distribution of X_k conditional on other relevant characteristics X_{-k} (and possibly other incomes Y_0). This can be done by regressing X_k on X_{-k} at dates t and t', as follows:

$$X_{kit} = X_{-kit} \mu_t + u_{kit} \quad [9]$$

where k is the variable, i is the individual, and t is the date. The vector of residuals u_{kit} represents the effects of unobservable characteristics (assumed to be orthogonal to X_{-k}) on X_k. The vector μ_t is a vector of coefficients capturing the dependency of X_k on the true exogenous variables X_{-k}, at time t. For the sake of simplicity, let us assume that the error terms u are normally distributed with a mean of zero and a common standard deviation σ_t.

The same equation can, of course, be estimated at date t', generating a corresponding vector of coefficients $\mu_{t'}$, and a standard error of the residuals given by $\sigma_{t'}$. We are then ready to simulate the effect of a change in the conditional distribution of X_k from t to t', by replacing the values of X_{kit} observed in the sample at time t, with:

$$X^*_{kit} = X_{-kit} \mu_{t'} + u_{kit} \frac{\sigma_{t'}}{\sigma_t} \quad [10]$$

The contribution of the change in the distribution of the variable X_k to the change in the distribution of incomes between t and t' may now be written as:

$$R_{tt'}^{x*} = D[\{X_{kit}, ^*, X_{-kit}, Y_{oht}, \Omega_{ht}\}, \beta_t, \gamma_t]$$
$$- D[\{X_{kit}, X_{-kit}, Y_{oht}, \Omega_{ht}\}, \beta_t, \gamma_t] \quad [11]$$

In this paper, we perform four regression estimations such as [9], and hence four simulations such as [10]. The four variables estimated are $X_k = \{n_{0-13}, n_{14-65}, n_{>65}, ed\}$. In the case of the education regression, the vector of explanatory variables $X\text{-}_{kit}$ was $(1, age, age^2, G_d,$ regional dummies).[13] In the case of the regressions with the numbers of household members in certain age intervals as dependent variables, the vector $X\text{-}_{kit}$ was $(1, age, age^2,$ ed, ed^2, regional dummies), where age and education are those of the household head. The simulations permitted by these estimations allow us to investigate the effects of the evolution of the distribution of educational attainment and of demographic structure on the distribution of income. We now turn to the results of the estimation stage of the model.

IV

Estimating the model

The results of the OLS estimation of equation [2] for wage earners (formal and informal) are shown in table 5. The static results are not surprising. All variables are significant and have the expected signs. The coefficients of education and their squares are positive and significant. The effect of experience (defined as [age – education – 6]), is positive but concave. The gender dummy (female =1) is negative, significant and large.

The dynamics are more interesting. Between 1976 and 1996, the earnings-education profile changed shape.

TABLE 5

Brazil: Wage earnings regression for wage-employees (equation [2])[a]

Year	1976	1981	1985	1996
Intercept	4.350	4.104	3.877	4.256
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Education	0.123	0.136	0.129	0.080
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Education^2	0.225	0.181	0.283	0.438
(x 100)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Experience	0.075	0.085	0.087	0.062
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Experience^2	-0.105	-0.119	-0.121	-0.080
(x 100)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Gender	-0.638	-0.590	-0.635	-0.493
(1 = female)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
R^2	0.525	0.538	0.547	0.474

Source: Calculations by the authors, based on the national household survey - PNAD (IBGE, 1976, 1981, 1985, 1996).

[a] The values in parentheses are P-values.

[13] The regional dummies were for the five Brazilian regions: North, Northeast, Centre-West and South, with Southeast being the default category.

After rising in the late 1970s, the linear component fell substantially from 1981 to 1996. Meanwhile, the coefficient of squared years of schooling fell up to 1981 but then more than doubled up to 1996, ending the period substantially above its initial 1976 value. Overall, the relationship became more convex, suggesting a steepening of marginal returns to education at high levels. However, when we plot the parabola which models the partial earnings-education relationship from equation [2], the lowering of the linear term dominates. The profile shifts upward from 1976 to 1981, and again to 1985, before falling precipitously (while becoming more convex) up to 1996 (see figure 2). The net effect across the entire period was a fall in the cumulative returns to education (from zero to t years) for the entire range, despite increasing marginal returns at high levels of education. The implications for poverty and inequality are clear, with the education price effect leading, ceteris paribus, to an increase in the former and a decline in the latter.

Returns to experience also increased from 1976 to 1981, and from 1981 to 1985, with a concave pattern and a maximum at around 35 years of experience (figure 3). From 1985 to 1996, however, there was a substantial decline in cumulative returns to experience, even with respect to 1976, up to 50 years of experience. The relationship became less concave, and the maximum returns moved up to around 40 years. Over the entire period, the experience price effect was mildly unequalizing (it contributed to increases in inequality until 1985, after which the effect was reversed) and had serious repercussions in terms of increased poverty.

The one piece of good news is the reduction in the male-female earnings disparity. While female earnings, controlling for both education and experience, remained substantially lower in all four years, suggesting that there
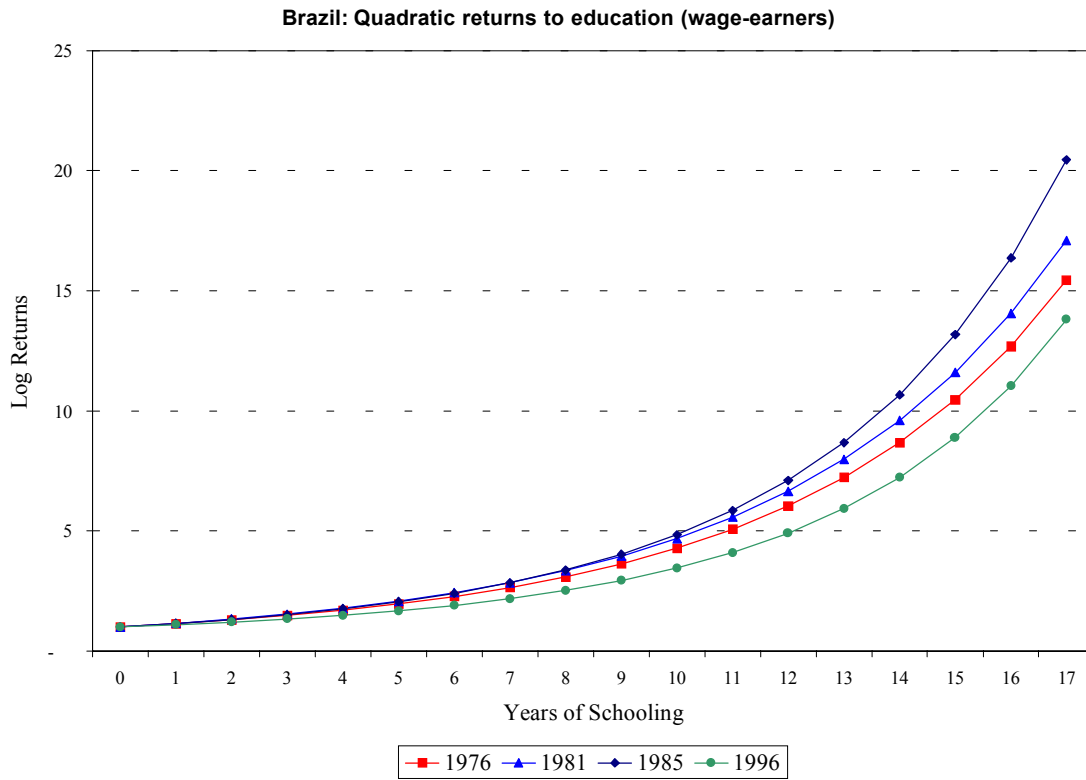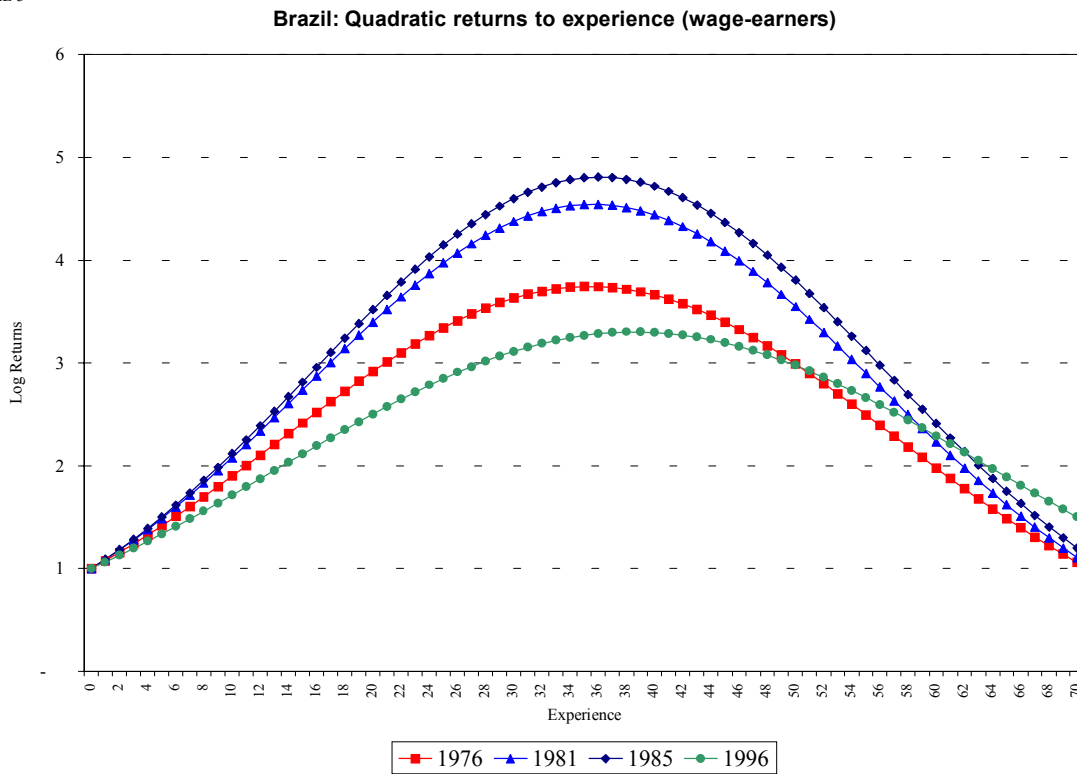
FIGURE 2

Brazil: Quadratic returns to education (wage-earners)

FIGURE 3

Brazil: Quadratic returns to experience (wage-earners)

may be some labour market discrimination, there was nevertheless a decline in this respect between 1976 and 1996. The effect of this, as we will see from the simulations reported in section V, was both mildly equalizing and poverty- reducing. Let us now turn to equation [3], which seeks to explain the earnings of the self-employed with the same set of independent variables as equation [2]. The results are reported in table 6, which shows that education is also an important determinant of incomes in the self-employed sector. The coefficient of the linear term has a higher value in all years than for wage-earners, but the quadratic term is lower. This implies that, ceteris paribus, the return to low levels of education might be higher in self-employment than in wage work, but would eventually become lower as years of schooling increase. This will have an impact on occupational choice, estimated through equation [4]. Dynamically, the same trend was observed as for wage-earners: the coefficient of the linear term fell over time, but the relationship became more convex.[14] The coefficients of experience and experience squared follow a similar pattern to that observed for wage earners. Once again, the cumulative return to experience fell over the bulk of the range from 1976 to 1996, contributing to the observed increase in poverty. The effect of being female (other things being equal) is even more markedly negative in this sector than in the wage-earning sector, though it too fell over the 1976-1996 period, despite a temporary increase in disparity in the 1980s.

A word of caution is in order before proceeding. All of the estimation results reported above refer to equations with total earnings as dependent variables. The changes in coefficients will therefore reflect changes not only in the hourly returns to a given characteristic but also in any supply responses that may have taken place. The analysis should be understood in this light.

Let us now turn to the estimation of the multinomial *logit* in equation [4]. This was estimated separately for household heads and for others, since the set of explanatory variables was slightly different in each case (see the description of vectors Z_1 and Z_i in section III above).[15]

For household heads, education was not significantly related to the likelihood of choosing to work in the wage sector vis-à-vis staying out of the labour force, at any time. The dominant effect on the occupational choices of urban household heads over this period was a sub-

TABLE 6

Brazil: Total earnings regression for the self employed (equation [3])[a]

Year	1976	1981	1985	1996
Intercept	4.319	4.192	3.853	4.250
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Education	0.196	0.148	0.165	0.114
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Education^2	−0.206	0.021	0.012	0.219
(x 100)	(0.0001)	(0.4892)	(0.6545)	(0.0001)
Experience	0.074	0.079	0.084	0.063
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Experience^2	−0.101	−0.108	−0.111	−0.082
(x 100)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Gender	−1.092	−1.148	−1.131	−0.714
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
R^2	0.431	0.434	0.438	0.336

Source: Calculations by the authors, based on the national household survey - PNAD (IBGE, 1976, 1981, 1985 and 1996).

[a] The values in parentheses are P-values.

stantial decline in the constant term affecting the probability of participating in either productive sector, as opposed to remaining outside the labour force or being unemployed. Since it is captured by the constant, this effect is not related to the educational or experience characteristics of the head, or to the endowments of his or her household. We interpret it, instead, as the effect of labour market demand-side conditions, leading to reduced participation in paid work.[16] This effect will be shown, in the occupational choice simulations reported in the next section, to increase both inequality and poverty.

For other members of the household, education did seem to raise the probability of choosing wage work vis-à-vis staying out of the labour force, with the relationship changing from concave to convex (albeit weakly so) over the period. It also enhanced the probability of engaging in own-account activities rather than staying outside the labour force in both periods, although this relationship remained concave. The number of children in the household significantly discouraged participation in both sectors, although more so in the wage–earning one. The change in the constant term was much smaller than for household heads, suggesting that negative labour market conditions hurt primary earners to a greater extent. Consequently, the effect of the occupational choices of other household members on poverty and inequality will be seen to be much milder than in the case of heads of household.

---

[14] In this case, it actually switched from concave to convex.

[15] Space constraints prevent the presentation of the tables reporting these estimations. They may be found in Appendix 3 to Ferreira and Paes de Barros (1999).

[16] In terms of the 'occupational choice' framework, these are changes in the constraints with respect to which those choices are made.

The results of the estimation of equation [9], with the educational level of individuals ten years old or older as the dependent variable, regressed against the vector $(1, age, age^2, G_d, regional dummies)$, is also given in Ferreira and Paes de Barros (1999). Over time, there is a considerable increase in the value of the intercept, which will yield higher predicted values for educational attainment, controlling for age, gender and regional location. Furthermore, the gender dummy went from large and negative to positive and significant, suggesting that women have more than caught up with males in educational attainment in Brazil over the last twenty years. The effect of individual age is stable, and regional disparities, with the South and Southeast ahead of the three central and northern regions, persist.

Regressing the number of household members in the age groups 0-13, 14-65 and over 65 (respectively) on the vector $(1, ed, ed^2, age, age^2, regional dummies)$ yields the finding that the schooling of the head has a large, negative and significant effect on the demand for children, so that as education levels rise, family sizes would tend to fall, other things being equal. Additionally, some degree of convergence across regions in family size can be inferred, with the positive 1976 regional dummy coefficients for all regions (with respect to the Southeast) declining over time and more than halving in value by 1996.

V

The simulation results

Having estimated earnings equations for both sectors of the model (wage-earners [2] and the self-employed [3]); participation equations for both household heads and non-heads [4]; and 'endowment' equations [9] for the exogenous determination of education and family composition, we are now in a position to carry out the decompositions described by Bourguignon, Ferreira and Lustig (1998). These simulations, as discussed earlier, are carried out for the entire distribution. The results are summarized in table 7, in terms of the mean household per capita income $\mu(y)$, four inequality indices (the Gini coefficient, the Theil-L index ($E(0)$), the Theil-T index ($E(1)$) and ($E(2)$)), and the standard three members of the Foster-Greer-Thorbecke class of poverty measures, $P(\alpha)$, with $\alpha = 0, 1, 2$, computed with respect to two poverty lines: an indigence line of R$30 per month and a poverty line of R$60 per month (both expressed at 1996 São Paulo metropolitan region prices).[17]

Table 7 contains information about a large number of simulated economic changes, always by applying combinations of 1996 coefficients to the 1976 population. In order to address the two puzzles posed in the Introduc-

tion –namely the increase in extreme urban poverty between 1976 and 1996 despite (sluggish) growth and (mild) reduction in inequality and the coexistence of a deteriorating labour market with stable 'headline' poverty– we now plot differences in the (logarithms) of incomes between the simulated distribution (of household incomes per capita) and that observed for 1976, for a number of the simulations in table 7.

Figure 4 plots the combined price effects (α and β) separately for wage-earners and the self-employed. As can be seen, these effects were negative (i.e., would have implied lower income in 1976) for all percentiles. The losses were greater for wage earners than for the self-employed, and for the latter they were regressive. These losses are exactly what one would have expected from the downward shifts of the partial earnings-education and earnings-experience profiles given in figures 2 and 3.

In figure 5, we adopt a different approach to the price effects, by plotting the income differences for each individual price effect simulation (for both sectors combined) and then aggregating them. As we would expect from figures 2 and 3, the returns to education and experience both tend to increase poverty. Only the change in partial returns to education is mildly equalizing (as may be seen from table 7). The change in the partial returns to experience tends to increase both inequality and poverty. The change in the intercept, calculated at the mean values of the independent variables, was also negative throughout. This proxies for a 'pure growth' effect, capturing the effects on earnings of processes not captured by edu-

---

[17] Table 7 and the remaining figures in this chapter refer to the simulation of applying the coefficients estimated for 1996 to 1976. Similar exercises were conducted for 1981 and 1985, and are reported in Ferreira and Paes de Barros (1999). Likewise, the 'return' simulation of applying the 1976 coefficients to 1996 was conducted, and the directions and broad magnitudes of the changes confirmed the results presented here.

TABLE 7

**Brazil: Simulated poverty and inequality for 1976,
using 1996 coefficients**

	Mean per capita income	Inequality				Poverty[a]					
						Z = R$30 / month			Z = R$ 60 / month		
		Gini	E(0)	E(1)	E(2)	P(0)	P(1)	P(2)	P(0)	P(1)	P(2)
1976 observed	265.101	0.595	0.648	0.760	2.657	0.0681	0.0211	0.0105	0.2209	0.0830	0.0428
1996 observed	276.460	0.591	0.586	0.694	1.523	0.0922	0.0530	0.0434	0.2176	0.1029	0.0703
Price effects											
α, β for wage earners	218.786	0.598	0.656	0.752	2.161	0.0984	0.0304	0.0141	0.2876	0.1129	0.0596
α, β for self-employed	250.446	0.597	0.658	0.770	2.787	0.0788	0.0250	0.0121	0.2399	0.0932	0.0490
α, β for both	204.071	0.598	0.655	0.754	2.190	0.1114	0.0357	0.0169	0.3084	0.1249	0.0673
α only, for both	233.837	0.601	0.664	0.774	2.691	0.0897	0.0275	0.0129	0.2688	0.1040	0.0545
All β (but not α) for both	216.876	0.593	0.644	0.736	2.055	0.0972	0.0303	0.0143	0.2837	0.1114	0.0590
Education β for both	232.830	0.593	0.639	0.759	2.691	0.0779	0.0234	0.0110	0.2531	0.0953	0.0488
Experience β for both	240.618	0.600	0.664	0.771	2.694	0.0851	0.0265	0.0125	0.2592	0.1000	0.0525
Gender β for both	270.259	0.595	0.649	0.751	2.590	0.0650	0.0191	0.0090	0.2160	0.0797	0.0404
Occupational choice effects											
γ for both sectors (both heads and other members)	260.323	0.609	0.650	0.788	2.633	0.0944	0.0451	0.0331	0.2471	0.1082	0.0671
γ for both sectors (only for other members)	265.643	0.598	0.657	0.757	2.482	0.0721	0.0231	0.0119	0.2274	0.0867	0.0454
γ, α, β for both sectors	202.325	0.610	0.649	0.788	2.401	0.1352	0.0597	0.0402	0.3248	0.1466	0.0902
Demographic patterns											
μd only, for all	277.028	0.574	0.585	0.704	2.432	0.0365	0.0113	0.0063	0.1711	0.0554	0.0264
$\mu d, \gamma, \alpha, \beta$ for all	210.995	0.587	0.577	0.727	2.177	0.0931	0.0433	0.0321	0.2724	0.1129	0.0677
Education endowment effects											
μe only, for all	339.753	0.594	0.650	0.740	2.485	0.0424	0.0136	0.0073	0.1593	0.0567	0.0287
$\mu d, \mu e$ for all	353.248	0.571	0.584	0.688	2.320	0.0225	0.0078	0.0049	0.1131	0.0359	0.0173
$\mu e, \mu d, \gamma, \alpha, \beta$ for all	263.676	0.594	0.600	0.727	1.896	0.0735	0.0374	0.0296	0.2204	0.0913	0.0561

Source: Based on the national household survey - PNAD (IBGE, 1976 and 1996).

[a] Z = poverty line.

cation, experience, gender, or the unobserved characteristics of individual workers. It is intended to capture the effects of capital accumulation, managerial and technical innovation, macroeconomic policy conditions, and other factors which are likely to determine economic growth but are not included explicitly in the Mincerian equation. Its negative effect in this simulation suggests that these factors increased poverty in urban Brazil over the period studied.

Once again, the only piece of good news comes from the gender simulation, which reports a poverty-reducing effect, as a result of the decline in male-female earnings differentials shown in tables 5 and 6. This effect, however, was far from being sufficient to offset the combined negative effects of the other price effects. As the thick line at the bottom of figure 5 indicates, the combined effect of imposing the 1996 parameters of the two Mincerian equations on the 1976 population was to substantially increase poverty.

Figure 6 plots the (logarithm) of the income differences between the distribution given by imposing the 1996 occupational choice parameters (the γ vector from the multinomial logit in equation [4]) on the 1976 population, and the observed 1976 distribution. It does so both for all individuals (the lower line), and for non-heads (the upper line). The effect of this simulated change in occupational choice and labour force participation behaviour is to greatly increase poverty and inequality, as the relevant indices in table 7 confirm. It suggests the existence of a group of people who, by voluntarily or involuntarily leaving the labour force, swelling the ranks of the unemployed, or being relegated to very ill-remunerated occupations, probably in the informal sector, are becoming increasingly impoverished.

By combining the negative price and occupational choice effects, we get an idea of the overall effect of Brazil's urban labour market conditions over this period. This is shown in figure 7, where the lowest curve plots the differences between the household per capita incomes from a distribution in which all αs, βs and γs change, and the observed 1976 distribution. It shows the substantially poverty-augmenting (and unequalizing) combined effect of changes in labour market prices and occupational choice parameters on the 1976 distribution.

FIGURE 4

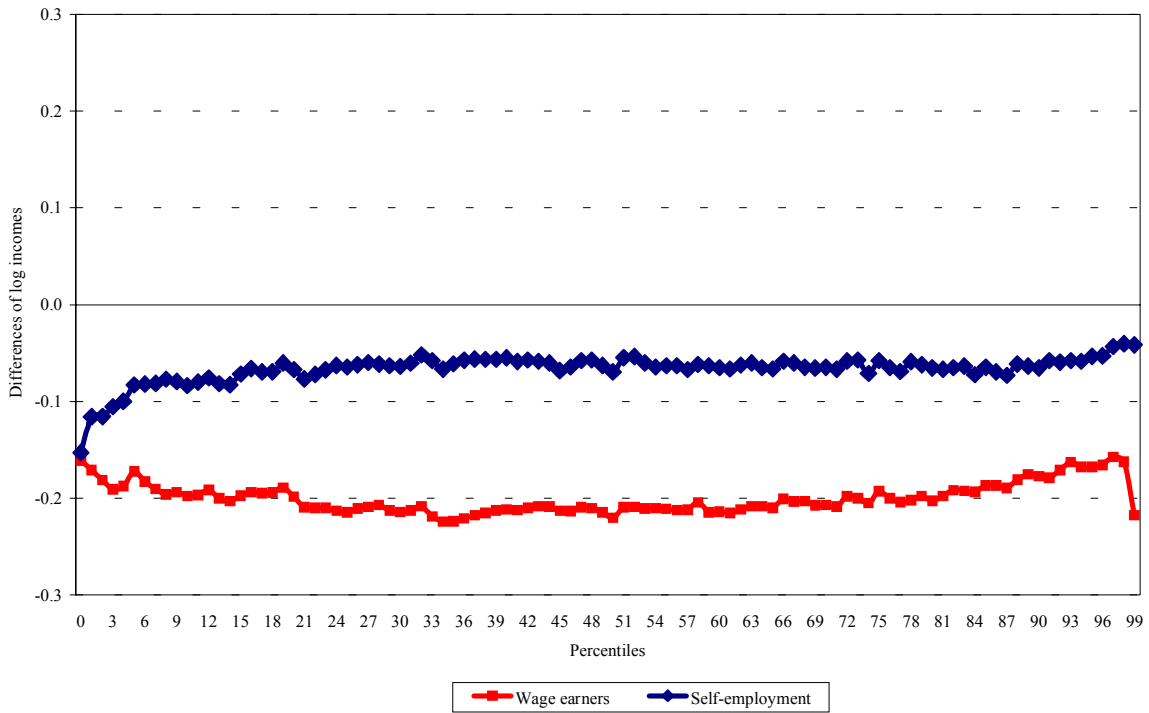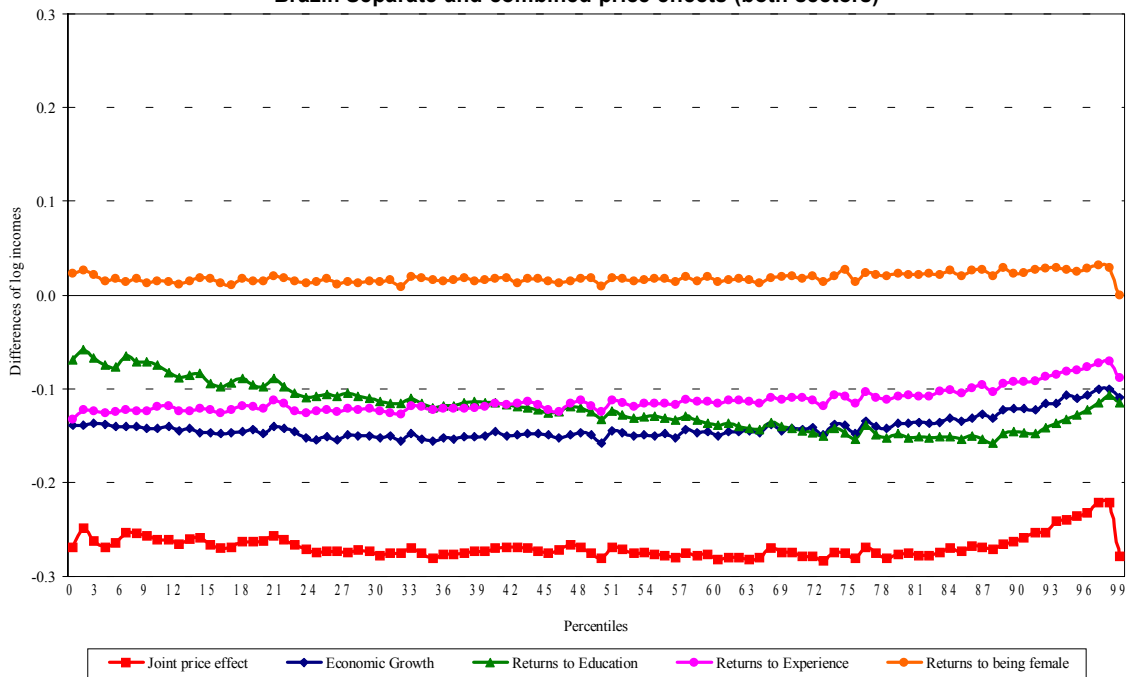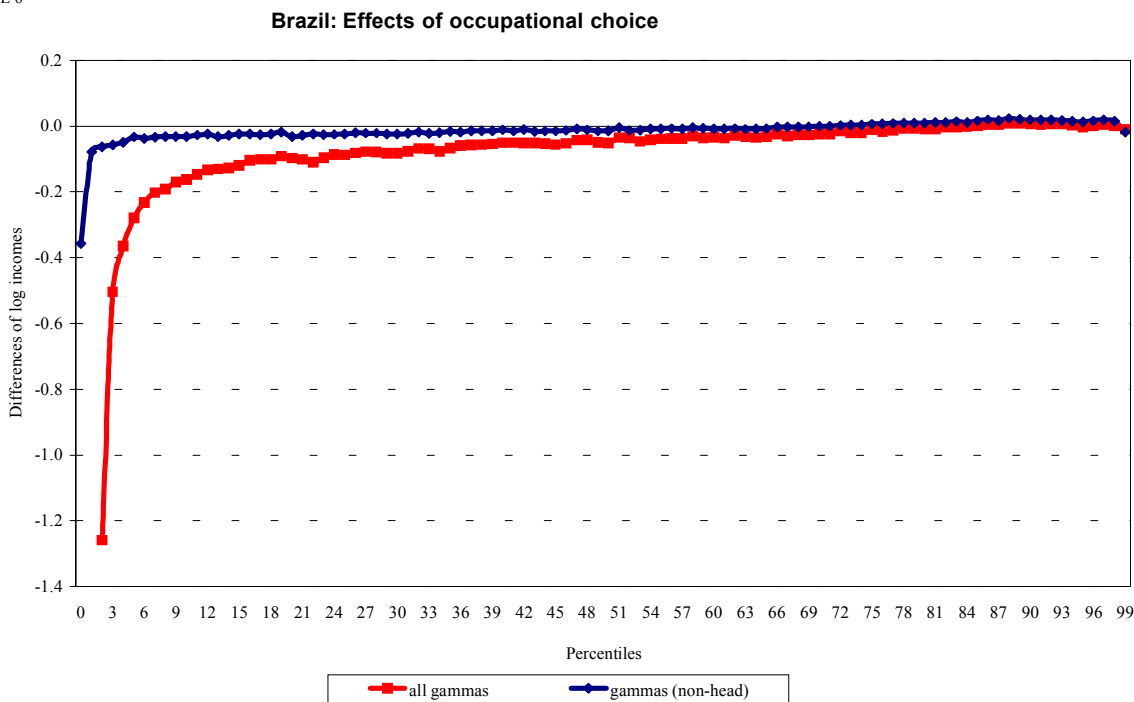**Brazil: Combined price effects, by sector**

FIGURE 5

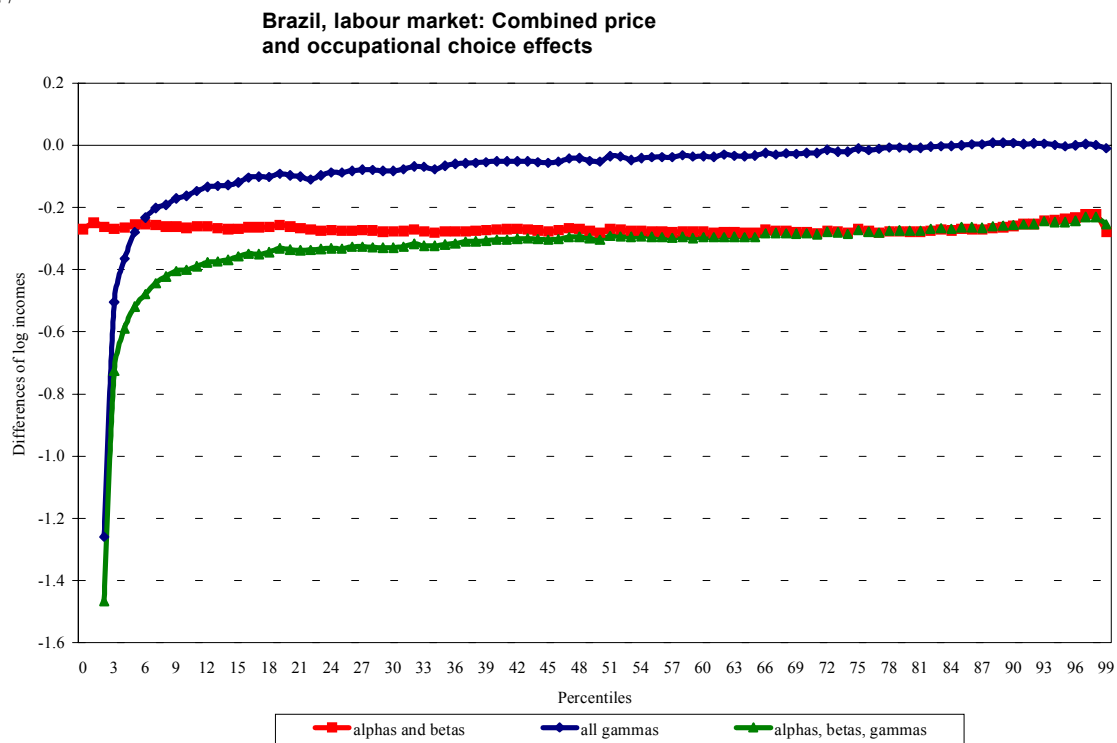**Brazil: Separate and combined price effects (both sectors)**

Source: IBGE, 1976 and 1996.

FIGURE 6

Brazil: Effects of occupational choice

Source: IBGE, 1976 and 1996.

FIGURE 7

**Brazil, labour market: Combined price
and occupational choice effects**

Source: IBGE, 1976 and 1996.

At this point, the second puzzle can be stated clearly: given these labour market circumstances, what factors can explain why mean incomes rose, headline poverty did not rise, and inequality appears to have fallen slightly? The first part of the answer is shown graphically in figure 8, where the upper line plots the differences between the (log) incomes from a distribution given by imposing on the 1976 population the transformation (10) for the demographic structure of the population. The changes in the parameters μ_d (and in the variances of the residuals in the corresponding regression) have a positive effect on incomes for all percentiles, and in an equalizing manner. However, when combined in a simulation in which the values of all αs, βs and γs also change, it can be seen that the positive demographic effect is still overwhelmed. Nevertheless, it is clear that the reduction in dependency ratios, and hence in family sizes, in urban Brazil over this period had an important mitigating effect on the distribution of incomes.

There remains one final piece of the puzzle which is needed in order to explain why the deterioration in labour market conditions did not have a worse impact on poverty. This, as should be evident from the increase in mean years of effective schooling registered in table 1, is the shift to the right in the distribution function of education. This is shown in figure 9, which reveals that gains in educational attainment were particularly pronounced at lower levels of education, and thus presumably among the poor.

A gain in educational endowments across the income distribution, but particularly among the poor, has both direct and indirect effects on incomes. The direct effects are through equations [2] and [3], where earnings are positive functions of schooling. The indirect effects are both through the occupational choices that individuals make, and through the further impact that education has on reducing the demand for children, and hence family size. A simulation of the effect of education is thus quite complex.[18] After it is completed, one observes (figure 10) a rather flat improvement in (log) incomes across the distribution (i.e., a scaling effect). When this is combined with changes in the parameters of the demographic equations, however, the effect becomes more marked, and is not only more poverty-reducing, but also mildly equalizing. The bottom line

---

[18] Note that the different effects are not simply being summed. The effect of greater educational endowments is simulated through every equation where it appears in the model, thereby affecting fertility choices and occupational statuses, as well as earnings.

FIGURE 8

Brazil: Demographic effects
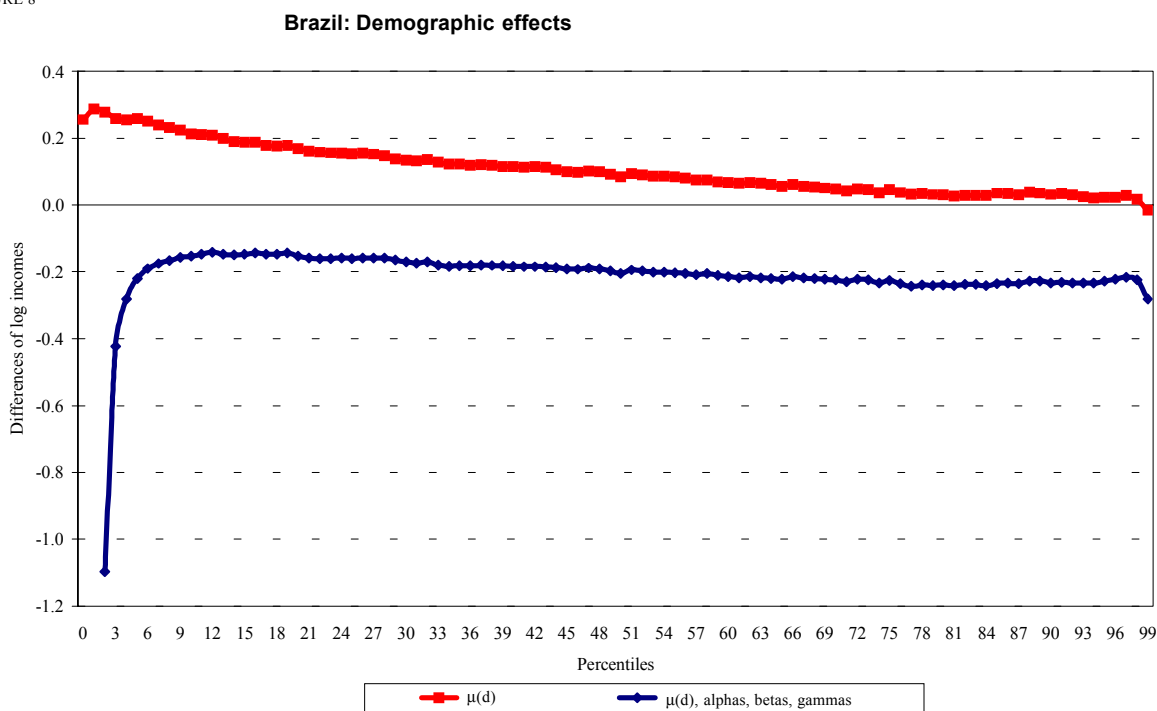
Source: IBGE, 1976 and 1996.

FIGURE 9

Brazil: Shift in distribution of schooling, 1976-1996

Years of Schooling

1976　1996
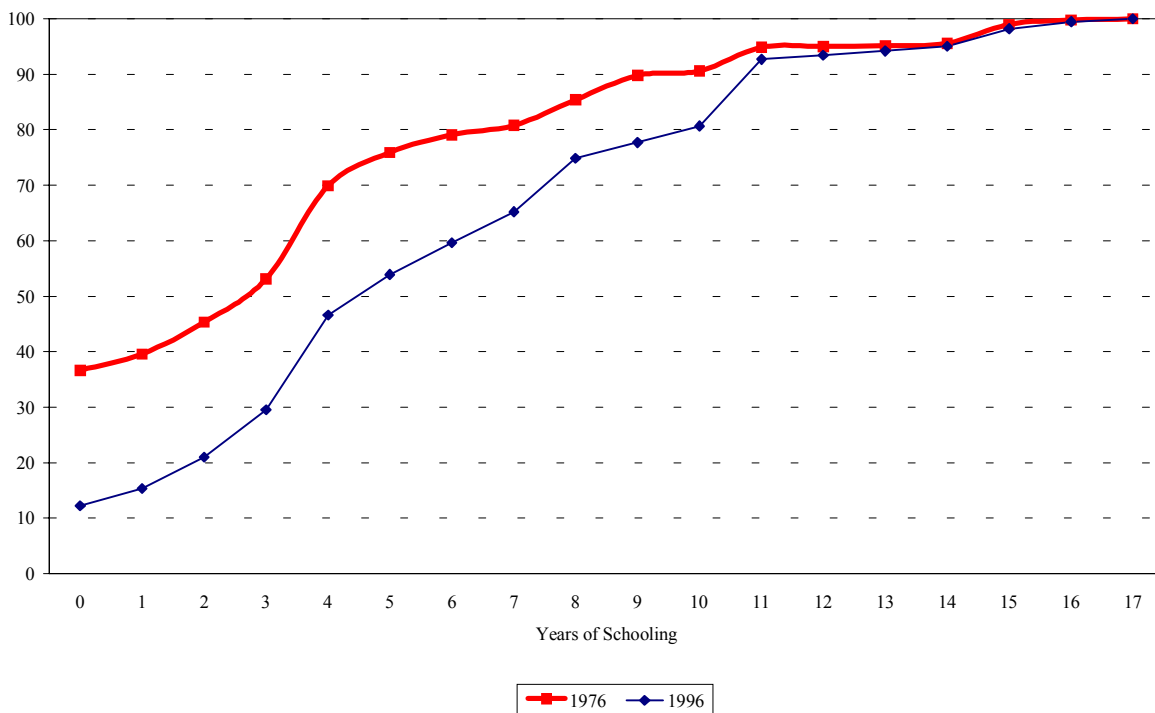
FIGURE 10

Brazil: Educational endowments and demographic effects

Differences of log incomes

Percentiles

μ(e)　　μ(d) and μ(e)　　μ(d), μ(e), alphas, betas, gammas
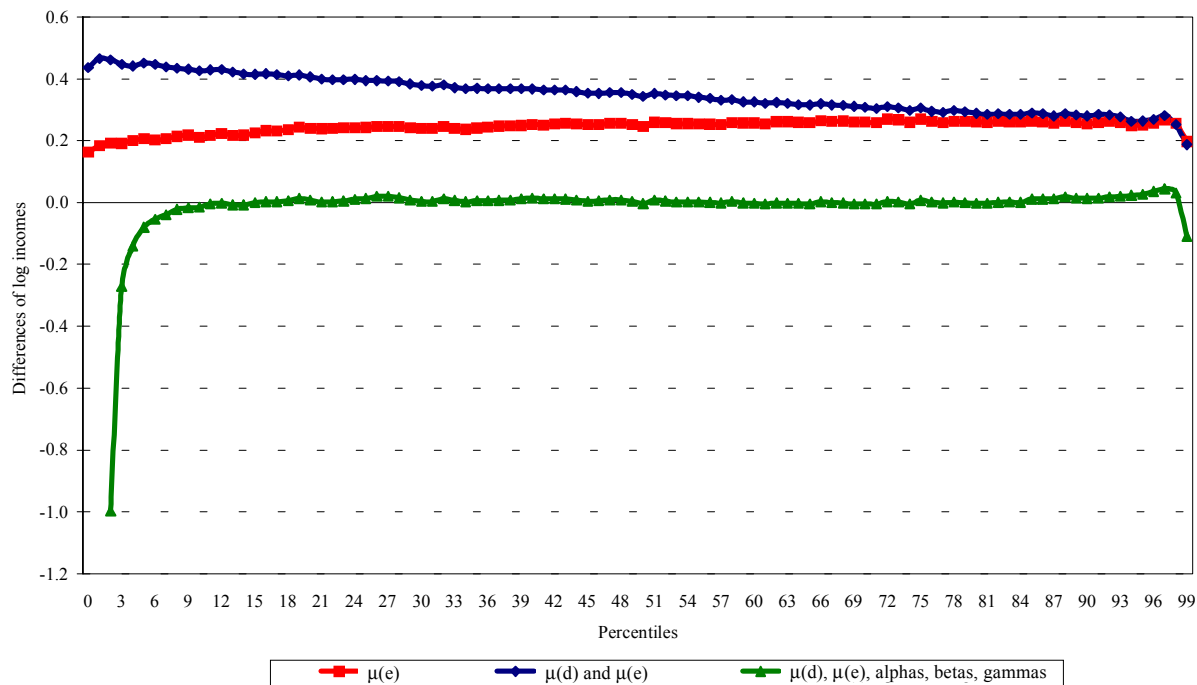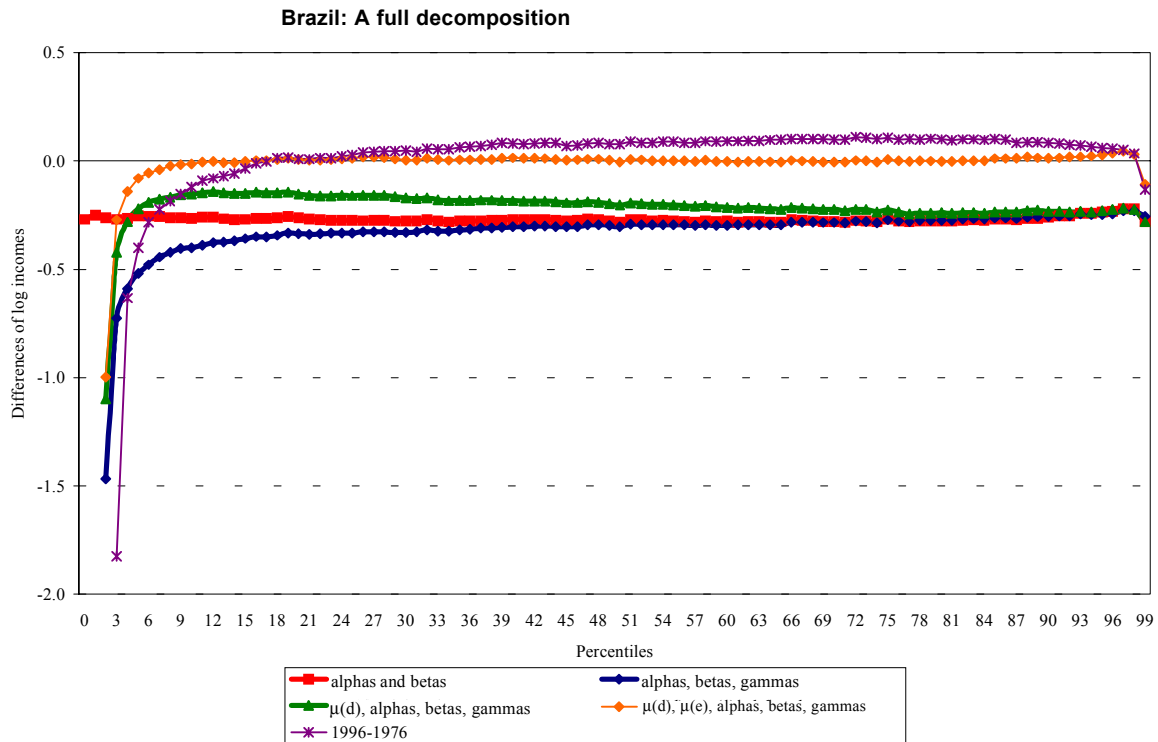
Source: IBGE, 1976 and 1996.

in figure 10, in keeping with this pattern, combines both of these effects with the changing αs, βs and γs. The result is striking: this complex combined simulation suggests that all of these effects, over twenty turbulent years, cancel each other out almost exactly from the 15th percentile up. This explains the smallness of the changes in headline poverty. From around the 12th percentile down, however, the simulation suggests a prevalence of the negative occupational choice (and to a lesser extent, price) effects, with substantial income losses. These account for the rise in indigence captured by the R$30/month poverty line.

The bottom line in figure 10 is, in a sense, the final attempt by this methodology to simulate the various changes leading from the 1976 to 1996 distribution. Figure 11 is a graphical test of the approach. Here, the line denoted "1996-1976" plots the differences in actual (log) incomes between the observed 1996 and the observed 1976 distributions. Along with this, we also plotted every (cumulative) stage of our simulations. First the impoverishing (but roughly equal) price effects; then these combined with the highly impoverishing occupational choice effects; then the slightly less bleak picture arising from a combination of the latter with the parameters of the family size equations, and finally, the curve plotting the differences between the incomes from the simulation with all parameters changing, and the observed 1976 figures. As may be seen, the last line appears to replicate the actual differences fairly well. Of course, the point of the exercise is not to replicate the actual changes perfectly, but rather to learn the different effects of different parameters, and possibly to infer any policy implications from them. Nevertheless, the success of the last simulation in approximately matching the actual changes does give some extra confidence in the methodology and in any lessons we may derive from it.

FIGURE 11

Brazil: A full decomposition

Source: IBGE, 1976 and 1996.

VI

Conclusions

In the final analysis, does this exercise help us improve our understanding of the evolution of Brazil's urban income distribution over this turbulent twenty-year period? Whereas many traditional analysts of income distribution dynamics might have inferred, from the small changes in mean income, in various inequality indices, and in the incidence of poverty,[19] that there was little –if anything- to investigate, digging a little deeper has unearthed a wealth of economic factors interacting to determine substantial changes in the environment faced by individuals and families, and in their responses.

In particular, we have found that, despite a small fall in measured inequality (although the Lorenz curves cross) and a small increase in mean income, extreme poverty has increased, whether measured by sufficiently low poverty lines or sufficiently high poverty aversion parameters. This seems to have been caused by outcomes related to participation decisions and occupational choices, in combination with declines in the labour market returns to education and experience. These changes are associated with greater unemployment and informality, as one would expect, but more research into them seems necessary. While we seem to have identified the existence of a group excluded from both the productive labour markets and any substantive form of safety net,

we have not been able to fully interpret the determinants of their occupational choices. Issues of mobility – exacerbated by the fact that the welfare indicator is based on current monthly income– also call for further research in this context. The policy implications would seem to indicate the desirability of targeted labour programmes, or other safety nets, but it would be foolhardy to go into greater detail until the profile of the group which seems to have fallen into extreme poverty in 1996 is better understood.

Secondly, we have found that, even above the 15th percentile, where urban Brazilians have essentially 'stayed put', this was the result of some hard climbing along a slippery slope. They had to gain an average of two extra years of schooling (which still leaves them under-educated for the country's per capita income level), and substantially reduce fertility, in order to counteract falling returns in both the formal labour market and in self-employment.

It may well be, as many now claim, that an investigation of non-monetary indicators –such as access to services, or life expectancy at birth– should lead us to consider the epithet of 'a lost decade' as too harsh for the 1980s. Unfortunately, however, we find that if one is sufficiently narrow-minded to consider only welfare measured in terms of money, urban Brazil has in fact experienced not just one, but two, lost decades.

(Original: English)

---

[19] With respect to the already low R$ 60/month poverty line, by historical standards for Brazil.

Bibliography

Almeida dos Reis, J. G. and R. Paes de Barros (1991): Wage inequality and the distribution of education: A study of the evolution of regional differences in inequality in metropolitan Brazil, *Journal of Development Economics*, vol. 36, N° 1, Amsterdam, The Netherlands, Elsevier Science Publishers, B.V.

Bacha, E. and L. Taylor (1980): Brazilian income distribution in the 1960s: "Facts", model results, and the controversy, in L. Taylor, E. Bacha, E. Cardoso and F. Lysy, *Models of Growth and Distribution for Brazil*, Washington, D.C., World Bank.

Bonelli, R. and G. L. Sedlacek (1989): Distribuição da renda: evolução no ultimo quarto de seculo, in G.L. Sedlacek and R. Paes de Barros, *Mercado de trabalho e distribuição*

da renda: uma coletanea, Serie monográfica, N° 35, Rio de Janeiro, Institute of Applied Economic Research (IPEA).

Bourguignon, F., F. Ferreira and N. Lustig (1998): The microeconomics of income distribution dynamics in East Asia and Latin America, *World Bank Research Proposal*, Washington, D.C., World Bank, April.

Buhmann, B., L. Rainwater, G. Schmaus and T. Smeeding (1988): Equivalence scales, well-being, inequality and poverty: Sensitivity estimates across ten countries using the Luxembourg income study database, *Review of Income and Wealth*, vol. 34, N° 2, New Haven, Connecticut, International Asociation for Research on Income and Wealth.

Deaton, A. and C. Paxson (1997): *Poverty Among Children and the Elderly in Developing Countries*, Working paper, N° 179, Princeton, New Jersey, Princeton University, Research Program in Development Studies.

Duryea, S. and M. Szekely (1998): *Labor markets in Latin America: A supply-side story*, paper presented at the IDB/IIC Annual Meeting, Cartagena de Indias, Colombia, Inter-American Development Bank (IDB)/Inter-American Investment Corporation (IIC).

Ferreira, F. and R. Paes de Barros (1999): The slippery slope: Explaining the increase in extreme poverty in urban Brazil, 1976-1996, *The Brazilian Review of Econometrics*, vol. 19, N° 2, Rio de Janeiro, Brazilian Econometrics Society, November.

Fishlow, A. (1972): Brazilian size distribution of income, *The American Economic Review*, vol. 62, N° 2, Nashville, Tennessee, American Economic Association.

Foster, J., J. Greer and E. Thorbecke (1984): A class of decomposable poverty measures, *Econometrica*, vol. 52, N° 3, Evanston, Illinois, Econometric Society.

Hoffman, R. (1989): Evolução da distribuição da renda no Brasil, entre pessoas e entre familias, 1979/86, in G.L. Sedlacek and R. Paes de Barros, *Mercado de trabalho e distribuição da renda: uma coletanea*, Serie monográfica, N° 35, Rio de Janeiro, Institute of Applied Economic Research (IPEA).

IBGE (Brazilian Geographical and Statistical Institute) (1976): *Pesquisa Nacional por Amostra de Domicilios*, Brasilia, Fundação IBGE.

——— (1981)**:** *Pesquisa Nacional por Amostra de Domicilios*, Brasilia, Fundação IBGE.

——— (1985): *Pesquisa Nacional por Amostra de Domicilios*, Brasilia, Fundação IBGE.

——— (1996): *Pesquisa Nacional por Amostra de Domicilios*, Brasilia, Fundação IBGE.

Juhn, C., K. Murphy and B. Pierce (1993): Wage inequality and the rise in returns to skill, *Journal of Political Economy*, vol. 101, N° 3, Chicago, Illinois, The University of Chicago Press.

Langoni, C.G. (1973): *Distribuição da renda e desenvolvimento econômico do Brasil*, Rio de Janeiro, Expressão e Cultura.

Mookherjee, D. and A. F. Shorrocks (1982): A decomposition analysis of the trend in UK income inequality, *The Economic Journal*, vol. 92, N° 368, Oxford, U.K., The Royal Economic Society.

Ramos, L. (1993): *A distribuição de rendimentos no Brasil: 1976/85*, Rio de Janeiro, Institute of Applied Economic Research (IPEA).

Structural changes and *productivity in Latin American* industry, 1970-1996

Jorge Katz

Director,
Division of Production,
Productivity and Management,
ECLAC
jkatz@eclac.cl

This article analyses the structural changes in Latin American industry, which speeded up in the 1990s with the consolidation in the region of the external openness programmes, the deregulation of many markets, and the privatization of major sectors of industrial activity which had previously been dominated by State enterprises. The branches of manufacturing which have turned in the best relative performances over the last twenty years are natural resource-based industries producing staple industrial commodities, industries assembling computers, video equipment, television sets or clothing, and the motor industry, which has been given preferential treatment in government economic policy. In contrast, industries producing labour-intensive final goods, those making intensive use of technological knowledge and new product design engineering, or those producing heavy capital goods have been losing relative weight. The pattern of production specialization and the places occupied in world markets for manufactures have clearly been changing, with greater emphasis on utilization of the natural comparative advantages of the region (i.e., its abundant natural resources) or on sectors which have been given special treatment in industrial policy. The article explores the behaviour of the Latin American industrial structure in terms of productivity, comparing it with that of the developed countries. Using the case of the United States as a reference universe, it estimates the labour productivity gap with respect to that country and evaluates the performance of countries and industrial branches in the region as a function of that parameter.

I

A general overview

In previous studies I already addressed the big changes undergone by the Latin American industrial structure in the 1980s and 1990s[1]. It is clear from those studies that that process of structural change speeded up in the 1990s with the consolidation in the region of the external openness programmes, the deregulation of many markets, and the privatization of major sectors of industrial activity which had previously been dominated by State enterprises. Although the studies in question do not analyse the details of each national case, they clearly show that two clearly differentiated patterns of production specialization and insertion in the world markets for manufactures have been growing up in the region. On the one hand, in the southern part of the region, and especially in Argentina, Chile, Brazil and Uruguay, the natural resource-based industries producing staple industrial commodities (such as iron and steel, petrochemical products, non-ferrous minerals, fish meal, vegetable oils, pulp and paper, etc.) have turned in a better relative performance over the last twenty years. On the other hand, in Mexico and the smaller Central American countries the pattern of production specialization has decisively favoured assembly industries (such as those assembling, computers, video equipment, television sets and clothing).

In the first case, the industrial branches in question produce highly standardized intermediate goods for which great locally-owned conglomerates –together with a small number of foreign corporations– have installed extremely modern capital-intensive processing plants using process technologies matching up to the best international levels. On the basis of such plants, the countries in question have turned themselves within the space of only a few years into major exporters of industrial commodities, which they sell on highly competitive markets where the Latin American firms act as price-takers, have little bargaining power, and have only small unit profit margins.

In the second case –that of assembly industries– the assembly plants are also extremely modern and close to the best international levels of the industry, operate in accordance with complex "just in time" production logistics, and are essentially designed to supply the United States market. In this case, the main comparative advantage is the low real wages paid in the countries of the region, which make it possible to successfully compete with Japan and South Korea in the United States domestic market.

In contrast with the above two cases, there has been a decline in the relative position occupied in manufacturing by industries producing labour-intensive final goods (footwear, clothing, etc.) and those that make intensive use of technological knowledge and new product design engineering to produce such items as capital goods, pharmaceutical and chemical inputs and scientific instruments. In the case of labour-intensive goods, the opening-up and deregulation of the region's economies have meant that they must compete with similar goods from countries where wages are much lower (such as China), while knowledge-intensive goods have not been able to compete in areas of production where technological obsolescence is rapid and goods have a short useful life but demand heavy outlays on technological research. At the international level, these industrial sectors have been rapidly incorporating both the use of microprocessors, digitalization and numerical control and new types of knowledge based on genetics and biotechnology: all fields where the relative backwardness of Latin American firms has become more marked in recent years, so that their international competitiveness has gone down even further.

The motor industry, for its part, has enjoyed special tariff and fiscal treatment both in the Southern Cone countries and those around the Gulf of Mexico. This industry has also made significant progress in technological modernization and has come to have greater relative weight within the production structure.

We have seen that the pattern of production specialization and the form of insertion in world markets for manufactures have favoured highly capital-intensive industries, such as those exploiting the natural comparative advantages of the region (its abundant forestry, fishery, oil, gas, mineral and tourist resources), or else

[1]See Katz (1997) and Katz, Benavente, Crespi and Stumpo (1997).

labour-intensive sectors which, as in the case of the assembly industries, reflect the global competitive strategies of some of the major transnational corporations which have decided to develop a big export platform in the area around the Gulf of Mexico. This "new" pattern of production and trade specialization has many pros and cons –for example, as regards the generation of new jobs and its impact on the external accounts of the economy– but these will not be dealt with in particular in the present article.

The above-mentioned changes have been taking place side by side with equally complex changes in the morphology, behaviour and relative weight of the different branches of production within each individual economy. Profound changes are taking place both in the regulatory and institutional framework of each sector of production (including rights of ownership over natural resources, patent legislation, and labour laws) and in the sets of actors participating in the process of structural change. Many firms have moved out of the market, there have been innumerable purchases and mergers of enterprises, and those which have managed to survive have been undergoing marked changes in the organization and planning of their production, their market strategies, and their technological and international marketing capabilities. The new production organization models display less vertical integration (that is to say, they involve more local and international subcontracting) and they are more conditioned by the external environment, as many firms now work "on line" with their international technology suppliers and process licensors and hence make use of fewer local engineering efforts.

This process, which began to take shape in the 1980s, in the midst of the external debt crisis and the slowdown in domestic market-based growth, gathered momentum in the 1990s in line with the increasingly competitive climate in the countries of the region and with the consolidation of a new structure of relative prices in each economy. This structure may be considered as being closer to the "true" opportunity cost of the domestic resources used than the structure that prevailed during the import substitution period and, in this respect, less likely to favour the use of local technological capabilities.

This process led to a major change in the sources and nature of the technological changes that the various sectors of production were incorporating. Technological advances of outside origin and private expenditure on research and development came to predominate over domestic efforts. Whereas the import substitution model had promoted the local production of capital goods and the use of technologies and engineering services of do-

mestic origin, the opening up of the economies to the exterior made imported capital goods cheaper and made it easier to obtain licences and technical assistance from abroad: a situation which became still more marked with the strengthening of intellectual property rights, which increased the propensity of the local agents of production to fill their needs with equipment and technology from the developed countries and the interest of foreign firms in licensing their products and production technologies.

As a result, foreign direct investment flows have significantly increased in the areas of both manufacturing and services and have been an important means of access to new technologies in the fields of products, processes and organization of production. Outstanding among the new foreign players who have entered on the regional scene in recent years are some big firms with extensive experience which were generally public enterprises in their home countries, where they operated in such fields as telecommunications, energy, transport, water supply, etc. Their incorporation into the local production environment has been accompanied by technological modernization processes and improvements in productivity, both in the areas in which they themselves operate and in the sectors using their services (i.e., upstream and downstream of their place in the production structure). On the other hand, the labour-saving bias implicit in the new technologies has become still more marked, thus influencing the industrial sector's growing difficulty in generating new jobs in line with the growth of the economically active population.

In addition to the foregoing, it may be noted that the external trade balance of the new manufacturing sector has been becoming increasingly negative because of the growing weight of imports of machinery and production equipment, vehicles, and items in the area of electromechanics and electronics, in contrast with the slower growth rate of exports of industrial commodities.

This concludes our very brief description of past developments. We shall now see how this "new" structure of production has performed in terms of labour productivity, and especially, how it has performed in comparison with the developed world, using the United States as a "reference universe" against which to weigh the performance of the countries and industries of the region.

The ideal situation would be to work with indicators of total productivity: i.e., the whole set of production factors used in the economy. Because of the absence of data on gross capital formation at the level of industrial branches, however, we have opted to concentrate

on what occurred in the area of labour productivity. The information presented here corresponds to nine countries of the region and 27 industrial branches, defined at the three-digit level of the International Standard Industrial Classification (ISIC) for the period from 1970 to 1996. We will give special attention to the 1990-1996 stage, since those years witnessed an increase in the efforts at external openness, deregulation and privatization of economic activities in the various countries of the region. We will begin at the aggregate level –that is to say, analysing the evolution of labour productivity for manufacturing as a whole in each country– and then go on to analyse the different sectoral situations in order to identify differences between the performance patterns of the different countries and industrial branches, as compared with the "reference universe".

As we shall see below, the data show that behaviour was heterogeneous both between countries and between branches of activity. Some countries and industrial branches show clear signs of coming closer to the labour productivity of the United States, whereas others have markedly lost ground.

In section II below we will present the basic estimators obtained in the course of this study, while in section III we will seek to find a possible explanation for what happened. This leads us to the study of the behaviour of the agents of production, the changes in this behaviour over time, and the historical and institutional context in which the agents have acted.

The current literature offers various different approaches to the study of the factors determining improvements in factor productivity. On the one hand, there are explanations of a neoclassical nature expressed through "growth accounting", based on the contributions of R. Solow and others in the 1950s (Solow, 1957) and liberally re-used by economists more recently in the context of modern growth theory (see for example Barro and Sala-i-Martin (1996) and Rommer (1986 and 1992)).

In these authors, the analysis is based on conventional neoclassical assumptions of perfect information, fully specified generic production functions, "well-behaved" firms, equilibrium paths, perfectly competitive markets, and factors paid according to their marginal productivity.

In this type of conceptual framework, the explanation of how productivity grows with time is extremely stylized and does not allow of any differences of behaviour between firms in the same sector, does not offer any opportunity for the discussion of alternative business strategies in the context of a given set of exogenous data, and does not allow for the presence of mar-

ket failures or different processes of accumulation of experience and learning by firms in a given branch of production. The model is an extremely simple specification of microeconomic behaviour, the operation of the markets, and the long-term growth of a given society. In this stylized version, historical and institutional variables play almost no part, except through their influence on relative factor prices (Katz and Kosacoff, 1998). It is precisely this drastic simplification of real conditions, with the elimination of all market imperfections, of the uncertainty and asymmetrical information that can affect the behaviour of economic agents, and of the differences between generations as regards their propensity to save and consume (Solow, 1988), which makes it desirable to add individual equilibrium forms of behaviour in order to obtain macroeconomic investment, production or consumption functions with which to describe the aggregate behaviour of the economy.

Modern studies on growth theory, however, incorporate the possibility of increasing returns to scale and externalities at the branch level, without this involving any contradictions with the basic logic of the competition model, since it is possible to continue assuming that each agent complies of his own free will with the basic requisites of the behaviour model, but there are also circumstances foreign to each particular agent to which the existence of increasing returns to scale and externalities may be attributed. Even so, the institutional complexities and the question of uncertainty underlying the behaviour of enterprises in terms of technology and innovation are still not properly identified by the theory, as shown in various recent studies by "evolutionist" authors such as Nelson (1997).

However, in the course of the last two decades a number of economists (following a different analytical path from that referred to above) have tried to explore the issues of innovation and productivity in a different conceptual framework of a more classical and evolutionary nature (Nelson and Winter, 1982; Dosi (ed.), 1988; Freeman, 1994, and Metcalfe, 1997, among others). These authors seek to explain improvements in productivity not only as the result of forms of behaviour that maximize equilibrium, but also as the consequence of a process of "natural selection" in which competition plays a crucial role. The success or failure of different enterprises, imperfect information, uncertainty, and differences in strategy between enterprises competing in the same market are central features of a long-term process of "purification" involving both "mutations" of an almost genetic nature (Nelson, 1997) and historical, cultural and institutional factors –genotypes and fenotypes,

in Nelson's "biological" language– which shape the path of productivity growth over time. Here, growth is not seen as the result of equilibrium-based forms of behaviour but forms a "cultural construct" allowing of more Schumpeterian-type processes of "creative destruction" which, by definition, goes beyond the static Paretian framework of the conventional neoclassical model. The explanatory mechanism has a historical and institutional background, a basic substrate of uncertainty and market flaws, which does not fit into the neoclassical metaphor. Whereas the latter is firmly rooted in conventional price theory, the "evolutionary" analytical discourse has a large component of "cultural anthropology" which goes far beyond the ambit of our typical textbook microeconomics.

The entry and exit of firms from the market, mergers, changes in strategy by firms, and the gradual changes taking place in what we will call here the "sectoral competitive regime" are key elements for understanding the reasons why the average productivity of a sector of production improves over time. The historical and institutional context plays a central role, conditioning what the agents of production want to do, know how to do and can do, unlike the context of the neoclassical model, in which those agents always know everything they need to know and are perfectly aware of what they should do. In this latter context, forms of conduct are automatic responses to a given set of exogenous data. In the first-named context, however, they are "adaptative": that is to say, they are based on trial and error, and competition acts as a "selection mechanism" which rewards certain decisions and punishes others.

The present study comes within this latter analytical field and seeks to study the changes undergone by labour productivity in Latin American industry in the context of the structural changes and changes in patterns of specialization referred to earlier: that is to say, in the context of the changes towards industries processing natural resources, assembly industries, and sectors –such as the motor industry– which have managed to obtain preferential forms of treatment from the economic authorities which have protected them from the otherwise across-the-board openness of the economy to the exterior.

In particular, we are interested in analysing the evolution of labour productivity in industry in the 1990s and the differences registered in the various countries and industrial branches in this respect compared with the past: that is to say, compared with the import substitution stage.

Throughout this study, the analysis will be essentially at the meso-economic level, to which end we will examine a set of inter-temporal and cross-sectional estimates of industrial activity at the three-digit level of aggregation. This will enable us to offer a first explanatory hypothesis of what happened over the 1970-1996 period and of the rising trend in labour productivity in the region –although not in every one of the countries or every one of the industries– in the 1990s. We believe that although this analysis at the meso-economic level is useful and necessary, it is not of itself sufficient: it is only the first step in an explanatory chain which must necessarily be continued at the level of the individual firms, although for reasons of space this will not be possible in this study. It seems to us to be essential to progress towards a more complex and detailed level of explanation which will help us to understand, for example, the role played by various big domestic economic groups and also various transnational corporations in the different countries. Ultimately, this will enable us to understand the reasons for the different long-term performances of the countries of the region and the particular features of the process of restructuring of production which have been taking shape in each case after the recent efforts at greater external trade openness and market deregulation.

The sectors of production which have shown the greatest capacity to survive in this external openness process are those which come closest to the static comparative advantages of each economy, and also those which, through political lobbying, have secured special treatment from governments (such as the motor industry or assembly industries). In these sectors, levels of investment were maintained even against the background of the severe drop in aggregate investment in the 1980s. This led to the formation of a new pattern of production specialization and new forms of insertion in the world markets for manufactures which emphasize the production of highly standardized industrial commodities and a few branches of light industry (assembly activities) in which various transnational corporations have played a leading role as sources of the new product designs, new process technologies and international marketing channels which have allowed a number of countries of the region to compete on developed country markets.

Our explanatory hypothesis suggests that the new pattern of production specialization of the countries of the region has been determined by their static comparative advantages, the lingering effects of the industrial policies applied in the import substitution period and, finally, by the strategies of a few big transnational corporations which have decided to use a number of countries located around the Gulf of Mexico as export platforms to the United States market.

II

Evolution of labour productivity in

Latin American industry, 1970-1996

This section gives estimates of the growth rates of Latin American labour productivity, first for the manufacturing sector as a whole in nine countries of the region, and then for 27 industrial branches (at the three-digit ISIC level) in five countries: Argentina, Brazil, Chile, Colombia and Mexico. In the calculations we will use the PADI[2] data base, which includes series for value added at current and constant prices, employment, wages paid, labour productivity (measured as value added per man/year), unit labour costs and gross unit margins, for the period from 1970 to 1996. On the basis of this information it is possible to make some previously non-existent comparisons of labour productivity between countries of the region and between them and the United States.

1. Labour productivity in the industrial sector

We will begin by presenting labour productivity indicators for the manufacturing sector as a whole in nine countries of the region and the United States, first covering the period from 1970 to 1996 and then the 1990-1996 stage, when there were clear indications of a significant speeding-up of the growth rate of the product per person employed in several of the countries examined. For the period as a whole (i.e., 1970-1996), we see that only three countries of the region (Argentina, Colombia and Mexico) attained growth rates of labour productivity in industry higher than those registered in the United States manufacturing sector (table 1).

This indicates that, although in absolute terms the distance is still large, the relative labour productivity gap in manufacturing between those three countries and the United States has tended to narrow: after starting in the 1970s from absolute levels which were only about 30% of average United States industrial labour productivity (even less in the case of Colombia), we see that by the end of the period Argentine industry had almost doubled its productivity, while Colombia and Mexico registered somewhat smaller but nevertheless significant improvements (table 2).

Of the other countries studied, Brazil registered substantial improvements in the growth rate of labour productivity between 1990 and 1996, although it had made little progress in closing the relative labour productivity gap in the previous two decades. In the five remaining countries (Chile, Costa Rica, Jamaica, Peru and Uruguay) this gap has either remained more or less constant or even tended to widen. Although Chile doubled the growth rate of its industrial labour productivity after opening up its economy, its position over the period as a whole remained more or less static, without much change in comparison with its initial situation. Peru, Uruguay and Jamaica, for their part, clearly lost ground in terms of labour productivity compared with the United States.

2. Labour productivity by branches of manufacturing in Argentina, Brazil, Chile, Colombia and Mexico

The fact that the relative gap in manufacturing labour productivity narrowed somewhat in Argentina, Brazil, Colombia and Mexico at the aggregate level does not mean, of course, that the performance in this respect was the same in each and every one of their industrial branches. In order to analyse the differences in behaviour between industrial branches in greater detail for these five countries, we made identical estimates of their relative labour productivity compared with similar branches in the United States, at the three-digit level of the ISIC (table 3).

A coefficient greater than one indicates that the industrial branch in question tended to narrow the relative labour productivity gap with the corresponding branch in the United States between the base year (1976) and the last year in the series studied. It also gives an indication of the extent to which the gap was reduced. A coefficient of less than one indicates a relative setback in this field. As we can see, there is a great deal of heterogeneity both between branches in the same country and between countries.

This information shows us, for example, that in Argentina 12 branches displayed what we might call a "successful performance" which allowed them to significantly

TABLE 1

**Latin America (nine countries) and United States: Labour productivity
indicators in manufacturing, 1970-1996 and 1990-1996**

	1970-1996			1990-1996		
	Industrial product	Employment	Labour productivity	Industrial product	Employment	Labour productivity
Argentina	1.18	−2.62	3.80	4.87	−3.15	8.02
Brazil	2.81	0.95	1.86	2.26	−6.41	8.67
Chile	2.76	1.51	1.25	6.40	3.49	2.91
Colombia	3.98	1.24	2.74	3.52	−0.22	3.74
Costa Rica[a]	4.39	4.83	−0.44	...	...	...
Jamaica[a]	0.11	1.66	−1.55	...	...	...
Mexico	3.79	0.91	2.88	2.27	−0.03	2.30
Peru	1.17	2.85	−1.68	5.09	1.97	3.12
Uruguay	0.61	0.37	0.24	−1.46	−8.58	7.12
United States	2.39	0.35	2.04	5.04	0.30	4.74

Source: PADI data base of the ECLAC Division of Production, Productivity and Management.

[a] Data only up to 1992.

TABLE 2
**Latin America (nine countries): Evolution
of the relative labour productivity
gap in manufacturing between
Latin America and the United States**
(Manufactures as a whole) (US = 1.00)

Country	1970	1980	1990	1996
Argentina	0.42	0.41	0.55	0.67
Brazil	0.28	0.26	0.29	0.37
Chile[a]	0.25	0.24	0.23	0.20
Colombia[b]	0.29	0.25	0.37	0.34
Costa Rica[b]	−	−	0.15	0.14
Jamaica[b]	0.26	0.16	0.16	0.13
Mexico[c]	0.32	0.30	0.44	0.38
Peru	0.33	0.25	0.16	0.15
Uruguay[a]	0.35	0.22	0.20	0.22

Source: PADI data base.

[a] Data only up to 1995. [b] Data only up to 1992. [c] Data only up to 1994.

narrow the relative labour productivity gap with the United States. We consider as a "successful performance" a coefficient close to or greater than 2, which would give the branch in question a relative labour productivity indicator at the end of the period which was approximately double that of 1970. There were seven such branches in Brazil, five in Colombia, but only three in Chile.

If we look at particular branches rather than countries as a whole, we can see, for example, that in the case of branch 384 (transport equipment) Argentina and Colombia register large relative improvements in labour productivity, Brazil showed more modest gains, and Chile registered a clear setback. In branch 371 (iron and steel) Argentina, Colombia and Brazil register extremely pro-

nounced improvements in their relative position (Argentina even surpassed average United States productivity at the end of the period), whereas although Chile and Mexico did make advances, these were much less spectacular than in the first three countries.

Behind each of these sectoral situations there is a particular market morphology, a specific regulatory and institutional situation, and different competition strategies between firms which make up what we will call a special "system of competition and innovation" for each sector and country. It is essential to understand how this system works in each country, and how it differs between countries and industrial branches, in order to gain an idea of how the Latin American production apparatus has been gradually restructured. Although such a study is outside the scope of this article, it is important to note that the results presented here should be complemented in the future with detailed studies of the various sectoral systems of competition and innovation in order to gain a better understanding of what has happened in Latin American industry in the last two decades.

3. The increase in the growth rate of labour productivity in the 1990s

Except in the case of Mexico, where the growth rate of labour productivity in the 1990s was lower than that of the 1970-1996 period as a whole, table 1 shows that there was a considerable increase in this growth rate in all the other countries studied, including the United States. However, the figures also reflect the growing inability of the industrial sector in the various Latin American

TABLE 3

**Latin America (five countries): Labour productivity gap
of Latin American industrial branches compared
with similar branches in the United States, 1970-1996**
(27 industrial branches)

ISIC group		Argentina	Brazil	Chile[a]	Colombia	Mexico[b]
311	Food	1.10	1.14	0.67	0.93	1.21
313	Beverages	1.04	0.72	0.91	0.79	0.83
314	Tobacco	0.74	0.21	0.76	0.28	0.38
321	Textiles	1.67	1.43	0.77	1.23	0.75
322	Clothing	1.17	1.20	0.75	1.30	1.85
323	Leather and leather products	1.38	0.93	0.45	0.58	0.97
324	Footwear	0.78	1.13	0.65	1.03	0.72
331	Wood and wood products	0.55	0.87	0.97	0.94	0.94
332	Furniture	2.69	1.40	1.13	0.85	0.96
341	Paper and paper products	0.99	1.26	1.10	1.12	1.03
342	Printing and publishing	1.21	0.86	1.43	0.89	1.03
351	Industrial chemicals	1.92	1.18	1.79	1.09	0.88
352	Other chemical products	1.98	0.60	0.97	0.86	0.58
353	Petroleum refineries	1.22	1.57	3.35	0.28	0.30
354	Coal products	1.85	2.07	2.14	2.10	1.55
355	Rubber products	1.55	2.55	0.41	1.36	1.24
356	Plastic products	0.81	1.25	0.51	1.50	1.25
361	Pottery, china and earthenware	1.33	1.20	0.45	2.24	2.47
362	Glass and glass products	1.91	1.92	1.67	1.57	1.60
369	Non-metallic mineral products	2.35	1.28	1.68	1.36	1.39
371	Iron and steel	2.54	1.97	1.33	2.82	1.54
372	Non-ferrous metals	1.28	2.50	0.43	1.92	2.39
381	Metal products	2.07	1.78	1.22	1.79	1.39
382	Machinery except electrical	1.91	1.12	1.31	0.75	0.72
383	Electrical machinery	2.68	1.97	0.94	0.99	1.76
384	Transport equipment	2.00	1.33	0.76	2.07	1.81
385	Professional and scientific equipment	1.29	1.48	1.22	3.27	3.81
390	Other manufacturing industries	0.52	0.76	0.92	1.26	2.19

Source: PADI data base.

[a] 1970-1995. [b] 1970-1994.

countries to generate new jobs or even keep up the levels of employment registered at the beginning of the decade. We thus see that during the period from 1990 to 1996 there was a sharp drop in employment, in absolute terms, in Argentina, Brazil, Colombia and Uruguay. In all these cases, it is this decline in manufacturing employment, rather than any significant increase in the physical volume of production, which explains the marked rise in labour productivity in the region in recent years.

Chile, which began its external openness process a good deal earlier than other Latin American countries and, when it did so, registered an increase in the number of bankruptcies of industrial firms and in the levels of open unemployment, represents an exception to this rule because in the 1990s, although it achieved (modest) improvements in labour productivity, these were not at the expense of employment. It might be assumed that this country is at a more advanced stage of its macroeconomic stabilization and structural reform programme,

whereas the other countries are still at an earlier stage in the structural adjustment process and are thus suffering a stronger impact of the "destructive" forces described by Schumpeter. According to this line of reasoning, we might assume that when the level of macroeconomic uncertainty prevailing in those societies is reduced, together with a decline in interest rates and an increase in domestic saving and investment, part of the structural unemployment phenomenon will tend to correct itself, even though there will still be a certain degree of structural inability of the manufacturing sector to generate new jobs at the rate demanded by the growth in the economically active population.

After having thus analysed the evolution of labour productivity, in the next section we will set forth an explanatory hypothesis of what happened, involving an evolutionary dynamic of successive phases in the process of stabilization, structural reform and defensive and pro-active investments which occurred over time. The

intra-sectoral restructuring process, through the entry and exit of firms in the market, and the change in the relative weight of the different branches of industry in manufac-

turing as a whole, are central elements in this explanatory hypothesis for the developments registered in the region in this field over the last two decades.

III

An explanatory hypothesis

The process we are about to describe consists of at least two different evolutionary dynamics which should be separated and distinguished from each other: on the one hand there is the dynamic inherent in each branch of industry, and on the other there is the dynamic for the structure as a whole. Both of these will be explored in this section. With regard to the first dynamic, it should be noted that the changes in a branch of production over time reflect i) the entry into the market of new enterprises which bring with them new technology; ii) the exit of older and relatively less efficient enterprises from the market, and iii) the improvement of those enterprises which remain in the market, through physical investments, organizational changes and other measures. Thus, in each branch of production, a process of selection takes place which causes some firms to gain ground, others to lose it, and some even to disappear altogether.

With regard to the second dynamic, in addition to the foregoing there is a process of change in the relative weight of the different industrial activities in manufacturing as a whole, which may also be seen as a reflection of the different demand elasticities faced by the various production activities.

In other words, the intra-sectoral selection process among firms in the same line of production and the changes in the relative weight of the different industrial activities over time constitute the two structural components of the evolutionary dynamic we are about to describe. Let us now deal with these two matters separately.

1. The process of selection among firms within a given branch of production

Every branch of industry contains firms with different levels of production efficiency. The capital goods and organizational technologies used by different enterprises differ significantly, even between close competitors, thus leading to a very heterogeneous structure of unit production costs and levels of operating profitability within each branch. In practice, these differences are much more

important than might be imagined from the standpoint of competitive balance, sometimes reaching ratios of 5:1 or even more.

The reduction in tariffs associated with the external trade openness process and the increase in competition in markets due to the arrival of foreign firms have acted in the present case as an important selection mechanism, obliging local producers to adapt to a new and much more stringent form of competition. Firms have reacted with different degrees of effectiveness to this challenge, or in some cases have not reacted at all, so that many of them have languished even to the point of physically disappearing.

The above patterns of conduct have not been solely the result of the information available to each enterprise and its corresponding reactions but have also been influenced by the morphology and behaviour of the factor markets in which each enterprise has had to operate. The existence of market flaws and imperfect access to long-term finance and the technological know-how needed to carry out major operational changes at the enterprise level are traditional obstacles which help to understand why some firms have carried out successful processes of adaptation and others not, after the recent structural reforms in the Latin American countries.

The existence of imperfect markets and incomplete information undoubtedly presents firms with a much more complex challenge than in the simple models described in conventional textbooks. In a context of asymmetrical information and highly imperfect capital and technology markets, it is not necessarily the most inefficient firms that disappear from the market. On the contrary, it may well be –and indeed the empirical evidence proves this– that in a highly turbulent and uncertain macroeconomic situation enterprises whose strategies concentrate on short-term actions with a high speculative content may have a better chance of surviving the period of structural reforms.

In short, the process we are trying to describe involves a dynamic of change in industrial branches which

is marked by the entry and exit of firms from the market, hostile takeovers, the relative success of some firms and the failure of others, against a background of market flaws and a dynamic of "creative destruction" which has not so far been addressed very much in the studies on the industrial development of the region.

2. Changes in the composition of industrial production

After having so far dealt with the dynamics of intra-sectoral change, we will now turn to the changes in the relative weight of the different branches of industry in the aggregate product.

Let us assume, for the sake of argument, that the manufacturing sector comprises only two such branches of industry. Let us assume, for example, that it consists of one capital-intensive branch processing natural resources and one metal products and machinery branch producing capital goods and consumer goods. In the initial situation, v and z describe the product per person employed and the capital per person in each of the two branches in question, while point d represents the corresponding weighted product. Let us now assume that between t_0 and t_1 the two branches grow at different rates, so that the indicators move to v' and z' respectively. The position of point d' will now depend on how much each branch has grown over its particular trajectory and also on the relative weight of each of them at the end of the period within the aggregate industrial product (figure 1).

When the situation is set forth in this way, it is not hard to see why the macroeconomic stabilization and structural reform programmes of the 1970s and 1980s had a highly differential impact on different sectors of production. The industries processing natural resources had little difficulty in turning over to export activities on

FIGURE 1

**Latin America: Changes in the composition of industrial production**

a large scale by replacing their domestic sales with the export to world markets of widely used industrial commodities such as pulp and paper, iron and steel, aluminium, etc. In contrast, however, the branches in the metal products and machinery sector producing capital goods and consumer durables for the domestic market were simultaneously faced with a sharp drop in that market and a growing inflow of imported substitutes which now flooded local markets as a result of the reduction of tariffs. This has been the situation which has prevailed in the first phase of the external openness process in countries such as Argentina, Brazil, Chile and others, where local capital goods and consumer durables producers have had to simultaneously cope with a drop in domestic demand and a massive inflow of imported substitutes. It is this situation which leads, at the completion of the openness process, to a structure of production which is much more devoted to the processing of natural resources and is less specialized in the production of capital goods and consumer durables for the domestic market.

If we now take both these phenomena together – that is to say, on the one hand the intra-sectoral changes associated with the different capacities of adaptation of the firms in a given branch, and on the other the different rates of expansion of the branches as a function of the elasticity of demand and the growth rates they have achieved– we can visualize a complex evolutionary dynamic in which the macroeconomic, meso-economic and microeconomic aspects act together to condition the planning horizon of firms, their degree of uncertainty about the future, the regulatory framework, the demand conditions in which they operate, their access to factor markets, their real perception of what is going on in their economic and social environment, and the various flaws in the factor markets they must use. This interaction between the macroeconomic, meso-economic and microeconomic aspects in a context of macroeconomic turbulence and extremely imperfect production factor markets underlies the macroeconomic stabilization and structural reform programmes carried out in recent years in the countries of the region.

Each industrial branch and each firm metabolizes the changes in the system of incentives in a different way, as a function of its past history, its incomplete perception of what is going on, and its (imperfect) capacity to adapt to the new circumstances. Although the empirical information available on the above-mentioned set of variables is only fragmentary and is extremely dispersed, we will nevertheless try to present in the following pages various items of information which will help to reconstruct an evolutionary situation of the type described here.
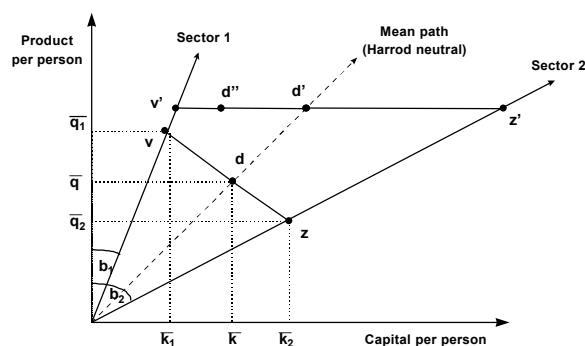
IV

The empirical evidence

1. Intra-sectoral restructuring: the entry and exit of firms from the market

The available data show that the external openness processes have been accompanied, on the one hand, by an increase in the rate of disappearance of firms, and on the other, by the entry of new firms into the market which have brought with them more modern technologies which are a clear improvement on the average technological practices which had previously prevailed in the industries in question.

In Chile, the opening-up of the economy in the 1970s was associated with a marked increase in the rate of closure of firms. It is calculated that over 4,000 firms left the market between the second half of that decade and the early 1990s, especially in such branches of industry as textiles, clothing, footwear, and metal products and machinery industries such as consumer durables and capital goods (figure 2). These closures of firms were particularly numerous among medium-sized enterprises (by Chilean standards), but did not hit the smaller and larger firms so hard (Mizala, 1992). The few studies available on this subject show that the closures of firms were not necessarily connected with indicators of operational efficiency: in a highly turbulent and uncertain macro-economic situation speculative forms of conduct tended to prevail over attempts to rationalize production, so that the business successes or failures were not due to the presence or absence of organizational or technological excellence but, in many cases, to speculative forms of conduct on the financial level.
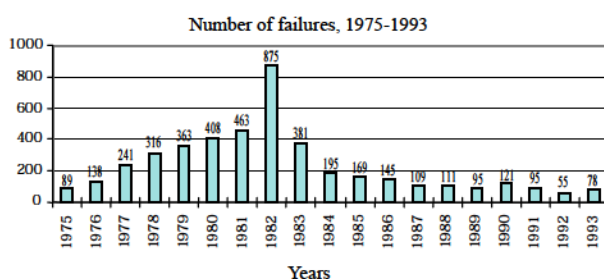
With regard to the entry of firms into the market, it is interesting to note, for example, that the restructuring of the vegetable oils industry in Argentina was associated with the entry of a new generation of production facilities of much larger scale and higher capital density per worker, where labour productivity was practically double that of the typical plants of the 1970s (table 4). The industry is moving towards production processes of a chemical nature in which the technological and organizational know-how is very different from that of the old edible oil industry which existed in the country before. The changes are to be seen not only in the production technology used but also in the institutional base and system of competition prevailing in this branch. New actors and new forms of vertical integration both within the industry and towards collateral branches (transport, port silos, docks) form part of the sectoral restructuring process in this case (De Obschatko, 1996). Several sectoral studies reveal similar patterns of disappearance and creation of jobs (Katz, 2000 and Katz (forthcoming)).

In short, the empirical evidence strongly confirms the idea that the industrial restructuring process we are describing here is associated with the exit from the market of "marginal" firms (although in this case marginal does not necessarily mean more inefficient) and the entry of a new generation of industrial establishments closer to the general international level of the industry in question. These, by necessity, are much more capital-intensive, more labour-saving and more internationally competitive than the generation of industrial plants they are replacing.

2. Changes in the relative weight of the different industrial branches

As already noted in the introduction to this article, there has been an increase over time in the relative weight of the industrial branches which process natural resources (table 5). This can be clearly seen from the table in question, which also shows how the group of traditional industries (footwear, clothing, wood products, furniture, textiles, printing, etc.) have contracted and lost ground in relative terms between 1970 and 1996. Finally, in the metal products and machinery branches we considered

FIGURE 2
Chile: Failures of industrial firms

Number of failures, 1975-1993

Source: A. Mizala, 1992.

TABLE 4
**Argentina: Number of plants, employment
and labour productivity in the Argentine
vegetable oils industry, 1973-1974 and 1993-1994**

Year	Number of plants	Number of employees	Volume of production *(thousands of tons)*	Output per plant *(thousands of tons)*	Output per worker *(tons)*
1973-1974	67	6 895	1 740	26	252
1993-1994	59	4 943	12 220	207	2 472

Source: De Obschatko (1996).

it necessary to separate the motor industry (which received special treatment in industrial policy) from the rest of the industries producing consumer durables and capital goods. In the case of the latter, the contraction or stagnation associated with the increased openness to external trade can clearly be seen in the various cases analysed. However, the rapid expansion of the motor industry in Argentina, Brazil, Colombia and Mexico in the 1990s more than offset the contraction in the rest of the metal products and machinery branch.

3.   The labour productivity gap and employment

Table 6 shows the relation between the relative labour productivity gap and employment in Argentina, Chile and Mexico. The blocks refer respectively to three-digit ISIC industrial branches which have reduced the relative labour productivity gap with the United States or have lost ground in that respect and at the same time have reduced or increased their level of employment. Weighing the various industrial sectors by their relative share in the industrial product at the end of the period studied, we see that in Argentina, for example, 70% of industry is located in the sub-group of branches which have gradually come closer to average United States pro-

ductivity but have reduced their levels of employment in absolute terms. Only 20% of industry narrowed the relative gap with United States productivity yet at the same time generated additional jobs.

In contrast, in the case of Chile 12% of its manufacturing industry narrowed the productivity gap while reducing employment levels, but 33% did so in an expansionary manner which generated additional jobs. This undoubtedly reflects the markedly different structural characteristics between the two industrial restructuring processes. As suggested in earlier pages, it may be supposed that the differences between the two cases may be explained by the different phases of the macroeconomic adjustment process that the countries in question were passing through, with the Chilean case representing a situation of greater stability at the aggregate level in which fresh investment played a more significant role than in the case of Argentina.

The Mexican case would appear to come midway between these two situations. In Mexico, 37% of its industry narrowed the relative labour productivity gap with the United States while simultaneously generating new jobs, while 30% did so through restructuring processes which reduced the level of employment. Undoubtedly the assembly industries have acted as generators of fresh employment, whereas there have been dismissals of employees in the non-assembly sectors of Mexican industry.

The Argentine case is the most extreme of the three. In the last section of this article, when we examine some of the new structural problems that the region has to face as a result of the industrial restructuring process, we will show how the greater difficulty of the new industrial structure in generating new jobs is one of the main items on the new agenda of public policy issues that the governments of the region must now deal with.

TABLE 5

**Latin America (five countries): Changes over time in the relative weight
of the different industrial branches in the global manufacturing product**

	Argentina			Brazil			Chile			Colombia			Mexico		
	1970	1990	1996	1970	1990	1996	1970	1990	1996	1970	1990	1996	1970	1990	1996
I[a]	15.6	14.3	13.1	18.8	22.9	22.8	14.9	10.1	10.2	10.7	9.6	10.5	13.3	12.3	13.9
II[b]	9.9	8.5	12.1	9.9	7.0	8.7	7.7	2.3	2.0	2.9	4.3	6.5	5.5	9.5	10.8
III+IV[c]	36.2	46.7	45.7	35.8	39.6	42.4	43.2	55.5	56.2	45.7	51.1	51.2	46.8	46.8	46.5
V[d]	38.2	30.5	29.0	35.5	30.5	26.1	34.2	32.0	31.6	40.7	34.9	31.8	34.4	31.4	28.8

[a] Metal products and machinery industry, excluding the motor industry (ISIC groups 381, 382, 383, 385).

[b] Transport equipment (ISIC group 384).

[c] Foodstuffs, beverages and tobacco (ISIC groups 311, 313, 314) plus industries processing natural resources (ISIC groups 341, 351, 354, 355, 356, 371, 372), excluding group 372 in the case of Chile.

[d] Traditional labour-intensive industries (ISIC groups 321, 322, 323, 324, 331, 332, 342, 352, 361, 362, 369, 390).

TABLE 6

**Argentina, Chile and Mexico: Relative labour
productivity gap and absorption of labour, 1970-1996**

A. Argentina

	Employment increased	Employment reduced		
Gap closing	353 355 371	311 321 332 342 351 352	354 362 369 372 381 382	383 384 385
Gap widening	313 323 324	314 322 331 341	356 361 390	

B. Chile

	Employment increased	Employment reduced	
Gap closing	331 341 369 371 372 381	342 351 353 362	
Gap widening	313 323 324	314 321 355 382 383	384 385 390

C. Mexico

	Employment increased			Employment reduced
Gap closing	323 331 332 341 355 361	362 369 371 372 381 384	390	311 322 342 354 356 383 385
Gap widening	313 314 321 324 351 352 353	382		

Source: PADI data base of the ECLAC Division of Production, Productivity and Management.

Now that we have analysed various pieces of empirical evidence which complement our overview of the industrial restructuring process taking shape in Latin America, the final pages of this article will seek to give a broad picture of the situation and make a general assess-ment of possible weak points in the present industrialization model which make its long-term sustainability open to question. Employment and the foreign trade balance would appear to be the weakest points in this respect and are worthy of detailed analysis.

V

Final reflections

This study seeks to make some progress towards the construction of a micro/macro explanatory hypothesis for the complex restructuring process currently being undergone by Latin American industry. The way in which the interplay between the macro, meso and micro levels conditions the growth path and structural transformation process of each industrial branch and industry as a whole has been little explored by economists and undoubtedly requires more work in terms of conceptualization and study. In this last section we will briefly address both these matters.

1. Some features of the micro/macro context underlying the industrial restructuring process in the region

We will begin by assuming that the conduct of firms in investment matters depends primarily, though not exclusively, on the opportunities and risks as perceived by each entrepreneur. This perception shapes their profit expectations and conditions their propensity to invest. We will assume, then, that it is the pursuit of profits (or avoidance of the loss of previous gains) which causes entrepreneurs to act, "defending" prior investments if this is economically feasible and/or embarking on new investment projects which will expand the production capacity they control.

In a simple balanced competition model, three factors would influence the decisions of this type of entrepreneur: i) the present and expected future price of the product manufactured, ii) the present and expected future prices of the production factors needed for this, and iii) the amounts of factors needed per unit of production (in the light of the available technology).

Let us first consider a simplified situation: the firm manufactures a single homogeneous product, there are no imported substitutes, the factor markets are "well behaved", the firm possesses the necessary technology, and there is no macroeconomic uncertainty, as the economy operates in conditions of external and fiscal equilibrium. How does the replacement of an old plant with a new one take place in this type of context? In other words, what is the conceptual model based on conventional price theory that economists use to explain how to determine the optimum rate of births and deaths of firms and the replacement of one production technique by another one which is better (that is to say, one which can produce the same good at a lower unit cost)?

In the experimental conditions thus assumed, it is the market price of the product and the factor prices, together with the savings on factor use made possible by the new technologies, which determine the optimum rate of replacement of old plants and technologies by new ones. The arrival of a new firm (or a new technology) causes the market price of the product in question to go down, and when one of the existing firms is unable to cover its variable production costs with the sale price of its products it must necessarily leave the market.

Let us now consider how this stylized model of microeconomic behaviour would be affected if we admit the existence of factor market flaws, incomplete information available to the firm, volatility of the macroeconomic variables, and a specific regulatory and institutional framework for each sector which of itself affects the conduct of the entrepreneurs in addition to the possible effect of the economic variables. This exercise allows us to take account in our analysis of the effect that increased openness, deregulation and privatization have had on microeconomic conduct and on the restructuring path that Latin American industry has been following, in a context of market flaws and imperfect availability of information for the firms.

Casual observation suggests that there is a first period –which we will call Phase I of the macroeconomic stabilization and structural reform programme– in which there is, among other things: i) exchange rate devalua-

tion, ii) rises in real interest rates, iii) a decline in the level of tariff protection, iv) a contraction in real wages, and v) elimination of price controls. All this involves a change in the structure of the "main" prices in the economy, with the ultimate objective of reducing the domestic absorption of goods and services in order to make it possible to achieve a new state of external and fiscal equilibrium. From the point of view of the individual firms, the above changes take the form of shifts and turns in their cost and demand curves (Katz and Vera, 1997) because, among other things, of declines in the import prices of goods that can take the place of those that they produce themselves, lower prices of the capital goods they use in their production, reductions in the real wages they have to pay their workers, and changes in the rates of interest they have to pay on their working capital. As well as taking account of these phenomena, which are due primarily to changes in the overall system of macroeconomic incentives prevailing in the society in question, entrepreneurs must also take account of data of a strictly sectoral nature which may or may not substantially alter the institutional context, in addition to changes in the macroeconomic variables. The strictly sectoral context is affected by such variables as the special tariff treatment that each sector may receive (as in the case of the motor industry, for example), promotional subsidies for particular geographical locations, rights of ownership of natural resources, export credits and insurance, etc.

The interaction between the macroeconomic and the sectoral levels means that, even when there is a generalized decline in aggregate investment, there may nevertheless be significant increases in investments in particular branches of industry, if strictly sectoral forces justify this.

The above changes set in motion a complex dynamic of death and disappearance of firms, takeovers, and the entry of new producers into the market. Business profit calculations are radically altered by the changes in relative prices and the fall in domestic demand (and increase in external demand) as a result of exchange rate devaluation. The increase in macroeconomic turbulence normally observed in Phase I of the stabilization programme means that the planning horizon of enterprises tends to be shortened and short-term investments tend to prevail over those connected with production and technology.

Not all macroeconomic stabilization programmes manage to restore the fiscal and external balance of the economy. When they do, the stabilization episode puts the economy on a new path of external and fiscal equilibrium, with a new structure of relative prices and income distribution, so that the country can enter on what

we may call Phase II of the stabilization programme, in which there may be a gradual reduction in real interest rates, revitalization of domestic demand and a rise in the growth rate of the economy.

In many cases, however, the programmes are not successful in stabilizing the economy, and the adjustment programme fails to reduce the imbalance prevailing at the macroeconomic level. In such an event, the macroeconomic turbulence tends to get worse, most seriously affecting the behaviour of the individual economic actors. The imbalance in the public accounts increases, as does the State deficit, and unless external financial support is obtained in order to stabilize the fiscal accounts the government is obliged to resort to domestic indebtedness, thus further increasing the already high real interest rates prevailing in the economy and leading to an even greater contraction in the level of activity than that due to the initial adjustment.

In all cases, however, it is clear that the change in the system of macroeconomic incentives affects the profit calculations of each firm, thus setting in motion an intra-sectoral restructuring process of the type described earlier.

In Phase I, in line with the changes in the system of macroeconomic policies, first of all domestic demand contracts and somewhat later –since it is necessary to establish the necessary import channels– imported substitutes arrive on the domestic market at lower local prices than many domestic firms are able to offer. Some of the latter are thus faced with imminent physical disappearance: an effect which is further accentuated by the tariff reductions that also form part of the macroeconomic stabilization programme.

In Phase II, under the structural adjustment programme, domestic interest rates tend to go down and domestic demand and public and private investment gradually recover, thus improving the potential profitability of investing in new installed capacity. New production plants thus appear which use more modern technology and are more capital-intensive.

It may be intuitively perceived that during the sequence in question two processes are simultaneously at work –an intra-sectoral restructuring process and another of an inter-sectoral nature– and resources move to the activities with the highest rates of profitability in the new relative price structure that prevails in the economy. Industrial sectors making intensive use of natural resources tend to gain ground, since they are engaged primarily in export activities and are benefited by the exchange rate devaluation. In addition, some other industrial branches also gain ground because they have managed to obtain

special treatment under the prevailing economic policy and thus do not have to adapt to the general conditions of greater external trade openness. The economy thus moves in the direction of greater capital-intensity in the various production activities and greater incorporation of labour-saving technologies.

The empirical information presented earlier suggests that over the last two decades the region has indeed gone through a process of intra-sectoral and inter-sectoral change of this type. Some firms have disappeared, while other new firms have entered the market which are much more capital-intensive and are closer to the international "state of the art". The production structure has tended to move towards the exploitation of natural comparative advantages, and sectors which enjoy relatively privileged regulatory frameworks (the motor industry) or non-tradeable service activities have managed to maintain or even increase their relative shares of the manufacturing product.

Can this process of restructuring and modernization of production be maintained over time? There are two points which stand out in this respect and give rise to some doubts as to whether the answer to this question can be affirmative. The first is the capacity of the industrial sector to generate new jobs at the rate demanded by the growth of the economically active population; the second is the endemic deterioration of the external trade balance, which seems to be associated with the new structure of production and the pattern of specialization in international trade that the countries of the region have been adopting. In both these cases there are serious grounds for doubt regarding the answer.

2. Employment and the external trade balance: two weak points in the present Latin American industrial development model

a) *The low rate of absorption of labour*

The stylized facts analysed earlier in this article clearly show that one of the main problems in the new industrialization model is its low capacity to generate employment. Although it is true that this problem is to some extent exaggerated, because subcontracting and the new tendencies towards less vertical integration of production processes give rather a false idea of the true effects of the labour-saving aspects of the new technologies in organization of labour, there can be no doubt that computerized production, based on numerical control and "real time" operation of production facilities, does lead to the elimination of production line workers and administrative personnel previously employed in the planning and organization of work.

This is also furthered by the relatively lower cost of capital goods and the rapid spread of information technology, which explain the growing use in recent years of computers, data transmission equipment, etc.

In line with the explanatory hypothesis presented earlier in this article, it could be concluded that the trend towards less use of labour would tend to be concentrated in Phase I of the macroeconomic stabilization and structural reform programmes, whereas in Phase II –with the increase in the rate of domestic saving, the revitalization of business ventures and the appearance of many new investments– there should be a renewed tendency towards increased demand for labour, especially in the case of human resources skilled in computer-based production technologies.

It would appear to be this process of successful macroeconomic stabilization and the renewed spirit of enterprise among entrepreneurs which enabled the Chilean economy to overcome the high indices of unemployment and closure of factories registered in the second half of the 1970s and first half of the 1980s. This reading of the process suggests the existence of endogenous micro/macro forces capable of partly correcting the structural unemployment phenomena associated with Phase I of the new technological paradigm being experienced by the countries of the region. These forces, however, only seem to operate in the long term, and only when there are sustained successes in terms of macroeconomic stability and recovery of growth and investment rates.

Another problem which is a source of concern is connected with the growing manufacturing trade deficit registered by the countries of the region.

b) *Towards a chronic trade deficit?*

We will now deal with some points of interest regarding the external trade balance for industrial products (table 7). The data show, first, that the negative external trade balance has been growing over time, and second, that this is closely connected with the industrial restructuring process and the new form of insertion in international trade which became established in the 1990s. The negative trade balances tend to accumulate in the metal products and machinery branches producing capital goods, agricultural machinery, consumer durables and scientific instruments (in table 7 these appear as Group I). More recently, negative balances have also begun to be registered in the Group II branches engaged in the production of vehicles and transport equipment. After having registered positive trade balances in Brazil and Mexico in the early 1990s, the recent import boom has

TABLE 7

**Latin America (six countries): Trade balance of
industrial branch groups I, II, III, IV and V[a]**

(Gross value of total production of manufactures, in millions of current dollars)

	Argentina				Brazil			
Group	1970	1974	1990	1996	1970	1974	1990	1995
I	−451.7	−411.0	−631.6	−7 545.6	−815.5	−2 797.1	−2 589.0	−10 463.6
II	−68.2	−6.0	−18.2	−1 842.4	−235.6	−432.4	1 641.2	−3 017.9
I+II	−519.9	−417.0	−649.8	−9 388.0	−1 051.2	−3 229.4	−947.9	−13 481.6
III	759.2	1 243.7	3 833.4	6 084.5	1 330.2	2 996.0	4 089.3	5 192.7
IV	−543.9	−1 269.7	610.4	−3 286.9	−622.9	−4 385.2	4 000.7	1 556.7
III+IV	215.4	−26.0	4 443.8	2 797.6	707.3	−1 389.2	8 090.1	6 749.3
V	−77.9	−39.8	798.4	−899.0	52.7	432.7	2 110.3	866.6
Net balance	−382.4	−482.8	4 592.4	−7 489.4	−291.1	−4 185.9	9 252.5	−5 865.6
Gross value of production	8 660.7	45 677.5	79 949.1	129 800.6	25 903.2	68 074.7	238 719.3	202 622.6
Net balance/gross value of production (%)	−4.4	−1.1	5.7	−5.8	−1.1	−6.1	3.9	−2.9

	Chile				Colombia			
Group	1970	1974	1990	1996	1970	1974	1990	1996
I	−320.4	−351.6	−2 625.8	−5 300.5	−260.8	−333.1	−1 689.6	−3 250.8
II	−117.4	−128.0	−646.4	−2 166.2	−158.8	−179.5	−516.1	−1 077.4
I+II	−437.7	−479.5	−3 272.3	−7 466.7	−419.6	−512.6	−2 205.7	−4 328.3
III	−33.2	−220.4	601.0	1 248.5	−0.1	56.6	135.7	−170.9
IV	848.8	1 673.4	3 297.6	3 726.2	−222.9	−548.4	−1 502.8	−1 751.6
III+IV	815.6	1 453.0	3 898.6	4 974.7	−223.0	−491.9	−1 367.1	−1 922.5
V	−70.4	−93.4	−284.6	−1 449.5	−35.3	138.1	761.9	56.7
Net balance	307.4	880.0	341.8	−3 941.5	−677.9	−866.3	−2 810.8	−6 194.0
Gross value of production	2 416.7	3 440.0	15 133.8	23 768.3	3 505.7	7 011.5	21 013.8	33 161.7
Net balance/gross value of production (%)	12.7	25.6	2.3	−16.6	−19.3	−12.4	−13.4	−18.7

	Mexico				Uruguay			
Group	1970	1974	1990	1994	1970	1974	1990	1995
I	−812.1	−1395.8	−7202.6	−7655.8	−47.0	−40.5	−329.1	−700.9
II	−389.8	−712.2	402.5	−391.7	−28.8	−23.5	−107.9	−273.9
I+II	−1201.9	−2108.1	−6800.1	−8047.4	−75.8	−64.0	−437.0	−974.9
III	208.1	253.1	−1913.7	−2482.4	93.1	163.2	503.5	448.2
IV	−333.8	−982.6	−1935.3	−8910.0	−54.4	−118.7	−221.0	−406.7
III+IV	−125.7	−729.5	−3849.0	−11392.4	38.6	44.5	282.5	41.5
V	−85.0	67.2	−1201.1	−3250.5	50.8	64.2	537.7	252.6
Net balance	−1412.7	−2770.4	−11850.2	−22690.3	13.6	44.7	383.1	−680.8
Gross value of production	21947.8	45312.5	128234.0	178528.6	1381.3	3862.9	10031.7	8564.0
Net balance/gross value of production (%)	−6.4	−6.1	−9.2	−12.7	1.0	1.2	3.8	−7.9

Source: PADI data base of the ECLAC Division of Production, Productivity and Management.

[a]   Groups: I = metal products and machinery; II = vehicles and transport equipment; III = foodstuffs; IV = widely used industrial commodities; V = labour-intensive goods.

turned the balance into a heavy deficit. In Argentina and Colombia it was never possible to reverse the negative trade balance, in spite of the expansion of the sector in recent years and the improvement in relative productivity achieved by it in the 1990s.

On the other hand, it may be seen from table 7 that in the case of Argentina, Uruguay, Chile and Brazil, which are important exporters of foodstuffs to markets inside and outside the region, positive trade balances tend to build up in the branches making up Group III. Colombia, however, and even more so Mexico are net importers of food. Finally, in Group IV, Chile and Brazil have net trade surpluses in the field of widely used industrial commodities.

Taking into account the systematic nature of the situation described, we may assert in conclusion that the new industrialization model involves a clear pattern of production specialization and international insertion in the fields of the processing of natural resources and the production of foodstuffs and industrial commodities with little local value added, while the model also involves growing dependence on the exterior in the fields of production machinery and equipment, scientific instruments and capital goods in general. Although decades have passed since ECLAC drew attention to the possible fragility of a pattern of production specialization and international insertion of this type, and although we know today that the production of both foodstuffs and industrial commodities can incorporate on a large scale the new technologies arising in such fields as genetics, biotech-

nology and mineralogy, the long-term pattern that seems to exist behind the figures presented here raises serious doubts about the possibility that the economies of the region can maintain an external balance when their industrial activities register growing trade deficits in the fields of metal products and machinery and capital goods.

Will this be a new Achilles' Heel that production and technological development policies will have to deal with in the medium and long term? All that we can say is that so far this problem has received little attention, and a new, forward-looking debate is needed if we are to gain a better understanding of the long-term sustainability of the external trade openness and economic deregulation programmes undertaken in recent years by a number of countries of the region.

(Original: Spanish)

Bibliography

Barro, R. J. and X. Sala-i-Martin (1995): *Economic Growth*, New York, McGraw-Hill.

Beckerman, W. (1975): *The British Economy in 1975*, London, The National Institute of Economic and Social Research.

De Obschatko, E. (1996): *Industrialización basada en recursos naturales*, Santiago, Chile, Economic Commission for Latin America and the Caribbean (ECLAC), *mimeo*.

Dosi, G. (ed.) (1988): *Technical Change and Economic Theory*, London, Pinter Publishers.

Freeman, C. (1994): The economics of technical change, *Cambridge Journal of Economics*, vol. 18, No. 5, Cambridge, Massachusetts, Academic Press Limited.

Guadagni, A. (1997): *Panorama actual de la economía argentina*, Colección Temas Económicos y Sociales, Buenos Aires, Fundación Omega Seguros.

Kaldor, N. (1966): *Causes of the Slow Rate of Growth in the United Kingdom*, London, Cambridge University Press.

Katz, J. (1969): *Production Functions, Foreign Investment and Growth*, Amsterdam, The Netherlands, North-Holland Publishing Company.

——— (1997): *Cambios en la estructura y comportamiento del aparato productivo latinoamericano en los años 1990: después del Consenso de Washington ¿Qué?*, "Desarrollo Productivo" series, No. 65, Santiago, Chile, ECLAC.

——— (2000): *Reformas estructurales, "regímenes sectoriales" y desempeño industrial de América Latina en los años 1990*, Santiago, Chile, ECLAC/IDRC.

——— (forthcoming): *Reformas estructurales, productividad y conducta tecnológica*, Santiago, Chile, ECLAC/Fondo de Cultura Económica.

Katz, J., J. M. Benavente, G. Crespi and G. Stumpo (1997): New problems and opportunities for industrial development in Latin America, *Oxford Development Studies*, vol. 25, No. 3, Oxford, U.K.

Katz, J. and B. Kosacoff. (1989): *El proceso de industrialización en la Argentina. Evolución, retroceso y perspectiva*, Buenos Aires, ECLAC Office in Buenos Aires.

——— (1998): Aprendizaje tecnológico, desarrollo institucional y la microeconomía de la sustitución de importaciones, *Desarrollo económico*, vol. 37, No. 148, Buenos Aires, Instituto de Desarrollo Económico (IDES).

Katz, J. and H. Vera. (1997): The ongoing history of a Chilean metal products and machinery firm, *CEPAL Review*, No. 63, LC/G.1986-P, Santiago, Chile, ECLAC.

Metcalfe, J.S. (1997): *Innovation systems and endogenous growth theory*, Manchester, U.K., November, *mimeo*.

Mizala, A. (1992): Las reformas económicas de los años setenta y la industria manufacturera chilena, *Colección Estudios CIEPLAN*, No. 35, Santiago, Chile, Economic Research Corporation for Latin America (CIEPLAN).

Nelson, R. and S. Winter (1982): *An Evolutionary Theory of Economic Change*, Cambridge, Massachusetts, Cambridge University Press.

Nelson, R. (1991): Why do firms differ, and how does it matter?, *Strategic Management Journal*, vol. 12.

——— (1997): How new is new growth theory?, *Challenge*, vol. 40, No. 5, New York, M.E. Sharpe, Inc.

Rommer, P. (1986): Increasing returns and long-run growth, *Journal of Political Economy*, vol. 94, No. 5, Chicago, Illinois, University of Chicago Press.

——— (1992): Two strategies for economic development: Using ideas and producing ideas, *Proceedings of the World Bank Annual Conference on Development Economics, 1992*, Washington, D.C., World Bank.

Salter, G.W. (1960): *Productivity and Technical Change*, London, Cambridge University Press.

Solow, R. (1957): Technical change and the aggregate production function, *Review of Economics and Statistics*, vol. 39, No. 3, Cambridge, Massachusetts, Harvard University Press.

——— (1988): Growth theory and after, *The American Economic Review*, vol. 78, No. 3, Washington, D.C., American Economic Association, June.

Technological change

and industrial dynamics

in Latin America

Gustavo Burachik

Department of Economics,
Universidad Nacional del Sur,
Bahía Blanca, Argentina
burachik@criba.edu.ar

The subject-matter of this article lies at the crossroads be-
tween the literature on technological change and that on in-
dustrial dynamics. The analysis centers on the links between
the form of accumulation of technical know-how in an in-
dustry and the likelihood that the innovation in question can
become a vehicle for the entry of new enterprises into the
sector. The studies on the developed countries tackle this
matter through two approaches: that of technological regimes
and that based on the life-cycle of industry. Both these con-
cepts are of an evolutionary nature and are set forth in sec-
tion II below. After analysing whether the different sectoral
forms of innovation are associated with different rates, char-
acteristics and survival prospects for firms entering an in-
dustry, section III seeks to determine what conclusions on
dynamics can be drawn from the literature on technological
change in the manufacturing firms of the main countries of
Latin America. Lastly, section IV offers some final reflec-
tions. The main contributions made by this study are the
following: i) it offers a different perspective for interpreting
technological change in Latin America and seeks to develop
a concept equivalent to that of the "innovative advantage"
used in studies on the developed countries; ii) it suggests
that, in a context in which enterprises mainly innovate
through the incorporation of know-how developed by other
organizations, established enterprises tend to enjoy advan-
tages for the incorporation of technical progress, and iii) in
view of this, it may be assumed that in those activities where
product or process innovation creates competitive advan-
tages for established enterprises, "innovative entry" will not
be a frequent phenomenon.

I

Introduction

The improvement in the statistical bases of the industrialized countries has now made it possible to tackle new aspects of the phenomenon of the birth of enterprises, giving rise to numerous empirical studies on the birth of new enterprises and their performance (i.e., their survival and/or growth) after their establishment.[1]

The researchers of those countries mainly approach the problem from the standpoint of the problems of the industrial economy or one of its subject-areas such as market theory, the labour market (rotation of employment), etc. Many studies –especially the neo-Schumpeterian ones– also incorporate an approach centered on the repercussions of entrepreneurial demography on economic development, but they all take it for granted that there is a competitive dynamic in which (albeit to a different extent in different sectors) innovation is a central element in differentiation between enterprises. Interest is focussed on the primary elements of innovation and its special features both at the sectoral and enterprise level. Thus, the literature on technological regimes and the life cycle of industries offers schemes and concepts which seek to explain i) what factors determine the differences between sectors as regards the characteristics of innovation and the conditions of entry, survival and exit of enterprises, and ii) how these two elements are combined in the different types of activities.

For the semi-industrialized countries (SICs), in contrast, the studies on entrepreneurial demography open up a new area of reflection on the very nature of structural change. The possibility thus arises of integrating within a single conceptual framework the study of the determinants of entrepreneurial renewal in the most dynamic sectors, the prospects for the appearance of new activities, and the entry of firms into the market as a means of competitive pressure which will improve the innovation capacity of the firms which are already established, among other aspects. However, the theories and stylized empirical facts given in the studies on more advanced economies need to be rethought to some extent before they are applied to the study of semi-industrialized structures.

Section II of this article looks at the conceptual schemes that link up inter-sectoral differences of dynamics (patterns of entry and survival) with the nature of the technological innovation process. The central element in these arguments (which may be considered as parts of evolution theory) is the analysis of the factors that determine which enterprises –new entrants or firms that are already established– will find it easier to introduce technological innovations. This seems a suitable way of progressing towards a comparative analysis of the dynamic attributes of production structures of unequal economic and institutional development.

Section III, which forms the main part of this study, seeks to progress in the analysis of industrial dynamics in SICs. To this end, the concepts which have been crucial elements in the relation between technological innovation and industrial dynamics in the developed countries (technological opportunity and the appropriability and accumulability of knowledge) are reviewed in the light of the literature on technological change in SICs, especially those of Latin America.

The rereading of the studies on technical change in SICs in the light of the theories presented in section II forms a useful framework for the analysis of the dynamic attributes of production structures in such countries. Finally, section IV presents the main conclusions of this study.

[1] See, for example, Arrighetti (1994), Audretsch (1995a), Baldwin and Gorecki (1991), all the articles in *International Journal of Industrial Organization* (1995), Dunne, Roberts and Samuelson (1988 and 1989), Mata (1994), Mata and Portugal (1994), and Wagner (1994).

II

Conceptual schemes that help to explain

the inter-sectoral differences in industrial dynamics

Two different theoretical approaches to the inter-sectoral differences in industrial dynamics may be distinguished. One of them is based on the study of certain features of technology and market structure.[2] The other centers on the differences due to the different ways of generating technological know-how, and is the approach followed in the present study.

The disparities in industrial dynamics caused by the characteristics of the technological innovation process may be explained through two main schemes: that based on technological regimes, and that based on the life cycle of industries, both of which are connected with the neo-Schumpeterian literature on technological change.

The summary given below sets forth these two approaches because they are the conceptual basis for all the empirical studies (using extensive, generally longitudinal, data bases) which have been made on the relation between technology and industrial dynamics in the developed countries. This "unicity" of the range of approaches discussed is thus due to methodological rather than theoretical considerations. Other (alternative and/or complementary) approaches could very well interconnect with the ideas presented below. However, no attempt will be made in this article to refer to and compare other approaches, since it merely aims to organize a collection of theoretical and empirical material which is quite recent and has been very little analysed in the SICs, and then put forward some ideas on possible conclusions that might be drawn from it and might be useful for those countries.

Before presenting our summary of the two approaches, it is worth specifying clearly the assumptions regarding entrepreneurial conduct that implicitly or explicitly underlie them. In spite of their common Schumpeterian roots, the authors consulted may differ somewhat in their assumptions about entrepreneurs' profitability goals. Nevertheless, all the schemes can easily be conceived on the basis of active learning models (Ericson and Pakes, 1995): i) the agents detect opportunities for profit in the economy, and in order to check

them out they must invest and enter the market; ii) in any activity or moment in time the active enterprises have different levels of efficiency, and iii) once they are in an activity, enterprises make investments ("actively explore the economic environment they are operating in") in order to increase their capacity to make profits; those whose "exploratory actions" are not successful suffer a decline in their profitability which eventually brings them to a situation where exit from the market is the best response.

1. Technological regimes

The starting point for modern analysis of the relation between technological change and industrial change is the work of Schumpeter. It is well known, however, that Schumpeter himself gradually altered his view of industrial dynamics in his successive studies on this subject, in line with the changes undergone by industrial capitalism in the developed countries from the beginning of the century up to the Second World War. Thus, the process of "creative destruction", carried out fundamentally by new entrepreneurs, which he posited in his *Theory of economic development* (Schumpeter, 1912) was replaced in *Capitalism, socialism and democracy* (Schumpeter, 1942) by a scheme in which the generation of innovations tends to take place within big corporations that allocate resources and apply procedures specifically for that purpose.

Possibly (at least partly) as a reaction to the gloomy outlook that Schumpeter envisaged for capitalism in his 1942 study, neo-Schumpeterian authors studying innovation theory rescued the concept of "creative destruction" and placed it side by side with the alternative pattern (big corporations as the main source of innovation) in a unified conceptual scheme.

The idea of technological regimes is closely linked with the evolutionary view of technical change at the enterprise level. This approach criticises the orthodox view, whereby firms have unrestricted access to production facilities which include a list of all the available technologies, so that they only have to choose the technology that fits in best with the price relations of the production factors involved. Evolutionary

---

microeconomics, in contrast, suggests the incorporation of other facets into economic analysis, especially the organizational aspect, in which the technological decision-making process is merely the last visible link. In this context, the technical options open to the enterprise are not determined exogenously: on the contrary, they are "idiosyncratic" in so far as they are the result of the firm's own experience and, in particular, the successes and failures in its "exploratory actions" (Nelson and Winter, 1982, parts I and II).

The exploratory actions (through expenditure on research and development, for example), for their part, immediately suggest the need to recognize that there is a variety of "sources" of technical information and innovative ideas that the agents use to bring their techniques in line with their goals in terms of profitability and ranges of production. New knowledge can come from outside the firm or it can be the result of processes of in-house accumulation and research activities carried out by its own members. The role of the different sources of information needed to gain access to technical innovations varies significantly in the different manufacturing sectors and types of technology, however. It is this wide variety of technological and innovative environments within the production structure which gives rise to the concept of technological regimes.

According to Winter, "Along with differences in the relative importance (however measured) of the different sources, there are differences in a variety of related aspects, including such matters as the intrinsic ease or difficulty of imitation, the number of distinguishable knowledge bases relative to a productive routine, the degree to which successes in basic research translate easily into successes in applied research (and vice versa), the size of the resource commitment typical of a 'project', and so forth. To characterize the key features of a particular knowledge environment in these various respects is to define a 'technological regime'" (Winter, 1986, p. 205).[3]

The next step is to understand the two Schumpeterian innovation patterns as expressions of different underlying technological regimes. Thus, in the "entrepreneurial regime" new firms are the vehicle for innovative progress and the rotation is very intense because the established firms never manage to put their advantages on a solid, lasting basis that can withstand the disruptive capacity

of the new entrants. In the "routinized regime", in contrast, the opposite pattern is observed: the established firms are always in a better position to innovate because they have access to a growing set of innovation opportunities and are in a position to cash in to the full on the technical advantages thus obtained. This is the situation of activities where innovation comes basically from the research and development laboratories of the established (generally dominant) firms. The innovative advantages of these firms are further heightened if technological opportunities are expanded and the conditions for private appropriability are strengthened (patents, difficulty of imitation, etc.). The innovative position of the new entrants, for their part, is improved if they have ready access to relevant technical information coming from sources outside the established firms.

The contributions by Malerba and Orsenigo (1995 and 1997) are along the same lines, emphasizing the wide range of forms that the organization of innovative activities can take in the different industrial sectors and suggesting that they can nevertheless be adequately covered by the two Schumpeterian archetypes.

These authors have made empirical studies in which they analyse the evolution of the list of innovative enterprises[4] in the different industrial sectors. In the sectors grouped together under the heading of Schumpeter Mark I, technical progress takes place through *widening*: the list of innovative enterprises constantly expands with the entry of new organizations (generally quite small), and the innovation race among enterprises is very even, with the order of the leading successful firms continually changing. Here, a process of "creative destruction" is at work. The key features in this pattern are ease of entry and the central role played by new firms in innovation.

In contrast, in the sectors grouped together as Schumpeter Mark II the technical process advances through *deepening*: innovative activity is dominated by a small group of firms which continually introduce improvements, making use of the technological capacity they have built up over time. Here, the industrial dynamic is represented by a process of "creative accumulation". The big firms have institutionalized the innovation process with the establishment of research and development laboratories and the recruitment of researchers, technicians and engineers.[5]

---

[3] The concept of technological regimes was already described in Nelson and Winter (1982, chapters XI and XII). However, the model presented there does not provide for the entry of new firms and is thus not capable of exploring the inter-sectoral diversity as regards the role of innovative entry. Winter (1986) shows substantial progress in this sense.

[4] Generally on the basis of information on patents applied for.

[5] Naturally, these two patterns must be seen as the extremes in a broad range of situations. Furthermore, as we shall see below, a production activity or technology may undergo changes in the way its innovative activities are organized in the course of its life cycle.

Malerba and Orsenigo find that these archetypes reflect systematic differences between the different technologies; generally speaking, each industrial branch always tends to be listed under the same pattern in each of the different national cases. This link between the characteristics of the innovation process in the different sectors and the Schumpeterian patterns is in line with the concept of technological regimes; this notion provides a synthetic representation of some of the most important economic properties of technologies and the characteristics of the learning processes involved in innovative activities (Malerba and Orsenigo, 1997, pp. 84-85). The basic hypothesis is that there are factors related with the forms of accumulation of knowledge which have a critical influence on the way innovative activities are organized in a given class of technology. The technological regime would be the result of a particular configuration of i) certain features of the technology, such as the conditions of opportunity and appropriability and the degree of accumulativity of the technological knowledge in question, and ii) certain characteristics of the knowledge base.[6]

Thus, the two Schumpeterian archetypes can be reformulated in terms of these concepts. The "creative destruction" pattern of innovation reflects conditions of high opportunity but low appropriability and accumulativity. The low appropriability means that firms can build up appreciable advantages of innovation, but not on a lasting basis. The low accumulativity suggests that the technological advantages are not associated with experience in the activity, which clears the way for innovative entry. The "creative accumulation" innovation pattern, for its part, reflects conditions of high opportunity, appropriability and accumulativity.[7]

The concept of technological regimes is also adopted by Audretsch (1991, 1995a, 1995b and 1997) in order to explain inter-sectoral differences in the rates of birth and subsequent performance of new firms.

Audretsch (1995a) considers that the main role in the process of technological change is played by the individual (not the firm) who possesses innovative knowledge. He defines the problem of appropriability as the search for an organizational solution which offers the most profitable conditions for developing the new idea commercially. The agent must choose between selling his knowledge to an existing firm or starting a firm of his own. Here, the key question is whether there is i) uncertainty about the economic value of the potential innovation and asymmetrical access to information by the parties involved, and ii) agency costs associated with the development of the idea within an existing organization.

If there are, the organizational solution to the problem of appropriability will depend on the existence of economies of scale or scope. If there are serious problems of information, however, the inventor and the enterprise will tend to differ in their appraisal of the expected profitability of the project. Furthermore, if there are problems of asymmetrical information this will mean that the enterprise will only be able to incorporate the project within the framework of a follow-up and control scheme, which means that there will be agency costs. Audretsch says that it is the appearance of these frictions which often make it more difficult to transfer a new item of knowledge from an individual to an established firm and thus open the way for the establishment of a new innovative firm.

The presence of these factors varies from one industrial branch to another, as a function of the underlying knowledge conditions. This is so "not only because the relative importance of innovative activity varies systematically across industries, but also because the opportunities for new [or established] firms to generate that innovative activity also vary from industry to industry" (Audretsch, 1995b, p. 448). In some activities, the search for innovation is routinized and the new knowledge generated can be processed within the framework of the typical organizational structures of large enterprises. In others, however, innovations usually originate in items of knowledge which are not accumulated in a routine manner and are therefore difficult to absorb in this way.[8] In this latter case, the existence of the problems of information and agency costs in question will oblige an individual who has a potential innovation to form a new enterprise in order to try to exploit it commercially. These two situations reflect the conditions referred to by Winter (1986): an entrepreneurial regime (which favours innovative entry but discourages the introduction of technical progress by established firms) and a routinized regime (which has the opposite characteristics).

---

[6] For example: generic or specific, tacit or codified, complex or simple, independent or systemic.

[7] The existence of high opportunity is a necessary condition for there to be widely differing levels of performance by firms. Only in an environment of this type can there be firms with broad and potentially lasting advantages in their production operations and innovative activity.

[8] Audretsch cites the examples of Xerox and Apple, which were created as a result of this type of process (Audretsch, 1995a, pp. 54-55).

Thus, for Audretsch the factors which determine the pattern of innovation prevailing in a given production activity are: i) the origin of the innovative knowledge (routine research within established firms or sources external to them), and ii) the degree to which uncertainty and asymmetrical information will give rise to agency costs if established firms decide to develop projects generated by outside inventors.

It is possible to give a more precise description of Audretsch's version regarding the differences between activities associated with the two types of innovation regimes. Thus, there are three levels on which important differences may be observed:

i) *Technology.* In a static perspective, there are certain features of the production function which take on importance according to the technological regime. In particular, the existence of substantial economies of scale and the use of highly capital-intensive technologies are associated with the routinized regime. From the dynamic point of view –that is to say, in terms of the evolution of technology– an important role is played by the forms of accumulation of knowledge (routinized or not) and their relation with the two innovation regimes.

ii) *Market structure.* Highly concentrated activities are associated with the existence of a routinized regime, while those where small firms have a substantial share of the market display a form of operation closer to the entrepreneurial regime.

iii) *Dynamic features.* As regards the entry of new firms, the branches whose form of innovation is similar to that of an entrepreneurial regime register a high rate of birth of new firms. As explained earlier, individuals who have new ideas will find it difficult to transfer them to established firms, so they will tend to channel them through the formation of new firms of their own. The branches with entrepreneurial-type regimes may be viewed as environments where it is generally relatively easy to generate innovations and, in particular, where innovative entry is a frequent phenomenon. In contrast, this will not be so frequent in activities with a routinized regime, because in them technical progress tends to be generated in the research and development laboratories of the (generally large) enterprises of the sector.

The entry conditions are also affected by the interaction of the following features of an industry: the existence of growing returns to scale, and the relative "generosity" of the sector in question in terms of the growth opportunities for new entrants and small firms. Audretsch identifies the latter feature with the degree of innovativeness of the industry. The probability that a new firm will remain active in a given period and sector is determined by the interaction between, on the one hand, the existence and intensity of the cost disadvantages involved in operating at a smaller scale than the minimum efficient level (which is the typical situation of most newly-created firms) and, on the other hand, the receptiveness of the environment to the innovative initiatives of the new entrants.

This position has some implications for entry conditions, such as: i) entry may not be seriously discouraged, even when economies of scale are important, if the potential entrepreneurs perceive that once they are operating they will have opportunities for expansion and innovation, and ii) the environments of entrepreneurial-type regimes in terms of competition are perceived by new or potential entrepreneurs as being more abundant in these types of opportunities.

Nevertheless, according to Audretsch (1995a), this greater attraction exerted by highly innovative entrepreneurial-type regimes has ambiguous effects. On the one hand, the receptiveness of the industry in terms of the opportunities for innovation by small firms stimulates the entry and retards the exit decisions even of firms which are far below the minimum efficient scale, but on the other hand, in environments where a process of intense technological change is under way and there is a high level of uncertainty as regards technical matters and the level and structure of future demand, there is less likelihood that an individual firm will be able to innovate successfully.

Audretsch draws the following conclusions from these arguments (which he tests empirically in the studies referred to) regarding the probability of exit or growth of the firms that survive:

i) in industrial branches marked by the existence of economies of scale the probability of survival will be low (because of the strong competitive affect of the entry of firms of sub-optimal scale), but the growth of those that survive will be high because their survival will depend precisely on whether the typical new entrants (of small scale) expand until they can reverse their cost disadvantages, and

ii) something similar will occur in branches where small firms have access to opportunities for growth and innovation. The likelihood that an individual firm will succeed in its innovative plans is low, and so also is its probability of survival. Those firms which achieve successful innovation (and thus survive) will grow fast, however.

With regard to the features of the growth processes, it must be asked whether these processes are a horizontal phenomenon (entry and exit of firms of similar ages)

or a vertical one (differences in the average ages of the firms that enter and exit). Audretsch suggests two metaphors for getting an idea of the different forms of industrial dynamics in keeping with the different technological regimes. In some industries the rotation of firms is similar to the growth pattern of the trees in a forest: the new firms displace the older ones, in this case because the new ones introduce novel and better ideas, whereas the oldest organizations become increasingly rigid and are unable to adapt to a constantly changing environment. This metaphor of the forest applies to activities whose innovation conditions correspond to those of an entrepreneurial regime, because the high probability of introducing an innovation favours the survival of the new entrants and increases the vulnerability of the older firms, which are already committed to a technical paradigm. In other industries (the great majority), however, the rotation of firms is more like that of a conical revolving door, whose broad base revolves as a function of the rate of entry of small firms which exit from the market quite soon after entering it, while its upper part, where the older and more firmly established enterprises are located, hardly revolves at all. Only a small proportion of the new firms manage to survive and subsequently grow. Environments of a routinized nature, in which the established firms have advantages in terms of innovation, are usually associated with this second type of rotation.

2. The life cycle of an industry

In the economic literature, the concept of the life cycle of a product is associated with the work of R. Vernon, especially in one of his articles (Vernon, 1966). Broadly, the article in question shows how the changes that take place in production methods and the characteristics of the product from the time when it is first marketed give rise to a process of relocation of production units from the original innovative country (the United States, in the basic pattern used by Vernon), first to other developed nations and later to semi-industrialized regions. On the basis of a highly stylized division of this cycle into periods, Vernon held that the processes of manufacturing new products are marked at the beginning by:

i) high requirements for flexibility (in so far as the techniques are not definitively specified);

ii) demand conditions which are determined by differentiation of the product, with, in the final analysis, some tolerance for moderate differences in costs between producers, and

iii) the need for efficient means of communication with clients, suppliers and even competitors, because of the high degree of uncertainty about the real size of the new market, the efforts that rivals will make to capture portions of it, the specifications of the inputs required, and the relative performance of the different versions of the product.

Generally speaking, as time passes and demand for the product increases, both the product itself and the methods for its manufacture become more standardized, thus altering the cost structure and market conditions. Finally, this opens the way for the spatial relocation of the production activities (Vernon, 1966).

The literature on technological change and industrial demography has adopted the life-cycle approach but focuses it on the evolution of entry and exit rates and the total number of firms in a specific market, from the first appearance of a product on the market until its maturity. Very briefly, we can say that it identifies five stages in this process:

I. A firm or small group of firms initiates the activity;

II. The net rate of entry and hence the total number of producers rises abruptly;

III. The flow of new entrants then slows down until it is completely offset by the outward flow of firms and the net rate of entry sinks to approximately zero, while a peak is reached in the number of producers;

IV. A very low rate of entry and an increase in the rate of exit give rise to a negative net rate of entry, and

V. The rate of exit then declines until a rate of entry around zero is restored, with an intermediate number of producers in the market (compared with the minimum level in stage I and the maximum in stage III).

The forces behind this process stem from certain features of the industry which evolve together with the market structure in the course of the life cycle, particularly the level of uncertainty, the intensity of innovative activities, and the specific forms of innovation (of products or processes) and hence the sources of knowledge and the innovative agents.[9]

We will first of all set forth a stylized version of the basic elements in this sequence. After the initiation of a market, at quite an early moment in its evolution there will be an intensive entry process. Usually, the new entrants seek to open up a space in the new activity by introducing different versions of the basic product, so that there is a proliferation of product innovations. With time, the entry rate slows down, while the exit rate in-

---

[9] Another of the features of the industry which evolves with the life cycle is the degree of specialization at the plant level. This aspect will not be dealt with here, however (see Klepper, 1997).

creases, so that at first the two rates are balanced, but later the outward flow of firms even exceeds the entry rate, so that there is a pronounced reduction in the total number of producers (known as a "shakeout"),[10] stabilization of the market shares of the survivors, and a shift in technological change from product to process innovations.

Klepper (1997) gives a more detailed account of this. In an initial, embryonic stage of the market, there is a high level of uncertainty, product design is primitive, and the manufacturing process is carried out with non-specialized machinery. Numerous firms enter the market, and there is intense competition based on product innovations. In an intermediate or growth stage, production expands rapidly, product design begins to stabilize (that is to say, there is a gradual reduction in the frequency and importance of product innovations), and the production process becomes more sophisticated (human labour is gradually replaced by specialized production equipment). The entry rate goes down, as does the total number of producers.[11] Finally, in the stage of maturity, production grows slowly, the entry rate declines still further, and the market shares of the main firms become stabilized, while at the same time innovations of all types become more sporadic, since management, marketing and production techniques have reached a high degree of refinement.

In addition, the changes that accompany each of these transitions give rise to a redistribution of innovative advantages among the different types of producers. In the initial stages of the development of a new product, there is a great deal of uncertainty about the results of the new technology and activities tend to be subject to trial-and-error procedures, so in these circumstances the big firms do not have any special advantage in terms of innovation. The absence of advantages of scale, together with the prospects of high profits for successful innovation, attract a large number of new entrants. With time, however, two processes take place: first, as the technological frontier shifts, the search for innovations becomes more sophisticated and is divided up into well-defined tasks, giving rise to economies of scale in research and development spending, and second, the stock of knowledge built up by producers increases, raising the amount of such spending that new entrants must make in order to attain the level of capabilities of the

established firms. Both these processes are equivalent to entry barriers.

It is interesting to note that the origin of technological information is a key factor in this evolution. There are two types of sources of information whose relative weight changes in the course of the life cycle of an industry (Gort and Klepper, 1982):

i) firms which are already established in the activity in question (I_1): information from this source is made up of transferable and non-transferable knowledge. The latter is generated through practice (learning by doing), tends to accumulate with time, and functions as an entry barrier, and

ii) alternative sources of information (I_2), such as firms operating in technologically related industries, independent inventors, equipment producers, etc.: the existence and size of these sources have a positive effect on the entry rate by reducing the value of practical experience in the activity in question –an effect which is further strengthened by the difficulty of carrying out market transactions of such knowledge. As explained earlier, when agents outside the established firms possess valuable information on a new product, the best solution for taking advantage of such knowledge is probably the formation of a new enterprise.

The hypothesis put forward by Gort and Klepper (1982) may be summed up as follows: in the first stages of the cycle most innovations come from I_2, but as from stage III the balance favours I_1 and the stock of this type of knowledge accumulated by the established firms begins to operate as an entry barrier.

The kinship between this hypothesis and that based on technological regimes is clear. Indeed, to a large extent the life cycle may be interpreted as the history of a transition from an entrepreneurial regime to a routinized regime.

Finally, we may say that many, although not all, of the products dealt with in the empirical studies display the pattern of industrial evolution described above. Klepper and Miller (1995), for example, classify a set of products in terms of whether or not their stylized cycles contain the intensive shaking-out process typical of stage IV. Klepper (1997) also analyses industries which have displayed other patterns, marked basically by the fact that the entry flow does not stop and there is not therefore a pronounced reduction in the number of firms.

3. Sectoral patterns of innovative entry

In short, if we combine the approach based on technological regimes with that based on the life cycle of in-

---

[10] For alternative views on the origin of shakeouts, see Klepper and Miller (1995) and Klepper (1996).

[11] A typical shakeout involves the exit of 40-50% of the number of producers that existed at the highest point of stage III (Gort and Klepper, 1982).

dustries, we see that the two situations in which the established firms have advantages in their access to knowledge and are therefore in a position to impede the entry of new firms are as follows:

i) in certain activities in which technological change is very important, routinization occurs when innovations can only be achieved through big investments in physical and human assets (research and development departments, laboratories, etc.) and the gradual accumulation

of knowledge that these assets generate (the pharmaceutical industry is one example of this), and

ii) in many industries, after an initial period when there is a considerable increase in the number of producers and product versions, a phase is reached in which technology stabilizes around a dominant product design and a set of production methods and the firms which are already using those methods are in a better position to make gradual improvements and refinements in them.

III

Technology and industrial dynamics
in the semi-industrialized countries

In this section, we transfer the theoretical analysis of the links between technological change and industrial dynamics to the particular conditions of the manufacturing sectors of developing countries, especially those of Latin America. In order to do this, we need to link up the findings of the literature on technological change in those countries with the main concepts on the links between technology and industrial dynamics, between the opportunity and sources of technological change, and between the accumulation of knowledge and the appropriability of the benefits of the technical progress generated by the firms. The studies on technological change in Latin America and other semi-industrialized countries (SICs) provide valuable information on these concepts, although they do not usually relate them with industrial dynamics phenomena such as the presence or absence of "innovative" advantages in new firms, or else in established ones.

At this point, we must briefly analyse a methodological question that cannot be ignored when dealing with the matters described in the following pages. The notion of changes in the (economic and institutional) environment on which the theoretical schemes and models conceived in the developed countries for identifying the patterns of microeconomic and industrial behaviour are based naturally reflects the particular experience of those economies. It is easy to imagine the potential difficulties that would arise if the concepts and theories developed for that environment were applied indiscriminately to the SICs, where most of the agents are often confronted with economical and political changes of a particularly intense and unpredictable nature. This is especially so because, in a competitive environment which has suf-

fered a succession of disturbances of this type in the past, the agents must adapt their individual behaviour by developing decision-making rules and strategies in keeping with this situation. In so doing, they develop special patterns of aggregate response (at the meso-economic and macroeconomic levels) which, like the rules for their individual behaviour, are different from those observed in countries with more stable economic and institutional environments. These matters have been under discussion for some years now in Latin American academic circles.[12] What we should note here is that the implications with respect to industrial dynamics that can be drawn from the literature on technological progress in the SICs are usually based on a supposedly stable economic and institutional environment.

Finally, before entering on the heart of the analysis, the following subsections: i) provide some details of the specific forms that technological change can take in the SICs and the reasons why they are of interest to this study; ii) analyse the main sources of knowledge on which technological progress in manufacturing is based in those countries, and iii) deal with the link between the institutional structure (in which there are dramatic contrasts between the SICs and the most highly developed economies) and industrial dynamics, through the role played by that structure in providing technical information from sources outside the established firms.

---

[12] See, for example, Fanelli and Frenkel (1996). Kosacoff and Ramos (1998) deal with these matters from the point of view of the industrial policy debate in the semi-industrialized countries.

1. Technological change in the semi-industrialized countries

The term "technology" refers to the activities involved in the processing of inputs into products or the technical information package containing a list and description of those activities. Thus, technological change must be understood as the introduction of changes in production activities or –which amounts to the same thing– in the set of technological information on which they are based. The knowledge base underlying this progress is fed through the formal or informal research activities carried out by the firm in question and the learning process stemming from its own production experience.

However, industrial enterprises in SICs generally only engage in certain types of innovative activities, which are basically the following:[13]

i) introduction of products and processes which are new to the local economy;

ii) adaptation of new products and processes to local conditions, and (possibly)

iii) introduction of improvements in process performance and/or in the products manufactured.

This classification is based on the specific content of innovative activities. If they are regrouped according to their effects on industrial dynamics, this gives two major sets of activities of this type which give rise to different industrial evolution processes.

The first of these is linked with the inauguration of local production of a good. An agent or small group of agents imports a technology package (product and process engineering and industrial organization scheme) and adapts it to local conditions, thus initiating local supply and very likely displacing imported products to some extent. The nature and intensity of this first wave of entrants will depend on a wide range of factors related to demand (size of market, profitability, expected growth, trade and tax policy) and supply (number of firms operating in related activities for which the new line of business represents an attractive means of diversification or integration, existence of other sub-groups of agents possessing the necessary capabilities for engaging in the new activity, ease of access to the necessary factors and technical and financial resources, etc.).

The second one involves the introduction of changes in the production process or product characteristics in an existing market. An agent or group of agents introduces a technical novelty, generally developed by a firm located abroad. In this case, the industrial dynamic takes place in one of the following alternative ways: the local agent making the innovation consists of one or more already existing firms which, among other objectives, seek to increase their profits and expand their share of the market, or the innovation represents the means by which one or more new local firms hope to take over a part of the market currently supplied by existing firms.

It is this second type of innovative episode which is of interest to us here. Consequently, the analysis must be aimed at determining if the established firms or the potential new entrants display systematic differences in terms of their access to the technological know-how needed to introduce the type of innovations made in SICs. This in turn makes it necessary to adjust the notion of innovative advantage on which the literature on the developed countries is based, in order for it to reflect the different nature of innovative activities in the SICs. It is suggested here that the capacity of firms in the latter countries to undertake this type of technical changes depends on the degree of progress they have made in two types of technological activities:

i) those which are related with imitation (search, assessment, implementation and adaptation) and which determine the capacity of firms to assimilate technical progress generated by other agents, and

ii) those associated with the achievement of incremental improvements of processes or products within a given type of technology.

The existence in SICs of systematic differences between new and established firms in terms of access to these types of capabilities is the central element in the following sub-sections. Studies on technological change in SICs usually concentrate on identifying episodes of such change in local plants, on its organizational (in-house) and institutional (links with external agents) sources, and on its impact on productivity and export performance. In contrast, the relationship between the form of accumulation of technical knowledge, innovation and the rotation of firms at the sectoral level has been given little attention.[14]

---

[13] See Fransman (1985), Katz (1976 and 1987) and Teitel (1987).

[14] This is probably due to the features which predominated in the competitive environment of SICs up to a few years ago. The relationship between innovation and industrial dynamics is only important in so far as the effectiveness in introducing new products or the level of technological excellence determine the patterns of profitability and/or survival among the population of firms. It seems clear that, in the present conditions of competition defined by the structural reforms resulting from the Washington Consensus, the capacity of innovation and imitation possessed by firms, as well as market structure and dynamics, have become more interdependent.

2. External sources of knowledge and the capacity for imitation

As already explained, industrial activities differ as regards the form of their processes of creation of new knowledge. The differences in technological opportunity and conditions of appropriability help to define what we described earlier as technological regimes and have also given rise to the well-known taxonomy developed by Pavitt (1984). According to this classification, the types of manufacturing activities that exist in SICs mostly fall into three categories: "supplier-dominated", "scale-intensive" and "specialized suppliers".[15] However, whereas in the advanced countries the firms in the last two groups are normally generators of innovations, in the relatively less developed countries they are typically imitators of imported technology. Thus, it is usually considered (Cooper, 1991) that, because of the way they incorporate technological change, most of the manufacturing enterprises in SICs actually fall into the category of "supplier-dominated" activities. The firms classified in this group are marked by their adoption of innovations incorporated in production equipment and/or the intermediate inputs they use. They are imitators of innovations made by other firms, so that the intensity of their "innovative-imitative" activities is associated with the rate at which they incorporate the most modern equipment and materials, generally developed in the most advanced countries.

It was argued earlier in this article that in the developed countries the industries where there are abundant sources of technological know-how outside the established firms offer greater opportunities for innovative entry. It should be noted that the literature on technological regimes and the industrial life cycle in the developed countries mostly presupposes that the new entrants have the capacity to generate and implement innovations, that is to say, to translate the available information into novelties with economic value. In the SICs, however, innovation consists mainly of imitation, and the capacity for such imitation cannot be taken for granted.[16]

This is so because the technological know-how incorporated from the outside is largely complementary to

that which the firms possess themselves, so that, as Nelson (1987) notes, imitation is not by any means a trivial technological activity. This complementarity takes two forms (Bell and Pavitt, 1993):

On the one hand, the acquisition of technology may be only the first step in a more extensive process of technological change which also includes the achievement of improvements in the performance of the new methods or products, and the efficacy with which these subsequent phases are carried out depends on the technological capabilities already accumulated by the firm (this aspect will be further developed in subsection 3 below).

On the other hand, the experience accumulated through the repeated execution of imitative activities gives the management and shop-floor staff a substantial advantage when it comes to selecting new acquisitions in the future and adapting them to local technical and economic conditions: the search for new products and processes requires that the firm should possess the necessary technological capabilities, so that in order for a firm to acquire additional know-how through such a search it needs a certain amount of prior knowledge (Fransman, 1985, p. 584).

The complementarity between external know-how and that which the firms have built up themselves is due to two main reasons.

The first is the need to adapt the technologies incorporated (Fransman, 1985; Katz, 1976 and 1987; Teitel, 1987). A firm in a less-developed country cannot invest in a technique used in a developed country without modifying it in one way or another: it cannot use that technique (begin to produce with it) without some kind of modification.[17] Indeed, the set of small adaptations made by an imitator located in an SIC can give rise to a new production function: once the various limitations and problems of the original technological designs from the developed country have been overcome through 'minor' local innovations, the new technology package will necessarily be different (and in many respects more 'appropriate') than that originally acquired abroad (Katz, 1987, pp. 46-47). This means that even if there is abundant technological know-how coming from sources outside the firms already established in an industry, this will not result in a situation where the new firms have a permanent stimulus to adopt the latest versions or try new permutations of such know-how in order to enter the market and compete with the established firms.

---

[15] For a more detailed explanation, see Pavitt (1984).

[16] The actual mode of access to technological know-how usually takes various forms. Although this aspect is not analysed in the present article, the different ways of access to knowledge are probably associated with different technical capability requirements on the part of the recipient and therefore also have different implications in terms of industrial dynamics.

[17] Nelson, quoted in Teitel, 1987, p. 113, footnote 10.

The second reason derives from the implicit and tacit nature of the technological know-how transferred. As the information received by the buyer is always less complete than that possessed by the seller, the capacity of the former to incorporate new principles and production practices will depend on his skill in deciphering the instructions and turning them into a set of effective and efficient routines and procedures. The more completely "embodied" the know-how is, however, the less vital this complementarity is likely to be. In many manufacturing sectors there are firms which have little technical experience but have access to the most modern equipment. However, their capacity to reach (or even surpass) pre-set levels of performance or to subsequently make incremental improvements (the second component in the capacity for imitation) will depend in part on the technological experience they have accumulated.

According to Fransman, "one important aspect of the 'selection environment' [the features of the environment that affect the selection of production technology in the SICs] is the knowledge possessed by the selector and the costs of alternative ways of obtaining knowledge. Furthermore, technological capabilities will play an important role in facilitating adequate choice of technology and in making the technique work in a satisfactory manner, once chosen" (Fransman, 1985, p. 583). In other words, "importing technology and generating technology domestically are usually not mutually exclusive alternatives" (Fransman, 1985, p. 615).[18]

Evenson and Westphal (1988) concur that the tacit nature of most technological know-how means that the effectiveness of technology transfers to firms in less developed nations depends on the technological capabilities of the recipients. These capabilities are also crucial for overcoming what those authors call the sensitivity of technology to circumstances,[19] that is to say, for carrying out the adjustments and adaptations needed for it to operate in circumstances different from those for which it was developed.[20]

Most of the authors also emphasize that the capacity of the manufacturing firms of SICs to imitate technology successfully cannot be taken for granted.

According to Bell and Pavitt, the difference with the firms of developed countries lies in the fact that there "those adopting and using the technology transferred normally already possess (albeit to different degrees) the particular type of knowledge and skills needed to play a creative technological role. In the developing countries, however, these capabilities usually need to be built up in order to take full advantage of the dynamic advantages of the spread of technology" (Bell and Pavitt, 1993, p. 162).

Lall expresses similar views, stating that in SICs technological change is more "localized" than in the more advanced countries, because the SICs have a less complete knowledge of the range of technological options available. He adds that "firm-level differences in technical efficiency persist everywhere, but firms in developing countries generally display both wider dispersions and lower average levels of efficiency in given activities than firms in developed countries" (Lall, 1994, p. 2). Subsequently, the same author explains that, unlike the industrialized countries, in the less developed countries the capacity to handle the existing technologies cannot be taken for granted but, on the contrary, is in fact the goal of technological activity.

For Lall, the endogenous accumulation of know-how has a crucial impact on the capacity of firms to imitate technologies right from the start of the transfer process: "Part of the difference [between the markets for technology and for goods] resides in the inherent difficulty of valuing the product and in the unequal distribution of knowledge between buyer and seller. In developing countries, this is exacerbated by the inadequacy of buyer skills and knowledge in both buying and implementing technologies" (Lall, 1994, p. 14).

In short, the capacity for imitation is associated with the past accumulation of know-how within the firms. This means that in order to be able to take advantage of the technical information available in the world at large it is necessary first of all to build certain technical capabilities. Thus, firms with experience in a particular production activity will be at an advantage for successfully incorporating exogenously generated technical progress

---

[18] Naturally, the degree of complementarity between the know-how acquired and that already possessed (the amount of in-house know-how needed to incorporate external know-how) will vary as a function of the characteristics of the technology (its complexity, the degree to which it contains implicit elements, etc.). It is also very likely that the degree of complementarity will be different for each of the two components of the capacity for imitation (incorporating technology developed by other agents or generating one's own flow of minor innovations).

[19] I.e., its sensitivity to changes in the environment in which it is used.

[20] The relative weight of the two factors differs from one sector to another: in general, sensitivity to circumstances particularly affects

transfers of technology in agriculture, whereas the tacit nature of know-how is relatively more important in flows of industrial know-how. "Nevertheless, more often than not, in order to reach adequate levels of productivity industrial processes must be adjusted to suit the particular circumstances in which they are to be used" (Evenson and Westphal, 1988, p. 2248).

and adapting it to local technical and economic conditions.

It must be clearly understood that while the innovative advantage which favours the established firms does not rule out the establishment of other new firms with relatively more advanced technologies, the first-named firms will systematically be in a better relative position to carry through the incorporation of technical progress. Innovative entry could adversely affect established firms financially, because of the sudden obsolescence of assets which must now be renewed but are not yet amortized, but –and this is the most important argument– it is much less likely that they will be at a technical disadvantage compared with new entrants in terms of identifying, evaluating, acquiring, assimilating and adapting the new techniques.[21]

3. Incremental improvements in products and processes

It has been said that innovative advantage is based on the capacity for the effective imitation of technology and on the skills needed to generate incremental technical improvements which will secure increases in productivity and/or modifications in the products. In the previous sub-section we showed that in order to take advantage of the technical information available in the environment (capacity for imitation) it is necessary first of all for the firms to have built up certain technical capabilities of their own. We will now analyse the origin of the knowledge needed in order to gain access to this second form of innovation.

It is important to bear in mind that, generally speaking, when a new product begins to be manufactured or a new production technique is used for the first time in an SIC the new technology has already ceased to undergo substantial modifications in the markets where it was originally developed and has entered on a mature stage in which any improvements and refinements are only marginal. Whatever the technological characteristics of the sector, however, these minor technological changes are precisely those that firms with production experience in the activity in question are most likely to introduce, and that new firms will find it most difficult to imitate (Agarwal and Gort, 1996; Gort and Klepper,

1982). In this sense, the literature on technological change in Latin America seems to suggest that incremental improvements are the result of the accumulation of know-how and endogenous efforts by established firms and that they are not therefore usually used by new firms in order to break into an activity.

As far back as the late 1960s and early 1970s, in a study on manufacturing in Argentina, Katz (1976) showed the importance of enterprise-level technological efforts in the SICs as sources of productivity gains. Firms had to build up their own technological capacity in order to make proper use of imported technology, and that capacity was responsible for a significant proportion of the efficiency gains obtained. Such gains were also aided by the increase in production (through the Verdoorn effect and/or spontaneous learning processes like "learning by doing") and were therefore also associated with the production history of the firms.

Katz admitted that in practice the firms studied did not generate "new and better" innovations, but many of them allocated resources to the provision of a certain amount of "minor but independently generated" innovations, not only in order to adapt but also to marginally improve production processes or product designs obtained from imports. While not overlooking the fact that the dynamism of innovation slackened in the course of the commercial history of the products concerned, Katz emphasized that the manufacture of mature products in SICs was associated with the introduction of improvements which were in turn the result of the learning process that took place in the firms involved.

A very extensive set of studies[22] on technological change in Latin American manufacturing plants comes to the same conclusion: endogenous technical efforts generate a flow of minor technical improvements which account for most of the increases in productivity.

According to Katz (ed.), 1987, after they have begun to operate firms pass through a sequence of technical learning processes. First of all they develop product-related innovative capabilities. Skills in the fields of process engineering and industrial organization (planning and control of the production process, etc.) begin to build up very slowly, in an informal manner, and only make a substantial leap forward a good deal later, generally as a result of some sudden important development

---

[21] However, the investment required in order to enter the market usually contains a proportion of sunk costs, and if these exceed the expected post-entry benefits the possibility of competing would be ruled out (Stiglitz, 1987). The need for this clarification became evident thanks to the observations of an anonymous referee.

[22] These studies form part of a very broad IDB/ECLAC/UNDP programme carried out in the early 1980s. A large part of the case studies are contained in Katz (ed.), 1987. The studies dealt in general with medium-sized and large industrial firms.

(such as the entry into the market of a serious competitor). Through these types of learning processes, together with their practical production experience, firms eventually acquire the capacity to generate a flow of minor innovations.

The set of studies in question also showed that most of the firms studied started off with technologies of low complexity and capital intensity and subsequently progressed, on the basis of the gradual accumulation of technical skills, to more complex and automated forms of production.

Similar views are expressed by Evenson and Westphal, who consider that technology always needs time to begin to operate at its highest level of productivity. The crucial point is that accumulated experience is a fundamental factor of differentiation: "the initial level of productivity, as well as the time and resources required to achieve the potential productivity, depend on the starting level of mastery" (Evenson and Westphal, 1988, p. 2262).

Taken together, the studies referred to above suggest that there is a process whereby the rate and quality of technical learning determine the overall present and potential technological possibilities of the plant, its products and its services. They also suggest that, in the total universe of established firms in a given activity, the solidity of their innovative advantages is positively associated with their age. In short, the established firms with experience in the activity in question not only enjoy advantages in terms of the incorporation and adaptation of new technology (imitative capacity) but are also in a better position than potential new entrants both to put variations of the standard product on the market and to operate with production methods more appropriate to local conditions or more efficient than those that agents outside the industry can import.

In some sectors (especially in the production of capital goods) there are other important sources of technological information. This is basically information stemming from inter-firm relations (Fransman, 1985): relations between the user and the producer of the equipment (a very important relation because, as already noted, in order for the SICs to use technologies developed in other countries they must adapt them to local conditions) and between the producer of the equipment and the supplier of parts and components. It seems clear that here too, in terms of access to these sources of technical information, established firms are bound to be in a better position than firms in related sectors or other potential entrants.

Nevertheless, a couple of additional elements must be borne in mind. These concern the less and less auto-

matic nature of "learning" through "doing". The two processes of change implicit in this formula (production experience which turns into know-how and know-how which evolves into technical change) are subject to the specific actions of the firms (Fransman, 1985, p. 595). Consequently, production experience *per se* can only be considered as a necessary but less and less sufficient condition for guaranteeing a flow of minor innovations which will have appreciable effects on productivity and on product qualities and performance.[23]

Cooper (1991) and Bell and Pavitt (1993) concur in finding that a characteristic feature of the present conditions of generation of new technical know-how in the developed countries is the growing gap between the type of information required by firms for using a technology and that required in order to modify it. It is increasingly difficult to build up innovative capacity without taking measures, and above all making investments, explicitly aimed at that purpose.[24] Thus, the skills and experience needed to generate and manage technical change are increasingly connected with the activities of specialized research and development laboratories, design offices, project management teams and production engineering departments. In the terminology used in the literature on industrial dynamics, this may be described as a tendency towards the "routinization" of technological change.

Increasingly, production experience *per se* can only serve to expand the capacity for innovation and/or imitation if it is backed up by measures and expenditure deliberately aimed at learning, thus adding another dimension to the analysis of innovation/imitation advantages in industrial dynamics. Thus, while it has already

---

[23] Whereas investment in operational know-how is a necessary condition for entry (and also, in certain conditions, for survival), investment for the development of technological capabilities is discretional. According to Bell and Pavitt (1993), as the return on such capabilities is uncertain or very difficult to evaluate, firms will tend to under-invest in them. Cooper (1991) finds that in SICs the learning process is very often a failure: firms content themselves with simply attaining the conditions needed for applying the technology and do not bother to familiarize themselves with its basic principles; the resources allocated to the assimilation of knowledge are insufficient, and this explains why the learning process fails more frequently than in the developed countries. Thus, the trade-off between the degree of advance of the technology acquired, its ease of assimilation and the amount of new local knowledge it can generate is solved by seeking the lowest level of risks: the technological objectives are modest and therefore relatively easy to attain, but the technical progress and productivity gains are only small.

[24] One of the reasons for this is the growing difficulty in carrying out reverse engineering of parts, sub-assemblies and components (this note was included in response to the observations of an outside referee).

been noted that there is a difference in terms of imitative capacity between firms of different ages (potential entrants and newly-formed firms as compared with experienced established firms), what is suggested now is that (in conditions of imperfect capital markets) there is an increasingly marked difference between firms of different economic size, because of the growing need to "formalize" the process of searching for new technology.

Thus, in short the capacity for imitation/innovation is linked with production experience and the deliberate learning efforts made by firms in the course of their production history, so that innovative entry is relatively difficult.

4. Institutions and learning

The recent literature on innovation in the developed countries highlights the importance of the depth and quality of the institutional infrastructure[25] linked more or less directly with industrial learning processes. From the point of view of industrial dynamics, institutions play two types of roles.

Firstly, in line with the concepts of the life cycle of industries and technological regimes, institutions can help –as the depositaries of relevant technological information– to offset to some extent the technological and innovative disadvantages of firms which are new entrants or have as yet little experience in the activity in question. The effectiveness with which the institutional infrastructure fulfills this function will depend on its capacity to limit the private appropriability of the knowledge generated by existing firms, thus facilitating the spread of technology to new firms and those with less production experience.

Secondly, in the context of the current debate on industrial and technological policy, institutions are acknowledged to have a leading role as facilitators of learning processes. In the developed countries, universities, public research laboratories and similar bodies complement the search for new knowledge being made by industrial firms. This relationship is complementary because of the nature of the interchange involved. Research bodies do not usually generate complete innovations that firms can appraise and possibly adopt, but rather provide some elements that firms can combine with the results of their own technological search processes. The example of Taiwan's technology policy cited by Bell and Pavitt (in Taiwan the public institutions "socialize" the learning process by acquiring foreign technology, making the initial assimilation effort, and then providing training services for disseminating it to private firms) clearly illustrates the kind of externalities that the institutional infrastructure can generate.

The precarious nature of the system of institutions connected with technological learning in Latin America has often been noted: the absence or weakness of such institutions can cause irreparable harm in this field. Although the allocation of specific resources is a necessary condition for proper learning processes in firms, the absence of suitable external institutional conditions can also inhibit the appearance of such processes (Cooper, 1991, p. 15).

Weakness of the institutional system appears to have implications in connection with the first of the aspects mentioned, in so far as the absence of an abundant flow of technological information from sources outside the established firms gives the latter an advantage in terms of innovation over firms that have little production experience.

---

[25] The term "institution" is used here with reference to the depth and quality of the system of intermediate organizations rather than the existence, efficacy and efficiency of rules and standards.

ease of entry in general. Firstly, because the entry of new firms represents the culmination of investment projects which must contain not only a given amount of technical and economic knowledge (whether novel or not) but also various additional elements, and the differences between sectors as regards the ease with which potential entrants can gain access to those elements, together with other structural features which differ from one industrial branch to another, open up the way for the study of a variety of dynamic regimes. Secondly, because the fact that the innovation that takes place in SICs is generally associated with a high degree of accumulativity of technological know-how (which generates unfavourable conditions for innovative entry) does not prevent the production activities from nevertheless having different underlying learning conditions and, hence, also differing in other aspects of their economic selection mechanisms.[26]

2. Effects of the entry of new firms on industrial dynamics

The most marked effect in this respect will depend on the role played by the appearance of new local firms within the broader evolution of the population of industrial enterprises in the SICs. If innovative entry is structurally impeded by the knowledge conditions underlying the type of innovations that take place in the SICs, then their role will tend to be linked with the relation between installed capacity and present and expected demand. In this case, the main contribution of local new entrants will be to add production capacity, since their technological packages will tend to be the same as those used by the established firms. With time, the arrival of new competitors may force the latter to speed up their technological modernization process, but it seems unlikely that the new entrants will be able to set off an episode of "creative destruction" which will force out all or part of the established firms. The two scenarios are radically different in terms of the profitability expectations of the potential new entrants. Thus, the "neo-classical case" (entry as a factor tending to bring the supply conditions into balance) would appear to prevail rather than the "Schumpeterian case" (entry as a vehicle for innovation).

This also means that the industrial dynamics will tend to be governed by the evolution of the quantitative

imbalances in the supply conditions, rather than by the introduction of innovations (with respect to the local economic environment) which would set off adjustment processes in the composition of the population of existing enterprises. Quite apart from the decisive weight acquired in it by the strategies of the leading firms in each sector, this pattern would appear to be particularly sensitive to macroeconomic fluctuations and to the particular way that each manufacturing activity links up with the different components of aggregate demand.

Some descriptions of the evolution of manufacturing enterprises in SICs after their initial establishment give at least a hint of what might be considered as the paradigmatic post-entry performance path, as do the passive and active learning models for developed countries.[27] Although it is necessary to bear very much in mind the differences between the present context and that which prevailed when he was making his study, Katz (ed., 1987) finds that from the moment of their establishment firms make gradual progress in terms of the type of technological capabilities they develop[28] and slowly deepen their technological commitment in terms of the degree of automation and continuity of the production process and technical complexity in general. He also finds that, after the entry of pioneering firms using relatively simple techniques, with little automation, with time "most of them, as well as newly arriving competitors in the same industries, opt for more capital-intensive technologies" (Katz, 1987, p. 30). In an industrial dynamics context like that described here, this passage suggests that the rate of technological change is dictated by the rate of the learning process of the firms which are already in the business rather than by the entry of new and innovative firms.

The better conditions of access to new equipment and foreign technology offered by the present macroeconomic and regulatory setting –after the structural reforms in Latin America in the 1980s and 1990s– have expanded the technological opportunities for established firms (rather than for potential new entrants), especially for those which are most active in their search for new technology.

It is also interesting to reflect on the ways in which the generic pattern of industrial dynamics described here

---

[26] For an exposition on the diversity of learning channels and their connection with different incremental technical change paths in United States industry, see Malerba (1992).

[27] The passive learning model most frequently cited is that of Jovanovic (1982). Extensions of that theory may be found in Frank (1988) and Hopenhayn (1992). Active learning models are presented in Ericson and Pakes (1995) and Pakes and Ericson (1998).

[28] In the 1980s the situation described by Katz was that first of all firms progress in the field of product technology and subsequently in the areas of process technology and organization of production.

inhibits or promotes the different forms of entrepreneurship.

Various different forms of entry into the market by new local manufacturing firms in the SICs may be envisaged.[29] One of them is pioneering entry, when a firm begins local production of a product which is new for the domestic market. Another is entry into an existing production activity, which may take two different forms, depending on whether or not the new enterprises bring substantially new technologies with which to challenge the market shares of the established firms.

It is precisely the form of entry which innovates on the technical, commercial and organizational practices of the established firms which is most inhibited by the better access of the latter to knowledge which would be useful for purposes of innovation. In other words, the characteristics of the process of innovation in the SICs appear to militate against the type of entry which would mean a competitive challenge (based on innovations) to the position of the existing firms.

Finally, there is a clear need to extend the analysis to another form of entrepreneurship which is consistent with the theories set forth here. This is the creation of innovative enterprises on the basis of staff who have left other organizations (generally large ones) already operating in a given line of business. This form of creation of new enterprises is an aspect of what has been called *intrapreneurship* (Wennekers and Thurik, 1999).

(Original: Spanish)

---

[29] As already noted, there are few cases where firms from SICs generate new products or processes that are also innovative with respect to the more advanced economies.

Bibliography

Agarwal, R. and M. Gort (1996): The evolution of markets and entry, exit and survival of firms, *The Review of Economics and Statistics*, vol. 78, N° 3, Amsterdam, The Netherlands, Elsevier Science Publishers B.V.

Arrighetti, A. (1994): Entry, growth and survival of manufacturing firms, *Small Business Economics*, vol. 6, N° 2, Massachusetts, Kluwer Academic Publishers.

Audretsch, D. (1991): New-firm survival and the technological regime, *The Review of Economics and Statistics*, vol. 73, N° 3, Amsterdam, The Netherlands, Elsevier Science Publishers B.V., August.

———(1995a): *Innovation and industry evolution*, Cambridge, Massachusetts, The MIT Press.

———(1995b): Innovation, growth and survival, *International Journal of Industrial Organization*, vol. 13, N° 4, Amsterdam, The Netherlands, North-Holland.

———(1997): Technological regimes, industrial demography and the evolution of industrial structures, *Industrial and Corporate Change*, vol. 6, N° 1, London, Oxford University Press.

Bain, J. (1956): *Barriers to New Competition. Their Character and Consequences in Manufacturing Industries*, Cambridge, Massachusetts, Harvard University Press.

Baldwin, J. and P. Gorecki (1991): Firm entry and exit in the Canadian manufacturing sector, 1970-1982, *The Canadian Journal of Economics*, vol. 24, N° 2, Toronto, Canadian Economics Association.

Bell, M. and K. Pavitt (1993): Technological accumulation and industrial growth: contrasts between developed and developing countries, *Industrial and Corporate Change*, vol. 2, N° 2, London, Oxford University Press.

Cooper, C. (1991): *Are Innovation Studies on Industrialized Economies Relevant to Technology Policy in Developing Countries?*, Working paper N° 3, Maastricht, The Netherlands, United Nations University (UNU)/ Institute for New Technologies (INTECH), June.

Dunne, T., M. Roberts and L. Samuelson (1988): Patterns of firm entry and exit in U.S. manufacturing industries, *Rand Journal of Economics*, vol. 19, N° 4, Santa Monica, California, RAND.

———(1989): The growth and failure of U.S. manufacturing plants, *The Quarterly Journal of Economics*, vol. 104, N° 4, Cambridge, Massachusetts, Harvard University.

Ericson, R. and A. Pakes (1995): Markov-perfect industry dynamics: a framework for empirical work, *Review of Economic Studies*, vol. 62, N° 1, Brighton, U.K., University of Sussex.

Evenson, R. and L. Westphal (1988): Technological change and technology strategy, in H. Chenery and T.N. Srinivasan (eds.), *Handbook of Development Economics*, vol. III, Amsterdam, The Netherlands, Elsevier Science B.V.

Fanelli, J. M. and R. Frenkel (1996): Estabilidad y estructura: interacciones en el crecimiento económico, in J. Katz (ed.), *Estabilización macroeconómica, reforma estructural y comportamiento industrial, estructura y funcionamiento del sector manufacturero*, Buenos Aires, Alianza Editorial.

Frank, M. (1988): An intertemporal model of industrial exit, *The Quarterly Journal of Economics*, vol. 103, N° 2, Cambridge, Massachusetts, Harvard University.

Fransman, M. (1985): Conceptualising technical change in the Third World in the 1980s: An interpretative survey, *The Journal of Development Studies*, vol. 21, N° 4, London, Frank Cass.

Geroski, P. A. (1991): *Market Dynamics and Entry*, Oxford, U.K., Basil Blackwell.

Gort, M. and S. Klepper (1982): Time paths in the diffusion of product innovations, *The Economic Journal*, vol. 92, N° 307, Oxford, U.K., The Royal Economic Society, September.

Hopenhayn, H. (1992): Entry, exit, and firm dynamics in long run equilibrium, *Econometrica*, vol. 60, N° 5, Evanston, Illinois, Econometric Society.

International Journal of Industrial Organization (1995): vol. 13, N° 4, Amsterdam, The Netherlands, North-Holland.

Jovanovic, B. (1982): Selection and the evolution of industry, *Econometrica*, vol. 50, N° 3, Evanston, Illinois, Econometric Society.

Katz, J. (1976): *Importación de tecnología, aprendizaje e industrialización dependiente*, Mexico City, Fondo de Cultura Económica (FCE).

———(1987): Domestic technology generation in LDCs: A review of research findings, in J. Katz (ed.), *Technology Generation in Latin American Manufacturing Industries*, London, Macmillan Press.

Katz, J. (ed.) (1987): *Technology Generation in Latin American Manufacturing Industries,* London, Macmillan Press.

Klepper, S. (1996): Entry, exit, growth, and innovation over the product life cycle, *The American Economic Review*, vol. 86, No. 3, Nashville, Tennessee, American Economic Association.

———(1997): Industry life cycles, *Industrial and Corporate Change*, vol. 6, No. 1, London, Oxford University Press.

Klepper, S. and J. Miller (1995): Entry, exit and shakeouts in the United States in new manufactured products, *International Journal of Industrial Organization*, vol. 13, N° 4, Amsterdam, North-Holland.

Kosacoff, B. and A. Ramos (1998): Consideraciones económicas sobre la política industrial, in B. Kosacoff (ed.), *Estrategias empresariales en tiempos de cambio. El desempeño industrial frente a nuevas incertidumbres*, Buenos Aires, ECLAC Office in Buenos Aires/Universidad Nacional de Quilmes.

Lall, S. (1994): Technological development, technology impacts and industrial strategy: a review of the issues, *Industry and Development*, N° 34, Vienna, United Nations Industrial Development Organization (UNIDO), June.

Malerba, F. (1992): Learning by firms and incremental technical change, *The Economic Journal*, vol. 102, N° 413, Oxford, U.K., The Royal Economic Society, July.

Malerba, F. and L. Orsenigo (1995): Schumpeterian patterns of innovation, *Cambridge Journal of Economics*, vol. 19, N° 1, London, Oxford University Press.

———(1997): Technological regimes and sectoral patterns of innovative activities, *Industrial and Corporate Change*, vol. 6, N° 1, London, Oxford University Press.

Mata, J. (1994): Firm growth during infancy, *Small Business Economics*, vol. 6, N° 1, Massachusetts, Kluwer Academic Publishers.

Mata, J. and P. Portugal (1994): Life duration of new firms, *The Journal of Industrial Economics*, vol. 42, N° 3, Oxford, U.K., Blackwell Publishers.

Nelson, R. (1987): Innovation and economic development: Theoretical retrospect and prospect, in J. Katz (ed.), *Technology Generation in Latin American Manufacturing Industries*, London, Macmillan Press.

Nelson, R. and S. Winter (1982): *An Evolutionary Theory of Economic Change*, Cambridge, Massachusetts, Cambridge University Press.

Orr, D. (1974): The determinants of entry: A study of the Canadian manufacturing industries, *The Review of Economics and Statistics*, vol. 61, N° 1, Cambridge, Massachusetts, Harvard University Press.

Pakes, A. and R. Ericson (1998): Empirical implications of alternative models of firm dynamics, *Journal of Economic Theory*, vol. 79, N° 1, Philadelphia, United States, Academic Press.

Pavitt, K. (1984): Sectoral patterns of technical change: towards a taxonomy and a theory, *Research Policy*, vol. 13, N° 6**,** Amsterdam, The Netherlands, North-Holland Publishing Company.

Schumpeter, J. (1912) *Teoría del desarrollo económico*, Mexico City, F.C.E., reprinted 1978.

(1942): *Capitalismo, socialismo y democracia,* Madrid/Mexico City/Buenos Aires, Aguilar, 1961 edition.

Stiglitz, J. (1987): Technological change, sunk costs, and competition, *Brookings Papers on Economic Activity*, N° 3, Washington, D.C., The Brookings Institution.

Teitel, S. (1987): Towards an understanding of technical change in semi-industrialized countries, in J. Katz (ed.), *Technology Generation in Latin American Manufacturing Industries*, London, Macmillan Press.

Van Dijk, M. (1998): *Technological regimes and industrial dynamics: the evidence from Dutch manufacturing*, paper presented at the Conference on The Economics of Industrial Structure and Innovation Dynamics, Lisbon, Portugal, INTECH.

Vernon, R. (1966): International investment and international trade in the product cycle, *The Quarterly Journal of Economics*, vol. 80, N° 2, Cambridge, Massachusetts, Harvard University.

Wagner, J. (1994): The post-entry performance of new small firms in German manufacturing industries, *The Journal of Industrial Economics*, vol. 42, N° 2, Oxford, U.K., Blackwell Publishers.

Wennekers, S. and R. Thurik (1999): *Understanding the Links between Entrepreneurship and Economic Growth*, Occasional Papers, N° 5, Indiana University, Institute for Development Strategies.

Winter, S. (1986): Schumpeterian competition in alternative technological regimes, in R. Day and G. Eliasson, *The Dynamics of Market Economies*, Amsterdam, The Netherlands, Elsevier Science Publishers B.V.

Creating capabilities
in local environments
and production networks

Gabriel Yoguel

Research Fellow,
Instituto de Industria Nacional,
Universidad Nacional
de General Sarmiento,
Argentina
yoguel@mail.retina.ar

In the new international setting, which is characterized by new technologies that make intensive use of information, globalization of markets, and the increased competitive pressures and uncertainty facing the agents, competitiveness is a systemic phenomenon. The endogenous capabilities of the agents, the degree of development of the environment they operate in and their integration in a production network have become key elements for developing capabilities and creating competitive advantages. It is being asserted more and more frequently that the competitive advantages of countries, regions and agents do not necessarily derive from their factor endowments but can be constructed through the development of endogenous capabilities and linkages with other agents. In the transition from static to dynamic advantages, the capacity to learn –conceived as an interactive process imbuing the whole of society– plays a key role. The present article analyses what the endogenous mechanisms for the creation of capabilities and the conversion of generic knowledge into specific know-how are, and what they depend on, at the level of the individual agents, production networks and the various local environments. Reference is made to the importance attached by economic theory in recent years to the relation between technology and learning processes, especially in the Schumpeterian and evolutionary approaches. The way in which the economic agents learn, transform generic knowledge into specific know-how and link up codified and tacit forms of knowledge is addressed, and finally it is emphasized that these processes are not the result of the natural linear development of production systems but are the consequence of a long evolutionary learning process.

I

Introduction

Important changes have taken place in the international setting over the last twenty years, especially the globalization of markets, the generalized spread of economic openness processes and the appearance of new technical and organizational models which involve intensive use of information. These changes have brought into question the idea of competitiveness as a purely macroeconomic and sectoral phenomenon, determined by static comparative advantages or by the resource endowment.

The volatility of demand, the segmentation of markets, the shortening of product life cycles, the strategic uncertainties associated with the new world situation and the possibility of combining economies of scale and of variety have meant a considerable increase in the competitive pressures that the economic agents must face. In addition to the traditional macroeconomic and sectoral factors which previously formed the key elements of competitiveness, there are now other elements which depend on the agents' conduct and on the nature of the economic and social environment in which they operate. As a result, competitiveness is now seen as a systemic phenomenon and both the agents' conduct and the degree of development of the local environment have taken on great importance in the creation of competitive advantages. The conviction is therefore growing up that the competitive advantages of countries, regions and agents do not necessarily depend on their factor endowment.

The new concepts are based on the idea that comparative advantages can be created and are therefore dynamic. In the transition from static comparative advantages to dynamic advantages, a key role is played by technology and by learning processes. The capacity to learn –conceived as an interactive process incorporated within the very fabric of society– and the development of "agent capabilities" will determine the economic success of firms, regions and countries (Ernst and Lundvall,

1997). Consequently, in this new context a factor of the greatest importance for the creation of comparative advantages is the reactions of enterprises aimed at conceiving, planning and implementing the development and improvement of products and processes, at introducing organizational changes, and at establishing new forms of linkages with the market. In other words, in the process of competition and the search for differentiation the agents seek to improve their technological capabilities (Lall, 1992), understood as their potential to turn generic knowledge into specific know-how through static and dynamic capabilities derived from formal and informal learning (Boscherini and Yoguel, 1996a). Static capabilities may be defined as the the set of formal and informal technological and organizational knowledge and skills that the agents generate in order to carry out their projects. They are not confined to information and equipment, but also include organizational capacity, patterns of conduct and routines affecting the decision-making process.

The present article aims to review the latest advances in this field described in the most recent literature, in order to understand what the endogenous mechanisms for the creation of capabilities and the conversion of generic knowledge into specific know-how are, and what they depend on, at the level of the individual agents, production networks and the various local environments. It is of the greatest interest to find out how the economic agents learn and how they turn generic knowledge into specific know-how: in short, how they innovate.

The studies analysed for this purpose were carried out in the developed countries and deal in particular with the new importance of learning processes in the creation of competitive advantages in industrial districts, industry clusters, and what is called the new territorial capability (Poma (forthcoming)). However, this analysis is also relevant to Latin America too. Even though in the 1990s the tendencies towards growing primary-sector domination of the production structure and the trade specialization profile became still more pronounced (ECLAC, 1996), it is obvious that reduction of the productivity differences with respect to the most highly developed countries, the sustainability over time of the models being applied, and the reduction of inequalities in income distribution will call for a higher level of complexity and

the incorporation of a higher proportion of both codified and tacit knowledge in the agents' production functions. It may be noted that in some agro-industries which have made notable advances in this new system of industrial specialization, in recent years the production processes have attained increasingly high levels of complexity, reflected in the growing importance of information-intensive technologies (Gutman, 1999).

In section II of this article, an analysis will be made of the role of learning processes in economic theory, with special emphasis on the new theoretical approaches. Section III looks at the role of tacit knowledge and the development of competitive advantages in the technical and labour-related capabilities of the agents, and section IV analyses learning processes in local systems and in industry clusters. Finally, section V presents some conclusions.

II

Technology, the firm, and the generation of knowledge: the latest theoretical approaches

Over the last fifteen years, the linkages between technology and learning processes have attracted renewed attention in economic theory, and this subject is constantly evolving. In this context, there is only a feeble relation between the theoretical framework of the technology commonly used and the appearance of these new processes in society. For example, neoclassical economic theory is based on a conception of the enterprise in which the economic agents always seek to maximize their performance, operate in conditions of perfect rationality and access to information, and are working in an environment where there is no uncertainty and there are rational expectations. In this context, capital is assumed to be malleable –the "jelly mould" theory– and technology is seen as a set of production techniques –a library containing all possible information and readily accessible to all the agents– which can be chosen and applied without difficulty at zero cost in terms of the relation between profit rates and wage rates (Solow, 1963). In order to select the best techniques, the agents weigh the latest advances in this set of techniques against the prevailing wage/profits ratio. When the wage/profits ratios go down from high levels to lower ones, the agents change labour-intensive techniques for others which are more capital-intensive. It is also assumed that the techniques are ordered in a manner functional to decreasing factor returns and that the price thresholds of the factors corresponding to each technique selected are crossed only once, in the direction corresponding to a production function with decreasing returns (Samuelson, 1962). Consequently, a labour-intensive technique with a high level of profitability cannot be used at a lower rate and the possibility of changing techniques is therefore ruled out.

The selection of techniques and their replacement when there are changes in the distribution conditions take place instantaneously and do not require a learning process.

In this standard neoclassical framework, which survived the old debate on the theory of capital in spite of the change of position by the participants,[1] it is assumed that technical progress is not incorporated into the production function and is independent of capital accumulation. Consequently, the acquisition of knowledge and the learning process of the agents do not constitute an explanatory argument connected with the production function and are considered to be independent of the social capital of the agents. Although some models incorporate the existence of learning curves to which the shift in the production function is ascribed –non-incorporated technical progress– this learning is considered to be exogenous and independent of the factors of production (especially social capital) and therefore does not affect them.

---

[1] Samuelson (1962) admits that the simple accounts of Jevons, Bohm-Bawerk, Wicksell and other neoclassical authors, according to which as the interest rate goes down as a result of abstention from present consumption in favour of future consumption, technology must in some way become more indirect, more mechanized and more productive, may not be universally valid. At the end of his work, he acknowledges that we owe a debt of gratitude to Pasinetti, Morishima, Garegnani and others for having shown that reversal is a logical possibility in any technology. In a noteworthy demonstration of intellectual honesty, Samuelson says that if all this causes dismay among those who yearn for the old parabolas of neoclassical theory, we should remember that academics were not born to enjoy an easy existence but must respect and give their due weight to the facts of life.

The inability of this theory to be applied empirically to explain the changes occurring in the structure of production gave rise to other interpretations which –in the light of the new international trade theory and growth theory– led to changes in some of the assumptions of the standard neoclassical model. Thus, for, example, Grossman and Helpmann (1992) developed a growth theory which endogenizes technical progress in the production function and shows that the possibility of strategic interaction between agents, research and development activities, the existence of economies of scale and learning processes is of key importance for the creation of competitive advantages and for explaining the agents' pattern of specialization in foreign trade. On the basis of a general equilibrium scheme, they consider that the agents' decisions to invest in research and development[2] and the benefits of the innovations arising from this depend on the conditions of operation of the market as a whole. Thus, the process of accumulation of knowledge endogenously generates the increases in productivity which sustain growth.

The most important change with regard to the explicit incorporation of knowledge into the theory, however, derives from the various neo-Schumpeterian and evolutionary schools of thought, which departed from the general equilibrium theory and implicitly accepted the existence of transactions effected under conditions of imbalance.[3] To put it in a very stylized manner, these theories start from a different conception of the enterprise and technology and assign a key role to the learning processes, both formal and informal, that the agents

use in the generation of competitive advantages. The enterprise theory takes account in its arguments of the limited rationality of the agents, their imperfect access to information, and the non-modelable uncertainty of the environment they are operating in. Uncertainty –which is an essential element in the analysis– is a parameter that the agents cannot express in probabilistic terms: the incomplete information cannot be completed, and the agents take their decisions under this drawback. Consequently, the agents' potential for turning generic knowledge into specific know-how –the development of their innovative capacity (Lall, 1992; Boscherini and Yoguel, 1996a)– decisively affects the possibilities of generating competitive advantages and partially reducing the strategic uncertainties that exist in the markets they operate on.[4] These theories start from the idea that a necessary condition for generating these learning processes is to have a certain minimum threshold level of capabilities (Rullani, forthcoming), which is lower in the case of the environments and countries where positive externalities are generated and the national innovation system works properly.

At the same time, as Dal Bó and Kosacoff (1998) point out, knowledge cannot be expressed entirely in explicit form, and hence cannot be transformed or converted into information as a tradeable good. This characteristic of knowledge introduced some highly specific aspects into the concept of technology, which must be considered not just as a stock of machines and production techniques,[5] but basically as a complex system of generation and diffusion of codified and tacit knowledge (Bell and Pavitt, 1995) accumulated by firms (Ernst and Lundvall, 1997).

This conception of technology includes two aspects which are not covered in neoclassical analysis (Metcalfe,

---

[2] In that theoretical framework, the knowledge generation process is limited to formal research and development laboratories.

[3] Although there is no desire here to place the work within any particular epistemological conception, it may be noted that the different elements in the body of neo-Schumpeterian theory –which breaks with the neoclassical analytical logic and in some cases returns to classical analysis– could converge towards a research programme along the lines suggested by Lakatos (1983). This would mean progressing towards a solid core consisting of i) an enterprise theory which assumes that the agents have limited rationality, imperfect information and suffer from uncertainty, and that they have to take their decisions in this context, and ii) a theory on technology and technical change which assigns a key role to the innovation process, understood as the process of transformation of generic knowledge into specific know-how and assimilation of both codified and tacit knowledge. The theory also makes new assertions that are as yet difficult to demonstrate empirically, especially the idea that experimental and tacit knowledge is decisive in developing the competitiveness of the agents. Within this theoretical framework, the interrelations between unproven and experimental knowledge (doxa) and proven knowledge (episteme) are excluded from most of the epistemology before and after Popper.

[4] According to Poma (1998), the new international setting also includes less possibility of control over uncertainties by the agents. Thus, for example, in the framework of Fordism uncertainty could be controlled both on the supply side (through the creation of automatic and repetitive phases in mass production) and on the demand side (through the creation of new needs among consumers). In the case of industrial districts, the uncertainty that existed in market conditions was tackled through the certainty of shared values and the system of personal relations, while technological uncertainties were dealt with through incremental innovations. In the new setting, in contrast, uncertainty has increased because the complexity of competition has also increased.

[5] This level of analysis is associated with the neoclassical view, whereby technology is the set of capital goods and production processes embodied in the machinery or fully transferable through handbooks. From this point of view, technological change is the process by which economies change over time as regards the goods they produce and the processes used to produce them.

1998). Firstly, it incorporates analysis of the capabilities of individuals and the capabilities generated within an organization, understood as the set of knowledge, routines, procedures, skills and practices available to it. These capabilities are more than just the total amount of technical and engineering knowledge, since they also include questions of organization and management connected with the functioning of the production process (David, 1985). According to this conception, technical progress is a succession of incremental innovations –much more significant than in the previous period– which complement radical innovations and indeed are more significant than them in some sectors. These improvements make it possible, using a given volume of resources, to produce more and better-quality goods, more efficiently. Secondly, the capacity to think and the implementation of applied learning processes also form part of technology. The literature on technological paths and paradigms (Pérez, 1985) considers that the starting point for the technical progress achieved by a firm is the equipment, the inputs, and the capabilities incorporated in individuals and organizations. These elements, together with the use of what Pérez (1983) calls "technical common sense", allow the agents to make either incremental innovations in the existing technology or radical changes in those branches where the ideal technological paradigm has not yet been reached.

Thus, firms do not select the best technology in the "library of technical knowledge" but, on the contrary, must make efforts of selection and adaptation which require certain minimum threshold levels of codified and, especially, tacit knowledge. While the codified component[6] of the learning process is basically tradeable, the tacit component[7] is firm-specific, cannot be bought on the market, and is a key factor in technological differ-

ences and in the specific competitive advantages of the agents (Lall, 1995). According to Poma (forthcoming), since tacit knowledge could be defined as the precipitation of the whole store of memories covering the sequences of operations that allow different objectives to be successfully attained, the use of a common (formal or informal) language is necessary in order for knowledge to be circulated and spread. In this sense, part of tacit knowledge must be placed within the framework of some formal parameters of thinking. Such a language can therefore be interpreted as an "institution" (whether formal or informal) that can facilitate or limit the learning process (Poma, forthcoming).

When access to codified knowledge is generally available to all the agents, such knowledge is not an element that leads to differences in forms of conduct and performance, but when not all the agents have access to codified knowledge because they have only imperfect information, they do not have the minimum capabilities needed, or they do not have the minimum tacit knowledge to assimilate it, then unequal access to codified knowledge can indeed become an element of differentiation.

According to Lall (1995), in order to create competitive advantages it is necessary not only to master the technology in a static sense –to reach the levels laid down in the handbooks– but also to carry out learning processes (with curves that cannot be predicted) which will enable firms to achieve better products and processes, to make changes in their organization, and to increase the complexity of the linkages with the local system. According to Ernst and Lundvall (1997), even to put codified knowledge into practice (interpretation of engineering and design handbooks, utilization of generic-type scientific and management knowledge, specification of quality criteria, etc.) an organization also needs tacit knowledge reflected in the organizational routines and collective experience of specific groups of the firm in the fields of research and development, management, production and marketing. Consequently, the development of tacit capabilities within the firm represents an intangible asset that is hard to transfer, can have a positive effect on the operating results, and can become a barrier to the entry into the market of agents who do not have this kind of knowledge.[8]

---

[6] Codified knowledge comprises the whole set of transmissible technological knowledge (embodied in materials, machines, components and final products) and organizational know-how (transmissible through such means of communication as the Internet, courses, etc.), which can be acquired through the market (Becattini and Rullani, 1993).

[7] Tacit knowledge comprises: i) know-how, not codified in handbooks, on applied technology as used in the work process; ii) general and conductual knowledge; iii) the capacity to solve non-codified problems, and iv) the capacity to link up situations and interact with other human resources. In short, tacit knowledge makes it possible to form a complex mental picture of the work process (Novick, 1998). These kinds of attributes which are demanded from workers (and which cannot be expressed specifically or completely formalized) are strongly influenced by the context (Mertens, 1996) and are acquired in various different ways such as on the shop floor or in associations or informal exchanges (Ducatel, 1998).

[8] As almost all the indicators of level of knowledge refer to formal education and research and development efforts, the image of the economics of learning is distorted and fails to reflect the importance of incremental innovation processes throughout the organization. In order to capture the degree of development of codified and tacit

According to Nightingale (1996), as the development of knowledge –which forms part of the social fabric– depends on the innate capacity of individuals to recognize similarities, the tacit elements are of decisive importance.[9] Learning does not mean simply accumulating more information, but recognizing different types of behaviour and connections between elements which are stored in the memory, and this depends on the knowledge accumulated through experience and the automatic capacity individuals have of linking up experience with knowledge.

In view of this, some authors suggest that the degree of mutual relation between codified and tacit knowledge decisively affects the efficiency that will be reached in a firm's learning process. These learning processes, which are generated in sectors with unequal degrees of technological development, involve not only formal training and research and development activities but also a set of informal activities which the agents do not always recognize as such (learning by doing, learning by interaction, learning by producing). These different types of learning processes are gradually accumulated during the active life of the agents and result in tangible and intangible assets of vital importance for competition. These assets are not eternal, however, because they must be weighed in the process of developing capabilities. While some are degraded and lose their value because they do not reach the minimum threshold levels demanded by the market, others are winners and gradually shape the elements in the predominant technological pattern.

In order to identify more precisely the different types of learning and their influence on the competitiveness of the agents, Johnson and Lundvall (1994) developed a typology in which four types of knowledge are distinguished, depending on their tacit or codified nature. The knowledge they call *know-what* may be assimilated to what would normally be classified as facts or information, while the knowledge they call *know-why* is of a scientific nature and refers to the principles and laws of movement of nature. Both are essentially codified knowledge and can be acquired on the market in the form of books, courses, data bases and similar sources. The types of knowledge they call *know-how* and *know-who*, in contrast, are of a tacit nature. The first of these refers to the skills acquired through direct experience in production and management activities. The second type is connected with the knowledge developed and maintained within an enterprise or research groups. Firms can gain access to this type of knowledge not only through their own activities but also through inter-firm cooperation and strategic alliances. In particular, access to know-who requires direct contact and communication between individuals and the development of relations of mutual confidence. In these circumstances, and as may be learnt in social practice, the agents with the greatest relative development (those of Silicon Valley, for example) establish links with formal and informal networks of scientists engaged in various fundamental research programmes potentially capable of practical application. This type of knowledge is therefore not usually transferred through formal information channels.[10]

As Ducatel (1998) suggests, the four forms of knowledge have strong mutual links. The development of codified knowledge (*know-what* and *know-why*) and the replication of experimental results depend on tacit knowledge and the recognition of the importance of *know-who* in the field of science and technology. Because of the limited rationality of the agents, in conditions of uncertainty (regarding future market conditions or the homogeneity of inputs, for example) a necessary condition for the development and incorporation of codified knowledge is often the existence of prior tacit elements which are not easy to codify. For its part, tacit knowledge –which has its origin in complexity and variations in quality, and which becomes particularly important in situations of uncertainty, where it is necessary to use different human capacities simultaneously and to relate different parameters with each other– likewise requires a certain minimum amount of prior codified knowledge. Although globalization and information technologies make access to codified options easier and cheaper through the Internet, they increase the strategic

---

knowledge which is not centered on specific research and development units, it is necessary to make use of other qualitative and quantitative indicators and embrace the idea that creative capacity is spread throughout the organization (Lassini, 1992). The progress made in this direction in Argentina is described in Boscherini and Yoguel (1996b), Boscherini, López and Yoguel (1997), Rearte, Lanari and Alegre (1997) and Moori-Koenig and Yoguel (1998).

[9] As Nightingale (1996) notes, tacit knowledge is of fundamental importance for understanding the meaning of the word "cut" in expressions such as i) cut a cake and ii) cut the grass. Although the same word is used in both cases, the meaning associated with "cut" is different in each context. The word "cut" thus has a different meaning which depends on the previously accumulated experimental knowledge (tacit knowledge), and this in turn differs between different individuals. In other words, the meaning of "cut" is linked with the prior tacit knowledge and is not immediately clear from the word on its own.

[10] This type of transfer of knowledge takes place through various forms of links between university research centres and enterprises.

uncertainties of the agents and thus provide even stronger reasons to develop tacit knowledge.

The importance for the long-term success of the agents assumed by the tacit elements of knowledge in the new context, which is reflected in their capacity to adapt to change (flexibility) and to make changes themselves (innovation), is in sharp contrast with the lack of these elements in the production function characteristic of the traditional majority of firms.[11]

Thus, in an analytical framework of imperfect information and rationality, firms are faced with a situation of uncertainty which they cannot foresee and where technology is not just the purchase of machines accompanied by codified information. Cognitive factors and formal and informal learning processes take on vital importance for the development of innovation capability and hence for the competitiveness of the agents.

This growing role of cognitive processes in the construction of tangible and intangible assets is thus found not only within organizations –"intelligent organizations" (Bessant, 1991)– but also in the environments in which firms operate ("intelligent regions"). Within the organizations, the idea of qualifying firms is beginning to gain strength, and a transition is beginning to take place from the demand for qualifications typical of Fordist organizations to a demand for capabilities, which is a new trend that is still coexisting side by side with earlier forms.[12]

III

The agents' technical and labour-related capabilities: tacit knowledge and the development of competitive advantages

In the development of technical knowledge within the enterprise and the possibility of the latter taking advantage of codified and tacit knowledge –within a context marked by limited rationality of the agents, imperfect information, uncertainty about the market situation and rapid technical change– the profile of the labour skills of the human resources involved is extremely important. These labour skills, understood as the total amount of knowledge of different origins and types possessed by the workers of an organization, have a number of characteristics which must be validated in the market and which exist in a state of uncertainty (Gallart, 1998). These characteristics refer to the capacity to solve problems, to learn and transmit knowledge to the organization on the basis of certain fundamental qualifications (Cariola and Quiroz, 1998), to manage resources and information, to develop interpersonal relations, to master technology (Mertens, 1996),[13] and to analyse and select options out of a range of alternatives (Novick, Bartolomé, Buceta, Millavares and Senén González, 1998). Unlike the traditional qualifications of human resources, which could be validated with a training certificate, these broader capabilities can only be validated in specific working situations (Novick, Bartolomé, Buceta, Millavares and Senén González, 1998).

---

[11] The growing importance of the tacit aspects of the learning process has reduced the usefulness of the traditional methods of measuring learning, based largely on proxy variables of the formal aspects of learning of the organizations studied (research and development laboratories, patents, etc.). See in this respect, among others, Malerba (1993), Acs and Audrescht (1988), Lassini (1992), Malerba and Orsenigo (1993) and Boscherini and Yoguel (1996b).

[12] The neo-Schumpeterian authors commented upon in this section consider that in learning processes the economic agents combine codified "scientific" knowledge with another type of knowledge that can include both rational and non-rational elements with a strong inductive content. They therefore depart, from the epistemological point of view, from Popper's ideas: in the course of their learning process the agents evolve in the manner proposed by Lamarck (they can correct their paths) rather than through a Darwinian process of natural selection (Gómez, 1995).

[13] Information management is the capacity of human resources to seek, evaluate, process, interpret and communicate information; systemic comprehension is the capacity to understand complex interrelations and understand and design systems; mastery of technology takes the form of the ability to select and adapt technologies, and the development of interpersonal relations is the capacity to interact with human resources inside and outside the organization, to work as a team, to teach and to learn.

As Mertens (1996) notes, workers are required to have a complex set of attributes, including in particular the capacity to assume more responsibility, to communicate, and to solve problems and learn, in addition to mental and manual skills. Within this set of qualities, the capacity to learn is the most important element, because of the greater complexity of innovation systems. Ducatel (1998) says that the capabilities required in the work process include i) the capacity to manage models in one's mind; ii) an understanding of the way machines work and interact; iii) the capacity to draw deductions from statistics; iv) oral and visual communication capability; v) willingness to accept individual responsibility for the work process and the product; vi) good judgement, and vii) skill in combining technical and business matters. In what this author calls the "learning triangle", the theoretical, vocational and experimental forms of knowledge interact, which demands a strong link between the work process and the educational system that is rarely found in practice. In so far as work is less and less capable of being directly observed but instead takes place largely in the worker's head (Hanser, 1995), labour skills involve new types of basic and technical, conductual and intellectual knowledge (Novick, Bartolomé, Buceta, Millavares and Senén González, 1998).[14]

In the process of developing capabilities, the agents can acquire tacit or general knowledge by using cooperation mechanisms that help it to circulate or creating a special form of organization of the work process which makes possible and stimulates the circulation, appropriation and generation of tacit knowledge. The capacity of firms to cope with the pressures of competition depends on the combination of codified and tacit knowledge that they develop within themselves. Thus, their potential to carry out this process depends on their initial capability and past performance, the degrees of freedom allowed by the technological patterns that apply, the local environment, their possible inclusion in networks in which these processes take place, and the way the work process and production are organized.

However, the process of learning by firms on the basis of codified and tacit knowledge can give either positive or negative results, starting off a process of "creative destruction" in which some firms generate solutions that enable them to stay in the market, other disappear, and still others enter the market for the first time. In a context where there are a variety of responses and degrees of freedom (Nelson, 1991) the selection of forms of conduct made by the market is imperfect, so that it is not always the best forms that survive.[15]

The generation and circulation of knowledge within a firm is a complex process whose intensity depends on: the need to solve concrete problems in a situation of uncertainty, which stimulates the demand for noncodifiable solutions; the degree of technical complexity of the equipment; the type of basic capabilities of the agents; the capacity to establish relations and work in a group, and the degree of utilization of the technical and organizational knowledge of the workers.

In particular, the generation and spread of tacit knowledge would appear to be associated with the characteristics of the human capital of the firm, the way the work process is organized,[16] and the degree of importance that the interpretation and adaptation of external codified knowledge has for the firm.[17] For the spread of tacit knowledge into the firm, the fundamental factor is the existence of networks and different types of linkages between the agents. The development of tacit knowledge within the firm is of a synergic nature, so that to a large extent the knowledge possessed by the individuals who form part of the organization is only of value within the firm and may have little value outside it. In other words, the stores of tacit knowledge possessed by individuals link together to form the competitive advantages of the organization and lose part of their value outside it.

Until tacit knowledge is disseminated and comes to be codified, the tacit elements possessed by firms form part of their capabilities and become an item of competitive advantage. For agent h the competitive advantage depends on the possession of items of tacit knowledge 1, 2 … m, and for agent g his competitive advantage will depend on the items of tacit knowledge n … $n+j$.

---

[14] In Latin America, there are two tendencies which run counter to this. On the one hand, the fact that the production structure is increasingly based on primary-sector activities means that there is less demand for highly skilled human resources, while on the other the external flexibility of the labour market militates against the development of capabilities of the type referred to.

[15] These flaws in the selection of forms of conduct are a key element which is taken into account in policies aimed at creating mechanisms to minimize such flaws.

[16] Degree of flexibility, type of hierarchical organization, existence of cells, extent to which the wages of individual workers are linked with the performance of the group they belong to.

[17] This aims to assess the amount of additional development that the firm carries out on the goods and services that it purchases or obtains in codified form and the human resources that it hires, turning them into something which is different, special or specific and not appropriable by other agents (such as changes in the layout of plant, adaptation of "soft" technologies, or adaptation of information for the development of products or processes).

For two agents *h* and *j*, their competitive advantages may be expressed as follows:

Competitive advantage $h = F(t1, t2, t3 ... tm)$
Competitive advantage $j = G(tn ... tn + j)$

The competitive advantage of *h* is greater than that of *j* if the tacit elements (1,2,3 … *m*) possessed by agent *h* delay longer in becoming codified than the elements *n, n+1, ... n+j* possessed by agent *j*. In that case, we can say that agent *h* will be able to appropriate his competitive advantages for a longer period of time. It may be noted that knowledge can be identified as a factor of production which has some special features that clearly take it out of the field of neoclassical analysis. On the one hand, the generation of knowledge (its production) increases with its consumption, thus clearly differentiating it from conventional factors of production. On the other hand, the synergies generated through the generation and dissemination of knowledge can mean that there are increasing returns to it, if there are virtuous linkages between public and private agents.

According to Ducatel (1998), the importance taken on by tacit knowledge shows up the shortcomings of the conventional educational system in terms of developing the capabilities of the agents. Formal education must be complemented with experience in order for the agents to acquire *know-how*. This consequently increases the importance of the interpersonal aspects of skills: the *know-who* side of the learning process. Mertens (1996) makes a detailed analysis of the elements that should be present in the formal educational system in order for it to be functional to the creation of the skills required in the production system. He considers that in the present-day world what is needed is not the senseless memorization of parallel subjects or the acquisition of relatively mechanical skills, but cross-ranging knowledge that can be updated in daily life and is reflected in the ability to solve problems different from those posed in the classroom.[18] Skills are a multidimensional concept based on physical and cognitive capacity and interpersonal relations which, although they are generally not provided by the formal educational system, nevertheless require it as a precondition for their proper development. If we assume that the learning process is a social process, then in order to ensure competitiveness the relations developed both within the firm and between it and the rest of the agents are of decisive importance.

IV

The practical manifestations of the learning process:

local systems and enterprise networks

It will be gathered from the foregoing that learning in an organization is something more than the application of conventional training processes. Nonaka and Takeuchi (1995) consider that a learning organization is one where inventing new knowledge is not a specialized activity, restricted to a small group, but a general form of behaviour extending to all the workers. What a learning organization needs is a broad variety of systems and procedures for capturing and mobilizing the *know-how, know-who, know-what* and *know-why* of its components in order to facilitate the various forms of conversion of knowledge.[19]

The aim of these systems is to involve the workers in learning practices and in the linking-up and application of knowledge. Achieving this requires the organization of team work, the existence of formal and informal training processes tailored to the needs of the firm and form-

---

[18] This is a major challenge for the educational system, which operates more as a body responsible for certifying qualifications than as a system for generating capabilities.

[19] A fundamental aspect of the work of these researchers is their definition of the dialectics of the four processes of conversion of knowl-

edge and the forms of organization which make this possible: i) the conversion of tacit knowledge into another form of tacit knowledge: a phase known as the "socialization of knowledge"; ii) the conversion of tacit knowledge into codified knowledge, or the "externalization of knowledge" phase; iii) the phase of the combination of codified knowledge; and iv) the conversion of codified knowledge into tacit knowledge, called the "internalization of knowledge" phase. In addition to these four processes (socialization, externalization, combination and internalization) they propose a form of organization different both from "top downwards" systems based entirely on the predominance of codified knowledge and "bottom upwards" systems based solely on the importance assigned to tacit knowledge.

ing part of a long-term approach, and the application of practical strategies designed to give the workers new skills so that they can carry out new tasks and take part in the planned rotation of jobs.

For Ducatel (1998), learning organizations display a certain set of characteristics. Thus: i) they have the ability to solve problems systematically; they usually start with a pre-established plan which they constantly review in order to consolidate their experience and use a type of management which expands the "critical mass" of ideas in the enterprise, encouraging the questioning of decisions, and which fosters the use of statistical techniques and the development of in-house thinking guided by considerations of intellectual rigour, discipline and precision; ii) they have the capability to experiment with new methods, to which end they promote systematic experimentation to test the new knowledge in operational programmes and projects and reward the taking of risks; this ensures that the lessons are effectively passed on to the entire organization and that its members learn to turn tacit knowledge (*know-how*) into codified knowledge that can be transferred both within the organization and to other agents; iii) they have the capability to learn from experience through the study of past errors and the systematic recording of results so that they can be made known to all the members of the organization, and iv) they have the capability to learn from others. In order to be able to do this, they are open to others, listen carefully, systematically analyse the results, carry out an ongoing analysis of the best practices, plan visits and interviews, get ideas from their clients on products, competitors and changes in preferences, observe their clients in action, and can transfer knowledge quickly and efficiently by preparing written, graphical and oral reports. This set of characteristics is to be observed in particular in the learning processes of agents who form part of local systems or enterprise networks.

1. Interaction between the local and global levels: the role of the environment in the learning process

Because of the systemic nature of competitiveness and the interactive character of innovation (Morgan, 1995), seen as a learning process in which new knowledge is introduced or existing knowledge is combined in order to generate new capabilities (Lundvall, 1992; Gregersen and Johnson, 1996), over the last twenty years the role of the local environment and its institutions in the development of the innovative capacity of enterprises has been taking on fresh significance. In the new international setting, the local environment and the global economy

are not antagonistic terms, as globalization owes its force to the complexity of the knowledge and the synergy produced through the competitive confrontation of different territorial levels and networks of agents.[20]

From this standpoint, the local environment may be understood as the set of local institutions and agents and their mutual relations; its characteristics are of decisive importance for developing the creative capacity of enterprises. This environment may be seen as a public space which, in its positive aspects, can give rise to phenomena of collective efficiency, defined as the competitive advantages derived from external economies and the joint action of the agents (Camagni (ed.), 1991; Bianchi and Miller, 1994).[21] The presence of a favourable environment can be reflected in the actions of the agents who make up civil society, who, thanks to their cooperation, capabilities and mutual pressures give rise to a collective tension which favours the development of innova-

---

[20] The interconnection of international financial markets associated with globalization helped to spark off these processes.

[21] Boscherini, Malet Quintar and Yoguel (1997) define a scale of theoretical local environments where the externalities differ as a function of the characteristics of the agents, the logic of the system, and the agents' strategies, as well as the degree of internal and external linkages of all the component elements. The environments in which the greatest positive externalities are generated are those which display the best characteristics in the three respects mentioned. In these environments there tend to be many agents (a large number of linkages, sectors and firms) with heterogeneous forms of production and little vertical integration, thus favouring inter-firm cooperation and the development of different forms of externalization. The degree of coverage of the (educational, information and services) institutions is high, and this is further heightened by their complementarity with each other. In this type of theoretical environment, a single type of logic may be identified, consisting of maintaining the core business, with business strategies tending towards the globalization, decentralization and in some cases vertical re-integration of the critical phases. The agents in these environments organize themselves without any visible coordination and form an advanced public space. The agents cooperate in production, trade and services activities and in problem-solving. The educational institutions carry out research and development work in conjunction with the firms, and the services they offer are in line with the agents' needs, within a strongly interactive framework. Finally, the relations with the world outside the local environment are based on links between groups of local institutions and enterprises and similar groups located outside the local system. At the other theoretical extreme, a negative type of environment generates the worst possible negative externalities for the firms. This type of environment is marked by the existence of only a small number of agents, with homogeneous forms of conduct and operating at a high level of vertical integration within an institutional system that fails to meet the minimum requirements for the development of individual and collective capabilities. No particular logic or strategy predominates in the functioning of the system. Without coordinating agents, the enterprises do not organize themselves as they did in the previous case. Finally, the agents do not have relations with the world outside the local system.

tive strategies and minimizes the differences between agents. International experience suggests that when these environments behave in a positive manner they act like a quasi-market operator which reduces dynamic uncertainties, offsets weaknesses in the quality of organization, furthers learning processes, provides the agents with the capabilities they lack, aids in the process of dissemination of codified and tacit knowledge and tends to reduce social inequality (Camagni (ed.), 1991).

Thus, positive environments promote the generation of "social capital" constructed on the basis of complex links that are reflected in the development of mutual confidence among the agents (Morgan, 1995), which helps to reduce uncertainty and to spread codified and tacit knowledge. This type of environment is the result of individual and collective learning processes. Although a positive environment tends to make the forms of conduct of the agents more similar, its presence does not mean automatic benefits for all of them. Thus, in order to be able to take advantage of the externalities of the environment firms must have more than a certain minimum threshold level of capabilities and of endogenous generation and transmission mechanisms, without which the process of learning and transformation of knowledge will not take place. The generation of such capabilities in local environments is also the consequence of an evolutionary process of creation and destruction of routines and conventions (Gregersen and Johnson, 1996). Consequently, learning also requires the discarding of some things previously learned. In this process, firms modify the environment through the internal transformation of knowledge and the way they link up with other agents. The development of learning processes in local environments depends on the path followed by the agents in their evolution, the technological pattern of the sector they operate in, and the degree of development of the environment (Camagni (ed.), 1991).

In order to establish the real value of their specific features (knowledge, capabilities), the local and national spaces must be compared with similar spaces at the world level. Rullani (forthcoming) has clearly described the interaction between the codified and tacit aspects of the process of generating knowledge, with special emphasis on the role played by the environment. As the tacit elements have a strong contextual and experience-related component, in the globalized world of today the territorial dimension has a leading role in the generation of knowledge. The contextual, experimental and evolutionary nature of knowledge leads not only to the generation of tacit knowledge but also to its codification so that it can be transformed and possibly transferred. In this way,

knowledge can circulate outside its original context and take on a non-territorial form, but in order to apply it effectively in another context it must be re-contextualized.

From the territorial point of view, then, there are two poles in the cognitive circuit: i) the global dimension, which covers the process of production of knowledge and its transfer and use outside its context (codification) and ii) the local dimension, which includes the process of learning and precipitation, when the knowledge becomes rooted in the territorial area in question. Thus, according to Rullani (forthcoming), the generation of the economic agents' knowledge takes place in a local and specific context. For it to be taken out of that context and transferred, it must first be codified. In this stage, knowledge takes on a global and abstract character. For this abstract knowledge to be of use, however, it must be re-contextualized and subjected to a process of adaptation which makes possible the creation of tacit knowledge.

This process of codification and decodification of knowledge which links together the local and global levels is a product of what Becattini and Rullani (1993) call "versatile integration"; they note that the efficacy of the various forms of integration (technological, organizational and communicational) depends not only on the efficiency of the codes employed by the different agents but also, and especially, on the capabilities and skills acquired by direct experience which cannot be expressed in standardized codes. An aspect which takes on vital importance in the process of generation and circulation of tacit knowledge in local environments is the way the formal and informal modes of language link up among the agents (Poma, forthcoming). Whereas codified knowledge is always learnt in the same way, tacit knowledge goes through a process of interpretation by the person learning it and is therefore somewhat different from the knowledge of the agent who transferred it: this discrepancy between transmitting/circulating and learning adds something to the knowledge and introduces innovations in it and also in the language (*ibid.*).

2. Learning processes in enterprise clusters

The aspects dealt with earlier in connection with the process of generation and circulation of knowledge within a given environment are reproduced in enterprise clusters, with the special feature deriving from the predominance of horizontal and vertical input-output relations among the agents that form part of them. As a considerable number of transactions take place outside the mar-

ket in these production networks or systems, tacit elements can take on even greater importance in them.

However, the intensity of the learning process in enterprise clusters will depend on the importance assigned to it both within the enterprise and by each of the agents that make it up, as well as on the degree of self-organization. The importance assumed by the tacit elements developed in the different sectors of production will depend on the depth of the exchange of experience and joint work, the importance of work in production cells, the circulation of information in each of the components of the production network, and the interaction with the local environment.

A very special feature of production networks is the synergy that can be generated in them by the learning process, which may be reflected in the generation of more numerous and more complex tacit elements, which become entry barriers to agents who do not already belong to the network. Thus, if agents h and j, whose competitive advantages were described in the previous section, belong to network Ti, the aggregate competitive advantage of the network will be greater than the sum of the functions of the agents that make it up, generating improvements in efficiency which may be assimilated to economies of scale and variety, taken together. Together with the joint development of tacit-type knowledge in enterprise clusters, however, there is also codification of tacit knowledge so that it can circulate through formal languages (Poma, forthcoming).[22]

The implementation of learning processes within enterprise clusters not only introduces changes in the training activities required but also in the way contracts are formulated and the implicit incentives. In the formal and informal relations that take place within a production network, the contracts and the fixing of prices involved in trading operations gradually acquire greater precision through the synergies generated in the joint learning process. While the minimum requirements of contracts (precision) would appear to be positively related to the level reached in the learning process, the rate of incentives would display an inverse relation with the level of learning: the greater the knowledge acquired, the fewer are the incentives needed to ensure that contracts work. The evolution of contracts assumes that the agents not only learn but also advance (there are changes in routines, new production processes are introduced, sunk investments are made, etc.); this process cannot be reversed and determines the initial positions in new contracts. In short, while the rate of risk is in inverse relation to the implementation of learning processes, the level of complexity of contracts is in direct relation to it. Consequently, as the rate of incentives needed for the formulation of contracts depends on the level of risk, the implementation of learning processes plays a decisive role –within the production network– in reducing the amount of incentives needed to make the system of contracts work.[23]

V

Some final remarks

As noted in the previous sections, in the new international setting the creation of competitive advantages by the agents depends on the implementation of learning processes of a systemic nature. These processes are affected not only by the set of individual characteristics of the agents, but also the degree of articulation of the local systems and the production networks of which they form part. From this point of view, the generation and circula-

tion of codified and tacit knowledge both within organizations and between them is powerfully influenced by: the complexity of the linkages that exist and the degree of formal and informal technological cooperation among the agents; the type of links between universities and research centres on the one hand and enterprises on the other; the training of human resources and the complexity of the educational and training system, and the de-

---

[22] According to Poma, this eliminates the discontinuities typical of the subjective interpretation of tacit knowledge and reduces the number of incremental innovations within it, giving rise to processes of more clearly-defined breaks which make possible the development of radical innovations.

[23] Volkswagen's continuous improvement group is an interesting example of an informal technical assistance and training system designed to develop capabilities on the basis of the specific tacit knowledge of the participating agents, thus making possible its mobilization, codification and valuation by the agents of the production system

gree of development of the intermediate agents (risk capital, specific technical services, professional associations, former students' groups, etc.) who also act as messengers (contacts) in the process of transmission and generation of information. As, in this new setting, knowledge is created and transferred in many different ways, informal areas of innovation grow up within enterprises and institutions which complement the activities of the research and development laboratories proper.

These processes are not the consequence of the natural, linear evolution of production systems, but require a long period of maturation. They should be interpreted as the result of a complex process of economic and social construction which takes time and has a historical form of evolution which is influenced by many different levels. In addition to formal and informal factors of a technological and economic nature, it is also influenced by other elements, especially the development of mutual confidence among the agents, the social valuation of the role of entrepreneur, and the degree of fulfillment of contracts. These are tendencies which, although still embryonic, are beginning to take root in the developed countries and also, with some delay, in the less developed nations.

---

in question. This generates new capabilities and knowledge which make it possible to resume the process of development of knowledge, starting from a higher level. In this process of technical assistance and training, the group operates as an intermediary (that is to say, as a technological adviser), which makes it possible to speed up the dissemination of knowledge. As it is based on the idea that he who really knows is the other party, this generates two-way integration of knowledge. In this sense, the training consists of directing knowledge of different types and levels, through a systematic process, to integrate higher-level technical knowledge with the knowledge possessed by those working directly in the production process. In the continuous improvement process, the training consists of generating a synergy of different types of knowledge to make possible external economies: thus, the synergy is greater than the knowledge contributed by each of the members of the system in question. The technological adviser programme applied in Argentina by the National Agency for the Promotion of Science and Technology of the Ministry of Science and Technology, which is aimed at small and medium-sized enterprises, is based on similar lines. The technological advisers assigned to groups of firms start from the level of the capabilities possessed by the agents and work together with them to develop their technological management capabilities. The idea is that, because of the flaws that exist in the technology market, what the enterprises want is not so much to obtain specific technical advice as to identify the key factors in developing the capabilities and innovative capacity of the firms (see Argentina, Presidencia de la Nación, Secretaría de Ciencia y Tecnología, 1997).

In order for the technological learning process to be successful in the developing countries, they need to acquire the necessary codified elements of the technology and develop the complementary tacit elements. In seeking to fulfill the first condition, they run up against flaws in the technology market, while in order to achieve the second they need first of all to build up capabilities in a prior evolution process which is often non-existent or incomplete. Consequently, the generation of competitive advantages linked with the intensive application of knowledge and more complex structures of production in the developing countries represents a challenge which goes beyond the actions of individual economic agents and calls for measures involving the public and private institutions and the other social agents as a whole.

Achieving these objectives requires, among other things: i) social revaluation of the processes of learning and education; ii) the creation of intermediary agents who will act as transmitters/translators between the different parts of the system and will spark off the learning processes of the agents and help to create the market; iii) the development of local environments that generate external economies; iv) the evolution from individual competitive advantages to competitive advantages based on enterprise clusters and local systems, and v) the establishment of effective links between the universities and the enterprises which boost the processes of codified and tacit learning.

The production specialization profile of most of the Latin American countries is biased towards the processing of natural resources or in-bond assembly activities and therefore hardly includes learning processes at all as a prime element for obtaining dynamic competitive advantages. In the future deepening of the development model (even if it continues to be based on the exploitation of natural resources), however, it will be necessary to pay greater attention to the issues mentioned above. Policy objectives such as strengthening added value chains, developing production capability for internationalization, integrating small and medium-sized enterprises in enterprise clusters and networks, and achieving a better territorial balance (Kosacoff (ed.), 1997) are directly linked to the development of the complex capabilities of the agents, which are difficult to obtain without a major learning process involving not only codified but also tacit knowledge.

(Original: Spanish)

Bibliography

Acs, Z. and D. Audrescht (1988): Innovation in large and small firms: An empirical analysis, *The American Economic Review*, vol. 78, Nashville, Tennessee, American Economic Association, September.

Argentina, Presidencia de la Nación, Secretaría de Ciencia y Tecnología (1997): *Plan Nacional Plurianual de Ciencia y Tecnología 1998-2000*, Documento, N° 1, Buenos Aires, Programa de Capacitación Tecnológica de Pymes Industriales.

Arrow, K. (1962): The economic implications of learning by doing, *The Review of Economics Studies*, vol. XXIX(3), N° 80, London, London School of Economics.

Becattini, G. and E. Rullani (1993): Sistema locale e mercato globale, *Economia e politica industriale*, N° 80, Milan, Italy, Franco Angeli.

Bell, M. and K. Pavitt (1995): The development of technological capabilities, in I. ul Haque (ed.), *Trade, Technology and International Competitiveness*, Washington, D.C., World Bank, Economic Development Institute (EDI).

Bessant, J. (1991): *Managing Advanced Manufacturing Technology*, Manchester, U.K., NCC Blackwell.

Bianchi, P. and L. Miller (1994): Innovation, collective action and endogenous growth: An essay on institutions and structural change, in F. Boscherini and L. Poma (eds.), *Territorio, conocimiento y competitividad en el espacio global*, Buenos Aires, Universidad Nacional de General Sarmiento.

Boscherini, F., M. López and G. Yoguel (1997): *Sistemas locales de innovación y el desarrollo de la capacidad innovativa de las firmas: un instrumento de captación aplicado al caso de Rafaela*, Documento de trabajo, N° 10, Buenos Aires, Economic Commission for Latin America and the Caribbean (ECLAC)/ Universidad Nacional de General Sarmiento, Instituto de Industrias.

Boscherini, F., N. Malet Quintar and G. Yoguel (1997): Consideraciones acerca del ambiente y el desarrollo de las capacidades innovativas de las firmas, *Segundas Jornadas Nacionales de la Red Pymes*, La Plata, Argentina, Centro de Estudios Bonaerense (CEB).

Boscherini, F. and G. Yoguel (1996a): *La capacidad innovativa y el fortalecimiento de la competitividad de las firmas: el caso de las PYMEs exportadoras argentinas*, Documento de trabajo, N° 71, Buenos Aires, ECLAC Office in Buenos Aires.

———(1996b): Algunas consideraciones sobre la medición de los procesos innovativos: la relevancia de los rasgos informales e incrementales, *Redes*, Buenos Aires, Universidad Nacional de Quilmes, Centro de Estudios e Investigaciones.

Camagni, R. (ed.) (1991): *Innovation Networks: Spatial Perspectives*, London, Belhaven Press.

Cariola, M. and A. Quiroz (1998): *Competencias generales, competencias laborales y curriculum*, Montevideo, Inter-American Research and Documentation Centre on Vocational Training (CINTERFOR), International Labour Organisation (ILO).

Cimoli, M. and G. Dosi (1994a): De los paradigmas tecnológicos a los sistemas nacionales de producción e innovación, *Comercio exterior*, vol. 44, N° 8, Mexico City, Banco Nacional de Comercio Exterior (BANCOMEXT).

———(1994b): *Technological Paradigms, Patterns of Learning and Development: An Introductory Roadmap*, London.

Dal Bó, E. and B. Kosacoff (1998): *Theoretical Approaches to the Microeconomic Evidence about Structural Change*, Buenos Aires, ECLAC Office in Buenos Aires.

David, P.A. (1985): CLIO and the economics of QWERTY, *The American Economic Review*, vol. 75, N° 2, Nashville, Tennessee, American Economic Association.

Ducatel, K. (1998): *Learning and Skills in the Knowledge Economy*, Working Paper N° 2, Aalborg, Denmark, Danish Research Unit for Industrial Dynamics (DRUID).

ECLAC (Economic Commission for Latin America and the Caribbean) (1996): *Quince años de desempeño económico. América Latina y el Caribe 1980-1995*, Santiago, Chile.

Ernst, D. and A. Lundvall (1997): *Information Technology in the Learning Economy: Challenges for Developing Countries*, Working Paper N° 12, Aalborg, Denmark, Danish Research Unit for Industrial Dynamics (DRUID).

Gallart, S. (1998): *Competitividad, redes productivas y competencias laborales*, Montevideo, CINTERFOR, ILO.

Gómez, R. (1995): *Neoliberalismo y seudociencia*, Buenos Aires, Editorial Lugar.

Gregersen, B. and B. Johnson (1996): Learning economies, innovation systems and European integration, *Regional Studies*, vol. 31, Oxford, U.K., Pergamon Press .

Grossman, G. M. and E. Helpman (1992): *Innovation and Growth in the Global Economy*, Cambridge, Massachusetts, MIT Press.

Gutman, G. (1999): *El sector agropecuario y el sistema agroalimentario: nuevos procesos*, Buenos Aires, Instituto Torcuato di Tella, Centre for Urban and Regional Studies (CEUR).

Hanser, L. (1995): *Traditional and Cognitive Job Analyses as Tools for Understanding the Skill Gap*, Berkeley, California, University of California, National Center for Research in Vocational Education.

Johnson, B. and B.A. Lundvall (1994): Sistemas nacionales de innovación y aprendizaje institucional, *Comercio exterior*, vol. 44, N° 8, Mexico City, BANCOMEXT.

Kosacoff, B. (ed.) (1997): *Estrategia de desarrollo empresarial*, Buenos Aires, ECLAC Office in Buenos Aires.

Lakatos, I. (1983): *La metodología de los programas de investigación científic*a, Madrid, Alianza Editorial.

Lall, S. (1992): Technological capabilities and industrialization, *World Development*, vol. 20, N° 2, Oxford, U.K., Pergamon Press.

———(1995): The creation of comparative advantage: the role of industrial policy, in I. ul Haque (ed.), *Trade, Technology and International Competitiveness*, Washington, D.C., World Bank, EDI.

Lam, A.(1998): *Tacit Knowledge, Organizational Learning and Innovation: A Societal Perspective*, Working Paper N° 22, Aalborg, Denmark, Danish Research Unit for Industrial Dynamics (DRUID).

Lassini, A. (1992): Il ruolo dell'attività di R&S per lo sviluppo della competitività delle PMI, in F. Onida, G. Viesti and A.M. Falzoni (eds.), *I distretti industriali: crisi o evoluzione?*, Milan, Italy, Egea.

Lundvall, B.A. (1992): *The learning economy: Challenges to economic theory and policy*, paper presented at the EAEPE Conference, Copenhagen, Denmark, European Association for Evolutionary Political Economy (EAEPE).

Malerba, F. (1993): National systems of innovation: The case of Italy, in R. Nelson (ed.), *National Innovation Systems*, Oxford, U.K., Oxford University Press.

Malerba, F. and L. Orsenigo (1993): L'accumulazione delle capacità tecnologiche nell'industria italiana (1969-1984), in L. Filippini (ed.), *Innovazione tecnologica e servizi alle imprese*, Milan, Italy, Franco Angeli.

Mertens, L. (1996): *Competencia laboral: sistemas, surgimiento y modelos*, Montevideo, CINTERFOR, ILO.

Metcalfe, S. (1998): *Evolutionary Economics and Creative Destruction*, London, Routledge.

Moori-Koenig, V. and G. Yoguel (1998): Capacidades innovadoras en un medio de escaso desarrollo del sistema local de innovación, *Comercio exterior*, vol. 48, N° 8, Mexico City, BANCOMEXT.

Morgan, K. (1995): *The Learning Region: Institutions, Innovation and Regional Renewal*, Cardiff, University of Wales.

Nelson, R. (1991): *Why Do Firms Differ and How Does it Matter?*, Working Paper N° 7, Berkeley, California, University of California, Center for Research Management.

Nightingale, P. (1996): *A Cognitive Model of Innovation Electronics*, Working Paper N° 11, Sussex, Science Policy Research Unit (SPRU).

Nonaka, I. and H. Takeuchi (1995): *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*, London, Oxford University Press.

Novick, M. (1998): *Estrategias y capacitación y aprendizaje de las firmas. Reflexión sobre empresas innovadoras en Argentina, Brasil y México*, paper presented to the ECLAC/GTZ Joint Project on Policies to Improve the Quality, Ef-

ficiency and Relevance of Vocational Training in Latin America and the Caribbean, Santiago, Chile, ECLAC.

Novick, M., M. Bartolomé, M. Buceta, M. Miravalles and C. Senén González (1998): *Nuevos puestos de trabajo y competencias laborales*, Papeles de la Oficina Técnica, N° 6, Montevideo, CINTERFOR, ILO.

Novick, M. and L. Gallart (1998): *Competitividad, redes productivas y competencias laborales*, Geneva, ILO.

Pérez, C. (1983): Structural change and the assimilation of the new technology in the economic and social system, *Futures*, vol. 15, N° 5, Surrey, U.K., IPC Science and Technology Press.

——— (1985): Microelectronics, long waves and world structural change: New perspectives for developing countries, *World Development*, vol. 13, N° 3, Oxford, U.K., Pergamon Press.

Poma, L. (1998): La nuova competizione territoriale, *Politica e organizzazione*, Quaderni adoc, N° 1, Bologna, Italy, Il Mulino

——— (forthcoming): La producción de conocimiento. Nuevas dinámicas competitivas para el territorio, conocimiento y competitividad en el espacio global, in F. Boscherini and L. Poma (eds.), *Territorio, conocimiento y competitividad en el espacio global*, Buenos Aires, Universidad Nacional de General Sarmiento.

Rearte, A., E. Lanari and P. Alegre (1997): *Sistemas de innovación y el desarrollo de la capacidad innovativa de las firmas: el caso de Mar del Plata*, Buenos Aires, Universidad Nacional de Mar del Plata.

Rullani, E: (forthcoming): El valor del conocimiento, in F. Boscherini and L. Poma (eds.), *Territorio, conocimiento y competitividad en el espacio global*, Buenos Aires, Universidad Nacional de General Sarmiento.

Samuelson, P. (1962): Parable and realism in capital theory: The surrogate production function, *Review of Economics Studies*, vol. XXIX(3), N° 80, London, London School of Economics.

Solow, R. (1963): *Capital Theory and the Rate of Return*, Amsterdam, The Netherlands, North-Holland Publishing Company.

The potential for hub ports
on the Pacific coast
of South America

Jan Hoffmann

Transport Unit,
Natural Resources and
Infrastructure Division,
ECLAC, Santiago, Chile
jhoffmann@eclac.cl

The external trade of a country is closely linked with its geographical location, with the transport services that cover the distance to markets, and the ports through which that trade passes. Recent advances in maritime transport, the growing international economic integration, and the privatization of ports in the countries on the Pacific coast of South America have given rise to expectations that ports could be developed that concentrate both domestic cargo and that of neighbouring countries for its subsequent redistribution: what are known as "hub ports". The main conclusion of the present study is that the potential for hub ports on the Pacific coast of South America is very limited. In the past, countries tried to prevent the foreign trade of their neighbours from using their ports to gain some kind of commercial benefit. Now, however, the situation has been reversed, and ports compete with each other for the trade of neighbouring countries. In itself, this competition is positive, but the problem is that in many cases it has been raised to a political level which has turned simple competition between ports into international competition between hypothetical future "hub ports". In view of the low degree of probability that the establishment of such ports on the west coast of South America will be a success, it might be more advisable to seek greater regional coordination of transport policies and of investments in port and land transport infrastructure, in order to promote integration between the countries of the Atlantic and Pacific coasts of South America.

I

Introduction

Hub ports ("puertos pivotes" in Spanish) are seaports that concentrate domestic and foreign cargo with different points of origin and/or destination for its subsequent redistribution.[1] They thus generate business for the local economy by transporting cargo that does not come from the actual hinterland of the port in question.

The question of whether or not there is potential for the emergence of such hub ports on the west coast of South America is important both for the economic integration of the South American countries and for their integration with other regions. For example, transport services between South America and the Asian Pacific Rim countries are crucial for the South American countries' participation in the Asia-Pacific Economic Cooperation forum (APEC), and port links are fundamental elements for connecting the bioceanic corridors with maritime transport services.

In more general terms, in recent years there have been many studies which analyse the relation between geographical aspects and the development of countries in the light of such variables as distance and transport. Radelet and Sachs (1998), for example, seek to identify the determinants of transport costs and then go on to investigate the relation between those costs and growth rates. The results show a clear negative relation between the two variables. In view of the importance of the maritime mode in international transport, improving its efficiency and reducing its cost should form part of any development policy.

In recent years, the maritime transport industry has undergone a marked process of concentration, including alliances and mergers between shipping companies, and there has been an increase in the transshipment of containerized cargo in ports.[2] At the same time, the Latin

American countries are opening up their economies and their international trade is growing faster than their product, giving rise to a big increase in the need for international transport services.

Both these tendencies –the advances in the maritime transport industry and the greater economic openness of the countries– have helped to create expectations that ports could sell their services to neighbouring countries. Traditionally, ports served almost exclusively the foreign trade of the countries where they were located, but there are now possibilities for them to provide services for cargo from other origins destined for third countries. Such expectations have arisen with respect to ports in the four South American countries with Pacific coastlines: Chile, Colombia, Ecuador and Peru.

These four countries also share the characteristic that their ports are being privatized and that they are seeking investors to improve port infrastructure and productivity. At first sight, it would therefore seem reasonable that the governments should seek investors not only to improve the services for their own cargo, but also to generate extra business through the export of port services.

In itself, the idea of offering port services for other countries' trade reflects a positive change of attitude. Thus, up to the early 1990s the idea was to avoid this happening, because exporters considered that the goods of neighbouring countries should not pass through their ports because they competed with domestic products, while farmers feared the entry of pests and diseases. Furthermore, the maritime authorities, which came under the respective navies, were against opening up their ports to countries with which they had border conflicts.

Today, however, in the context of greater regional political and economic integration and the progress made in privatizing ports, such opposition has lost its strength. Ports are competing for cargo and trying to attract pri-

---

[1] This definition is quite independent of the degree of industrialization of the port or its volume of traffic. We have tried to avoid using the term "megaport" because there is no generally accepted definition of this concept, and moreover its use is not necessary. The concentration of cargo may involve one or more modes of transport. If only maritime transport is involved, we speak of "transshipment". If cargo arrives from another country by land and leaves the port by sea, we use the term "transit".

[2] This transshipment involves two port movements: a container arrives on one ship, is stored temporarily in the port, and then leaves on another ship. It is used above all to take advantage of the econo-

mies of scale offered by bigger ships and to increase the frequency of services to a given destination. Transshipment traffic has greatly increased in recent years thanks to technological advances, the use of bigger ships, and increased use of containers.

vate investors, and their geographical location on the Pacific rim has opened up expectations of potentially big business reflected in the press in headlines such as "Megaports in South America: Conquering the Pacific" (*El Mercurio*, 1998, p. D1).

The present article will analyse whether there really is a potential for hub ports on the west coast of South America and whether the ports on that coast have comparative advantages for moving the trade between South America and the Asia-Pacific countries.

II

South America's trade and

its transport by sea

What is the relation between the geographical location of a country and investments in ports? Broadly, there are two possible interdependences:

i) The port would be a means of modifying trade flows: improvement of the ports could help to offset geographical disadvantages and promote the country's external trade; in this case, the country would invest in its ports as part of its trade policy.
ii) The trade flows and geographical location would be an opportunity for generating income through the supply of port services: the ports could take advantage of their privileged geographical location and offer their services for the foreign trade of their own and neighbouring countries; in this case, the country would invest in its ports in order to export port services.

Both these motives could play an important role in the potential development of hub ports in South America. The aim would be to reduce transport costs for the foreign trade of the country in question while at the same time attracting additional cargo from neighbouring countries, which would help the port to generate economies of scale and hence ultimately also reduce the costs of the country's own foreign trade.

1.   The port as a facilitator of foreign trade

Trade flows are influenced by the geographical location and distances between countries, as well as the presence or absence of transport services covering those distances. Countries which are close to each other have more bilateral trade than countries which are further apart. This is partly explained by historical, political, cultural and linguistic reasons, but also by transport costs and the time goods take to arrive. According to a regression made by Gallup and Sachs (1999), each 1,000 kilometres of dis-

tance between a country and its main markets raises the transport costs by one percentage point of the value of the goods.

In 1998, 99.75% of the total volume of the foreign trade of Argentina, Brazil, Chile, Peru and Uruguay with Asia, North America and Europe was transported by sea, and only 0.25% by air. The situation is somewhat different when the trade is analysed by value, however: since the goods of highest value and lowest weight tend to be transported by air, the share of sea transport in intercontinental trade goes down to 80.15% of the total value, while the share of air transport goes up to 19.85%.

Within the trade of the South American countries which is transported by sea, it may be noted that Chile, Ecuador and Peru, which are on the west coast, have relatively less trade with Europe than Argentina, Brazil and Uruguay, which are on the east coast (figure 1).

It may be seen from the figure that, together, the latter three countries have 4.5 times more intercontinental trade by sea than Chile, Ecuador and Peru. Within this trade, the east (Atlantic) coast countries' trade with Europe was almost three times greater than their trade with North America, whereas the Pacific coast countries' trade with North America was almost double their trade with Europe. Although in terms of total volume the three Atlantic coast countries had 3.5 times more trade with Asia, in relative terms the Pacific countries' trade by sea with Asia was equally important for them.

Are these trade flows by sea the result of the distances and shipping and port services involved? Later on, we will examine the comparative advantages of ports on the two coasts for trade with the various continents. For the moment, however, we may note that the relative weight of the intercontinental trade of the countries in question corresponds approximately to the distances between the South American coasts and the other conti-

FIGURE 1

Three South American countries on the Pacific coast and three on the Atlantic: Volume of their intercontinental trade by sea,[a] 1998

(Percentages and thousands of tons)

Per cent

Thousands of tons	Chile, Ecuador, Peru	Argentina, Brazil, Uruguay
☐ Europe	14,300	132,179
■ North America	26,873	45,563
☐ Asia	26,648	95,058
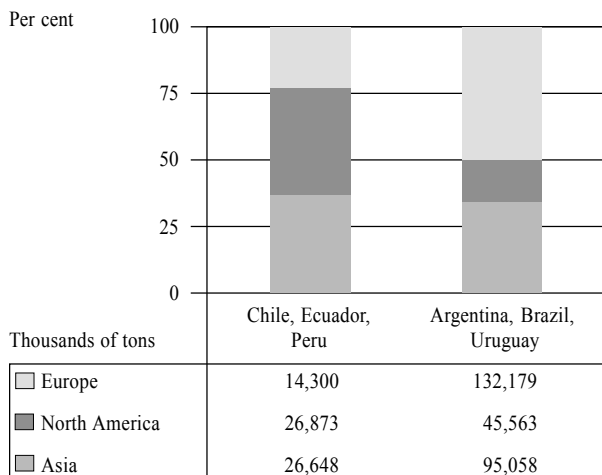
Source: International transport database of the ECLAC Transport Unit.

[a]   Includes imports and exports. Trade with Africa accounts for less than one per cent of total trade. Data for Ecuador include trade by air. North America comprises only the United States and Canada. Asia includes the Asian countries and also Australia and New Zealand.

nents. The MERCOSUR countries (Argentina, Brazil and Uruguay) are closer to Europe than Chile, Ecuador and Peru. In order to reach Europe, ships from the latter countries must go through the Panama Canal, which involves extra costs and delay. Both coasts of South America are approximately the same distance from Asian ports.

Can trade be promoted through investments in port infrastructure? The answer is affirmative, provided that such investment reduces costs and/or raises productivity. Such improvements reduce the "economic distance": i.e., they reduce the negative impact of the geographical distances involved. The recent (and ongoing) port privatization and modernization operations and maritime transport liberalization measures taken by the South American countries can be expected to give rise to a general increase in intercontinental trade.

Can trade be promoted if a specific region makes investments in its port infrastructure? In principle, the answer would be yes, but probably not in the specific case of South America, where all the ports serve trade with all continents, so that there do not seem to be any reasons to expect changes in direction of the main maritime trade flows to the various continents. The situation might be different if the road transport infrastructure were

changed, for example in order to facilitate the access of Chilean goods to Argentine ports, which would probably lead to an increase in Chile's trade with Europe.

To sum up, trade flows and international transport services influence each other mutually. Both are partly the result of the geographical location of the countries and the distances to the main markets. The impact of trade flows on the volume of port traffic is stronger, however, than the influence that greater port efficiency could have on the volume of trade.

2.   Geographical location as a factor for the establishment of hub ports

Do the ports on the Pacific coast of South America have sufficient comparative advantages to become hub ports? What are the possibilities of concentrating cargo in Chile, Colombia, Ecuador and Peru for its subsequent redistribution?

In its *Review of Maritime Transport* (UNCTAD, 1999, p. 93), UNCTAD notes that in South America there are a number of ports which are impatient to become hub ports, and on the west coast several Chilean ports will compete with Callao (Peru) or Guayaquil (Ecuador). According to *El Mercurio* (1998), "Chile and Peru are vying to establish megaports on their coasts which could link up with bioceanic corridors to become the leading port of the region for trade with Asia". Many articles in the specialized press highlight the "intense competition" between the ports along the west coast of South America (see, for example, *Schednet News*, 1999).

In Ecuador, Manta is being mentioned as an "international transfer port". The review CAMAE (1999), for example, describes its "geographical advantages" and "technical advantages", claims that "international megafirms need to have a port of this category on the South American coast", and highlights its potential for "serving as a port for unloading containers arriving in large ships from abroad and then distributing them to other ports in smaller vessels" and "Minimizing costs and maximizing the transport of cargo between different ports of Asia, Europe and the United States and South America".

In mid-1998 the United States Trade Development Agency (TDA) authorized the expenditure of US$ 362,000 on a prefeasibility study in this respect. According to CAMAE (1999), "the project was considered to be viable, so that the TDA included it among the 125 projects eligible for investment in South America and registered the Transfer Port project under code TRAN-39: "Ecuador - Expansion of the Port of Manta" and the "recommended

capital expenditure programme for the Port of Manta amounts to US$ 135,996,240".

In Peru, the port of Callao is that which has the highest hopes of becoming a hub port. According to a brochure designed to promote private investments, "Peru's strategic location in South America makes its seaports highly attractive as potential outlets for seaborne trade between Latin America and Asia. Furthermore, the foreign trade of Peru and other emerging economies of Latin America is expected to keep on growing, thus further increasing the demand for port services" (Comisión de Promoción de Concesiones Privadas, 1998). Another brochure, in this case published by the National Ports Corporation (Empresa Nacional de Puertos - ENAPU), states that "Peru's ports enjoy a privileged geographical location on the Pacific Rim which enables them to act as ports that link up with the countries of the Atlantic coast and interior of South America through a vast network of railroads, highways and navigable rivers suitable for intermodal transport" (ENAPU, undated). The specialized press, too, mentions that Callao is well located for taking transshipment cargo to and from the whole of the west coast of South America (Lloyds List, 1999).

Private investments amounting to some US$ 300 million are expected to be made under the concessions for the port of Callao. In 1999 a US$ 240 million soft loan from the Japanese government for its modernization was rejected. The granting of a concession for a container terminal, which was planned for 1998 or 1999, was postponed until the year 2000. One of the reasons for the postponement was the existence of doubts as to whether the port should not be divided into several terminals in order to increase in-port competition and avoid a monopoly. On the other hand, if Callao wanted to become a hub port it might be better not to divide it but rather to try to put it in the most competitive position possible compared with other ports.

In Chile, the port best known for its aspirations to become a hub port is Mejillones, north of Antofagasta. Indeed, in the local press it is usually called a "megaport". As far back as 1996, in a working paper of the regional government of Antofagasta entitled "Megaport of Mejillones" (Schellmann, 1996), it was claimed that the bay of Mejillones has "unrivalled natural advantages" and that "the megaport of Mejillones is a strategic point

where the hinterland of the great production areas of the Gran Chaco joins up with the Asia-Pacific Basin". According to *El Diario* (1999), a Chilean government representative said that "Mejillones is winning the battle to become a megaport of the South Pacific".

With regard to the amounts of investment involved, when the Mejillones project was begun it was estimated that the total investment would be some US$ 600 million and it was planned to grant the concession in late 1998. After various postponements, however, in late 1999 the concession was awarded to a consortium of Chilean firms which undertook to invest a total of around US$ 100 million by the year 2002. This consortium is currently seeking finance from commercial banks and multilateral financial institutions. This first phase of the project is mainly limited to the construction of installations for handling copper exports, although the documentation soliciting loans to finance the project continues to stress the long-term potential for attracting cargo from neighbouring countries and for the transshipment of containers.

Other Chilean ports with expectations of attracting more transshipment or transit traffic are, in particular, Arica, Iquique, Valparaíso, San Antonio and Talcahuano/San Vicente, although none of them are usually described as "megaports". According to the newspaper *Estrategia* (1998), the Mayor of Iquique "announced that the Ministry of Public Works had approved the deepening of the Northern port from 16 to 17 metres draft for a new berth that will take vessels with a capacity of 7,000 containers. 'The new vessels operating now have got bigger, and Iquique does not want to be left out of the world market. It is not only the megaport being built at Mejillones that has the right to receive such ships', he added". According to the Web page for the Iquique Free Zone (2000), that Zone is "in a strategic geographical position" and is "South America's principal place of business, where markets of the Pacific Basin and the Southern Cone of the American Continent connect".

In short, there are expectations of the possible establishment of hub ports in all the South American countries on the Pacific coast, based on the growth of trade, regional and world economic integration, the privatization of ports, and the perceived advantages of a strategic geographical location.

III

Transshipment centres in the world

In terms of volume, most maritime cargo is transported as liquid bulk (above all petroleum) and dry bulk (grains, coal, iron ore). In terms of the value of goods and freight charges, containerized cargo is more important.

Analysis of potential hub ports generally centers on their possibilities of concentrating containerized cargo transported by sea. This cargo is transported by regular liner services. Bulk cargo, in contrast, is generally transported in chartered vessels and is less suitable for transshipment operations.

1. A business decision

The selection of the mode of transport for a foreign trade operation generally depends on a mainly commercial decision: the goods must arrive at their destination as soon as possible and at the lowest cost and risk.

a) Journey time

Rapid delivery is increasingly important. The average value of each ton of merchandise is going up all the time, and this also raises the capital costs. Just-in-time delivery is becoming more and more common. The incidence of a transshipment operation on the total journey time depends on various factors: on the one hand, the transshipment operation in itself involves extra costs and time, and may also mean a diversion from the direct route in order to reach the transshipment centre. On the other hand, however, the goods may be loaded on a faster ship at that centre.

b) Frequency

A journey which is rapid in itself is not much use to an exporter if his cargo has to wait many days or even weeks for a direct transport service. One of the main advantages of passing through hub ports is that they concentrate cargo and make possible more frequent departures to the different destinations.

c) Cost

The extra cost of a transshipment operation may be partly offset by the advantage of being able to use bigger ships with lower operating costs. On the route between the Unite States and Asia, for example, it is estimated that

the use of the biggest ships (called "post-panamax" because they are too big to go through the Panama Canal) gives shippers a cost advantage of US$ 27 per container compared with "panamax" ships, which are the largest ones that can use the Canal (Drewry Shipping Consultants, 1996; Hoffmann, 1999). The ships currently serving the South American Pacific ports are only about half the size of panamax ships.

Quite apart from the possibility of consolidating cargo of different origin, the volume of trade between ports on a given route could itself justify the use of bigger ships on intermediate stages. For example, if we assume that there are 50 containers of bilateral trade from each of 12 ports (i.e., 11 stages), then on the last stage the ship would only be carrying 550 containers (the trade with the remaining 11 ports), whereas on the sixth stage (between ports 6 and 7) it would be transporting 1,800 containers. The general formula is:

$$\text{Number of containers on the ship} = k\,(n-k)$$
where n = total number of ports on the route and
k = number of the stage.

This example reflects quite realistically the case of trade between the west coast of South America and Europe or North America. There are various services which call in at 10 to 15 ports per voyage, and the number of containers unloaded in each port rarely exceeds 600.

d) Risk

Every transshipment operation involves the risk of loss or damage of the goods and delays due to errors or strikes. Insurance premiums therefore tend to be higher if transshipment services are used.

e) Volume

The journey frequencies and size of the ships used are naturally closely linked with the volume of the transactions that must be covered. If this volume is not large enough even to fill smaller ships running at a frequency of at least one departure per month, there will simply be no direct service at all, and it will be necessary to use feeder services that link the port with a hub port.

The traffic balances also depend on the volume of goods transported. If cargo is only available in one di-

rection, it is less profitable to establish a direct service, and it is more expedient to try to concentrate the cargo at places where the maritime transport flows can be balanced in both directions.

f) *The case of South America*

As a real example of the relation between journey lengths and frequencies, it may be noted that five different weekly services to Northern Europe depart from the MIT port on the Caribbean coast of Panama, whereas there are only three similar direct services from the west coast of South America, between one and three times per month. These services pass through the Panama Canal, and one of them calls in at MIT. On average, the journey between MIT and the Northern European ports on the five weekly services takes one day less than the Panama-Europe leg of the services from San Antonio and Callao. Altogether, the five weekly services departing from MIT also connect with a larger number of different ports in Europe.

Consequently, if for example a Peruvian exporter does not want to wait for the departure of one of the three direct services, he may be able to find another service that will take his goods to Panama, where they can take advantage of the next departure of one of the five weekly services to Europe.

Because of the increase in the number of mergers and alliances between shipping companies, such combinations of services are increasingly frequent. In the trade with Asia there are already a number of established services which link up North-South services from South America with East-West services in Panama or Los Angeles. There are even services which carry out transshipment at the Panamanian MIT port, on the Caribbean coast. In that case, the containers pass through the Canal twice.

Generally speaking, the connections between the west coast of South America and Asia are more suitable for the use of transshipment services, because the stage on which big ships can be used is much longer than in the services to Europe or North America.

In short, ports consolidate cargo so that it will reach its destination more cheaply and quickly. The decisions in this respect are eminently commercial and hardly involve political considerations.

2. Current transshipment centres

The biggest container ports are currently in Asia, the United States and Europe. There, the transshipment ports are located primarily at points where the main sea routes intersect. Tables 1 to 4 show the volume of container traffic in different regions of the world and the volume

TABLE 1

The five main container ports in the world: port traffic in containers, 1998
(in TEU)

Port	Port traffic
Singapore	15,000,000
Hong Kong	14,582,000
Long Beach/Los Angeles (United States)	7,478,218
Kaohsiung (Taiwan)	6,271,053
Rotterdam	6,010,000

Source: Cargo Systems, 1999.

TABLE 2

The five main container ports in Latin America and the Caribbean: Port traffic in containers, 1998
(in TEU)

Port	Port traffic
Buenos Aires (Argentina)	1,138,000
Cristóbal (Panama)	1,117,035
Santos (Brazil)	859,500
Kingston (Jamaica)	670,858
Puerto Cabello (Venezuela)	486,774

Source: ECLAC, 1999.

TABLE 3

The five main container ports on the west coast of South America: port traffic in containers, 1998
(in TEU)

Port	Port traffic
San Antonio (Chile)	415,001
Guayaquil (Ecuador)	407,434
Callao (Peru)	378,013
Valparaíso (Chile)	255,687
Buenaventura (Colombia)	143,420

Source: ECLAC, 1999.

TABLE 4

Main transshipment areas of the world: Port transshipment movements, 1998
(In TEU)

Area	Port movements
Southeast Asia	13,356,000
Far East	8,374,000
Northern Europe	6,312,000
Southern Europe	5,940,000
Middle East	3,077,000
Central America and the Caribbean	1,994,000
North America	1,623,000
Africa	1,215,000
South Asia	1,200,000
South America	230,000
Oceania	112,000

Source: Drewry Shipping Consultants, 1999.

of transshipment operations by regions, both measured by TEUs.[3]

At the world level, there were 185 million port movements during 1998, including movements of empty containers and transshipment operations. Of this total, 23% corresponded to transshipment movements (Drewry Shipping Consultants, 1999); the percentages by the main transshipment ports are given in table 5. The small amount of transshipment traffic currently registered in South America (table 6) is concentrated above all in Cartagena (Colombia) and Puerto Cabello (Venezuela).

To sum up, both the total volumes of cargo transported in South America and the percentage of transshipment within those totals are very small compared with other regions of the world. South America accounts for only 3.4% of world movements of containers in ports, and only 3.6% of this is transshipment traffic. Indeed, the South American region accounts for only 0.5% of the total transshipment operations in the world.

3.   Requisites for a hub port

a)   *Land links*

Many hub ports concentrate cargo by land, as for example in the case of those in Northern Europe and the United States. Hong Kong also receives most of its cargo by land. In order to be able to concentrate cargo in this way the port must naturally have links with other forms of transport, especially railways, which are important for obtaining high volumes of cargo. If the port is in an industrial area which offers other services for the cargo, this could be an additional advantage.

b)   *Maritime links*

The world's main transshipment centre (where the cargo arrives and leaves by sea) is Singapore. In recent times there has been a tendency to establish ports whish have almost no traffic of local origin and are devoted to transshipment traffic. The main ports in the region which serve as transshipment centres have also continued to grow because they have international maritime services and cargo from smaller ports must be transferred to them to connect with those services. The main basis for the viability of those centres is their geographical location.

In ports where intercontinental routes cross or connect, transshipment operations take place between ships serving two different routes. Examples of this are

---

[3] TEU = *twenty foot equivalent unit*: equivalent to a 20 foot container.

TABLE 5
**Main transshipment ports:
Transshipment as a percentage of
port container traffic, 1998**[a]

Port	Transshipment (%)
Malta	93
Damietta (Egypt)	90
Algeciras (Spain)	84
Singapore	82
Gioia Tauro (Italy)	80
Kingston (Jamaica)	75
Colombo (Sri Lanka)	70
MIT (Panama)	70
Dubai	50
Kaohsiung (Taiwan)	43
Rotterdam (Netherlands)	40
Bremerhaven (Germany)	30
Hamburg (Germany)	30
Felixstowe (United Kingdom)	28
Antwerp (Belgium)	25
Pusan (Korea)	21
Hong Kong	18
Kobe (Japan)	15

Source: Data from Drewry Shipping Consultants and direct information from the ports.
[a] The data are for 1998 or the last available year.

TABLE 6
**South American ports: Transshipment
as a percentage of port container traffic, 1999**

Port	Transshipment (%)
Cartagena (Colombia)	50
Puerto Cabello (Venezuela)	38
Callao (Peru)	6
Buenos Aires, Puerto Nuevo (Argentina)	3
San Antonio (Chile)	3
Guayaquil (Ecuador)	2
Santos (Brazil)	2
Rio de Janeiro (Brazil)	2

Source: Prepared by the author on the basis of various sources.

Algeciras (Africa-Europe route and North America-Europe-Asia route), Jamaica and Panama (South America-North America-Europe route and Europe-North America-Asia route) and Singapore (Europe-Asia-North America route and Australia-Europe route). Transshipment operations are also carried out at Gioia Tauro, Malta and Dubai between different parallel services linking North America, Europe and Asia. These ports operate with or without cargo from their local hinterland.

Some ports connect up a local market with an international route that passes through the region. Examples of this are Colombo (Indian subcontinent), Gioia Tauro and Malta (the Mediterranean), Jamaica and Panama

(Caribbean, Central America, west coast of South America), Miami (Caribbean) and Singapore (southeast Asia). These ports operate with or without cargo from their local hinterland.

There is a tendency to concentrate cargo from neighbouring islands and countries at ports located at the end of international routes. Examples of this are Gothenburg (connects with Scandinavian ports), Port of Spain (connects with other Caribbean ports), San Antonio (connects with ports in southern Chile), and Buenos Aires (connects with various river ports in the area). These are usually main regional ports whose cargo comes mainly from the local hinterland, and they are not usually called "hub ports" because the transshipment traffic tends to represent only a small percentage of their total port traffic.

Obviously, transshipment operations can be carried out for different purposes at each hub port. Exporters and importers in regions of the traditional north-south traffic generally have at their disposal direct services to the main markets in Europe, North America and Asia, but also an increasing number of services involving at least one transshipment operation. Examples of this are Australia (with transshipment at Singapore), India (with

transshipment in Sri Lanka), East Africa (with transshipment in the Middle East), West Africa (with transshipment at Algerciras), and South America (with transshipment in Jamaica, Panama, or North American ports).

Zohil and Prijon (1999) have analysed (for the Mediterranean area) the relation between the volume of port traffic generated by the port area itself, geographical location, and the volume of transshipment traffic. They conclude that the volumes of transshipment traffic of a port are a linear function of the volume of port traffic and an inverse linear function of the distance from the main line of transit. In other words, ships tend to prefer ports for which they have local cargo and take advantage of their presence there to engage in transshipment operations. The shorter the detour from the main route that the stopover involves, the more likely that port is to be chosen as a transshipment centre.

To sum up, in order to become a hub port a port must have ample land transport links, be located in a place where maritime routes connect or cross, or have big volumes of locally generated cargo. None of these conditions exist in the ports on the west coast of South America to the extent that they do in the hub ports already operating in the world.

IV

Maritime services in South America

What are the features of the regular lines offering maritime transport services in South America? How do the services on the west coast compare with those on the east coast? What sort of distances does a direct service by a charter vessel have to cover? These questions need to be analysed in order to determine if it is possible to concentrate cargo from South American countries and where, and whether the ports on the Pacific side have comparative advantages compared with those on the Atlantic.

1. Regular liner services

a) *Comparison between the east and west coasts of South America*

If we compare the regular liner services covering South American ports on the Pacific and Atlantic coasts (table 7), as well as the options open to South American exporters and importers, the following observations may be made.

There are almost twice as many ports with liner service on the east coast as on the west coast of South America, and 56% more regular services depart from east coast ports than from the main ports on the west coast. There are also more companies offering such services, which are also more frequent.

Ships sailing from the east coast arrive more quickly on the east coast of the United States, Europe and Southeast Asia (Singapore). The journey time to the continent of Asia (Hong Kong) is approximately the same in both cases. Ships sailing from the west coast arrive more quickly in Japan and the west coast of the United States, however. Indeed, there are no regular direct services between Los Angeles and the MERCOSUR countries, although there are regular services that transship cargo at Puerto Cabello (Venezuela).

Greater economies of scale are obtained on the east coast, which handles almost twice as many containers. Each of the regular services transports about 35% more

TABLE 7

**South America: Regular liner services from ports on the
Pacific (west) and Atlantic (east) coasts of South America[a]**

	West coast: Chile, Colombia[b], Ecuador, Peru	**East coast**: Argentina, Brazil, Uruguay
Ports: Number of ports with regular services as at 1 January 2000[c]	**Total** 13: Chile 7, Colombia 1, Ecuador 2, Peru 3	**Total** 25: Argentina 5, Brazil 19, Uruguay 1
Main ports: Number of regular services as at 1 January 2000[c]	San Antonio (Chile) 20 Callao (Peru) 20	Buenos Aires (Argentina) 25 Santos (Brazil) 27
Services: Total number of departures per month on each coast, as at 1 January 2000[c]	**Total**: 356 To Asia: 74 To North America: 221 To Europe: 61	Total: 556 To Asia: 98 To North America: 196 To Europe: 273
Volume moved per service: Estimated average number of TEUs moved by each service per year[d]	All services: 9195 To Asia: 8708 To North America: 8093 To Europe: 13,125	All services: 12,500 To Asia: 10,906 To North America: 18,121 To Europe: 9302
Journey time: Minimum duration of voyage from the main South American ports as at 1 January 2000	San Antonio - Singapore: 36 days San Antonio - Hong Kong: 33 days San Antonio - Yokohama: 26 days San Antonio - New York: 19 days San Antonio - Hamburg: 31 days Callao - Singapore: 34 days Callao - Hong Kong: 25 days Callao - Yokohama: 21 days Callao - New York: 14 days Callao - Hamburg: 25 days	Buenos Aires - Singapore: 25 days Buenos Aires - Hong Kong: 29 days Buenos Aires - Yokohama: 35 days Buenos Aires - New York: 16 days Buenos Aires - Hamburg: 19 days Santos - Singapore: 21 days Santos - Hong Kong: 25 days Santos - Yokohama: 31 days Santos - New York: 14 days Santos - Hamburg: 15 days
Frequencies: Number of days between departures on each direct service. Weighted average for each port as at 1 January 2000 (for example: a weekly service would give the number 7.0).	San Antonio - Asia: 10.2 San Antonio - N. America: 11.0 San Antonio - Europe: 13.8 Callao - Asia: 10.2 Callao - N. America: 11.0 Callao - Europe: 13.8	Santos - Asia: 9.6 Santos - N. America: 9.4 Santos - Europe: 9.3 Buenos Aires - Asia: 8.4 Buenos Aires - N. America: 9.7 Buenos Aires - Europe: 9.2
Shipping companies: Number of shipping companies offering liner services as at 1 January 2000[c]	To Asia: 8 To North America: 20 To Europe: 12	To Asia: 14 To North America: 30 To Europe: 23
Size of ships: maximum size of ships as at 1 January 2000	2200 TEU, with ship's own cranes	3428 TEU, without cranes on ship (used on the Asia service)
Trade in containers: Total number of shipping movements, imports and exports, in 1998 (in thousands of TEU). Includes non-regular services.	**Total**: 1131 Asia: 209 North America: 607 Europe: 315	**Total**: 2200 Asia: 349 North America: 1051 Europe: 800
Balances: Export/import balance of shipping movements in TEU in 1998	**Total**: almost in balance = 1.11 Asia: surplus = 2.48 North America: deficit = 0.79 Europe: surplus = 1.30	**Total**: almost in balance = 0.96 Asia: surplus = 2.23 North America: deficit = 0.68 Europe: almost in balance = 1.07

Source: Prepared by the author on the basis of American Shipper (2000); Datamar Consultores Asociados (2000) and World Sea Trade Service (1998).

[a] This table does not take account of the ports on the north coast of South America (Cayenne, Colombia, Guyana, Suriname and Venezuela). TEU = twenty-foot equivalent unit (a unit equal to a 20-foot container). Figures for San Antonio include services from Valparaíso. Many services are the result of cooperation among several shipping companies.
[b] Buenaventura.
[c] Only direct services; these may include stopovers in other ports, but without transshipment.
[d] Approximation, based on services offered in January 2000 and containers moved during 1998.

containers, and the vessels used are larger. There are services that use ships which do not have cranes of their own. This saves costs, as it is not necessary to transport dead weight and idle capital on the voyage. As most of the ports on the west coast do not yet have specialized cranes for handling containers, however, the services covering that coast have to have their own cranes.

Both coasts have a surplus of containers in their trade with Asia and a deficit in their trade with North America. Overall, however, South America's trade in containers is more or less balanced.

b) *Comparison of the services from the east and west coasts of South America to other regions*

A more detailed analysis of the different trade regions reveals that the relative weight of sea transport services to Asia is approximately the same for both coasts. The main difference is the predominance of services to North America on the west coast and of services to Europe on the east coast. Every day 4.5 times more ships leave for Europe from east coast ports than from those on the west coast. The Atlantic side also has 32% more departures to Asia, while there are 13% more departures for North America from the west coast. The final result is similar in terms of the total number of services. Figure 2 shows the simple sum of services from all the ports on each coast of South America to the three destination regions. Services which go from a South American port to two destination regions (for example, first Miami and then Hamburg) are included in both regions. The simple co-efficient of correlation between the number of services (figure 2) and the number of port departures (table 7) is +99%.

In addition to this quantitative comparison, it is also necessary to take into account the frequency of the services. There are weekly services to Asia and North America from ports in all the countries in question. In the case of services to Europe, however, the maximum frequency of those from Chile and Peru is only once every ten days, whereas there are lines from Argentina and Brazil which leave every five days.

In the case of the route to Europe, the ports on the east coast have a clear advantage over those on the west coast. Exporters and importers who have access to ports on both coasts should therefore prefer those on the Atlantic.

For services to the west coast of North America, the Pacific ports have an advantage, as there are no regular direct services between the east coast of South America and western North America. It should be borne in mind,

FIGURE 2
West and east coasts of South America:
Liner services by destination region, January 2000
(Percentages and number of services)
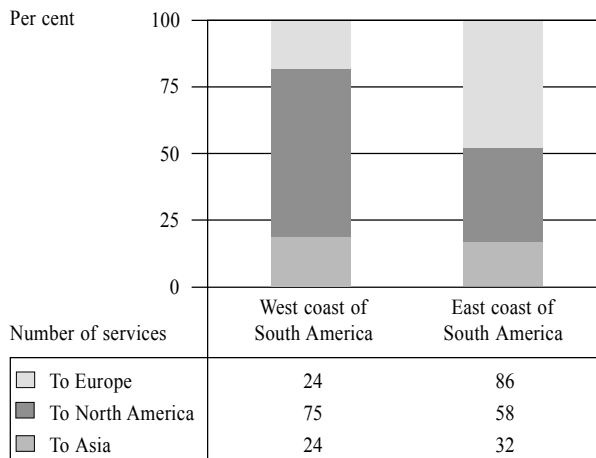
Source: Prepared by the author on the basis of *American Shipper* (2000).

however, that in order to reach any given destination inside the United States or Canada it is not necessary to enter by the east or west coast in particular, because in North America, unlike South America, there are intermodal connections providing efficient transport from any port in those two countries.

With regard to services to Asia, for most destinations it would be better to use the services sailing from Argentina, Brazil or Uruguay, provided the intermodal land connections within South America permit this.

In short, the services offered by the ports of both coasts basically reflect the needs of local trade. The larger volumes available on the east coast result in more services, bigger ships and higher journey frequencies. The greater relative importance of services to Europe from the east coast and to North America from the west coast basically reflect the needs of the corresponding foreign trade with those regions (see figure 1 above).

c) *Regular services from the Pacific coast of South America*

There are currently no regular international services from the main west coast ports of South America which involve transshipment of containers within South America. There are direct services, and also the possibility of using services involving transshipment in Central America or North America.

As regards regular services connecting the Pacific coast with Asia, there are direct services which include

the main ports in Chile, Peru, Ecuador and Colombia and cross the Pacific from Buenaventura, Callao or San Antonio (for example, those provided by the shipping lines CSAV, Nedlloyd, NYK, P&O Nedlloyd and Rickmers). There are also indirect services which involve transshipment in, for example, the MIT port in Panama or Los Angeles (provided by the shipping lines APL, CCNI, COSCO, Ecuadorian Line, Hapag Lloyd and Maersk).

Analysis of the regular services covering the South Pacific coast indicate that all the main shipping lines include Callao and San Antonio. Some use San Antonio instead of Valparaíso, but there are no services which include Callao but do not include either San Antonio or Valparaíso. Table 8 shows the number of regular services per port.

d) *The future outlook*

The shipping lines are tending to establish their main routes in an east-west direction, with transshipment services linking them up with north-south routes. The tendency towards more services involving transshipment is also reflected in the statistics which indicate that port traffic is growing faster than actual maritime movements: each movement of a container between the exporting and importing country constitutes only one maritime movement, but it may involve two, four or even more port movements, depending on the number of transshipment operations during the voyage.

Figure 3 illustrates a possible future pattern of regular liner services. We will surely not arrive at the extreme of having only indirect services from the west coast of South America, but there is a definite trend in this direction. In the case of various regular services on the Pacific coast of South America, the pattern shown in figure 3 is already almost a reality, with the main hub ports being Long Beach in the United States, the MIT port in Panama, and Kingston in Jamaica.

TABLE 8
Pacific coast of South America: Regular direct services, by ports

Port	Number of services		
	Asia	North America	Europe
Buenaventura (Colombia)	3	8	2
Guayaquil (Ecuador)	3	10	3
Callao (Peru)	4	14	4
Iquique (Chile)	4	8	1
Antofagasta (Chile)	2	4	2
Valparaíso and San Antonio (Chile)	4	14	4
Talcahuano and San Vicente (Chile)	2	4	1

Source: Prepared by the author on the basis of *American Shipper* (2000).

FIGURE 3
World: Possible future pattern of direct and indirect services[a]
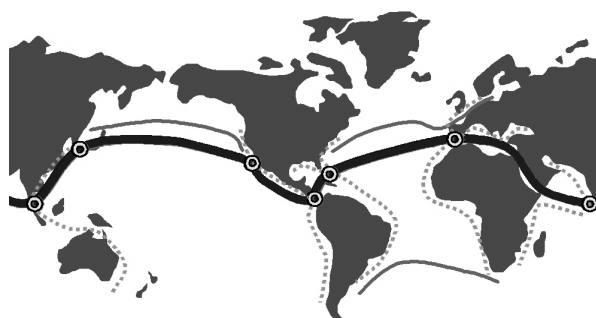
Source: Containerisation International (1999).

[a] The thick line shows east-west services around the world in ships of up to 15,000 TEU, assuming widening of the Panama Canal. Alternatively, there could be to-and-fro services sailing to and from the two coasts of the United States. The three thin solid lines show direct services between Asia and North America, North America and Europe, and South America and South Africa. The broken lines show feeder services connecting up with the main east-west service.

2. Distances for charter services

Not all trade is transported by regular shipping lines. Much of it –especially in the case of dry and liquid bulk cargo– is transported in ships which are chartered for specific voyages and do not follow established routes to cover a schedule of port visits but basically seek the shortest distance to their destination.

Comparisons of the distances involved should not be limited only to crossing the Pacific, but should also include the alternative of sailing via the Cape of Good Hope in South Africa. Furthermore, they should include not only the most easterly Asian ports, but also the biggest ports in that region, namely Hong Kong and Singapore (table 9).

Some of these comparisons may give surprising results:

– From any Brazilian port or Buenos Aires, the distance to Singapore is shorter than from any port on the west coast of South America.
– Rio de Janeiro is the same distance from Hong Kong as Antofagasta, which is on the same latitude.
– In order to reach Los Angeles from Buenos Aires, the voyage is shorter through the Straits of Magellan than through the Panama Canal.
– Valparaíso appears to be closer to Singapore than Callao, but the straight-line route would pass very close to the Antarctic and would probably not be viable for most voyages.

TABLE 9
Distances by sea

	Panama	Singapore	Hong Kong	Yokohama	Los Angeles	New York	Hamburg
Los Angeles	2 912	7 867	6 380	4 839	0	4 930 via Panama	8 012 via Panama
Panama (Colón)	0	10 504	9 194	7 725	2 956	1 972	5 005
Buenaventura	395	10 375	9 317	7 681	3 047	2 369 via Panama	5 440 via Panama
Guayaquil	892	10 726	9 505	7 987	3 228	2 872 via Panama	5 947 via Panama
Callao	1 387	10 676	10 018	8 558	3 654	3 367 via Panama	6 442 via Panama
Antofagasta	2 178	10 524	10 532	9 154	4 433	4 158 via Panama	7 233 via Panama
Valparaíso/ San Antonio	2 858	9 945	10 532	9 280	4 806	4 638 via Panama	7 713 via Panama
Recife	3 217	8 934 via S. Africa	10 220 via S. Africa	10 942 via Panama	6 173 via Panama	3 698	4 450
Rio de Janeiro	4 289	8 863 via S. Africa	10 149 via S. Africa	11.517 via Straits of Magellan	7 245 via Panama	4 780	5 535
Santos	4 565	9 035 via S. Africa	10 321 via S. Africa	11 335 via Straits of Magellan	7 521 via Panama	4 955	5 710
Buenos Aires	5 390	9 301 via S. Africa	10 587 via S. Africa	10 647 via Straits of Magellan	7 243 via Straits of Magellan	5 910	6 665

Source: Fairplay Ports Guide, 1998.

– New York is closer to Callao and Valparaíso than Los Angeles. Economically, however, the distance is greater than the number of miles would indicate, because in order to reach New York it is necessary to pass through the Panama Canal.
– Colombia, Ecuador, Peru and even Chile are closer to Europe than to Asia.
– Carrying out a transshipment operation at Los Angeles when sending goods between Asia and the west coast of South America involves practically no increase in the total distance. Thus, travelling from Guayaquil to Singapore via Los Angeles involves a detour of only 3.5%, Callao-Hong Kong via Los Angeles increases the distance by less than 1%, and Valparaíso-Yokohama via Los Angeles involves a detour of less than 4%.

We thus see that, in terms of distances by sea, the Pacific coast of South America does not offer any comparative advantage for South America's trade with Asia, but in comparison with the Atlantic coast it does offer such an advantage for trade with North America.

3. **Reserved cargo**

Although the traditional cargo reservation practices which hindered maritime trade throughout the region up to the 1980s no longer exist, a number of bilateral accords are still in force which prevent international shipping companies from serving intra-regional trade and domestic cabotage.

Thus, the international shipping lines connecting Chile with Europe, North America and Asia call in at the main ports on the Pacific coast of South America, such as Callao, Guayaquil and Buenaventura, but a number of them cannot transport cargo between San Antonio (Chile) and Guayaquil (Ecuador) or between San Antonio and Buenaventura (Colombia). For example, Mitsui OSK does not even mention that its ships call at Guayaquil and Buenaventura in the announcements of its services published in Chile, because if a Chilean client wanted to use Mitsui OSK's services on that route he would be prohibited from doing so.

The same thing occurs inside countries too. In Chile, ships not flying the Chilean flag cannot transport cargo

from Valparaíso to Antofagasta or Iquique even if they have unused space available –which they usually do. The opposition to the opening up of cabotage comes mainly from the trade unions representing the workers of shipping companies and also road transport firms. Both groups are aware of the danger of losing jobs if domestic sea transport is opened up to lines that do not fly the Chilean flag and therefore do not have to employ Chilean crew either.

This reservation of cargo is also an obstacle to port development, since every container that leaves the port by land transport instead of by sea increases urban traffic congestion and reduces the number of port movements.

In short, the concentration of cargo at hub ports is made more difficult if international shipping companies are excluded from these services, which are also made more expensive by insisting on the use of a carrier of a specific nationality.

4. Freight rates

Maritime freight rates, which fluctuate daily, depend on many factors, such as the type of product carried, the trade balances, the distance, the use of containers and the size thereof, the value of the goods, port productivity, and economies of scale. Table 10 shows the marine freight and insurance costs of the imports by sea of five South American countries, as recorded by the customs authorities. The disparities between the countries are largely explained by the different products imported. Thus, for example, Brazil imports large amounts of cereals from Argentina, which involve lower transport costs per ton than, for example, Argentina's imports of vehicles from Asia.

A more detailed econometric analysis which takes account of the different products, the distance, the value of the goods and the volumes involved shows that within a given group of products –fertilizers or vehicles, for example– every 1% increase in the volume means a re-

TABLE 10
South America (five countries):
Freight and insurance costs of imports
transported by sea or river, 1998

Country	Difference between CIF and FOB values	
	% of FOB value	Dollars per ton
Argentina	6.5	56
Brazil	6.2	27
Chile	8.4	47
Peru	9.2	40
Uruguay	8.3	51

Source: ECLAC, 2000.

duction of 0.1-0.2% in the transport costs per ton, due to the use of bigger ships and more specialized port equipment. Since Brazil has larger import volumes, it is only natural that it should pay less for their transport.

Freight quotations obtained from shipping companies confirm that in early 2000 freight rates for the export of containers from the MERCOSUR countries were lower than the rates for Chile, Ecuador or Peru. The sales staff of the different companies all agreed that this is because on the east coast of South America there is more competition, bigger ships are used, and the cost per ship in each port is lower if it can be spread over a larger number of containers. According to Sgut (1999) "freight rates to the Far East and Australasia are 30% lower from the Atlantic coast than from the Pacific". It should be noted, however, that the freight charged for each trade transaction depends on many different factors, and the fluctuations on the two coasts are not always similar.

In January 2000 the land freight between Buenos Aires and Valparaíso was US$ 1,650 per 20-foot container. According to non-official information from three shipping companies and a cargo forwarding agent, in the same month the sea transport freight for FAK (Freight All Kinds) cargo from Asia, North America or Europe to South American ports was between US$ 1,400 and US$ 2,000, the rates to Atlantic ports being lower than those to ports on the Pacific. When the land freight rates are compared with the maritime ones, not much cargo can be expected to be sent from a capital on one side of South America for export from a port on the other side.

The interdependence of the freight rates in different markets may be seen if we examine the rates in force between Asia and the west coast of South America. According to the data given in table 7 above, there is a surplus in the container trade in favour of South America, so the export freight rates might be expected to be higher than those for imports, since there is a shortage of empty containers in South America. However, in January 2000 the rates for imports from Singapore were approximately 40% higher than those for exports from South America to Singapore. The reason for this is the heavy imbalance in the trade between the United States and Asia, which gives rise to an overall surplus of empty containers on the west coast of the Americas.

To sum up, the freight rates for South America's foreign trade by sea are in line with what might be expected from the volumes, balances and products involved in the South American countries' trade with other regions. Sea freight rates are much lower than the rates for land transport, and the freight rates for the Atlantic coast are currently lower than those for the Pacific.

V

Ports and port traffic

1. Economies of scale

Port traffic costs are lower when the latest technology and best superstructure can be used. With increasingly high fixed costs and lower variable costs, an increase in volume leads to a reduction in the cost per container. This attracts additional cargo, and this in turn causes a reduction in unit costs, thus giving rise to a virtuous circle.

If we compare the Atlantic and Pacific coasts of South America, we see that the Atlantic coast has greater possibilities of obtaining economies of scale. In terms of tons of cargo moved, including cabotage (figure 4), the ports on the east coast transport five times more cargo than those on the west coast.

The port with the biggest volume of traffic in South America is Tubarão (Brazil), which mainly handles iron ore, while the port with the biggest volume on the Pacific is Balao (Ecuador), which mainly receives oil tankers. Of the 25 ports with the biggest volume of traffic in South America, 20 are on the Atlantic and only 5 on the Pacific (table 11).

In short, for liquid and solid bulk cargo the Atlantic ports have a comparative advantage because they move large volumes of cargo from their own hinterland and their unit costs tend to go down as a result of economies of scale.
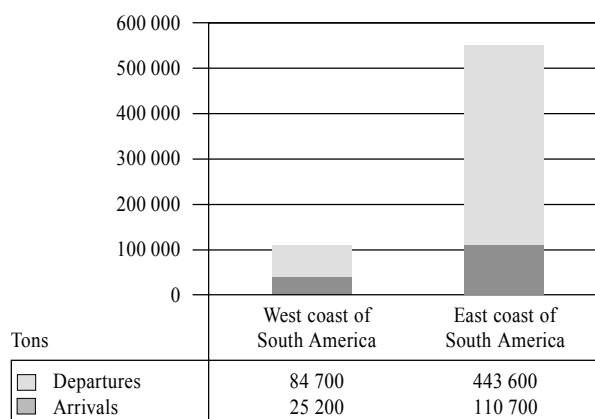
2. The danger of private monopolies

In South America, there has been a process of privatization of publicly-owned ports in recent years. Many specialized ports have always been privately owned, and new private ports are also being built.

This tendency towards greater participation by the private sector has created some fears that private monopolies may be established. Indeed, one of the main challenges in privatization is the need to keep a very close watch on these processes and avoid monopolistic abuses by the private sector after the State has ceased to operate the ports. Even if the ports are not privatized, however, there would still be the danger of public monopolies.

It is more difficult to avoid monopolies in the Pacific coast ports than in those on the Atlantic. Importers and exporters of many cities in Argentina and Brazil have easy access to several ports, whereas on the Pacific side there is usually access only to a single obvious port

FIGURE 4

South America: Forecast of port movements in the year 2000

(Tons)

Tons	West coast of South America	East coast of South America
□ Departures	84 700	443 600
▨ Arrivals	25 200	110 700

Source: UNCTAD, 1999.

TABLE 11

South America: The 25 main ports on the Atlantic and South Pacific, 1998[a]

(Millions of metric tons)

Port	Volume
Tubarão (Brazil)	69.6
Itaqui (Brazil)	50.3
São Sebastião (Brazil)	42.4
Santos (Brazil)	24.9
Aratu (Brazil)	19.1
Paranagua (Brazil)	19.1
Buenos Aires (Argentina)	18.8
San Lorenzo (Argentina)	18.6
Angra dos Reis (Brazil)	18.4
Praia Mole (Brazil)	15.4
Bahía Blanca (Argentina)	14.8
Balao (Ecuador)	14.1
Rio de Janeiro (Brazil)	11.3
Belem (Brazil)	12.6
Ponta Ubu (Brazil)	11.7
São Francisco (Brazil)	11.2
Rio Grande (Brazil)	11.5
Rosario (Argentina)	11.4
Callao (Peru)	10.2
Caleta Olivia (Argentina)	8.6
Buenaventura (Colombia)	7.3
San Antonio (Chile)	7.4
Caleta Cord. (Argentina)	7.3
San Vicente (Chile)	7.0
Quequén (Argentina)	6.5

Source: ECLAC, 2000.

[a] This list does not include ports in Colombia and Venezuela on the north coast of South America.

(Buenaventura, Guayaquil or Callao, for example). The inter-port competition is greatest between the ports of Argentina, Brazil and Uruguay: there, exporters and importers have access to railways, good highways and even river transport, so that they have more options to choose between different ports. In Colombia, inter-port competition is weaker on the Pacific side, where Buenaventura is the dominant port, than on the Caribbean, where Cartagena, Barranquilla and a number of smaller ports compete strongly for local and transshipment cargo.

In Chile, the situation is not so difficult as in the other three countries on the west coast, as there is strong competition between San Antonio and Valparaíso. There is also competition between the public ports of the former EMPORCHI (Empresa Portuaria de Chile, dissolved in 1999) and the 100% private ports that exist, for example, in the Concepción area, and between established ports and those that are currently being built (between Antofagasta and Mejillones, for example).

The need to regulate ports after their privatization makes it more difficult for them to become hub ports. The main transshipment centres in Latin America and the Caribbean –the MIT port in Panama, Kingston in Jamaica, and Freeport in the Bahamas– keep the rates they charge their clients confidential, thus making the supervisory work of the State regulators more difficult.

If a port is divided up into several terminals which compete with each other, this creates intra-port competition and thus avoids private monopolies. On the other hand, however, the division of a port like Callao which handles relatively low volumes of goods would leave each operator with very little trade, and this could discourage major investments.

Port charges are different for local cargo and transshipment cargo: the latter is charged much less than cargo leaving the port by land transport. This partly reflects the lower costs, and also the greater elasticity of demand for transshipment services.

To sum up, freedom to set port charges that are in line with the needs of the market is much more important –in fact, it is indispensable– for transshipment centres than for ports that only handle local cargo. Such freedom is more difficult (though not impossible) to ensure in privatized public ports if the State is afraid that monopolistic abuses may arise, as in the case of the main ports of the Pacific coast in Colombia, Ecuador and Peru.

3. Physical and geographical aspects

Because of the form of the tectonic platform of South America and its relative shift to the west, the Pacific coast is steeply sloping and has very few bays. Building a breakwater on it costs much more than on other coasts, and this too limits the potential of many ports for growing into hub ports. One of the exceptions to this limitation would appear to be the port planned at Mejillones, in northern Chile.

Several of the main ports on the west coast are also limited in their future growth by the fact that they are located inside cities. One of the reasons why San Antonio has grown more than its main competitor, Valparaíso, is that it has better land links with Santiago and more room for expansion.

Finally, along the whole Pacific coast the growth of ports may be limited by the danger of earthquakes and tidal waves. Antofagasta and San Antonio, for example, have suffered serious damage in recent decades. Quays, breakwaters and buildings all cost more to build in earthquake-prone areas than in other regions. Furthermore, the cost of capital for investors is always higher if it has to include a premium to cover the risk of tidal waves.

We thus see that tectonic aspects, urban congestion and lack of space for expansion are factors that limit many western South American ports' possibilities of becoming hub ports.

VI

The hinterland

1. The impact of the Andes

Above all in Chile and Peru, but also to some extent in Ecuador and Colombia, the mountain range of the Andes seriously impedes the possibility of attracting cargo from neighbouring South American countries. Transporting cargo for long distances by road –railways are not very important in the international trade of South America– and having to cross mountain passes 3,000 to 4,000 metres high would only be worth it if, on reaching the Pacific, the cost of this land transport was offset by other savings.

The proponents of new projects do not always seem to bear this in mind, as may be seen, for example, from the expectations expressed with respect to the hinterland of Mejillones (figure 5).

In a brief study on potential port privatization operations in Latin America (Hoffmann, 1997), it was considered that the impact of the Andes on trade between Chile and Argentina is equivalent to an additional distance of some 4,700 kilometres. In other words, although Chile and Argentina are neighbours, which in most parts of the world would mean greater bilateral trade, these two countries have a level of such trade that would normally correspond to countries separated by several thousand kilometres. This does not mean that crossing the Andes in a truck is as costly as travelling all those kilometres, but it does mean that, together, the impact of the closure of tunnels and passes in winter, the uncertainty about when they will be closed, the limited number of passes, the need to adjust motors for the great height, and the additional expenditure of energy has a negative effect on trade equal to a very long distance over level ground.

Improvements in the land transport infrastructure naturally help to reduce the negative impact of the Andes. Indeed, the current situation is not so extreme that no cargo from neighbouring countries passes through the ports of Colombia, Ecuador, Peru or Chile. Thus, fruit from western Argentina for the United States market already passes through Valparaíso and San Antonio. In 1999, 1.7% of San Antonio's total volume of cargo was to or from Argentina, while 0.6% was to or from other South American countries. The mineral products of northern Argentina and Bolivia can pass through Antofagasta

FIGURE 5
Chile: Expectations regarding the hinterland of a hub port at Mejillones
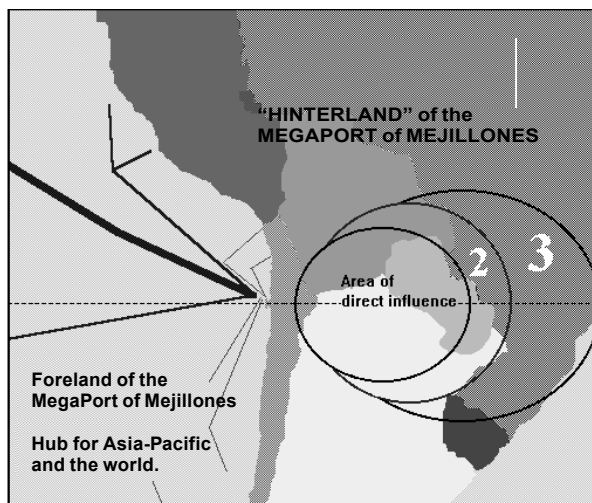
Direct area of influence: Businesses would clearly choose the megaport of Mejillones.
Area 2: The megaport of Mejillones would offer clear advantages.
Area 3: This area would pose a challenge for the megaport of Mejillones.

Source: Schellmann, 1996.

or Mejillones, and much Bolivian trade already passes through Arica. The Iquique Free Zone has successfully specialized in imports from Asia and the United States for Bolivia and the MERCOSUR countries. Various other Chilean and Peruvian ports are also moving cargo for Bolivia and even Paraguay. Future growth in mining production in northern Argentina and western Bolivia could result in bigger volumes of cargo for the ports of northern Chile and southern Peru.

One of the obstacles that persists among various member countries of the Andean Community is the restriction of land transport services to hauliers of the individual countries in question, which makes it necessary to transfer cargo from one truck to another at the frontier.

Increasing the trade flows of neighbouring countries which pass through Chilean or Peruvian ports depends above all on improving the land transport infrastructure.

When projecting the impact of such improvements, however, it must be borne in mind that land transport connections are also being improved in Argentina and Brazil. In MERCOSUR, the railways have been privatized, the Paraná waterway is being dredged, and the road infrastructure in general is being improved with private and public investments. At the end of 1999, rail transport of a container between the Argentine cities of Mendoza and Buenos Aires cost US$ 500, whereas transporting a container by road between Mendoza and Valparaíso cost between US$ 500 and US$ 800, depending on the weight (Cámara Marítima de Chile A.G., 1999). In other words, although Mendoza is three times as far from Buenos Aires as it is from Valparaíso, the transport to Buenos Aires costs less.

Sgut (1999) considers that projects for bioceanic corridors turn up on the desks of Ministers of Transport and subsequently of the heads of State themselves, and for political reasons these authorities are not in a position to reject any initiative of this kind outright, even if it is not economically feasible.

To sum up, South America is geographically divided by the second highest mountain range in the world, and this represents a disadvantage for ports on the west coast of that region which aspire to attract cargo from neighbouring countries.

2.   Economic growth and local cargo

The ports on the Pacific coast of South America have some growth potential, based above all on domestic cargo and that of the nearby hinterland. Whether this cargo arrives at its destination by direct services or with a transshipment operation in a foreign country is of little importance for the port itself or for foreign trade.

In the long term, Latin America's economic growth is expected to be above the world average. In the shorter term, however, the prospects do not seem so promising, especially for Ecuador and Colombia. At all events, all the countries need efficient ports for their own trade.

The port operators who are potential investors in the ports being offered under concessions along the coast have also confirmed their interest in local cargo, sometimes called "captive cargo", in which they hope to do good business. The transshipment of containers, however, seems to them to be a risky and volatile business. Nor can they base their multi-million dollar investments on uncertain expectations of bulk cargo coming from neighbouring countries, which may or may not materialize once the land transport connections have been improved.

At present, much of the port traffic on the west coast of South America is due to the domestic trade of each country, and this will probably continue to be so in the future.

VII

Concentration of cargo: a good idea, but where?

1.   By land or by sea

The cargo for a hub port can come from its hinterland, by land transport, or it can arrive by sea for transshipment in the port. We have seen that the possibilities for concentrating cargo by land on the west coast of South America are only limited. On a very long coastline, with cities close to the sea and a hinterland restricted by the Andes, it is difficult to follow the example of Rotterdam, Los Angeles and Hong Kong, or even of São Paulo or Buenos Aires on the east coast of South America.

At present, the main east-west corridor in South America is that connecting Valparaíso, Santiago, Mendoza and Buenos Aires. It is based on the bilateral trade between Argentina and Chile and has to cover a shorter distance than other corridors further north. For the areas of agricultural and mining production in west-

ern Argentina, it may be advantageous to use Chilean ports, especially for trade with the United States. However, there seems little reason to expect these transit flows through Chilean ports to change much in the future compared with their present levels.

The limitations on the concentration of cargo by land do not affect concentration by sea in transshipment centres. The trends observed in maritime transport mean that the percentage of containers transshipped at least once during their voyage is increasing. The question is whether this transshipment is to take place on the coast of South America or outside the region.

In short, it is unlikely that ports on the west coast of South America will be able to concentrate large volumes of cargo transported by land. Containerized cargo leaving by sea may pass through transshipment centres, but these may not necessarily be located in South America.

2. Outside South America

In a study commissioned for the Mejillones port project in northern Chile, it was estimated that there could be as much as 8 million tons of containerized cargo for transshipment at that port, near Antofagasta. Even if that figure is feasible, however, the location does not seem to be appropriate. At the moment, a large transshipment centre on the west coast of South America would not appear to be viable.

On voyages to Europe and the east coast of North America, transshipment in Panama, Jamaica or at Freeport (Bahamas) involves hardly any detour from the main north-south route. These ports have bigger volumes of cargo of their own and are located on the intersections of various east-west and north-south routes.

For voyages to Asia, transshipment in Panama or Los Angeles involves a detour of between zero and five per cent (see table 9 above). Los Angeles/Long Beach is one of the three main container ports in the world. The largest types of ships sail daily from it for Asia, and ships from South America can pick up or discharge cargo to and from North America there.

Among world orders for gantry cranes, there are almost none from ports on the west coast of South America. As at the beginning of 1999, Latin America and the Caribbean together accounted for 12% of world orders for such cranes (a total of 161). This percentage is higher than the figure that would correspond to the region in the light of its port traffic in containers and reflects the modernization processes that were under way in this respect in that year in Central America, the Caribbean and MERCOSUR. On the Pacific coast, only Buenaventura had orders for gantry cranes pending. If suitable investment and expansion programmes were carried out in terms of dredging, provision of cranes and more space for containers at the main ports along the coast, they could probably attract regular services with bigger ships that would eventually not need their own cranes. This would reduce the transport costs for the foreign trade of all the South American countries on the Pacific coast, regardless of whether or not the cargo in question was transshipped in countries outside the region.

The MIT port, on the Caribbean side of Panama, is one of the ports whose transshipment traffic has grown most in recent years, and it is expected to keep on growing. An increase of the volume of such traffic is also expected at Balboa, on the Pacific side, where operations were recently begun with the aim of attracting such traffic, and there are also other projects for the construction of new ports which will compete with Balboa on that side of the country. A concession has also just been granted for the operation of the railway that connects ports on the two sides of Panama, thus making possible transshipment operations between the two oceans. All these advances, together with the great advantage of having the Panama Canal, suggest that the role of Panamanian ports as transshipment centres will continue to grow in importance. These ports are better placed than the South American ports to connect regional markets such as Central America, the Caribbean subregion and the west coast of South America with the main east-west routes.

In short, in view of the tendencies observed in maritime transport and a future situation like that shown in figure 3, it may be expected that there will be an increase in the proportion of container movements by sea that involve one or more transshipment operations. However, these operations will take place at ports which have a bigger volume of traffic and are located closer to the main east-west trade routes, such as Los Angeles, the MIT port in Panama, Kingston (Jamaica) or Freeport (Bahamas).

3. On the west coast of South America

The viability of transshipment operations at some ports on the west coast of South America should not be ruled out altogether: as may be seen from table 6, for example, 6% of the movement of containers at Callao already consists of transshipment operations.

If cabotage services were liberalized and the reservation of cargo for national carriers between various countries on that coast were ended, it would be easier for the international lines themselves to establish their own feeder services. At the same time, such liberalization would also make direct services more efficient, because a line could make better use of its idle capacity by collecting cargo along the coast to help fill its ships.

The investments planned for the next few years in the ports of Colombia, Ecuador, Peru and Chile will make it more feasible to use bigger container ships, even possibly without their own cranes. At present, however, the main Chilean shipping company, Compañía Sudamericana de Vapores (CSAV), has on order several 3,100 TEU ships with their own cranes for handling containers, which indicates that it does not consider that the cranes planned in the ports are adequate. These ships are suitable for long trips, with relatively high volumes of cargo. Hub ports on the west coast of South America would require smaller ships with their own cranes, subsequently developing to serve larger ships with lower costs per TEU and a faster turnaround in port, which would necessarily mean using ships without cranes of their own

and carrying out transshipment operations with the port's cranes.

If some day there is greater concentration of containerized cargo on the west coast of South America, Callao would probably be the most suitable port for becoming a transshipment centre. Although this is unlikely, shipping lines could then have an incentive to stop sailing half-empty to Valparaíso and San Antonio. If Peru's economic growth were as fast or faster than that of Chile, Peru's larger population would make it possible (although not probable) in the long term for Callao to have a bigger volume of port traffic than Valparaíso and San Antonio together. In that case, thanks to economies of scale,

port charges could be lower and importers and exporters would have more frequent sailings on regular services at their disposal. As Callao is north of Chile, the detour from the main east-west route would be smaller.

To sum up, the viability of transshipment activities at some ports on the west coast of South America should not be ruled out altogether. The port with the biggest possibilities of carrying out these operations could be Callao. However, San Antonio and Valparaíso have better possibilities for attracting cargo from a broader hinterland, including Argentina, and could thus continue to justify direct services that do not pass through a transshipment centre in Peru.

VIII

Summary and conclusions

The decisions to use one mode of transport or another, to pass through one port or another, or to use services with or without transshipment are mainly taken on the basis of business considerations. The public sector must concern itself with the location of hub ports, however, because these require both private and public investments and it is the public sector which defines the conditions for private-sector participation in such ports.

A country's external trade is closely linked with its geographical location, the transport services that cover the distances to markets, and the ports through which that trade passes. This gives rise to "an interesting vein for public-private association, since the private sector should be responsible for management and contribute the resources for financing the necessary investments, while the State should establish a transparent legal framework which permits competition, as well as adopting a long-term approach which seamlessly links the port with its environment and its area of influence" (Lagos, 1997). It is therefore crucial to make an in-depth analysis of the types of transport services that the different trade flows require and the present and potential area of influence of the country's ports.

In all the countries on the Pacific coast of South America there are aspirations to develop hub ports, the idea being to concentrate cargo by land and sea from the country itself and its neighbours. These aspirations have led to extensive public investment programmes in port and land transport infrastructure and have also directly influenced the conditions laid down for port privatization processes.

The main conclusion of the present study is that the potential for hub ports on the Pacific coast of South America is very limited. This conclusion is based on an analysis of the trade of the South American countries; an examination of the situation of hub ports that exist in other regions of the world; an analysis of regular sea transport lines and the journey distances and freight rates involved and, finally, a study of the volumes of port traffic and the possibilities of attracting cargo from a broader hinterland. The results of these analyses are as follows:

i) After Australia, South America is the region with the least transshipment traffic. At present, the total movement of containers and the percentage of transshipment traffic within that total are lower than in other regions. South America has a share of only 0.5% in world transshipment traffic.

ii) The fundamental factors for a hub port are its location and the volume of cargo from its hinterland. In order to become a hub port, a port must have ample land transport connections, be located in an area where maritime routes connect or cross, or have high volumes of cargo generated in the area around the port. None of these conditions exist on the west coast of South America to the extent that they exist in the case of other hub ports already operating in the world.

iii) In the South American countries, the regular sea transport services primarily reflect the needs of each country's own trade. Because of the larger volumes of trade through the Atlantic ports of South America, the ports on that coast have more services, with bigger ships

and more frequent sailings, than on the Pacific coast. The greater relative importance of services to Europe from the east coast and to North America from the west coast corresponds to the external trade with those regions.

iv) The shipping services from Argentina, Brazil and Uruguay have advantages over the services sailing from ports on the west coast of South America. The countries on the east coast of South America have twice as many ports, with 56% more regular services that have more frequent departures and use bigger ships, and each regular service moves 35% more containers. The ports on the Atlantic coast move five times more cargo than those on the Pacific, and the sea freight rates are lower. These advantages are the result of the bigger volumes of trade of the MERCOSUR countries and should in no way be interpreted as a criticism of the ports or the port and maritime policies of any South American country. At all events, however, these differences mean that an exporter who has equally easy access to ports on both coasts will find various advantages if he opts for a port on the east coast.

v) Their geographical location gives the Pacific countries an advantage over those on the Atlantic in trade with North America. This is because of the sea distances involved and also the fact that most of the services linking the Pacific coast of South America with Europe and Asia automatically call in at North American ports. A detour via Los Angeles during a voyage from Chile, Colombia, Ecuador or Peru to Japan, Hong Kong or Singapore only involves a 1% to 5% increase in the distance travelled.

vi) Their geographical location is a disadvantage for the Pacific countries in trade with Europe, however. The distances from ports on the same latitude are much shorter in the case of the eastern coast of South America, and moreover the ships do not have to pass through the Panama Canal.

vii) Overall, their geographical location does not give the countries on the Pacific any advantage in trade with Asia. Yokohama (Japan) is closer to the Pacific countries, Hong Kong is the same distance from both coasts of South America, and Singapore is closer to Argentina, Brazil and Uruguay. Maritime services from Santos or Buenos Aires to Asia pass by South Africa and Sri Lanka. Services sailing from the west coast do, however, have the advantage of being able to connect in Los Angeles with services between North America and Asia, thus allowing them to take advantage of the bigger ships and more frequent sailings on that route. In other words, the only advantage of the ports of Chile, Ecuador and Peru in trade with Asia is the possibility of connecting with a transshipment centre outside South America.

viii) The hinterland of the Pacific ports is restricted by the Andes mountains, whose negative impact on bilateral trade is equivalent to an extra distance of several thousand kilometres of flat terrain. Transporting a container from Mendoza in western Argentina to Buenos Aires costs less than transporting it by road to San Antonio (Chile), although Buenos Aires is three times further. At the present time the main Pacific ports register very low percentages of transit traffic, and these trade flows are not expected to change much in the future.

ix) There will be no hub ports in western South America. Even if there were an adequate land transport infrastructure, on the maritime side there are not sufficient reasons to justify a bigger concentration of cargo transported by land in the Pacific ports. Although in general the use of containers and the percentage of transshipment operations is increasing, the transshipment centres for cargo from the western South American countries are outside the region, in the Bahamas, the United States, Jamaica and Panama.

These conclusions must seem pessimistic, and may be a disappointment to those who believe or believed in new possibilities of generating income and employment through the sale of port services to neighbouring countries. In no case, however, are we suggesting that investments should not be made in transport infrastructure or that ports should not continue to be modernized and privatized. On the contrary, the disadvantages described in this article should be seen as an incentive to make renewed efforts to improve transport services.

The idea of offering transport services for neighbouring countries' trade in itself reflects a very positive change of attitude. Up to a few years ago, countries sought to prevent the products of other countries from gaining in competitiveness through the use of their ports. However, it is not possible to forcibly influence the decisions of shipping companies or other suppliers of transport services to concentrate cargo in a particular place. The physical and administrative obstacles that prevent transport companies from finding the most cost-effective solutions for their clients can and must be reduced. Improvements in land transport infrastructure, the liberalization of cabotage, whether by land or sea, and the reduction of delays at border passes would directly benefit those wishing to import from or export to other regions.

Such advances would also benefit trade among the South American countries. Over half of the volume (in tons) of South American imports comes from within Latin America and the Caribbean (ECLAC, 2000). Furthermore, in recent years this intra-regional trade has grown more than interregional commerce. Investments in bioceanic

corridors designed to connect countries with hub ports that concentrate intercontinental cargo would appear to be less promising than investments in intra-regional trade corridors connecting the main economic centres of the region with each other.

In the past, countries competed in external trade by trying to prevent neighbouring countries from gaining benefits by using their ports. Today, however, countries are seeking to compete for the possibility of handling their neighbours' exports or imports. In itself, this competition is positive, but in many cases it has been raised to a political level which has turned simple competition between ports into an international struggle between hypothetical future hub ports. In view of the limited likelihood that the establishment of such ports will be a success on the west coast, perhaps it would be better to seek greater regional coordination of transport policies and infrastructure investments in order to promote integration between the countries on the Atlantic and Pacific coasts of South America.

(Original: Spanish)

Bibliography

American Shipper (2000): 3 January (www.compairdata.com).

Camae (1999): Guayaquil, Ecuador, October.

Cámara Marítima de Chile A.G. (1999): 24 October (camport@entelchile.net).

Cargo Systems (1999): November (www.cargosystems.net).

Comisión de Promoción de Concesiones Privadas (1998): PROMCEPRI: *Concessions of the Peruvian Sea Cargo Terminals*, Lima, June.

Containerisation International (1999): London, Informa Group, December.

Datamar Consultores Asociados (2000): 4 January (lorimer@amcham.com.br).

Drewry Shipping Consultants (1996): *Post-Panamax Containerships*, London.

———(1999): *Container Market Outlook*, London.

ECLAC (Economic Commission for Latin America and the Caribbean) (1999): *Perfil marítimo*, Santiago, Chile, December.

———(2000): *Perfil marítimo*, Santiago, Chile, January (http://www.eclac.cl/espanol/investigacion/transporte/perfil/index.htm).

El Diario (1999): Santiago, Chile, 2 November.

El Mercurio (1998): Santiago, Chile, 30 May.

ENAPU (Empresa Nacional de Puertos) (undated): *Puertos del Perú*, Lima.

Estrategia (1998): Santiago, Chile, 17 June.

Fairplay Ports Guide (1998): London, Fairplay Publications.

Gallup, J. and J. Sachs (1999): *Geography and Economic Development*, CAER II Discussion Paper N° 39, Cambridge, Massachusetts, Harvard Institute for International Development, March.

Hoffmann, J. (1997): An overview of Latin American ports, in Ministerio de Obras Públicas, *Seminario sobre avances en la tecnología portuaria y marítima mundial y su relevancia en Chile*, Santiago, Chile.

———(1998): *Concentration in liner shipping: Its causes and impacts for ports and shipping services in developing regions*, LC/G.2027, Santiago, Chile, ECLAC.

Lagos, R. (1997): *Corredores bioceánicos*, Santiago, Chile, August, *mimeo*.

Lloyds List (1999): London, 28 April.

Radelet, S. and J. Sachs (1998): *Shipping costs, manufactured exports, and economic growth*, paper presented at the Annual Meeting of the American Economic Association, Nashville, Tennessee.

Schellmann, M. (1996): *Megaport of Mejillones*, Preliminary Working Paper, Antofagasta, Chile, Gobierno Regional de Antofagasta, December.

Sgut, M. (1999): *Corredores bioceánicos*, Buenos Aires, *mimeo*, December (martin.sgut@attglobal.net).

Shednet News (1999): 16 January (www.shednet.com).

UNCTAD (United Nations Conference on Trade and Development) (1999): *Review of Maritime Transport 1999*, Geneva.

World Sea Trade Service (1998): *WSTS Review*, vol. 1, Lexington, Massachusetts, The McGraw-Hill Companies.

Zohil, J. and M. Prijon (1999): The MED rule: The interdependence of container throughput and transhipment volumes in the Mediterranean ports, *Maritime Policy & Management*, vol. 26, N° 2, London, Taylor & Francis Ltd.

Zona Franca de Iquique (2000): January (www.zofri.cl).

The tax regime for

micro-enterprises in Cuba

Archibald R. M. Ritter

Department of Economics and
School of International Affairs,
Carleton University,
Ottawa, Canada
archritter@pigeon.carleton.ca

The government of Cuba established a tax regime for micro-enterprises as soon as it legalized the sector in 1993. It was designed to function in a difficult context in which a tax-paying culture did not exist, in which widespread non-compliance was feared, and in which some micro-entrepreneurs' incomes were high. The tax regime included advance monthly lump-sum payments, a 10% maximum amount of total revenues which could be deducted as costs in calculating taxable income, and an escalating tax schedule. This tax regime has a number of weaknesses which make it inequitable, inefficient and ineffective in revenue generation. The analysis of this essay indicates that the actual incidence of the tax is higher than the official tax scale when actual production costs exceed the maximum allowable 10%, even exceeding 100% in some circumstances. The up-front lump-sum tax payments result in marginal tax rates of 100% for initial levels of revenue, followed by a rate of 0% until a level of taxable income is reached where taxes payable according to the tax scale are equal to the initial lump-sum payment. The tax regime discriminates against micro-enterprises which have costs of purchased inputs in excess of 10% of gross revenues. It is regressive in that the tax rates are higher for lower-income micro-enterprises up to a fairly significant level of true net income. The micro-enterprise sector also faces a more onerous tax regime than the foreign and joint venture sector of the economy. From the standpoint of its impact on the efficiency of resource use, the tax regime restricts the entry of new firms into the sector and forces some out of business or underground, thereby reducing production, raising prices, reducing employment and reducing the generation of income. Finally, to the extent that firms "go underground," refrain from starting up, go out of business, or evade taxes, the government loses revenue. A number of modifications in the tax regime are suggested in order to overcome its weaknesses. These changes could help the sector to make a more valuable contribution to Cuba in terms of employment and income generation, increased production at lower real costs and prices, and increased tax collection.

I

Introduction

In September 1993, in the depths of the economic crisis produced mainly by the termination of the special relationship with the former Soviet Union, Cuba liberalized self-employment, legalizing many of the economic activities which Cubans were already practising in clandestine form as part of their personal and family survival strategies. This permitted the surfacing of many activities previously conducted in the underground economy and an explosion of entrepreneurial creativity. The need for a tax regime to tap the incomes generated in the self-employment sector was immediately obvious, and a tax system was adopted in September 1993 and fully implemented in 1996.

The tax regime, as designed and implemented, has indeed extracted revenues from the sector. However, the design of the regime has a number of weaknesses which reduce its fairness, damage the efficiency of resource allocation, and limit its viability. The objective of this study is to analyse the impacts of the tax regime for the micro enterprise sector in terms of i) the equity of income distribution, both within the sector and in society generally; ii) the efficiency of resource use within the sector and in the broader national economy

and society, and iii) the viability, effectiveness and sustainability of the collection of revenues for the government. Following an analysis and evaluation of the micro-enterprise tax regime, a number of observations are made on how it could be modified in order to achieve more fully the objectives of fairness, efficiency and viability.

Section II gives a review of the emergence of the micro-enterprise sector, its structure and character. In section III, the regulatory and policy environment within which it operates is sketched briefly. The nature of the micro-enterprise tax regime and the rationale for its design are presented in section IV, together with the main quantitative analysis of the regime and its functioning, while in section V the implications of the tax regime in terms of equity, efficiency and viability are analysed at the levels of the individual micro-enterprise, the micro-enterprise sector, and the broader economy and society. Section VI then presents a number of specific recommendations for the improvement of the regime from the standpoint of the three key criteria, and finally, in the last section, the basic argument is summed up and the main conclusions are reiterated.

II

The micro-enterprise sector

Despite the nationalization of small and medium-scale enterprises in the Revolutionary Offensive of March 1968, a small self-employment sector continued to exist, mainly in the areas of personal services, transport, artisanal activities, and various types of manufacturing. In 1988, the number of officially registered non-agricultural micro-enterprises was placed at 28,600, with 12,800 employees, accounting for 1.8% of total employment (CEE, 1988). The underground economy also included many micro-enterprise activities, and it continued to operate on a large scale throughout the 1968-1993 period, although its size is virtually impossible to ascertain. In the 1989-1993 period, the volume of clandestine self-employment increased rapidly as people's family

survival strategies required that they earn real incomes beyond those provided by employment in the State sector, since the latter were increasingly insufficient to acquire the daily necessities which were no longer available through the rationing system (Ritter, 1998a, pp. 74-76).

The liberalization of self-employment on 8 September 1993 permitted many micro-enterprises to surface from the underground economy, and permitted new ones to be legally established. Decree-Law 141 and the accompanying resolution legalized 117 types of activities in six general areas, including transportation, house repair, agriculture-related areas, family and personal services, housing services and other activities. By early 1997, a total of 157 types of activities had been legal-

ized, including "gastronomic services" which encompassed street vending and private restaurants or "paladares". The legislation of 1993 limited self-employment to retirees, housewives and laid-off workers, but excluded professionals and enterprise managers. In time, professionals who were declared redundant in their professional activities were permitted to enter one of the self-employment categories, but not as self-employed professionals. In order to enter self-employment, regularly-employed workers in the State sector also required the permission of their work centres. A variety of other restrictions on micro-enterprise activity were imposed as well, as discussed below.

The number of registered and legal micro-enterprises rose rapidly, reaching 169,098 by December 1994 and peaking at 208,786 in December 1995 (table 1). The real level of employment was higher than the number of micro-enterprises, as family members were frequently employed, and others often provided some of the inputs needed by registered micro-enterprises. An estimate by the United Nations Economic Commission for Latin America and the Caribbean (ECLAC, Subregional Headquarters in Mexico, 1997) placed the total volume of private sector employment at 400,000 or 9.4% of the total in 1996, but the number may have been even higher.

Indeed, if one includes informal non-registered part-time self-employment, the number of self-employed could reach 30% or more of the labour force. This is because virtually all families in Cuba, except those re-

ceiving remittances from relatives or having access to dollars in other ways, such as foreign travel, must supplement their State-sector wage incomes, which are insufficient for necessary purchases from the higher-priced agricultural markets and dollar stores, with additional incomes from a secondary activity.

Registered self-employment and all the other unregistered self-employment activities have made a valuable contribution to the Cuban economy and people. They have created employment and helped families earn the incomes necessary for their survival. Although some self-employment activities have generated high incomes, the vast majority provide important but modest income supplements. These self-employment activities constitute a massive school for entrepreneurship. They also provide basic services and goods to satisfy essential needs of virtually all Cubans.

The sector also generates large volumes of revenues for the State through the taxation system and the imposition of fines, which are onerous and apparently frequent. The monthly tax payments made by one dollar-sector restaurant amount to $US 520,[1] or almost 50 times the average monthly salary in the State sector (214 pesos in 1997), at the exchange rate applicable to Cuban citizens, namely $US 1.00 = 20 pesos.

---

[1] US$ 300 for the basic monthly payment, plus US$ 100 for an alcohol licence, plus US$ 60 for each of the minimum two family employees required by law.

TABLE 1

Cuba: Registered micro-enterprises
(Number)

	December 1994	December 1995	March 1996	January 1998
Approved and registered	169 098	208 786	206 824	159 506
Applications for registration	248 552	390 759	439 268	(268 295)
Applications rejected	10 675	11 519	12 665	(108 789)
Applications withdrawn	65 586	148 491	195 023	
Applications being processed	4 193	21 963	8 789	

Source: Ministerio de Trabajo y Seguridad Social de Cuba, and ECLAC, Subregional Headquarters in Mexico (1997).

TABLE 2

Cuba: Taxes on "gastronomic services"
(Monthly lump-sum payments)

	Local economy (pesos)	Tourism sector (dollars)
Sale of food and beverages ("al detalle")	100.00	—
Sale of home-made food	200.00	100.00
Private restaurants ("paladares")	400.00	300.00
Sale of alcoholic beverages in private restaurants	100.00	100.00
Tax for the minimum of two family-member employees required in private restaurants	120.00	120.00

Source: Ministerio de Trabajo y Seguridad Social/Ministerio de Finanzas y Precios, 1995.

III

The policy environment

The tax regime for micro-enterprise is but one element of the general policy environment within which micro-enterprises operate.

The first element in the policy framework is the broad range of detailed regulations which affect their size and operation. The regulatory environment includes the following general features as summarized in Decreto-Ley 174 (*Gaceta Oficial*, 1997):

– prohibition of "intermediaries" of any sort: producers must sell their own output, sellers must produce what they sell with no specialized retailers, no wholesalers, and no producers selling their output to others for sale;
– prohibition of self-employment in all professional areas;
– restrictions on access to markets, with prohibition of sales to all State entities, except where specifically permitted;
– prohibition of publicity;
– no access to credit, foreign exchange at the official parity rate, or directly-imported inputs;
– prohibition of hiring labour outside the family;
– prohibition of associations of own-account workers;
– additional specific limitations on private restaurants, including a 12 seat maximum;
– inputs can be purchased only in dollar stores and agricultural markets, where prices are several times higher than input prices in the State sector, and limitations on the products available;
– additional specific limitations on the sale of food products in the streets, such as prohibition of the use of benches, seats or tables;
– prohibition of the use of any location for a micro-enterprise activity except one's own home or, in some cases, a rented market stall.

There are also a variety of regulations on health, sanitary and safety conditions for "gastronomic services" and private transportation. Some regulations in these areas are of course necessary. However, the definition of the relevant standards appears to be open-ended, so that whatever an inspector feels is appropriate could be the operational standard. The nature of fines and punishments also appears to permit discretionary punishment for in-

fractions, with the intensity of the punishments being at the discretion of the relevant inspectors (*Gaceta Oficial*, 1997). For the 28 micro-enterprise infractions of an economic operational character, the fines can be significant – up to 1,500 pesos, or about seven times the national average monthly income, although this is perhaps not unduly onerous for the most lucrative types of self-employment. This compares with traffic offenses, which vary from about 2 pesos for an illegal left turn to 10 pesos for a dangerous offence such as running a red light. The suspension of the licence of an offending micro-enterprise for a minimum of two years which is a possibility in many cases is a heavy punishment, as it destroys the livelihood of the offending individual and the relevant family. Some private restaurant owners argue that the licence suspension is in fact permanent. The possible seizure of equipment, instruments, machinery or materials amounts to the confiscation of the accumulated capital of the micro-entrepreneur, modest though it might be. Multiple or repeated infractions receive harsher punishments. If the fine is not paid within 30 days, it is doubled. If it is not paid in 60 days, the relevant authorities are empowered to seize the offender's wage, salary, pension or any other income, then his bank account, and then any movable property.

The regulations are enforced by a corps of inspectors, including those who enforce the economic regulations and others who enforce the health, safety, environmental and labour regulations. The inspectors can levy fines, seize equipment and cancel licences immediately and put any micro-enterprise out of business at any time, apparently with only minimal chances of a successful appeal.

These regulations have a variety of economic consequences. First, they constrain the normal growth and expansion of individual micro-enterprises, condemning them to a small size. The result of this is that the incomes of the entrepreneurs are indeed limited considerably, thus preventing them from providing serious competition to the State-sector enterprises in some areas such as food services, transport and taxis. Second, the micro-enterprises operate inefficiently as a result of the economic regulations. The prohibition of "intermediaries," for example, compels artisans or fabricators of various products not only to retail their products on a day to day

basis, but also to find time to produce the products they sell– and no other products. The prohibition of specialized retailers thus wastes the time of the artisans and reduces the quantity and quality of their production or, more likely, encourages various infractions of the rules in daily operations, despite the efforts of the inspectors to enforce the regulations. The overall result of the regulations is therefore to waste the energies of many of the entrepreneurs, to lower their productivity, and to reduce the quality and quantity of their production. The general impact for Cuba is thus to waste its human energies as well as its capital and material resources.

Another area where public policy appears to be aimed at the containment of parts of the micro-enterprise sector is the competition from State enterprises, which often have structures of input costs which are only a small fraction of those faced by the micro-enterprises (Ritter, 1998a).

The general political environment within which the micro-enterprise sector has operated has been uncertain since its legalization in 1993. The original legislation included a clause which could be used at any time to eliminate parts of the sector (Cuba, Gobierno, 1993). In a speech in April 1997, Castro criticized the reforms of 1993-94, on the grounds that they generated incomes "ten times, twenty times, or thirty times those earned by a worker" and stated that they were implemented: ". . . never imagining that we would have to learn to live with them for a period of time that is difficult to predict, and that depends on many factors" (Castro, 1997). This statement was widely interpreted to mean that the continued existence of the micro-enterprise sector was being questioned.

The micro-enterprise sector is also criticized by the State-controlled press. Articles critical of the sector have appeared every so often (Lee, 1996; Del Barrio, 1998; Mayoral, 1995 and 1997; Ricardo, 1998). The press usually emphasizes the illegalities perpetrated by the sector, the alleged high incomes received, the evasion of taxes, and the clear need for stronger enforcement, greater vigilance and tighter controls.

IV

The tax regime for micro-enterprises

1. Structure of the tax regime

The micro-enterprise tax regime consists of compulsory lump-sum fees paid each month to the tax authority, the Oficina Nacional de Administración Tributaria (ONAT), together with a self-administered correction for the annual tax payments carried out by each micro-enterprise. A key feature of the system is the 10% maximum allowable deduction from taxable income for the purchase of inputs.

The first element in the tax regime is the monthly fixed lump-sum payment, which was first imposed in 1993, when Cuba embarked on the first phase of micro-enterprise liberalization. The Ministries of Finance and Prices and of Labour and Social Security set minimum rates, but the Administrative Councils of the municipal governments are empowered to establish rates above these minimum levels, with the approval of the relevant ministries. The rates can be changed every six months, in January and July. The legislation permits the Councils to raise the rates if they consider that the incomes of the micro-enterprises are "excessive," although what "excessive" means is not defined (Cuba, Gobierno, 1993). Here again, the ambiguous character of the law increases the uncertainty for micro-enterprises. The law permits only increases, not decreases in the tax rates. This also adds to uncertainty and makes the system inflexible for micro-entrepreneurs, who must pay the full fixed lump-sum in both fat and lean months and years. However, the micro-entrepreneur may leave the activity at the beginning of any month, immediately ceasing the monthly payments.

The rates were set at relatively low levels in September 1993, but have been increasing steadily since then. Especially noteworthy is the distinction between those micro-enterprises which operate in the dollar economy and those in the peso economy, the former being taxed at a rate which is 20 times that of the latter (at the exchange rate applicable to Cuban citizens).

The second feature of the tax regime is that at the end of each year, micro-enterprises must pay a tax on their revenues on the basis of an escalating tax schedule. They can, however, deduct the total of the monthly fixed lump-sum payments from the amount of tax owed according to this schedule. The procedure is as follows:

(i)   The micro-enterprises add up their gross revenues;
(ii) They subtract 10% of the gross revenues (20% in the case of private transport) as an allowable deduc-

tion for inputs purchased, in order to arrive at net taxable income;

(iii) They calculate the tax owing according to the scale in table 3. The payments are cumulative, with each component of income falling within each bracket being taxed at the rate for that bracket;

(iv) They deduct the sum of the monthly lump-sum payments from the tax owed according to the tax scale or schedule;

(v) If the amount owed according to the tax scale calculations exceeds the amount already paid through the monthly lump-sum payments, they must pay tax equal to the difference;

(vi) If the amount owed is less than the amount already paid, they do not receive a rebate for the excess tax already paid.

The official tax scale (table 3) includes one scale for income and taxes in national currency and another for income and taxes in US dollars or "convertible pesos." The tax scale applies to income net of the 10% of gross revenues which is the maximum allowable for purchased inputs. The progressiveness of the tax scale for dollar incomes is not too far out of line with international standards. It also seems reasonable from a Cuban perspective. On the other hand, for peso incomes the scale increases from 5% for the first 3,000 pesos to 50% for the tax bracket for incomes exceeding 60,000 pesos. The peso tax scale does appear to be excessive, as the 50% marginal tax rate comes into effect at an annual income of 60,000 pesos, which is only about $US 3,000.

A third feature of the tax system is that a maximum deduction of 10% from gross income is permitted for the purchase of inputs when determining taxable income. In other words, net income for tax purposes or "taxable income" is always considered to be 90% of gross income, regardless of the real value of purchased inputs. The only exception is in transportation, where the maximum deduction for purchased inputs is 20% (ONAT, 1997, p. 6). In this article, this is termed the "10% Maximum Cost Deductibility Rule". This feature of the tax system is problematic in that those micro-enterprises which face high costs for purchased inputs are in effect being taxed on these purchases. This means that the true tax rate on value added by the firm in its actual net revenues can be much higher than the rates shown in table 3, which show the true rates only for a micro-enterprise with actual costs of purchased inputs amounting to 10% of gross revenues.

What proportion of micro-enterprises have net revenue equal to or more than 90% of gross revenue? Perhaps a significant proportion. Many micro-enterprises involve labour-intensive activities with minimal equipment or purchased material inputs. The following micro-enterprises may be of this type: messengers, bicycle guards, child care, building attendant, manicurist, masseuse, shoe repair, domestic service, sports or language instruction, and some repair services. On the other hand, other activities (such as food vending, shoemakers, artisans, flower cultivators and sellers, used book or record sellers) involve the purchase and processing of substantial amounts of material inputs. Some private restaurant operators estimate their input costs as being above 60% (Ritter, 1998b).

TABLE 3

Cuba: Scales of taxes applicable to personal income

	Scale for income in pesos				Scale for income in dollars			
	Tax bracket			Tax rate (%)	Amount of tax	Tax bracket			Tax rate (%)	Amount of tax
(a)	0	–	3,000	5		0	–	2,400	10	
(b)	3,000	–	6,000	10		2,400	–	6,000	12	
(c)	6,000	–	12,000	15		6,000	–	9,600	15	
(d)	12,000	–	18,000	20		9,600	–	13,200	20	
(e)	18,000	–	24,000	25		13,200	–	18,000	25	
(f)	24,000	–	36,000	30		18,000	–	24,000	30	
(g)	36,000	–	48,000	35		24,000	–	36,000	35	
(h)	48,000	–	60,000	40		36,000	–	48,000	40	
(i)	60 000		or more	50		48,000	–	60,000	45	
						60 000		or more	50	

Source: ONAT (1997, p. 1); Ministerio de Finanzas y Precios (1996).

2.   The rationale for the current tax regime

There are three factors which explain the adoption of this tax regime. First, when self-employment was legalized in September 1993, prices, revenues and net profits were often exceedingly high. This was the result of the excess purchasing power in the hands of citizens (because the government was financing a large deficit –about 28% of GDP in 1993– through the creation of money, while prices in the State sector were fixed at low levels) and the limited number of micro-enterprises which initially existed. The imposition of the tax regime and the escalation of tax rates were therefore designed to remove a substantial proportion of this income for equity reasons. It may be noted again, however, that a majority of the micro-enterprises are involved in the provision of simple goods and services for low-income Cubans in the peso economy and generate only modest incomes, albeit probably higher than the average for State employees.

Second, the micro-enterprise tax regime was established at a time when there was not yet a well-established administration for this type of taxation. Nor had there been an established and transparent tax system or a generalized habit of paying taxes, because the tax rates and payments had previously been hidden. Moreover, prior to the legalization of micro-enterprises, many such firms had operated on a clandestine basis, avoiding tax payments and other types of regulations. The tax system which was implemented seems to have been designed to enforce a high level of compliance in a context in which non-compliance (in the underground economy) previously had been the norm and in which there had been no established practice of paying taxes openly.

Third, there was a significant amount of theft of products from the State sector of the economy, especially in the 1990s. A proportion of these found their way to the micro-enterprise sector as production inputs. One of the novel elements of the tax system –the 10% limit on the deduction from taxable income for purchased inputs– appears to have been designed to handle this situation. If it were impossible to know for sure the true value of purchased inputs, it would be risky to permit micro-enterprises to calculate their own input costs for determining taxable income. It was administratively easier to simply declare a maximum of 10% of gross income for purchased inputs for all micro-enterprises, regardless of their true level of such purchases.

A popular view among own-account entrepreneurs is that the tax system is designed to punish them for ideological reasons and ultimately put them out of business.

This is indeed a possibility, and there are statements by the leadership which seem to support this view. The heavy reliance on regulations and the severe punishment for "infractions" also suggests that this is the case. It is not likely that the current micro-enterprise tax regime was designed to kill the micro-enterprise sector, however: in reality, it was probably designed to collect taxes in a difficult environment in which open and transparent taxpaying was not the established practice.

3.   A quantitative analysis of the micro-enterprise tax regime

There are two complications which are encountered when analysing the micro-enterprise tax regime: first, the advance lump-sum tax, and second, the requirement in the income tax act that net income for tax purposes must always be fixed at 90% of gross income. To analyse the implications of these factors for the micro-enterprise tax regime, a step-by step approach is used here. First, a brief mathematical description is given of the relationship between (i) the "true" tax rate and (ii) the value added by the micro-enterprise (or conversely the value of purchased inputs as a percentage of total revenue). The implications of the advance tax payments combined with the tax scale are then analysed, using a graphical presentation.

The relationship between the "true tax rate" which is imposed on the micro-enterprise and the value of the inputs purchased by the micro-enterprise (or its net value added) is analysed below, beginning with a mathematical derivation of that relationship:

Definitions:

c:	Percentage of gross income used for purchasing inputs
GY:	Gross income
$GYk(1-0.1)$:	Officially determined tax brackets:
$GYk(1-c)$:	True value added or net income for tax bracket
NYk:	Taxable income for tax bracket
$TNYk$:	True net income for tax bracket
Tk:	Tax rate for specific tax bracket
Mk:	Amount of tax payment in tax bracket
$NAYk$:	Net after-tax income in tax bracket
$Vk = Mk/TNYk$:	Tax rate as % of true net income in tax bracket

True net income is defined as gross income less the cost of purchased inputs, i.e., net value added:

$$TNYk = GYk * (1\text{-}c) \qquad [1]$$

Taxable income is defined by the tax legislation as gross income less 10% for purchased inputs (regardless of the true value of those purchases): i.e., net value added for tax purposes is always considered to be 90% of gross income:

$$NYk = GYk * (1-0.1) = GYk*0.9 \qquad [2]$$

The amount of the tax payable is therefore the tax rate for the relevant tax bracket times the taxable income for that bracket:

$$Mk = (Tk/100) * NYk \qquad [3]$$

From [2], we have:

$$Gyk = NYk/0.9 \qquad [4]$$

and, substituting [4] in [1],

$$TNY = (NY/0.9) * (1-c) \qquad [5]$$

or, $\qquad TNYk = (NYk/0.9) * (1-c) \qquad [6]$

Thus, for a given value of c there is a corresponding value of $TNYk$ for every value of NYk:

$$Vk = Mk/TNYk \qquad [7]$$

i.e., the actual tax rate is the tax payment divided by the true net income for each tax bracket.

Then, substituting [3] in [7]:

$$Vk = \{(Tk/100) * NYk\} / TNYk \qquad [8]$$

Again, substituting [6] in [8],

$$Vk = \{(Tk/100) * NYk\} / \{(NYk/0.9) * (1-c)\}$$

or, $\qquad Vk = \dfrac{(Tk/100) * 0.9}{(1-c)} \qquad [9]$

Using the relationship of equation [9], the official ONAT tax scale given in table 3 is adjusted so as to represent the true tax rates on true net income. This gives table 4, section A of which shows the revised tax brackets in terms of true net income and the tax rates as proportions of true net income. Then, when different values are assigned to c, this gives different tax rates and brackets. In table 4, the revised scales and tax rates are presented for values of c of 10%, 40%, and 60%.

The rapid escalation of the true marginal tax rate is apparent in table 4. For the case of true net value added of 40% (i.e., $c = 0.6$), the highest tax bracket reaches 112.5%. This occurs at tax bracket (i), at a level of taxable income of 60,000 pesos or 26,667 pesos in terms of true net income.

If the "10% rule" were abolished so that every micro-enterprise could deduct its real costs of purchased inputs from taxable income, then the official tax scale would represent the true burden of the tax system on the micro-entrepreneur. As it stands, however, the official scale understates the true tax burden for all those micro-enterprises with purchased inputs exceeding 10% of gross revenues.

The lump-sum tax payments can be examined using the three cases presented in sections B, C and D of table 4, with purchased inputs accounting for 10%, 40%, and 60% respectively of gross income. These three cases are illustrated in figures 1, 2 and 3. The third case is explained in detail below, with only a quick reference to the other two.

In figure 3, with gross income on the horizontal axis and the percentage of gross revenues on the vertical axis, the horizontal line at the 60% level indicates that 60% of gross income is always expended on the purchase of inputs. The true value added, then, is the 40% of gross revenues lying above the 60% line. The escalating marginal tax rates for the sequential tax brackets are represented by the step-like tax rate function. The area between this function and the 60% line represents the tax payable by the micro-enterprise, while the area above the tax rate function but below the 100% line represents the after-tax net revenues retained by the micro-enterprise. It can be seen that the tax rate function escalates rapidly and exceeds 100% at the last tax bracket (66,667 pesos per year and above). The tax rate function escalates quickly due to the fact that the tax is levied on 90% of the gross revenues, but the micro-enterprise only receives 40% of the gross revenues as net income.

The additional complication arises because the micro-enterprises must pay the lump-sum taxes in advance. We will consider here the example of a private restaurant in the peso economy. For this case, and on an annual basis, the amount of the lump sum tax could be 620 pesos per month, or 7440 pesos per year (400 pesos lump sum plus 100 pesos "alcohol tax" plus 120 pesos for a minimum of two registered workers, or 620 pesos per month) (see table 2). The area representing the lump-sum tax in figure 3 shows that a gross income of 18,600 pesos is required to generate enough net revenues (40% of the gross income) to pay the tax. However, the lump-sum taxes already paid can be deducted from the tax payable according to the tax scale. This means that no additional taxes need to be paid until the total tax pay-

TABLE 4

**Cuba: Calculation of tax rates for effective net incomes
with different levels of purchased inputs (or net value added)**

A. When *c* represents purchased inputs as % of gross income

	Tax brackets as effective net income (*pesos*)		Tax rate for effective net income (%)
	From	To	
(a)		$3{,}000 \times (1\text{-c})^{(-1)}\,0.9$	$0.05 \times (1\text{-c})^{(-1)}\,0.9$
(b)	$3{,}000 \times (1\text{-}c)^{(-1)}\,0.9$	$6{,}000 \times (1\text{-}c)^{(-1)}\,0.9$	$0.10 \times (1\text{-}c)^{(-1)}\,0.9$
(c)	$6{,}000 \times (1\text{-}c)^{(-1)}\,0.9$	$12{,}000 \times (1\text{-}c)^{(-1)}\,0.9$	$0.15 \times (1\text{-}c)^{(-1)}\,0.9$
(d)	$12{,}000 \times (1\text{-}c)^{(-1)}\,0.9$	$18{,}000 \times (1\text{-}c)^{(-1)}\,0.9$	$0.20 \times (1\text{-}c)^{(-1)}\,0.9$
(e)	$18{,}000 \times (1\text{-}c)^{(-1)}\,0.9$	$24{,}000 \times (1\text{-}c)^{(-1)}\,0.9$	$0.25 \times (1\text{-}c)^{(-1)}\,0.9$
(f)	$24{,}000 \times (1\text{-}c)^{(-1)}\,0.9$	$36{,}000 \times (1\text{-}c)^{(-1)}\,0.9$	$0.30 \times (1\text{-}c)^{(-1)}\,0.9$
(g)	$36{,}000 \times (1\text{-}c)^{(-1)}\,0.9$	$48{,}000 \times (1\text{-}c)^{(-1)}\,0.9$	$0.35 \times (1\text{-}c)^{(-1)}\,0.9$
(h)	$48{,}000 \times (1\text{-}c)^{(-1)}\,0.9$	$60{,}000 \times (1\text{-}c)^{(-1)}\,0.9$	$0.40 \times (1\text{-}c)^{(-1)}\,0.9$
(i)	$60{,}000 \times (1\text{-}c)^{(-1)}\,0.9$		$0.50 \times (1\text{-}c)^{(-1)}\,0.9$

B. When *c* = 10%

	Tax brackets as effective net income (*pesos*)		Tax rate (%)
	From	To	
(a)	0	3,000	5
(b)	3,000	6,000	10
(c)	6,000	12,000	15
(d)	12,000	18,000	20
(e)	18,000	24,000	25
(f)	24,000	36,000	30
(g)	36,000	48,000	35
(h)	48,000	60,000	40
(i)	60,000		50

C. When *c* = 40%

	Tax brackets as effective net income (*pesos*)		Tax rate (%)
	From	To	
(a)	0	2,000	7.5
(b)	2,000	4,000	15.0
(c)	4,000	8,000	22.5
(d)	8,000	12,000	30.0
(e)	12,000	16,000	37.5
(f)	16,000	24,000	45.0
(g)	24,000	32,000	52.5
(h)	32,000	40,000	60.0
(i)	40,000		75.0

D. When *c* = 60%

	Tax brackets as effective net income (*pesos*)		Tax rate (%)
	From	To	
(a)	0	1,333	11.3
(b)	1,333	2,667	22.5
(c)	2,667	5,333	33.8
(d)	5,333	8,000	45.0
(e)	8,000	10,667	56.3
(f)	10,667	16,000	67.5
(g)	16,000	21,333	78.8
(h)	21,333	26,667	90.0
(i)	26,667		112.5

Source: Section A is based on equation [9] and table 3. The calculations in sections B, C and D are based on section A.

FIGURE 1

**Effective tax rates with lump-sum payment
(of 620 pesos per month or 7,440 per year) with true net
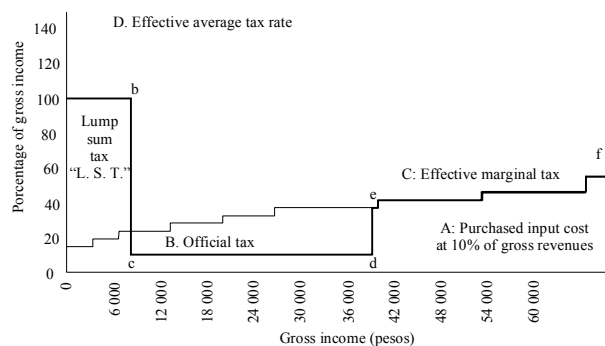value added at 90% of gross revenues**
(c=0.1)

FIGURE 2

**Effective tax rates with lump-sum payment
(of 620 pesos per month or 7,440 per year) and true net
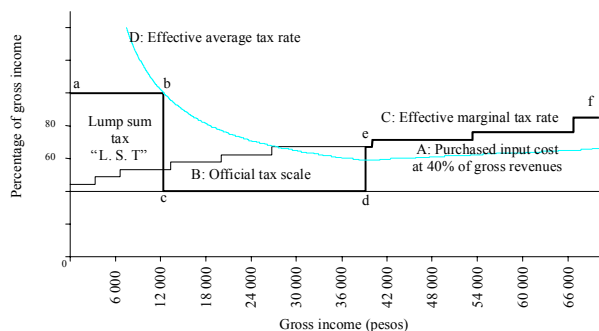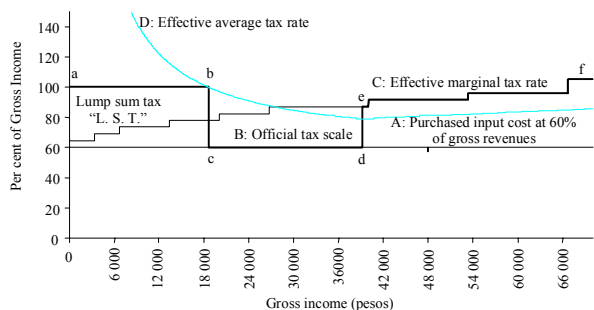value added at 60% of gross revenues**
(c=0.4)

FIGURE 3

**Effective tax rates with lump-sum payment
(of 620 pesos per month or 7,440 per year) and true net
value added at 40% of gross revenues**
(c=0.6)

able according to the scale reaches the lump-sum tax payment. In this case, the level of gross income at which the taxes payable equal the lump-sum payments is 39,743 pesos, at point *e*.

The coexistence of the deductible lump-sum tax with the tax payable according to the tax scale produces a curious pattern of marginal tax rates. Instead of the rates depicted by the step-like tax function, the effective marginal tax rate (MTR) function follows the line *abcdef*. For the first 18,600 pesos of gross income, the MTR on net income is 100% (this is an MTR in the sense that within this range of its gross income, the micro-enterprise must pay 100% of its net income as taxes, on a month by month basis). However, the MTR then falls to 0.0% from *c* to *d* for that portion of gross income on which no additional tax needs to be paid because it has already been paid via the lump-sum tax. Only after an income level of 39,743 pesos is passed, at point *e,* does the tax scale again become effective.

This pattern of marginal tax rates cannot fail to have a variety of impacts on micro-enterprise behaviour. When the lump-sum tax is included in the picture, as shown in figure 3, it can be seen that the net after-tax revenues in this case include only the area above line *bcdef* and below the 100% line. The possibilities for this micro-enterprise to earn sufficient net revenues to survive appear to be limited.

The calculations for the case illustrated in figure 3 are shown in table 5. In this table, column 1 shows the tax brackets defined in terms of gross income, less the allowable 10% deduction for purchased inputs. Column 2 then shows the true gross income applicable to the tax brackets. In column 4, the true net income for each tax bracket is presented, taking into consideration an assumed 60% cost for purchased inputs. With the tax rates of column 5, the tax payments are calculated in column 6, noting that the gross income (less 10% for purchased inputs) is what is actually taxed. The tax payments of column 6 are then presented as proportions of true net income (from column 4) to give the actual or true tax rate (shown in column 10). In column 7, net after-tax income is calculated as true net income less the tax payments made (column 6). The net after-tax income is then calculated in column 8 as a percentage of gross income (from column 2), and next, in column 9, as a percentage of true net income (the complications of the lump-sum payments are not included in this table).

In this case, the escalation of the marginal tax rates as proportions of the true net income is dramatic. With true net income equal to only 40% of gross income, the tax scale reaches high levels quickly. By the eighth tax

TABLE 5

**Cuba: Calculation for determining real tax rates for an
effective net income of 40% of gross income**[a]

(c = 60%)

1 Tax bracket	2 Gross income	3 Taxable income	4 Effective net income	5 Tax rate	6 Amount of tax	7 Net income after tax	8 Net income after tax, as % of gross income	9 Net income after tax, as % of effective net income	10 Effective rate, %
GYk *(1-.1)*	Gyk	Nyk	$TNYk = GYk\,(1\text{-}c)$	Tk	$Mk = TkxNYk$	$NAYk = TNYk\text{–}Mk$	$NAYk/ GYk$	$NAYk/ TNYk$	$Mk/ TNYk$
(a) 0 – 3,000	0 – 3,333	3,000	1,333.3	.05	.05x3,000=150	1,183.3	0.355	0.887	11.3
(b) 3,000 – 6,000	3,333 – 6,667	3,000	1,333.3	.10	.10x3,000=300	1,033.3	0.31	0.775	22.5
(c) 6,000 – 12,000	6,667 – 13,333	6,000	2,666.7	.15	900	1,766.7	0.265	0.663	33.8
(d) 12,000 – 18,000	13,333 – 20,000	6,000	2,666.7	.20	1,200	1,466.7	0.22	0.55	45.0
(e) 18,000 – 24,000	20,000 – 26,667	6,000	2,666.7	.25	1,500	1,166.7	0.175	0.44	56.3
(f) 24,000 – 36,000	26,667 – 40,000	12,000	5,333.3	.30	3,600	1,733.3	0.13	0.325	67.5
(g) 36,000 – 48,000	40,000 – 53,333	12,000	5,333.3	.35	4,200	1,133.3	0.085	0.21	78.8
(h) 48,000 – 60,000	53,333 – 66,667	12,000	5,333.3	.40	4,800	533.3	0.04	0.10	90.0
(i) 60,000 – 100,000	66,667 – 111,111	40,000	17,777.8	.50	20,000	–2,222.2	–0.05	–0.125	112.5

Source: Calculations by the author.

[a] Assuming that $c = 0.6$, i.e., that purchased inputs equal 60% of gross income. For definitions of the terms used, see page 145.

bracket, 96% of every peso earned goes to input-suppliers and the tax authority, and in the last tax bracket the tax payable reaches 112.5% of true net income, so that net after-tax income is negative.

The case of actual input purchases of 10% is illustrated in figure 1. This case presents no complications as the true net income is equal to the taxable income permitted by the official tax scale of the tax legislation, that is, allowing for a 10% deduction from gross income for purchased inputs. For this case, the tax rates applied by the tax scale are in fact the same as the tax rates for true net income. An intermediate case, in which input purchases equal 40% of gross income, is presented in figure 2.

Without presenting other possible cases, it can be concluded that the higher the level of purchased inputs by the micro-enterprise (or in other words, the lower the true net value added by the firm), the more onerous is the true tax burden. The tax rates do not appear abnormal for those enterprises with purchased inputs around 10% (true net income or value added as a proportion of gross income of 90%), and especially for those with even lower relative values of purchased inputs. However, as the level of value added or true net income as a proportion of gross income declines, the average tax rates begin to increase rapidly and ultimately become impossible.

When the lump-sum tax is included in the analysis, the average tax rate (ATR) is high at first, reflecting the advance nature of the payment (the ATR function is measured from the cost line upwards). As can be seen in each of figures 1, 2, and 3, the average tax rate declines up to the point where the official tax rate becomes effective at point *e* in each of the figures. The ATR schedule is especially important as a measure of the tax burden, as it shows what percentage of gross income is paid as taxes (from the cost line upwards). Thus the declining ATR and MTR indicate that the incidence of the tax system is regressive: i.e., the tax is higher for low income earners and declines as income rises up to *e* in the figures.

A final feature of the lump-sum tax, not illustrated in the figures, is that as the advance lump-sum tax payment increases in magnitude, the point at which the official tax scale becomes effective shifts towards the right in figures 2 and 3. This means that as the lump-sum payments increase, the tax system, which is regressive from the beginning, becomes even more so, with higher tax rates imposed on lower income earners.

V

Implications for equity, efficiency and viability

The micro-enterprise tax regime has a variety of impacts in terms of fairness, the efficiency of the economic system and the viability of the tax collection system itself. These impacts are explored here at the levels of the micro-enterprise, the micro-enterprise sector, and Cuba's society and economy more broadly.

1.   The micro-enterprise level

First, the lump-sum payments involve a marginal tax rate of 100% of true net earnings up to the level of income that covers them, at which point the marginal tax rate falls to 0% (point *c* in figures 1, 2 and 3). This 0% rate prevails until gross income reaches the level at which the tax payable under the official tax scale –but deductible in an amount equal to the lump-sum payments– is just equal to the total of the latter. At this point (*e* in the figures) the official tax scale becomes effective again.

When faced with this pattern of marginal tax rates, the micro-enterprise will very likely try to avoid paying the higher tax rates. It will try to locate itself as close as possible to point *d* in figures 1, 2, and 3 (the highest levels of income with a marginal tax rate of 0%) by earning gross income just below that point. This will be especially relevant for those micro-enterprises facing high levels of purchased inputs (the case analysed in figure 3), where the marginal tax rate jumps from 0% to 67.5% at point *e*. It can do this by restricting its output to this level and/or under-declaring income. If the firm restricts its output, society loses the goods and services forgone, the micro-enterprise sector loses income, and the government loses tax revenue. If, on the other hand, the micro-enterprise tries to under-declare its income, the government likewise loses tax revenue.

With the high barrier to entry created by the 100% initial marginal tax rate, some micro-enterprises may refrain from registering and may decide to operate on a clandestine basis. There can be no doubt that this has occurred in Cuba. Numerous tradesmen undertake house repairs; private car owners provide occasional taxi services, despite the risks involved; some households provide unofficial rented accommodation, and many people provide personal services unofficially. In this case, the government loses tax revenue, and indeed it may lose its tax base if the number of micro-enterprises diminishes.

The tax administration is not unaware of the lost revenues and uses inspectors and the "Comités por la Defensa de la Revolución" (CDRs) to try to prevent enterprises from operating in the underground economy. However, clandestine micro-enterprise is widespread and the members of the CDRs (which include most people) are often also involved in the underground economy, so that controlling it is difficult.

Despite the ostensible progressiveness of the official tax scale or schedule, the tax structure is in fact regressive for the lower and middle income micro-enterprises. Because the lump-sum tax creates a 100% marginal tax rate for the first amount of income earned, followed by a 0% marginal rate, relatively lower net incomes in this range are taxed more heavily than higher incomes. The average tax rate declines as income increases up to the level where the official tax scale becomes effective (point *e* in the figures). This result is inequitable, in that lower incomes are taxed more heavily than higher incomes in this range.

2.   The micro-enterprise sectoral level

The tax regime is also inequitable in terms of its impacts *among* micro-enterprises. Because of the "10% maximum cost deductibility rule", firms with higher levels of purchased inputs pay higher marginal and average tax rates in terms of their net income than other firms with the same net incomes but lower levels of purchased inputs. A second element of unfairness is that the tax structure discriminates against new entrants to any part of the sector, because the marginal tax rate is initially 100%, due to the fixed monthly lump-sum tax. New entrants must earn revenues immediately which would permit them to pay these taxes, whereas established micro-enterprises are more likely to have reached income levels which permit them to attain the range of gross income where the marginal tax rate is 0%. Thus they have an advantage which shields them in part from new competition by helping to keep out new firms.

From the standpoint of the micro-enterprise sector, the tax regime also results in a variety of inefficiencies. By creating a high barrier to entry because of the 100% marginal tax rate, the number of new firms entering the sector is reduced. High levels of uncertainty and risk for

new entrants also serve as a barrier keeping out potential new participants. Therefore, there may be less legal competition, higher prices, and lower volumes of production than if these barriers to entry were lower. On the other hand, for some other types of micro-enterprise where entrance into the clandestine economy is easier, the result may be increased competition from tax-evading suppliers of certain goods and services operating at lower costs.

The "10% rule" also has negative effects in terms of efficiency. Because of the high true taxation levels for those micro-enterprises which have higher levels of purchased inputs, the volume of production of goods and services from these firms is probably unduly low. The prices of their outputs also are probably unduly high, reflecting the high tax burden but not necessarily reflecting the real efficiency and true costs of production of such firms.

There is an additional negative effect of the "10% rule." As an economy evolves, there is normally increasing articulation of production among enterprises as they become more specialized and as inter-enterprise trade expands. With the "10% rule", however, there is a disincentive for a micro-enterprise to make purchases from other enterprises, as the tax treatment of such purchases is so harsh. The result is that enterprises try to avoid purchases of inputs from other firms. On an economy-wide basis, this would block the evolution of specialization among firms and inter-enterprise exchange.

3. The implications for society

Taxation of the micro-enterprise sector is necessary from the standpoint of equity. This sector, like all others, must pay its share of the cost of providing public goods and services. It was difficult to know what a fair level of taxation for the sector was, at a time when it was only just getting established. This is because, at least for a transitional period, the incomes earned in the sector were high for a number of reasons mentioned in section IV.2 . The high taxation rates in real terms –in relation to true net incomes in the sector– reflected the desire to tap these high incomes for both revenue-raising and social justice reasons.

From a social perspective, however, it is inequitable to tax those with lower incomes more heavily than those with higher incomes, yet this is precisely what the lump-sum tax does in the lower ranges of micro-enterprise incomes –those ranges where perhaps a large proportion of the firms are operating. It is also inequitable to tax micro-entrepreneurs or anyone else with the same net incomes at different rates, but this is what the 10% rule achieves.

Perhaps one of the most disturbing inequities of the micro-enterprise tax regime is that it is more onerous than the tax regime facing foreign companies operating in joint venture arrangements with Cuban State firms. The contrasting tax regimes are summarized in table 6, from which it is clear that while foreign enterprises in joint venture arrangements receive fairly normal tax treatment from a comparative international perspective, their treatment is much more favourable than that given to domestically-owned and oriented micro-enterprises.

The tax regime for micro-enterprises runs counter to the rationality of resource allocation and consequently lowers living standards. First, the onerous levels of taxation inherent in the tax regime lead enterprises to go out of business, or in some cases to restrict their own production in order to avoid higher tax brackets. Both of these results reduce the volume of goods and services

TABLE 6

Cuba: Comparison of tax regime for Cuban micro-enterprises with that applicable to foreign firms operating under joint ventures

	Micro-enterprise sector	Joint ventures
Effective tax rates	Can exceed 100% of net income	30% of net income (50% for mining and petroleum sectors)
Effective tax base	90% of gross income	Net income after deduction of costs
Possibility of deducting investments	Not deductible from taxable income	Fully deductible from taxable income
Lump-sum taxes	Compulsory advance lump-sum tax payments	None
Return of excess tax paid	No returns	Not applicable
"Tax holidays"	No	Yes
Repatriation of profits	No	Yes

Source: Analysis by the author.

produced in the sector and raise prices. Second, the micro-enterprise tax regime restricts entry into the sector, thereby restricting competition, reducing output and raising prices. Third, the fact that some micro-enterprises go underground or that many self-employed micro-entrepreneurs avoid registering and paying taxes also results in inefficiencies of resource use. Such enterprises must operate on a very small scale and continuously "under cover". This probably lowers the quality of their products and increases their prices in comparison with those firms operating legally (even allowing for the impact of taxes on the prices charged by legal micro-enterprises).

By blocking the entry of new micro-enterprises and promoting the exit of existing ones, the tax regime also reduces productive employment and the generation of income (open unemployment was stated to be 6.8% in 1996, while the open-unemployment equivalent of "underemployment" was estimated by the United Nations Economic Commission for Latin America and the Caribbean at 27.3% in the same year (ECLAC, Subregional Headquarters in Mexico, 1997, p. 378). This latter figure emphasizes the loss to the nation due to labour under-utilization).

One surprising result of the micro-enterprise tax regime is the discrimination inherent in it against domestically-oriented economic activity or domestic "value added". Micro-enterprises have little or no access to imported inputs except those acquired from recycled materials, or purchases from the dollar stores or special State sellers of inputs. The effective exchange rate for them in 1998 was approximately 20 pesos for each US dollar of imported inputs (plus taxes of 140% normally). For enterprises in the State sector, on the other hand, one US dollar of imported inputs costs only 1 peso, when the importation has been approved by the planning authorities. This means that State enterprises can obtain access to imports much more cheaply than micro-enterprises. The result of this is that often micro-enterprises use domestically available inputs to a higher degree than the State enterprises. An example of this occurs with the private restaurants, where close to 100% of all of the inputs are of Cuban origin, in contrast to the State fast-food chains such as "Burgui" (a Burger King look-alike), which import almost everything: tables, potatoes, hamburgers, chicken, trays, paper cups, and many specialized building materials. The current tax regime, in discriminating against micro-enterprises, thus discriminates against domestic value added and in favour of the more import-intensive State firms.

4. Implications for tax system viability

The micro-enterprise tax regime has some important implications for the functioning and viability of the general tax system, since it leads to reduced revenues in comparison with a more normal type of tax regime:

i) Because the true tax rates are so high for micro-enterprises with higher levels of purchased inputs, some firms are put out of business, leading to loss of revenue.
ii) When faced with high real tax rates, some micro-enterprises cease legal operation and enter the underground economy.
iii) Because the true level of the marginal tax rate is so high when it becomes applicable to those micro-enterprises with higher levels of purchased inputs, they have an incentive to under-declare their gross income, leading to revenue losses for the government.
iv) Because of the high lump-sum taxes, when contemplating the establishment of a new micro-enterprise or re-registering an established one, some potential micro-entrepreneurs may enter the clandestine economy or refrain from starting or continuing their business.

The perception that the micro-enterprise tax regime is unfair and excessively onerous has resulted in high levels of non-compliance. Indeed, some micro-enterprises can only survive through non-compliance, especially if they are unable to retreat to the underground economy due to a location or profile incompatible with clandestine operations. Some micro-enterprises may under-declare their incomes.

The result of the character of the tax system and the non-compliance it engenders is that the tax system lacks credibility and respect. Rather than leading to the gradual development of a "tax culture" in which people willingly and honestly pay their taxes, the system has provoked the habit of cheating. To some extent this is part of people's survival strategies in the difficult circumstances of the late 1990s. The nature of the tax regime leads some people to think that tax evasion is not unethical even if it is illegal.

In the long term, it may be difficult to change the current "culture of tax evasion" to one of compliance. This could continue to be a problem even after a reasonable micro-enterprise tax regime has been established.

VI

Policy recommendations

There are a number of changes in the tax regime which, if adopted, could reduce or eliminate some of the more harmful effects of the micro-enterprise tax regime. Some of them are quite simple and could be implemented quickly. Others are more complicated and would take longer to generate positive results.

Recommendation 1:
Establish a normal tax base

The first change might be to establish net income rather than gross income as the tax base. This means abolishing the "10% rule", which limits the allowable deduction from taxable income to a maximum of 10% of gross income. This would improve the equity of the system and end the discrimination against those micro-enterprises with purchased inputs exceeding 10% of gross income. By making the tax regime fairer in this way, non-compliance would be reduced and tax collection facilitated. This change would permit an expansion of output of the micro-enterprises which are more dependent on larger proportions of purchased inputs.

Recommendation 2:
Eliminate the advance lump-sum tax payments

The elimination of the fixed monthly payments would have a variety of positive results. It would remove a barrier to the entry of new enterprises, thereby increasing the number of firms, intensifying competition, increasing output and reducing prices in the sector. This would also, in time, lead to an increase rather than a decrease in tax revenues, which would grow as more micro-enterprises entered the sector and as more micro-enterprises surfaced from the underground economy because of the elimination of the high cost, risk and uncertainty created by the lump-sum taxes. Moreover, elimination of these tax payments would also eliminate some of the more obviously unfair aspects of the tax regime. No longer would the incidence of the tax in the lower and middle income ranges be regressive, and the non-refundability of tax over-payments would no longer exist. Correction of these two inequities would help to establish the credibility of the tax regime, facilitate tax compliance and reduce tax evasion.

Recommendation 3:
Build up the credibility of the tax system

At present, the tax regime is perceived to be unfair and unreasonable, so that cheating and non-compliance seem to be morally acceptable as well as necessary for survival. Any policies which would help to build up the credibility of the tax system would therefore be desirable. Some such changes could include reasonable financial penalties for tax evasion and the refunding to micro-entrepreneurs of overpayments of tax. The return to legality of some economic activities which had previously gone underground or remained in that situation would also help make the system fairer by reducing the number of unregistered micro-enterprises able to completely evade taxes.

Recommendation 4:
Lower the barriers to entry into the micro-enterprise sector

A further set of changes concerns the large number of rules and regulations mentioned in section III. These were designed to contain the sector, to limit incomes in the sector, and perhaps to provide protection to some State enterprises. This policy environment provides an incentive for micro-enterprises to remain underground, with a variety of harmful consequences as discussed earlier. The solution, however, is not to impose still more regulations. Instead, it would probably be wiser to lower the barriers to entry into the legal micro-enterprise sector. This would involve dropping many of the restrictions on micro-enterprises and permitting most applicants for entry into the sector to effectively enter it. It would involve changes in the structure of incentives, which currently encourage micro-enterprises to remain in the underground economy or go out of business.

If the entry barriers are lowered and the incentive structure is changed so as not to penalize legality, the number of micro-enterprises will increase. With increased competition in the sector, production will increase and prices will decline, while the average income level of the micro-entrepreneurs will approach the national average. Employment in the sector should increase, while it should go down in the underground economy. Finally, the tax revenues actually collected by the tax system should increase.

VII

Summary and conclusions

Cuba's tax regime for micro-enterprises has a number of shortcomings which make it inequitable, which damage the efficiency of resource allocation in the economy, and which reduce the system's effectiveness in collecting revenue.

The problems with the tax regime stem mainly from the "10% rule" and the character and size of the lump-sum tax payments which must be made each month regardless of actual income. The analysis made in this article shows that the actual incidence of the tax is much higher than the official tax scale when actual production costs exceed the 10% maximum cost deductibility rule. Indeed, for micro-enterprises with costs of purchased inputs in the 60% range, the true marginal tax rate on true net income (rather than gross income less the maximum allowable 10%) can reach levels in excess of 100%. Moreover, the advance lump-sum tax payments result in what are in effect marginal tax rates of 100% for initial levels of revenue, after which the marginal tax rate falls to 0% until the level of taxable income is reached where taxes payable under the tax scale are equal to the lump-sum payment. This creates a pattern of declining marginal and average tax rates in the lower and middle income ranges, so that the tax in this range is regressive. The tax rates in the official tax scale are also excessive for the vast majority of micro-enterprises which operate in the peso rather than the dollar economy.

From the standpoint of equity, the tax regime discriminates against those micro-enterprises which have costs of purchased inputs in excess of 10% of gross revenues. It is regressive in that the tax rates are higher for lower income micro-enterprises. It discriminates against potential new entrants compared with established firms because the former must pay the advance lump-sum taxes before they have even started to earn revenues. Much of the micro-enterprise sector also faces a more onerous tax regime than foreign firms in joint venture arrangements.

From the standpoint of its impact on the efficiency of resource use in Cuban society, the tax regime has unfortunate results. It restricts the entry of new firms into the sector and puts some firms out of business, thereby reducing production, raising prices to Cuban citizens, reducing employment in the sector and probably reduc-

ing the generation of income. It discriminates against domestic value-added in the micro-enterprise sector which uses inputs almost entirely of domestic origin, in contrast with the lower-taxed State sector which is highly dependent on imported inputs and permits profit repatriation.

Although the tax regime is designed to collect revenue from micro-enterprises, there are a number of effects which actually reduce the volume of taxes collected. First, the lump-sum tax payment arrangement and the high tax rates for those with purchases of inputs exceeding 10% of gross revenues encourage firms to "go underground" or to refrain from formally registering. The size of the clandestine economy is immense, and the loss of revenues comparably large. Thus, the nature of the tax regime has bred a culture of tax evasion rather than of compliance.

Some suggestions for improving the micro-enterprise tax system are presented in this article. First, a normal tax base should be established by eliminating the "10% rule" and basing taxable income on gross revenues less actual costs of production. Second, the advance lump-sum tax payment should be cancelled. Third, compliance with the tax regime in the long term should be fostered by making the regime as fair and reasonable as possible. Fourth, artificial policy-generated entry barriers against new micro-enterprises should be dropped. A more effective tax administration would also be necessary to prevent abuse of an arrangement permitting the deduction of costs from taxable income and the elimination of the advance lump-sum tax payments.

Generally, while the micro-enterprise sector must pay its share of taxes, it should also have a tax regime which is not more onerous than that facing joint ventures and foreign firms and which is also equitable among micro-enterprises themselves. A well-designed micro-enterprise tax regime could help permit the sector to play a more valuable role in the Cuban economy in terms of employment, income generation, and the production of needed goods and services at lower relative prices. It could also generate more revenue than the current system, which tends to drive micro-enterprises into clandestine status and tax evasion.

(Original: English)

Bibliography

Castro, Fidel (1997): Speech of April 4, 1997, *Granma International*, Havana, 23 April.

CEE (State Statistical Committee) (1988): *Anuario estadístico de Cuba*, Havana.

Cuba, Gobierno (1993): *Acerca del ejercicio del trabajo por cuenta propia,* Decreto-Ley N° 141 del Consejo de Estado, Havana, 8 September.

Del Barrio, E. (1998): Sin extremismo, pero con toda firmeza, *Granma*, Havana, 12 November.

ECLAC (Economic Commission for Latin America and the Caribbean), ECLAC Subregional Headquarters in Mexico (1997): *The Cuban Economy in the Nineties: Structural Reform and Economic Performance*, Mexico City.

Gaceta oficial (1997): *De las contravenciones personales de las regulaciones del trabajo por cuenta propia*, Decreto-Ley N° 174, Havana, 30 June.

Lee, S. (1996): Trabajo por cuenta propia: una reflexión necesaria, *Granma*, Havana, 13 September.

Mayoral, M. J. (1995): La ilegalidad no tiene futuro, *Granma*, Havana, 14 February.

———— (1997): Debate sobre el trabajo por cuenta propia en la capital, *Granma*, Havana, 14 April.

Ministerio de Finanzas y Precios (1996): *Declaración jurada. Divisas,* Instrucción N° 11, Havana.

Ministerio de Trabajo y Seguridad Social/Ministerio de Finanzas y Precios (1995): Resolución conjunta N° 4, *Granma*, Havana, 14 June.

ONAT (National Tax Administration Office) (1997): *Declaración jurada. Impuesto sobre ingresos personales moneda nacional*, Havana.

Ricardo, R. (1998): Continúa enfrentamiento por ilegalidades en la vivienda, *Granma*, Havana, 14 February.

Ritter, A. (1998a): Entrepreneurship, micro-enterprise and public policy in Cuba: Promotion, containment or asphyxiation, *Journal of Inter-American Studies and World Affairs*, vol. 40, N° 2, Miami, University of Miami, School of International Studies.

———— (1998b): Interviews with shoemakers, artisan workers, taxi operators, carpenters, house repairmen, "tricicletistas," restaurant operators and other micro-entrepreneurs, Havana, February, May and November, *mimeo*.

Employment and labour
markets in the
Dominican Republic:
an overview of the literature

José R. Sánchez-Fung

Department of Economics,
University of Kent,
United Kingdom
jrs6@ukc.ac.uk

This article reviews the literature on employment, wages and the labour markets in the Dominican Republic. Although this will primarily be of interest to economists, it may also be useful to social scientists in general and to those responsible for formulating policies affecting the functioning of the labour market at both the national and the Latin American regional level. The study begins by presenting the main sources of statistical information on labour, after which it focuses on studies on the overall functioning of the Dominican labour market. It then goes on to deal with studies exploring specific labour markets, such as those for micro-entrepreneurs and women, and the literature on wages. This review reveals that there are few studies carried out with scientific precision and aimed at economic aspects proper: most of the studies are of a descriptive, sociological or multi-disciplinary nature. The article concludes with some proposals for possible lines of future research in this field, especially the analysis of the existing labour statistics and the establishment of a suitable system for their collection and dissemination.

I

Introduction

This study seeks to organize the literature published on the behaviour of the labour factor and its associated elements in the Dominican Republic. A review has therefore been made of the studies on employment and unemployment, wages and the various labour markets. The study will primarily be of interest to economists, but the items covered may also be useful to social scientists and to those responsible for designing policies relating to the functioning of the labour market, both in the Dominican Republic and in Latin America in general.

This study is not entirely in the nature of a critique. It deals mainly with the studies which are most important from the economic point of view, although it also summarizes studies which relate to other disciplines or are of a multi-disciplinary nature. Only the most recent of the quantitative studies are reviewed in detail, since they have the greatest coverage and their data bases can probably still be consulted now.

After studying the labour statistics in the main sources of data for the analysis of labour in the Dominican Republic, the studies analysing total employment are summarized first, followed by analyses of the various specific labour markets. After a review of the main studies on wages, the article concludes with some recommendations on the line that could be taken by future studies in this field.

II

Labour statistics

The sources consulted for studying the question of labour in the Dominican Republic are of different methodological structures and cover different periods. There is no source in that country which has collected information on labour systematically over a really long period of time, although many efforts have been made to achieve that objective.

1. The primary sources

The population censuses carried out by the Oficina Nacional de Estadística (National Statistical Office - ONE) in 1920, 1935, 1950, 1960, 1970, 1981 and 1993 are the starting point *par excellence* for analysing the evolution of the population in general and certain characteristics of the members of the labour market, as well as for designing samples for other studies of a similar nature, such as surveys.[1]

Apart from the censuses, a number of surveys have also been made, among them the population and health survey (ENDESA), a multi-purpose study carried out in 1986, 1991 and 1996 (see ONAPLAN/CESDEM/PROFAMILIA, 1997) which includes some details of the work of those interviewed. The information from the expanded household questionnaire used in this survey has served as the basis for analysing various specific aspects of the Dominican labour market (see Ramírez, 1993).

The main statistics currently available on this market are those from the labour force survey carried out since 1986 by the Central Bank of the Dominican Republic.[2] Although initially this survey was carried out

[1] For a compact and up-to-date summary of the history, methodology and organization of the seven censuses carried out in the country by the ONE, see Nicasio and De La Rosa (1998).
[2] The methodological details of this survey, such as the sample design and the definitions used, may be consulted in BANCENT (1997, pp. 9-24), and Gregory (1997, pp. 1-2).

four times a year, since 1993 it has only been effected annually, for budgetary reaons. The survey is prepared by the Central Bank of the Dominican Republic rather than by the National Statistical Office (ONE), likewise mainly for budgetary and human resources reasons. The same thing has happened in the past with other surveys, such as the household surveys, and with the price and national accounts statistics.

The Central Bank's labour force surveys collect information on the number of persons employed and unemployed, average incomes, and hours worked, by income deciles. As from 1996, their scope has been expanded to collect information on the size of the firms that persons work for and on other types of income they receive, which facilitates the classification of employed persons according to whether they work in the formal or informal sector of the economy. The survey also collects information on income from remittances sent by other family members and retirement payments.

Among the main studies which have been made only once is the ENMO-9[3] nationwide labour survey (see FUNDAPEC, 1992) which was limited to the collection of information on firms in free zones, tourism and agro-industries and on students in the fourth year of the *bachillerato* course and technical training courses. For these sectors, this survey recorded: i) the characteristics of the labour market; ii) the stock of labour; iii) future labour needs, including training needs;[4] iv) the views and expectations of students, workers and entrepreneurs on the situation and prospects of the labour market; and v) the development of free zones, agro-industry and tourism and the relation between these sectors and the development of the economy as a whole.

The National Statistical Office (ONE) recently made a census of public employees in the Dominican Republic (ONE, 1998). According to this census, these employees represent 11.64% of the economically active population and total 303,478 persons, 41.1% of whom are women. Furthermore, almost half (40.8%) are aged between 35 and 49. One of the most interesting figures given by the census is that only about one-fifth of all public employees have university degrees (figure 1).

---

[3] This was made in 1991 as part of the technical cooperation programme to strengthen the technical and vocational education system in the Dominican Republic carried out by the APEC Educational Credit Foundation Inc. (FUNDAPEC) with resources from the Inter-American Development Bank.

[4] The training needs were estimated on the basis of forecasts of the number of workers who will enter the economic sectors studied each year.

FIGURE 1
Dominican Republic: Public servants, by level of education, 1997
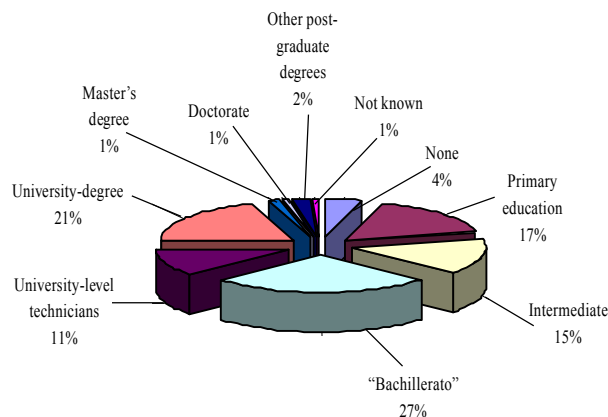
Source: National Statistical Office (ONE), 1998.

2. Primary sources of more limited scope

Del Rosario and Hidalgo (1986) developed and applied to the case of the Dominican Republic a methodology for calculating nominal and real wage indexes. Unfortunately, the collection of these indexes was not continued.

With regard to the rural sector, the labour characteristics were analysed on the basis of the survey prepared in 1989 by the Agricultural Studies Unit of the National Council for Agriculture (Santana, 1990).

The studies carried out by the Micro-enterprise Finance Fund (FONDOMICRO) were very useful for the analysis of micro- and small-scale enterprises, since they contain, among other data, statistics on labour in that sector (Cabal, 1992 and 1993; Moya and Ortiz, 1994 and 1995; Ortiz, 1996 and 1998).

Other interesting sources of information were the statistics on different aspects of trade union activities in the country (ONE, 1993) and the *Boletín de estadísticas laborales* of the Ministry of Labour (see Dominican Republic, Secretaría de Estado de Trabajo, 1999), which contains statistics prepared both by that Ministry and by other institutions.

The *Boletín de estadísticas laborales*, of which four issues have already appeared, provides detailed information on the registration of companies, new jobs generated, by sex and wage, branch of activity, size of firm and production regime, and casual jobs generated each year.

This information is collected by the Ministry of Labour through Form DGT-3, "Permanent staff payroll", and DGT-4, "Changes in permanent staff or payroll", which must be completed, according to the Labour Code, by all industrial, commercial and services establishments registered in the Labour Department of that Ministry.

Although the amount of data available in the bulletins of the Ministry of Labour is not sufficient to make econometric analyses based on time series, it would be possible to gain access to the original micro-sources and process them to carry out cross-sectional or longitudinal studies. It would also be interesting in the future to compare the statistics generated by this Ministry with those of the labour force survey carried out by the Central Bank.

The national survey of political and democratic culture (DEMOS) could be used as a source of data for the analysis of labour relations (see Duarte Brea and Tejada, 1998), while the national survey of young people between 15 and 24 (ENJOVEN-92) provides data on that

group's participation in the labour market (Tejada, Herold and Morris, 1998).

There are also sources of statistical information on two of the most dynamic sectors in the Dominican economy: the free zones and tourism. The periodical publications of the National Council of Export Free Zones provides information on the free zones and on employment in them.

For historical information on employment in the tourism sector, a useful source is *Turismo en cifras* (Dominican Republic, Secretaría de Estado de Turismo, 1996).

Other surveys relating to the labour factor which have been carried out in the country are the survey made by the National Planning Office (see ONAPLAN, 1968), that of the Central Bank of the Dominican Republic, the National Statistical Office and the United States Agency for International Development (BANCENT/ONE/USAID, 1968), and the survey carried out by the latter agency and the Ministry of Agriculture (USAID/Dominican Republic, Secretaría de Estado de Agricultura, 1977).[5]

III

Global analysis of the labour market

Most of the studies on the labour market in the Dominican Republic are based on the analysis of surveys, mostly without continuity, thus limiting the possibility of making systematic historical comparisons of the evolution of various variables or ratios.

For example, figure 2 shows the level of unemployment in the Dominican Republic over the period from 1968 to 1998, according to estimates obtained from a number of different sources.[6] This phenomenon apparently persisted throughout the period in question, at levels ranging from approximately 15% to 28%, without displaying any marked trends over time or volatility.

FIGURE 2
**Dominican Republic:
Unemployment, 1968-1998**
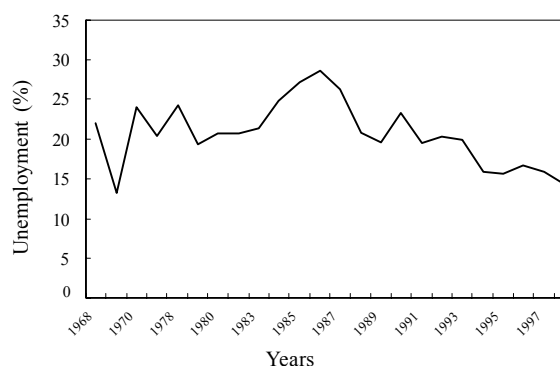
Source: For the years 1968-1980 and 1982-1983: García and Valdivia (1985), table 24, p. 65; for the years 1981 and 1984-1990: Ceara (1992), table 1, p. 14; for the period from 1991 to 1998: data from the Central Bank of the Dominican Republic.

[5] In addition to the sources mentioned, there are also secondary sources of statistics which have been or could be used to study various labour-related elements in the Dominican Republic. Among them are Ramírez, Tatis and Germán (1983) and Quiterio and Polanco (1990), who prepared population and labour projections for various periods, based mainly on population censuses. Other important sources are the time series on the minimum wage in nominal and real terms in the public and private sectors and the statistics and indicators on staff employed by the central government, collected by Adolfo Martí (1997, pp. 339-347). Finally, the *Bulletin of Labour Statistics* published by the ILO (see, for example, ILO, 1998) is another important source of data.

[6] Not all the data in figure 2 come from comparable sources, so they should be interpreted with caution.

Of the many employment studies made in the country, only a few can be considered sufficiently exact and therefore useful for reference purposes. This is true of studies made one or more decades ago.

The report entitled *Generación de empleo productivo y crecimiento económico: el caso de la República Dominicana* (International Labour Office, 1975) may be considered the most complete and frequently cited in its class; it includes both microeconomic and macroeconomic analyses.[7] After an opening section giving a summary and conclusions, the main body of the text is divided into five major parts: unemployment and poverty: the present situation; recent trends in the supply and demand of labour; details of employment in three basic sectors, and institutional factors. There are also several annexes.

One of the main objectives of the ILO study is to estimate the impact of rapid economic growth on employment and, on the basis of this relationship, to analyse how development could result in an increase in employment. It seeks to do this by entering into the fields of both positive and normative economics.

A shortcoming of the study is that the large number of statistical analyses presented, on which most of its conclusions are based, are basically of a descriptive nature. The study does not apply econometric techniques to particular hypotheses that could bear out its conclusions scientifically. Moreover, as the study itself warns, the figures used must be taken with caution.

One of the main conclusions of the analysis is that since 1969 the Dominican economy has been growing at the extraordinarily rapid rate of 10.9% per year, yet this rapid economic expansion does not seem to have reduced unemployment in Santo Domingo (ILO, p. 5). As we shall see later, this situation still persists and poses difficult questions to the government authorities, economists and social scientists in general.

García and Valdivia (1985) made a general analysis, sponsored by the ILO, of the situation of the Dominican economy in 1980-1983. This analysis deals with the subject of employment in some detail: specifically, the authors analyse the process of adjustment of the Dominican economy to the 1980-1983 international crisis and its repercussions on the labour market. The section entitled "Employment and incomes" is of particular interest for the purposes of the present study. In it, the authors analyse the growth of the labour force, employment and unemployment in urban areas, the medium-term evolution of real wages, and the adjustment of the labour market in 1980-1983.

Like the ILO, García and Valdivia (1985, pp. 67-68) note that a structural feature which is present even in the most dynamic periods of the Dominican economy is the tendency towards marked under-utilization of the labour force, both in rural and urban areas, as reflected in the high rates of open unemployment and under-employment detected by a number of sources. Thus, the 1980-1983 crisis came on top of tendencies which were already deeply rooted in the development of the Dominican Republic, further accentuating the serious problems that the country was already suffering.

These authors also point out the tendency of real wages to go down during the period from 1974 to 1983, noting that this decline occurred in spite of the subsidies applied to various components of the consumer price index during much of the period in question. Their study also shows that the differences in income between Santo Domingo and the other urban centres of the country tended to increase in this period.

More recently, Gregory made a study of the Dominican labour market which was based mainly on the labour force survey prepared by the Central Bank, although he did not limit himself to that source (Gregory, 1997).[8] In it, he analyses both the structure and the results of the Central Bank surveys. After examining the methodology of those surveys, he comes to the conclusion that the high rates of unemployment they report are due at least in part to the measure of unemployment used: they classify as unemployed persons who have not been actively seeking work during a reasonable period of time, so that a person who has been passively looking for work for several months can thus be classified as unemployed.

Table 6 of Gregory's study presents various different scenarios, based on the statements of unemployed persons about their search for work. The difference between the measurements is considerable. On the basis of

---

[7] A version of this report was published by the National Planning Office (ONAPLAN, 1974).

[8] It is interesting to note that Gregory's study was inspired by Arnold Harberger's comments (Harberger, 1997, pp. 12-14) on the economic growth and level of unemployment of the Dominican economy. Harberger noted that the country's healthy growth rates were not in keeping with the high level of unemployment registered in the Central Bank surveys. Similar observations were made in such studies as International Labour Office (1975), García and Valdivia (1985), Cabral (1975), Duarte (1983), Lozano (1987), Liz (1993), Weller (1993 and 1998), Alemán (1994) and Vargas (1995). In studies such as Alemán (1996a and b), Aristy (1996) and Santana (1996), specific strategies are proposed for increasing the capacity of the Dominican economy to generate employment, in view of the apparent persistence of the situation described by Harberger and many other authors.

the data for the expanded sample used in the Central Bank survey for April 1997, the rate of unemployment was 15.7% if all those declaring that they would accept work if offered were considered as unemployed, but only 6.5% if based on those actively seeking work (less than a month looking for work).

In the past, estimates of unemployment had also been made on the basis of the length of time seeking work. García and Valdivia (1985, tables 24 and 25) adjusted a series for the unemployment variable obtained on the basis of several surveys. As was to be expected, the measures of unemployment varied when all those who declared that they had been seeking work for a year or more during the reference period were excluded from the open unemployed and the economically active population. Thus, the unemployment rate went down from 24% to 20.8% (adjusted) in 1970 and from 21.4% to 15% (adjusted) in 1983. It would have been interesting if these authors had also calculated unemployment rates based on shorter periods of seeking work, as the Central Bank did.

From the studies reviewed so far it may be inferred that what is known of the Dominican labour market is largely uncertain. The main studies (International Labour Office, 1975; García and Valdivia, 1985; Gregory, 1997) base their analyses on statistics which are not completely reliable, as well as not using strictly scientific methods which take the fullest advantage of the available data.

Márquez (1998), for example, notes the difference in magnitude between unemployment in Latin America on the one hand and the Caribbean on the other, the level being historically lower in the former. He points out, however, that recent studies such as those by Rama (1995) for Jamaica and Gregory (1997) for the Dominican Republic indicate the possibility that the high rates of unemployment registered in those countries may be affected by measurement problems.

The primary sources of statistics for the Dominican Republic have rarely achieved continuity, and so far only one of them - the Central Bank's labour survey - is specifically designed to collect information on the labour market.

What can we say about the global behaviour of the Dominican labour market? In the words of Gregory (1997) "The general consensus about employment conditions in the Dominican Republic is that they are terrible. Unemployment affects between 20% and 60% of the labour force, with a modal value close to 30%. Continual reference to unemployment figures at these levels over the last 30 years or more has become a dogmatic routine. From this point of view, however, there is a marked conceptual difference between the views of the public authorities and those of economists specializing in labour issues as to what constitutes unemployment. As far as we could gather from conversations and other sources, the view of the public authorities is that anyone who is not working for a modern firm or for the government is unemployed. Obviously, this is very far removed from the standard definition of unemployment, which is considered to be the situation of individuals who have no paid employment and are *actively* looking for work" (Gregory, 1997, p. 33).

IV

Specific labour markets

1. Micro-enterprises

Some characteristics of employment in the micro-enterprise/small enterprise sector –one of the most dynamic sectors of the Dominican economy in the recent past– have been analysed in detail in various studies. The participation of women in this sector, for example, was studied by Cely (1993) on the basis of the first national survey of micro-enterprises and small enterprises made by FONDOMICRO in March 1992.

Cely analyses specific characteristics of these enterprises, such as the kind of workers employed in them, their structure, the activities they are engaged in, and their methods of finance. Among the most interesting findings of the study is the fact that at the time the survey was made around 45.8% of the enterprises in the sector belonged to women, although they only generated 33.4% of total employment in it.

Cabal (1996), likewise basing his research on the surveys made by FONDOMICRO, analyses the changes registered in the size of employment in the micro-enterprise/ small enterprise sector between 1992 and 1995. He defines the change in aggregate employment in the sector over a given period of time as the result of two basic

components, i.e., the net change in employment due to enterprises that start up and those that close down, and the change in employment due to the net expansion of the enterprises that survive. This definition is expressed in equation [1]:

$$\frac{\Delta e}{e_t} = \left[\frac{en}{e_t} - \frac{em}{e_t}\right] + \left[\frac{\Delta ep^x}{e_t} - \frac{\Delta ep^c}{e_t}\right] \qquad [1]$$

where *en* represents the number of jobs generated by new enterprises, *em* the jobs lost due to the closure of firms, Δep^x the increase in employment due to firms that expand, Δep^c the change in employment due to firms that contract, and e_t is total employment in the sector in period *t*.

On the basis of formula [1], Cabal estimates that the net growth of employment in the micro-enterprise/small enterprise sector in the periods 1992-1993, 1993-1994 and 1994-1995 was 8%, 1.5% and 0.6% respectively. The results of the study seem to indicate that employment in the sector has grown most when the economy has grown least, and vice versa: i.e., the variable displays a counter-cyclical form of behaviour (figure 3). This conclusion should be taken with caution, however, as it is based on the analysis of only a few observations.

Cely (1996) also used formula [1], but applied it to micro-enterprises and small enterprises belonging to women in the 1992-1995 period. The results obtained were different, as employment in these enterprises belonging to women went down by 1.4%, 0.9% and 8.8% respectively in the periods 1992-1993, 1993-1994 and 1994-1995, whereas in enterprises belonging to men net employment *increased* by 16.5%, 8.7% and 3.9%.[9]

More recently, Ortiz (1998) explored some characteristics of micro-enterprises and their labour force. According to this study, in 1998 the workers in such enterprises represented 28.4% of the economically active population (i.e., 955,683 workers), with an average age of 28 and a balanced breakdown by sex. This study also reports that, on average, the employees of the micro-enterprises and small enterprises surveyed have a higher academic level than the owners.
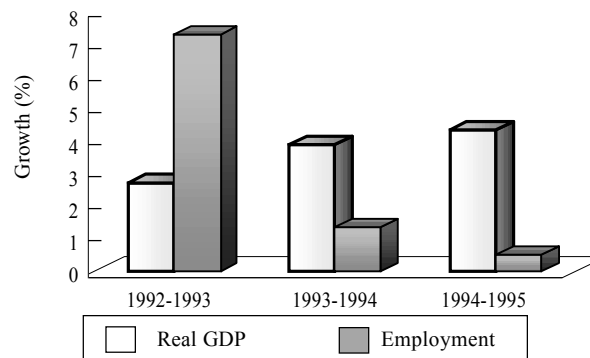
2. Women

An important starting point for analysing women's employment in the Dominican Republic is the bibliogra-

FIGURE 3

Dominican Republic: Growth of the economy and of employment in micro-enterprises and small enterprises, 1993-1995

Source: Central Bank of the Dominican Republic, and Cabal (1996).

phy prepared by Sherrie Baver (The CUNY Dominican Studies Institute, 1997, pp. 33-40), which lists the main studies, researchers and research centres which have analysed the integration of women into the Dominican labour market. The study by Jean Weisman (The CUNY Dominican Studies Institute, 1997, pp. 41-56) also analyses important aspects of women's employment in the country.

Gurak, Kritz, Mota and Ortega (1979) present a study of women's employment in the Dominican Republic which focuses in particular on the important effect of fertility on women's participation in the labour market. The study shows that, among married women, fertility is strongly related with education and employment: it is negative with respect to the former and ambiguous with respect to the latter. For the universe of employed women, the study finds that the probability of employment is positively related with the education of both the woman herself and her parents and negatively related with age and married status. Education, however, is negatively related with work in the services sector in general and domestic service in particular. In the case of non-working women, both future intention to work and the fact of seeking work are positively related with the education of the father and negatively with the existence of a sentimental union.

In *Forum,* No. 21 (Moya, ed., 1986), some aspects of the socio-economic conditions of working women in the Dominican Republic are analysed from a basically sociological standpoint. Among the studies contained in this compendium is a set of statistics presented by Clara Báez (pp. 49-55) on the participation of Dominican women in the labour market. These statistics indicate that

[9] Cely (1996, chapter 5) gives a detailed explanation of the behaviour of each of the components involved in the calculation of the variation in net employment in micro-enterprises and small enterprises, by sex, in the periods in question.

the largest sector of female employment in the country is domestic service. Women are at a disadvantage compared with men as regards wages and status, have little participation in the rural agricultural labour market, but are predominant in the rural non-agricultural labour market.

Forum, No. 21 also contains a study by François Pou (pp. 83-107) which analyses the national survey of rural women carried out between December 1984 and January 1985 by the Female Action Research Centre (CIPAF). Pou concludes that in the rural sector women tend to engage primarily in home-related activities and only to a lesser extent offer their services on the labour market.

The studies prepared by Báez and Matrillé (1994) and by Martínez (1994) analyse some aspects of the labour market for women trained in non-traditional occupations[10], on the basis of information generated from individual surveys, especially for the Los Mina Norte sector of the city of Santo Domingo.

Martínez notes that, in spite of the material and social barriers against the participation of women in occupations traditionally reserved for men, women are nevertheless willing to do this kind of work if necessary. For such participation to be effective, however, it is necessary to change the deeply rooted beliefs of women and of society as a whole about the masculine nature of the occupations in question. Báez and Matrillé, for their part, advocate more training of women in non-traditional occupations, which should stimulate increased participation by them in such trades.

3.   Work by minors

Work by minors has been studied by various authors, including Duarte (1979), Duarte, Gómez and Ariza (1991), Lizardo (1994), Silié (1996) and Alemán (1996a). Alemán gives a summary of the main studies which have been carried out, analyses some particular aspects, including the change in the socio-economic structure of households, and notes that the increase in the participation of minors in the labour market has been accompanied by a reduction in the participation of adults (see Alemán, 1996a, table 19, p. 65).[11]

Duarte, Gómez and Ariza (1991) present statistics showing that the main cause of work by minors is the precarious situation of their homes rather than the wishes of their parents (see table 4.5 on page 131 of the study in question).

4.   Free zones

A number of studies have investigated the question of work in the export free zones. Dauhajre, Riley, Mena and Guerrero (1989, pp. 67-109)[12] have analysed the economic impact of free zones on employment in the Dominican Republic. Using a descriptive analysis of the statistics on the direct and indirect employment generated and the wages paid, the study in question also presents econometric estimates according to which a devaluation of the official exchange rate would have a positive impact on employment in the sector, whereas an increase in the minimum wage would have the opposite effect.

Gregory (1997, pp. 24-25), however, notes that "Firms in the free zones also report incomes that amply exceed the minimum wages laid down for the sector. The average monthly income of women –RD$ 2,590 in the first half of 1996– is 54% higher than the legal minimum wage, while in the case of men the corresponding monthly income –RD$ 3,624– is more than double the minimum".

Taking a multidisciplinary approach, Pérez Sáinz (1995a) makes a comparative analysis of the way governments, entrepreneurs and trade unionists interact with regard to labour problems in the assembly industries and free zones of Costa Rica and the Dominican Republic. The same author (Pérez Sáinz, 1995b) also compares various factors concerning assembly industries and free zones in Central America and the Dominican Republic: specifically, aspects connected with the process of generation of employment by those activities.

5.   Migrants and Haitian labour

One of the features of the Dominican labour market has been migration, the most important component of which is the migration of Dominicans to the United States, generally through Puerto Rico.

With regard to this subject, Rodríguez (1999) notes that "According to figures of the United States Immigration and Naturalization Service, the Dominican Republic, which had a little over seven million inhabitants in

---

[10] According to the criterion adopted in the studies in question, non-traditional occupations for women would include, for example, working as bricklayers, cabinet-makers or auto mechanics.

[11] From 45.2% to 43% and from 95.3% to 89.8% between 1981 and 1991, for female and male family heads, respectively.

[12] See also Corten and Duarte (1986).

1990, is on the list of the ten countries with the largest number of immigrants admitted to the United States as permanent residents between 1981 and 1996 ..... In relative terms, the number of Dominicans admitted was much greater than in the case of Mexicans, since the former represented 7% of the population of the Dominican Republic in 1990, whereas the latter were only equivalent to around 4% of the population of Mexico".

Ramón Hernández and Narey López (The CUNY Dominican Studies Institute, 1997, pp. 59-78) give an excellent overview of the problem and a full bibliography of what has been published on this matter.

Immigration of Haitian labour, for its part, has mainly been to the sugar cane-growing areas and the construction sector. A significant number of Haitians are also engaged in trade on the frontier between that country and the Dominican Republic. According to the General Directorate of Migration, over 20,000 Haitians enter and leave the country every day for trading purposes (*Listín Diario*, 1999, p. 20A).

With regard to immigration by Haitians, it would be interesting to know how large it is. How many Haitians work or live (thus representing potential labour) on Dominican territory? Corten and Duarte (1994) give an ex-

tensive bibliography on this subject. According to them, about half a million Haitians and Dominicans of Haitian origin live in the Dominican Republic, although the actual composition of the figure has not been clearly determined.

It would also be interesting to know if Haitian labour displaces Dominican workers, or if it is simply an expression of the free play of supply and demand for labour. The latter seems reasonable in view of the fact that historically Dominicans have rejected the kind of work that Haitians do, such as cutting sugar cane. The problem is complex and has been analysed from a number of angles (Martínez, 1999).

From the economic point of view, a reasonable approximation of the problem of (in)migration can be made, in principle, through a model like that proposed by Harris and Todaro (1970).

Although the problems described are important for the development of the Dominican economy and society, there are no studies of an economic nature which make a strict analysis of their dynamics. In contrast, there are a considerable number of studies from the standpoint of other branches of the social sciences (generally sociology).

V

Wages

1. Income and schooling

The relation between education and the labour market in the Dominican Republic has been analysed in many studies, including for example *Bases para formular una política de empleo en la República Dominicana* (ONAPLAN, 1974) and *Los recursos humanos y el empleo en la República Dominicana* (Education Development Center, 1975).[13]

With regard to the specific relation between income and schooling, Santos (1998) presents econometric estimates based on cross-sectional data from the 1991 and 1996 population and health surveys.[14] This study analyses the impact of secondary education on income, also taking into account other details such as the type of school

attended, geographical area of residence, sex, level of income, and social class.

According to the results obtained in this study, which are similar for both 1991 and 1996, there is a positive relation between the level of schooling and income, with the fact of having studied in a public school being the only one of the factors considered which has a negative impact on income. In contrast, the empirical evidence collected in the study shows that residing in an urban area, coming from the middle or upper class, and being a male positive influence income levels.

2. Minimum wages[15]

Among the few studies relating to minimum wages in the Dominican Republic is that recently presented by

---

[13] These studies are analysed in Cabral (1976).

[14] The determinants of the educational levels reached in the Dominican Republic have been analysed by Lizardo and Guzmán (1999).

[15] For some aspects of minimum wages in the free zones, see section IV.4 above.

the Ministry of Labour (República Dominicana, Secretaría de Estado de Trabajo, 1997). This study analyses the evolution of minimum wages, the relation between those for men and women, and the relation between the minimum wage and various macroeconomic aggregates of the Dominican economy. It notes that only 4.1% of all those employed receive the minimum wage.

The authors conclude that the minimum wage apparently does not have any effect on the level of inflation but could be an important factor in the services sector, because of its high share in the cost structure of the sector.

Dauhajre, Achecar and Swindale (1994) also investigate the relation between minimum wages and income distribution. They conclude that the real minimum wage has a positive effect on general well-being, because when it is increased it seems to improve income distribution, as measured by the Gini coefficient.

Dauhajre, Riley, Mena and Guerrero (1989, pp. 67-109) present econometric estimates according to which devaluation of the official exchange rate would have a positive impact on employment in the sector, whereas an increase in the minimum wage would have the opposite effect. The explanation would appear to be that, other things being equal, devaluation reduces the cost of local labour expressed in foreign currency.

As noted earlier, however, Gregory (1997, pp. 24-25) points out that the wages in firms in the free zones are considerably higher than the minimum wages laid down for the sector (in the first half of 1996 they were 54% higher for women and over 100% higher for men). He also states (*ibid.*, p. 21) that the multiplicity of minimum wages that exists in the Dominican economy complicates any effort to quantify the impact of the legal minimum wage on workers' real incomes.

VI

Conclusions and possible lines
of future research

The main aim of the present study has been to review the literature on labour-related matters in the Dominican Republic. Although a significant number of studies were identified, there was a notable paucity of studies prepared from an economic standpoint with a strictly scientific approach. Instead, most of the studies were of a descriptive, sociological or multi-disciplinary nature.

Future work in this field could be aimed, among other directions, at the preparation of econometric studies which could help verify the consistency of the statistics currently available (primarily the Central Bank survey on the labour force).

Through an analysis of this type, it would be interesting to find strictly scientific answers to such questions as: Does a significant level of unemployment really exist? What is its incidence? What kind of connection is there between vacancies and unemployment? How long do spells of unemployment last, on average? What are the rates of entry into the ranks of the unemployed and exit from them? What impact do minimum wages have on the development of the economy as a whole or specific sectors of it? What factors determine investment in human capital? How does technological development affect the well-being of workers? Are there notable dif-

ferences between the wages paid to workers with similar skills simply because they work in different industries? What role is played by the work of minors? Are there relative inequalities affecting the job opportunities for women and young people? What is the role of the government as an employer? How are wages and employment in the public sector determined?

The results of studies along these lines could be used by those responsible for designing and implementing policies relating to the labour factor in the country and also for improving the collection of statistics. It is essential to build up a system capable of compiling labour statistics that are reliable and, above all, periodically updated.

There are also other lines of research which could be fruitful, such as the detailed analysis of a number of institutional factors like, for example, the economic consequences of the present Dominican labour regulations (Law No. 16-92).[16] Likewise, there are almost no studies on the various ways in which trade unions and em-

---

[16] Dauhajre, Aristy, Riley, Mena, Féliz, Guiliani and De Castro (1996, pp. 459-472) have made a succinct analysis of these regulations.

ployers interact to determine the level of employment and wages.[17]

No less important is the question of the establishment of a social security system which is soundly based and, above all, self-sustaining over time. The key component

in a system of this nature –suitable legislation– is currently being debated in the National Congress (see Senado de la República Dominicana, 1999).[18]

(Original: Spanish)

Bibliography

Alemán, J. L. (1994): Empleo y apertura externa en la República Dominicana, *Estudios sociales*, No. 97, Santo Domingo, Centro de Investigación y Acción Social de la Compañía de Jesús.

———(1996a): *Políticas y programas sociales: menores en dificultad, generación de empleo y pobreza crítica*, Santo Domingo, Pontificia Universidad Católica Madre y Maestra (PUCMM)/Centro Universitario de Estudios Políticos y Sociales (CUEPS)/Editora Centenario.

———(1996b): Lucha contra la pobreza, enfatizando la generación de empleo, PUCMM/CUEPS, *Agenda nacional de desarrollo. Planes de acción para las diez principales prioridades*, vol. II, Santo Domingo, Grupo de Acción por la Democracia (GAD)/Editora Centenario.

APEC Educational Credit Foundation, Inc. (FUNDAPEC) (1992): *Encuesta nacional de mano de obra ENMO-91*, Santo Domingo.

Aristy, J. (1996): Políticas encaminadas a la reducción de la pobreza y a la generación de empleo, PUCMM/CUEPS, *Agenda nacional de desarrollo. Planes de acción para las diez principales prioridades*, vol. II, Santo Domingo, GAD/Editora Centenario.

Baéz, C. and M. Matrillé (1994): *Mercado laboral para las mujeres formadas en ocupaciones no tradicionales: Caso Los Mina Norte*, Santo Domingo, Inter-American Development Bank (IDB)/APEC Educational Credit Foundation, Inc. (FUNDAPEC)/ Editora Palma.

Cabal, M. (1992): *Microempresas y pequeñas empresas en la República Dominicana, resultados de una encuesta nacional*, Santo Domingo, Fondo de Desarrollo para la Microempresa (FONDOMICRO).

———(1993): *Evolución de las microempresas y pequeñas empresas en la República Dominicana,1992-1993*, Santo Domingo, FONDOMICRO.

———(1996): *Cambios en el tamaño y el empleo de las microempresas y pequeñas empresas en la República Dominicana, 1992-1995*, Santo Domingo, FONDOMICRO.

Cabral, M. J. (1975): La inflación dominicana, distribución del ingreso y empleo, Academia de Ciencias de la República Dominicana, *Anuario*, vol. 1, pp. 431-462, Santo Domingo, Instituto Tecnológico de Santo Domingo.

———(1976): Algunas observaciones sobre la educación y el empleo en la República Dominicana, *Eme Eme*, No. 5, Santo Domingo.

Cassá, R. (1990): *Movimiento obrero y lucha socialista en la República Dominicana, desde los orígenes hasta 1960*, Santo Domingo, Fundación Cultural Dominicana/ Editora Taller.

Ceara, M. (1992): Empleo, deuda y crecimiento económico: un modelo macro, *Ensayos de Economía*, Santo Domingo, Economic Research Centre for the Caribbean (CIECA), April.

Cely, P. (1993): *Microempresas y pequeñas empresas de mujeres en la República Dominicana: resultados de una encuesta nacional*, Santo Domingo, FONDOMICRO.

———(1996): *Dinámica de las microempresas y pequeñas empresas de mujeres en la República Dominicana, 1992-1995: una perspectiva dinámica*, Santo Domingo, FONDOMICRO.

Central Bank of the Dominican Republic (1997): *Mercado de trabajo 1991-1996*, Santo Domingo.

———(various years): *Boletín trimestral* (various issues), Santo Domingo.

Central Bank of the Dominican Republic (BANCENT)/National Statistical Office (ONE), United States Agency for International Development (USAID) (1968): *Estudio sobre presupuestos familiares*, Santo Domingo.

Corten, A. and I. Duarte (1986): Procesos de proletarización de mujeres: las trabajadoras de industrias de ensamblajes, in I. Duarte (ed.), Trabajadores urbanos: ensayos sobre fuerza laboral en la República Dominicana, *Estudios sociales*, Santo Domingo, Centro de Investigación y Acción Social de la Compañía de Jesús.

———(1994): Quinientos mil haitianos en República Dominicana, *Estudios sociales*, No. 27, Santo Domingo, Centro de Investigación y Acción Social de la Compañía de Jesús.

Dauhajre, A., J. Achecar and A. Swindale (1994): *Estabilización, apertura y pobreza en República Dominicana, 1986-1992*, Santo Domingo, Fundación Economía y Desarrollo.

Dauhajre, A., J. Aristy, E. Riley, R. Mena, S. Féliz, H. Guiliani and J. De Castro (1996): *Programa macroeconómico de mediano plazo para la República Dominicana: 1996-2000*, Santo Domingo, Editora Corripio.

---

[17] De Peña (1977) and Cassá (1990) have however done some useful work in this field. The first-named reviews historical issues which are important for analysing the trade union movement in the Dominican Republic, while the latter author makes a historical-sociological analysis of workers' movements. Hernández (1998) addresses the question of collective negotiation in the Dominican Republic from a legal standpoint.

[18] See also, by way of illustration, Santana and Rathe (1993), PUCMM/CUEPS (1996, chapter VI, pp. 327-368) and Hernández and Hernández (eds.), 1996: this latter work deals primarily with legal aspects of the matter.

Dauhajre, A., E. Riley, R. Mena and J. A. Guerrero (1989): *Impacto económico de las zonas francas industriales de exportación de la República Dominicana*, Santo Domingo, Fundación Economía y Desarrollo.

De Peña, J. (1977): *Breve historia del movimiento sindical dominicano*, Santo Domingo, Ediciones Dominicanas Populares.

Del Rosario, G. and T. Hidalgo (1986): *Metodología para calcular el índice de salario nominal y su aplicación en la economía dominicana*, Santo Domingo, Central Bank of the Dominican Republic.

Duarte, I. (1979): La fuerza de trabajo infantil en Santo Domingo, *Estudios sociales*, No. 12, Santo Domingo, Centro de Investigación y Acción Social de la Compañía de Jesús.

———(1983): Fuerza laboral urbana en Santo Domingo, 1980-1983, *Estudios sociales*, No. 16, Santo Domingo, Centro de Investigación y Acción Social de la Compañía de Jesús.

Duarte, I., R. Brea y R. Tejada (1998): *Cultura política y democrática en la República Dominicana, resultados de la segunda encuesta nacional de cultura política y democrática (DEMOS-97)*, Santo Domingo, PUCMM/Programa de Apoyo a las Iniciativas Democráticas (PID)/Asociación Pro Bienestar de la Familia (PROFAMILIA)/ Instituto de Estudios de Población y Desarrollo (IEPD), March.

Duarte, I., C. J. Gómez and M. Ariza (1991): *Menores en circunstancias especialmente difíciles en la República Dominicana*, Santo Domingo, United Nations Children's Fund (UNICEF)/ Association for Family Welfare (PROFAMILIA)/ Instituto de Estudios de Población y Desarrollo (IEPD).

Education Development Center (1975): *Los recursos humanos y el empleo en la República Dominicana. Informe y recomendaciones*, Newton, Massachusetts.

FUNDAPEC (Fundación APEC de Crédito Educativo (1992): *Encuesta nacional de mano de obra ENMO-91*, Santo Domingo.

García, N. E. and M. Valdivia (1985): *Crisis externa, ajuste interno y mercado de trabajo: República Dominicana, 1980-83*, Monografía sobre empleo No. 49, Santiago, Chile, International Labour Organisation (ILO), Regional Employment Programme for Latin America and the Caribbean (PREALC).

Gregory, P. (1997): *Empleo y desempleo en la República Dominicana*, Santo Domingo, Central Bank of the Dominican Republic, December.

Gurak, D.T., M.M. Kritz, V. Mota and M. Ortega (1979): *Female employment in the Dominican Republic*, paper presented at the Population Association of America Meetings, Philadelphia, 26 April.

Harberger, A.C. (1997): *Economic policy and performance in the Dominican Republic*, Santo Domingo, November, *mimeo*.

Harris, J. and M. Todaro (1970): Migration, unemployment and development: A two-sector analysis, *The American Economic Review*, vol. 60, No. 1, Nashville, Tennessee, American Economic Association.

Hernández, L. (1998): *La negociación colectiva en Santo Domingo*, Santo Domingo, Fundación Institucionalidad y Justicia (FINJUS).

Hernández, L. and C. Hernández (eds.) (1996): *La seguridad social en Santo Domingo*, Santo Domingo, Instituto de Estudios del Trabajo, Inc./Editora Corripio.

International Labour Office (1975): *Generación de empleo productivo y crecimiento económico: El caso de la República Dominicana*, Geneva.

ILO (International Labour Organisation) (1998): *Bulletin of Labour Statistics*, No. 57, Geneva.

Listín Diario (1999): Santo Domingo, 18 August.

Liz, R.E. (1993): *Crecimiento económico, empleo y capacitación*, paper presented at the seminar on concerted policies for social reform and the reduction of poverty, Santo Domingo, United Nations Development Programme (UNDP)/PUCMM, 23-24 November, *mimeo*.

Lizardo, J. (1994): *Condiciones de los menores en la República Dominicana*, Santo Domingo, Institute for Population and Development Studies (IEPD)/UNICEF, October.

Lizardo, M. M. and R. M. Guzmán (1999): Niveles de escolaridad y sus factores determinantes: una cuantificación econométrica, *Ciencia y sociedad*, vol. 24, Santo Domingo, Instituto Tecnológico de Santo Domingo, April-June.

Lozano, W. (1987): Desempleo estructural, dinámica económica y fragmentación de los mercados de trabajo urbanos: el caso dominicano, *Ciencia y sociedad*, vol. 12, No. 3, Santo Domingo, Instituto Tecnológico de Santo Domingo.

Márquez, G. (1998): *El desempleo en América Latina y El Caribe a mediados de los años 90*, Documento de trabajo No. 377, Washington, D.C., IDB, Office of the Chief Economist, August.

Martí, A. (1997): *Instrumental para el análisis de la economía dominicana. Base de datos (1947-1995)*, Santo Domingo, Editora Búho.

Martínez, L. (1994): *Actitudes femeninas frente a los oficios no tradicionales*, Santo Domingo, CE-MUJER/IDB/APEC Educational Credit Foundation, Inc.

Martínez, S. (1999): From hidden hand to heavy hand: Sugar, the estate, and migrant labor in Haiti and the Dominican Republic, *Latin American Research Review*, vol. 34, No. 1, Yale, University of New Mexico.

Moya, F. (ed.) (1986): Condiciones socioeconómicas de la mujer trabajadora en la República Dominicana, *Forum*, No. 21, Santo Domingo.

Moya, F. and M. Ortiz (1994): *Microempresas y microempresarios en la República Dominicana, 1993-1994*, Santo Domingo, FONDOMICRO.

———(1995): *Indicadores de las microempresas en la República Dominicana, 1994-1995*, FONDOMICRO, Santo Domingo.

National Planning Office (ONAPLAN) (1968): *Plataforma para el desarrollo económico y social de la República Dominicana*, Santo Domingo.

———(1974): *Bases para formular una política de empleo en la República Dominicana*, Serie PLANDES, No. 19, Santo Domingo.

ONAPLAN/CESDEM/PROFAMILIA (National Planning Office/Centre for Social and Demographic Studies/ Association for Family Welfare (1997): *Encuesta demográfica y de salud 1996*, Santo Domingo.

National Statistical Office (ONE) (1993): *Boletín de registro general de organizaciones sindicales en la República Dominicana*, Santo Domingo.

————(1998): *Censo de servidores públicos*, Santo Domingo, *mimeo*.

Nicasio, I. and J. De La Rosa (1998): *Historia, metodología y organización de los censos en República Dominicana, 1920-1993*, Santo Domingo, Editora Búho.

Ortiz, M. (1996): *Perfil de las microempresas en la República Dominicana y características de las microempresas de Santiago, 1995-1996*, Santo Domingo, FONDOMICRO.

————(1998): *Microempresas y fuerza laboral en la República Dominicana, 1997-1998*, Santo Domingo, FONDOMICRO.

Pérez Sáinz, J. P. (1995a): *Actores y cultura laboral en la industria de exportación: una comparación entre Costa Rica y la República Dominicana*, Documento de trabajo No. 6, San José de Costa Rica, ILO.

————(1995b): *Impacto socio-laboral de maquilas y zonas francas en Centroamérica y la República Dominicana*, Documento de trabajo No. 16, San José de Costa Rica, ILO.

PUCMM/CUEPS (Pontificia Universidad Católica Madre y Maestra/ Centro Universitario de Estudios Políticos y Sociales)(1996): *Agenda nacional de desarrollo. Planes de acción para las diez principales prioridades*, vol. II, Santo Domingo, GAD, Editora Centenario.

Quiterio, G. y J. Polanco (1990): *República Dominicana: estimaciones y proyecciones de la población económicamente activa y demanda de empleos por regiones y subregiones, 1980-2000*, Santo Domingo, National Planning Office (ONAPLAN)/ Institute for Population and Development Studies (IEPD), October.

Rama, M. (1995): *Issues and non-issues in the Jamaican labor market*, Washington, D.C., World Bank, *mimeo*.

Ramírez, N. (1993): *La fuerza de trabajo en la República Dominicana: análisis de los resultados del cuestionario de hogar ampliado, encuesta demográfica y de salud 1991, ENDESA 1991*, Santo Domingo, Institute for Population and Development Studies (IEPD).

Ramírez, N., A. Tatis and D. Germán (1983): *Población y mano de obra en la República Dominicana: perspectivas de la fuerza de trabajo y del empleo-desempleo en el período 1980-1990*, Santo Domingo, Institute for Population and Development Studies (IEPD).

República Dominicana, Secretaría de Estado de Trabajo (1997): *Estudio sobre salario en la economía dominicana*, Santo Domingo, Ecocaribe S.A., May.

————(1999): *Estadísticas laborales*, Boletín No. 4, Santo Domingo, March.

República Dominicana, Secretaría de Estado de Turismo (1996): *Turismo en cifras*, Santo Domingo, July.

Rodríguez, M. (1999): República Dominicana: los migrantes y el futuro del país, *Listín diario*, Santo Domingo, 24 September.

Santana, I. (1990): *Informe sobre la mano de obra rural*, Santo Domingo, Consejo Nacional de Agricultura, Proyecto Análisis de Política Agropecuaria, Unidad de Estudios Agropecuarios.

————(1996): Estrategias y políticas de lucha contra la pobreza y generación de empleo, PUCMM/CUEPS, *Agenda nacional de desarrollo. Planes de acción para las diez principales prioridades*, vol. II, Santo Domingo, GAD, Editora Centenario.

Santana, I. and M. Rathe (1993): *Reforma social, una agenda para combatir la pobreza*, Santo Domingo, Ediciones de la Fundación Siglo 21.

Santos, A.U. (1998): *Análisis de los efectos de la educación secundaria en el individuo y el rol representado en el mercado de trabajo*, Santo Domingo, World Bank, *mimeo*.

Senado de la República Dominicana (1999): *Proyectos de ley de seguridad social, documento base para las vistas públicas provinciales y del Distrito Nacional*, Santo Domingo, Comisión Permanente de Seguridad Social, January.

Silié, R. (1996): *El trabajo infanto-juvenil en las ciudades dominicanas*, Santo Domingo, Latin American Faculty of Social Sciences (FLACSO)/UNICEF.

Tejada, R., J. Herold and L. Morris (1998): *República Dominicana: encuesta nacional de jóvenes 1992 ENJOVEN-92. Informe final*, Santo Domingo, Institute for Population and Development Studies (IEPD).

The CUNY Dominican Studies Institute (1997): *Dominican Studies: Resources and Research Questions*, Dominican research monographs, New York.

USAID/Dominican Republic, Ministry of Agriculture (1977): *Empleo, República Dominicana. Serie de documentos sobre análisis sectorial agropecuario*, Documento estadístico de trabajo No. 1, Washington, D.C.

Vargas, L. (1995): *Fuerza de trabajo: desempleo, subempleo y salarios*, paper presented at the seminar "Problemas de Población en República Dominicana", Santo Domingo, National Planning Office, Secretariado Técnico de la Presidencia, 31 May- 2 June, *mimeo*.

Weller, J. (1993): *La situación del empleo en la República Dominicana: Informe de consultoría*, San José de Costa Rica, ILO.

————(1998): *Los mercados laborales en América Latina: su evolución en el largo plazo y sus tendencias recientes*, "Reformas económicas" series, No. 11, Santiago, Chile, Economic Commission for Latin America and the Caribbean (ECLAC), December.

Recent ECLAC publications

Periodical publications

Statistical Yearbook for Latin America and the Caribbean, 1999 Edition, LC/G.2066-P, United Nations publication, Sales No. E/S.00.II.G.1, ECLAC, Santiago, Chile, February 2000. Bilingual English/Spanish version, 766 pages.

The 1999 edition of the *Statistical Yearbook for Latin America and the Caribbean* contains a selection, updated to the beginning of December, of the main statistical series available on economic and social trends in the countries of the region.

This year's edition does not differ from the 1998 edition in terms of its structure, but minor changes continue to be made in the tables included in the chapters on social development and welfare and on social conditions as improvements are made in the processing of the available household surveys. The tables given in the chapter on the balance of payments continue to be based on the approach to its analysis used in the fifth edition of the *Balance of Payments Manual* published by the International Monetary Fund in 1993.

Part One of the *Yearbook* consists of derived social and economic indicators (growth rates, ratios or coefficients) which provide an overview of each area of interest, along with the background material needed to enable the information to be used in specialized analyses.

Part Two provides historical series in absolute figures which can be used for a wide variety of purposes. The majority of the statistical tables give figures on a single topic, organized in such a way as to facilitate comparisons among countries as well as between individual countries and regional totals or averages. The tables on the balance of payments and national accounts are the only exceptions in this respect, since they have been prepared on a country-by-country basis.

In most of the tables the countries appear in alphabetical order, excluding those for which there are no data or for which the amounts are zero or negligible.

In Part One of the *Yearbook*, most of the indicators correspond to the years 1980, 1985 and 1990, and to the period from 1993 to 1998. In cases where the data have not been updated, the most recent figures available for each country are given. Some of the indicators that are based on census data are given only for the years when the respective censuses were conducted. The statistical series in Part Two, both country figures and regional estimates, give data for 1980, 1985 and the period from 1990 to 1998.

In view of the favourable response from users, as well as the degree of accuracy achieved in previous versions, preliminary estimates for the year of issue of the *Yearbook* (in this case, 1999) are given once again. These estimates are the result of an effort made during the last two months of each year to inform the international community about the macroeconomic trends observed in the countries of the region during the period covered. In this connection, it is important to note that due to the different closing dates, in certain cases the historical figures given for recent years in the tables of the *Yearbook* may differ slightly from those given in the *Preliminary Overview of the Economies of Latin America and the Caribbean, 1999*.

Other publications

Estrategias de política económica en un mundo incierto. Reglas, indicadores, criterios (Economic policy strategies in an uncertain world. Rules, indicators, criteria), LC/IP/G.123-P, "Cuadernos del ILPES" series, No. 45. United Nations publication, Sales No. S.00.III.F.1, Santiago, Chile, April 2000, 198 pages.

This publication refers to the long-standing debate on the use of rules in economic policy and presents the numerous different opinions that exist about the various issues involved.

Chapter I contains a brief summary of the main theoretical explanations for macroeconomic fluctuations, from those that see them as phenomena typical of the market or as faults of the government, up to those that attribute them to flaws and imperfections in monopoly markets. The use of rules is considered to be important in both interpretations, albeit with different approaches. In the first case, the use of rules is a coercive measure suitable for avoiding overintervention; in the second case, the rules act as guidelines and the aim is to introduce self-regulating mechanisms which will stabilize markets and reduce the chaos caused by "hidden hands". In the first case the rules are passive and designed to avoid intervention, while in the second they are active and involve intervention.

The authorities have four simultaneous objectives: to reduce unemployment or keep it down to a low level; to ensure sustained growth; to control inflation, and to keep watch for potential imbalances which could arise on the fiscal or external sides. The art of economic policy is to create a virtuous circle within this magic square. This task must be carried out in an uncertain environment, so economic policy cannot be seen as a mere optimization programme. Pragmatic, anticipatory strategies would appear to be the best response in a stochastic world.

These matters are analysed in chapter II, as are the links between trend growth, the cycle and economic policies. The factors explaining long-term growth are reviewed, and it is noted that, empirically, there are only a few macroeconomic variables which are really important. Comparative growth is explained above all by the rate of investment, the initial GDP (inversely: the countries with the highest per capita income grow relatively less), the rate of primary schooling, and life expectancy. Variables regarding the size of the State sector are not generally significant. There is, however, a negative correlation between the average growth rate of economies and its standard deviation, which shows that economic policies can be an important factor in long-term growth. It seems empirically important to direct the authorities' actions towards the objective of short-term macroeconomic stabilization.

The imposition of restrictions merely serves to reduce the authorities' degree of freedom. The adoption of explicit (final or intermediate) objectives can help to improve institutional organization, provided that some degree of discretionality is maintained in the system and policies are always applied in a transparent and symmetrical manner. The idea of discretionality within overall rules is designed to give the economic agents a certain amount of security in an uncertain world. Within the context of these guiding criteria, an analysis is then made of some specific aspects of the dilemma between the flexibility afforded by discretionality and the credibility created by the imposition of rules.

Chapter III analyses the costs and benefits of alternative strategies in open economies, as a function of the objectives proposed and the instruments available. In this respect - even more than in others - there are no absolute truths, so that neighbouring countries may adopt completely different exchange rate and capital flow regimes, yet with an equal degree of conviction. The preference for one regime over another depends not only on objective factors, such as the structural characteristics of the economy in question and the type of shocks it is facing, but also on subjective factors such as conviction (when choosing between regulatory or market-based policies) and motivation (whether or not to pursue the objective of autonomy of national policies) of the authorities. In this respect, the reference to the new international architecture is an excellent way of underlining the need to design a set of rules and procedures which will make it possible to improve cooperation and prevent the periodic appearance of national balance of payments crises.

Chapter IV analyses the desirability of using rules and indicators in monetary policy management, both from the positive and from the normative point of view. The experience of countries which have been successful in combating inflation is reviewed, with emphasis on the diversity of the strategies and institutional designs which have made it possible to attain this objective. Rather than the degree of independence of the Central Bank and the use of strict rules, the important thing seems to be the legitimacy of the policies applied, which is achieved through a proper structure of the domestic decision-making process, suitable coordination with the other authorities, and some degree of discretionality in decision-making.

The fifth and last chapter addresses the same issues, but from the perspective of the public finances. Discipline and flexibility are seen as two fundamental principles of fiscal policy in the present process of economic and financial globalization. Fiscal discipline is essential for giving credibility to monetary policy, while flexibility is necessary in order to cope with the unexpected situations inherent in a highly variable macroeconomic environment. Within this context, an analysis is made of the dilemmas that arise between policy objectives, sustainability and macroeconomic regulation, and the possible strategies for achieving these aims, either through discretionality or through the use of rules.

Beyond the question of good or bad policy recipes, the challenge is to advance towards greater fiscal governance through the development of instruments which will guide the budgetary process towards a scheme in which expenditure adapts to the long-term availability of resources, clearly identifying the transitory factors and ensuring that a medium-term view is taken in the adoption of fiscal policy decisions. Although there is no question of establishing universally valid criteria, it is obvious that the adoption of medium-term fiscal policy rules is more flexible and hence also more realistic than a simple balanced budget rule. In a world of continuous changes and uncertain situations, economic policies must be very active in seeking to reduce excessive fluctuations, and fiscal policy undoubtedly has an important part to play in attaining this objective.

Equity, development and citizenship, LC/G.2071 (SES.28/3), document presented at the 28th session of the Economic Commission for Latin America and the Caribbean (Mexico City, 3-7 April 2000), ECLAC, Santiago, Chile, 15 March 2000, 355 pages.

This document sets forth in their entirety ECLAC's views on the challenges raised by the development of the region in the world of today, which are associated not only with the growing demands due to the globalization of markets but also the need to extend the application of universal values, especially human rights, social development, gender equity, respect for ethnic and cultural diversity, and protection of the environment. The region is facing these challenges with a mixed legacy of capabilities connected not only with the economic reforms but also with the still not fully corrected sequels of the debt crisis and a long-standing situation of poverty, exclusion and social inequality.

In the face of these challenges and inherited conditions, the document highlights the importance of building on what has already been achieved in terms of the reduction of fiscal deficits and inflation, the development of exports, economic integration, and participation of the private sector in development. It also emphasizes, however, the need to tackle once and for all the complex questions of equity facing the region - where advances in this field are the basic yardstick for measuring the quality of our development - and to do this while consolidating the still-insufficient recovery of economic growth and strengthening democracy.

According to the document, equity calls for integral social policies based on three main principles - universality, solidarity and efficiency - and two "master keys": education and employment. Education is the key to simultaneous advances in equity, development and citizenship, but its effects will only become a reality if there is a dynamic process of high-quality job creation. The document therefore addresses in detail the various challenges in these fields, such as the universalization of secondary education and the improvement of the quality of education at all levels, on the one hand, and the need to make both workers and enterprises more adaptable to technological change and economic cycles. It also mentions the need to extend social safety nets and make them more solidary and efficient.

With regard to development, the document proposes a style of macroeconomic management which gives due attention not only to inflation and the external balance but also to stability of economic growth and employment. It also says that macroeconomic policy should incorporate the objectives of economic growth more forcefully, on the basis of sound fiscal systems, moderate real interest rates, and competitive exchange rates, complemented with active production development policies. The latter policies should be aimed at maintaining competition and regulating non-competitive markets, correcting flaws in the production factor markets, and promoting innovation and full use of the complementarities between different enterprises, which further systemic competitiveness. Emphasis is also placed on the need to take a definitive step towards the full incorporation of the items on the sustainable development agenda, promoting forms of competitiveness based on the accumulation of capital in the broadest sense: human, physical and natural capital.

The citizenship agenda highlights the need to approach the issue of development with equity in a democratic context and notes that, in addition to the problems of equity, our region displays growing problems of social cohesion, so that it is important to make joint efforts to create "more society": that is to say, a more widespread awareness of individuals' duties to society, more spaces for discussion and accord-building and, in general, a culture of collective co-existence and development based on tolerance for differences and the negotiated solution of conflicts.

The report summarized above has four main parts. Part I has two chapters: the first chapter provides a global view of the document as a whole, while at the same time going into some detail about the proposals put forward, and the second chapter summarizes the legacy of the 1990s. Part II deals with issues relating to equity: the principles of social policy and the struggle against poverty; the development of education; employment; social security, and public

social spending and the goals that should be pursued in terms of social services coverage. Part III deals with economic issues in the light of the social dimensions of the economy and sustainable development, successively addressing stability and economic growth, production development, the specific problems of smaller economies, the regulation of public utilities, and the consolidation of spaces of sustainable development. The fourth and final part of the document has only a single chapter, containing reflections on citizenship and social cohesion.

Foreign investment in Latin America and the Caribbean, LC/G.2061-P, United Nations publication, Sales No. E.00.II.G.4, Santiago, Chile, January 2000, 208 pages.

This study attempts to contribute to a better understanding of foreign direct investment (FDI) in Latin America and the Caribbean by examining the plentiful but disparate statistical information available on the subject within the context of a conceptual framework based on an analysis of corporate strategies.

In 1998, FDI inflows amounted to US$ 76.7 billion and thus surpassed the record level of the previous year; it was estimated that they would be even larger in 1999. FDI is a major economic variable in the performance of the Latin American and Caribbean economies and therefore needs to be understood more fully. In 1998 over 40% of the region's total inflows of FDI went to Brazil, where the liberalization and deregulation of financial services and the privatization of State assets in the telecommunications and electricity generation and distribution industries opened up enormous opportunities for foreign investors interested in purchasing existing assets. In contrast, the creation of new assets has been a more prominent feature of foreign investment activity in Mexico.

In addition to detailed, up-to-date statistical information, this report provides comprehensive analyses of Mexico's position as a recipient of FDI, of Spain's status as a major new investor, and of the apparel industry, where FDI has generated substantial export flows to the North American market. Each of these analyses affords a valuable insight into the nature and impact of FDI in the region and contributes to a fuller understanding of this phenomenon.

The Caribbean in the decade of the 1990s, LC/CAR/G.600, ECLAC Subregional Headquarters for the Caribbean, Port of Spain, Trinidad and Tobago, 31 January 2000, Port of Spain.

Most of the small island States of the Caribbean have a number of features which, although they are shared with other larger developing countries, usually put these small countries in a more unfavourable position for achieving economic and social development. Among the main factors affecting the economic results, sectoral specialization and vulnerability of small island States are their limited resources and problems connected with their size and remoteness. They depend to a large extent on trade, and in many cases the ratio of such trade to their GDP is over 100%. The markets of these countries are small, and their exports are limited to a narrow range of products or services and markets. Because of their geographical location, they are specially vulnerable to hurricanes, and their fragile ecosystems are seriously threatened by population pressures and the need to carry on economic activities in a very small area. In recent years, these countries have been developing in the context of an increasingly integrated world economy, marked by a steady increase in trade and finance flows and in competition.

In the 1980s and 1990s, the Caribbean countries have undertaken far-reaching economic and social reforms. On the macroeconomic level, they have taken measures to reduce the fiscal and balance of payments deficits which arose in the late 1970s and early 1980s and to stabilize their economies. They have also put into effect trade and financial reforms to strengthen the opening-up of their economies and take advantage of the possibilities offered by an increasingly liberalized world economy. They have also redoubled their efforts to solve problems connected with social equity by taking measures to promote employment, improve access to health and education, and reduce poverty, especially in those countries of the subregion which registered low growth rates or situations of political instability in the 1990s.

Efforts have also been made to reduce gender inequity through the implementation of policies and programmes designed to improve the social and economic status of women. Because of migration and population growth, the population characteristics of the Caribbean countries have been gradually changing, and for this reason it has been necessary to redouble the efforts to improve the supply of social services such as education, health and social security. In 1994 the Action Programme for the Sustainable Development of Small Island Developing States was adopted. This Programme drew the attention of the governments of the subregion and the international community at large to the need to preserve the fragile environment of these States, most of which are islands, and to adopt policies which address economic, environmental and social problems at one and the same time.

Privatización portuaria. Bases, alternativas y consecuencias (Port privatization. Bases, options and consequences), LC/G.2045-P, by Larry Burkhalter. United Nations publication, Sales No. S.99.II.G.59, ECLAC, Santiago, Chile, December 1999, 248 pages.

Economic globalization has brought with it a profound reorientation of the economic policies of the Latin American and Caribbean countries, and with it a process of rethinking the question of port modernization, which is addressed in this study.

In the import substitution model that prevailed in previous decades, some degree of inefficiency could be tolerated in the ports, and might even be considered functional in so far as it protected domestic industry. Today, however, with a more open economic model, the driving force of economic development lies in promoting the export of goods and services where the country in question has comparative advantages. The competition prevailing in the new world setting demands that the prices of products on foreign markets should be brought down as low as possible, and one of the components of those prices is the cost of port services. At the same time, port costs are also important for imports, since in today's globalized world the amounts paid for imported inputs form a significant proportion of the cost of any export. The old model of State monopoly port authorities is incompatible with an increasingly competitive world. It is not just a question of privatizing the ports and replacing a public monopoly with a private one, however. There are other matters which are no less important and which are analysed in this book: competition among ports, regulation, and the social dimensions of changes in the field of labour. This study will be a very useful instrument for addressing the changes in this area, especially as it will enable many countries to take advantage of the experience of others which are more advanced in the process of port reform.

Nuevas políticas comerciales en América Latina y Asia. Algunos casos nacionales (New trade policies in Latin America and Asia. Some national case studies), LC/G.2053-P, United Nations publication, Sales No. S.99.II.G.47, ECLAC, Santiago, Chile, December 1999, 583 pages.

This study analyses the experience of the two regions in question in order to identify the fundamental elements of the State's role in the promotion of industry and trade.

It also analyses the new role of the State in the economy and the mechanisms open to it for playing this role, in the light of the present international trade regime that emerged from the Uruguay Round and the regional agreements which are in effect.

In each case the following aspects are considered:

i) the sources of the economic growth of each country, in order to analyse the causal relations between growth and exports and to determine whether development has been generated by exports or whether exports have been generated by economic growth itself or by investments;

ii) the trade and industrial policies of the country, in order to determine whether they were selective or neutral before the end of the Uruguay Round, with details of their positive or negative aspects;

iii) the changes in the system of incentives, the design and control of the instruments for improving the competitiveness of the industrial and export sectors, and the timetable of the application of incentives or protective measures;

iv) future policies in the fields of trade and industry, within the framework of the restrictions imposed by the new trade system and the modifications or changes it is proposed to make, or which have already been made, in national legislation on various new areas (services, intellectual property rights and trade-related investment measures (TRIMs)), as well as in various aspects of trade policies (consolidation and reduction of tariffs and the reduction and elimination of non-tariff barriers, safeguard measures, subsidies, anti-dumping measures, etc.), in order to bring them in line with the agreements of the World Trade Organization (WTO);

v) mechanisms to strengthen domestic production capacity which are compatible with the WTO accords, especially the policies to be applied in such areas as human resources, infrastructure, trade finance, labour force training, quality control, and regulation and promotion of small and medium-sized enterprises;

vi) existing institutions responsible for policy-making in the areas of trade and industry, including such aspects as relations between the private sector and the State, the independence of the civil service, coordination and links between macro-, meso- and micro-

policies, and the creation of networks among the various State bodies, and

vii) the role played by regional and subregional integration agreements in determining the nature and scope of industrial and trade policies.

An important conclusion emerging from the studies made is that although the developing countries will have to face stricter limitations on their possibilities of applying selective trade and industrial policies, they still have a certain amount of leeway in this respect and can still make use of a number of policies in order to achieve their economic objectives in these areas without contravening the commitments assumed under multilateral and regional accords.

Economic profiles of the Caribbean countries, LC/CAR/G.572, ECLAC Subregional Headquarters for the Caribbean, Port of Spain, Trinidad and Tobago, 19 October 1999, Port of Spain.

The Subregional Headquarters for the Caribbean of the Economic Commission for Latin America and the Caribbean (ECLAC) has prepared these economic profiles of the 26 countries and non-self-governing territories of the Caribbean subregion: Anguilla, Antigua and Barbuda, Aruba, Bahamas, Barbados, Belize, Bermuda, British Virgin Islands, Cayman Islands, Cuba, Dominica, Dominican Republic, Grenada, Guyana, Haiti, Jamaica, Monserrat, Netherlands Antilles, Puerto Rico, Turks and Caicos Islands, St. Kitts and Nevis, St. Lucia, St. Vincent and the Grenadines, Suriname, Trinidad and Tobago and U.S. Virgin Islands.

Each profile seeks to give the reader a succinct picture of the economy of the country or territory in question, along with a selection of economic variables. The profiles are divided into several sections: basic data, economic background, economic activity, tax system, external debt and debt service, and balance of payments. Special attention is given to rates of unemployment and inflation and exchange rates, and data are given on present income and expenditure and the budgetary situation. The section on the external debt and debt service analyses the debt/GDP relation and the debt service coefficient, while in most cases the section on the balance of payments analyses the trade balance, the current account balance, transfers, the capital and financial account balances, and the overall balance.

The document does not try to analyse or examine the economic prospects of the countries or territories, but simply describes the economic situation of each of them in the 1993-1998 period, on the basis of the available data. It is hoped that this publication will help to obtain a better understanding of the economic situation of the countries and territories of the subregion.

Recent ECLAC publications