CEPAL
REVIEW

ECONOMIC COMMISSION FOR LATIN AMERICA AND THE CARIBBEAN

UNITED NATIONS

ECLAC

N° 102
DECEMBER • 2010

ISSN 0251-2920

CEPAL

REVIEW

ECONOMIC
COMMISSION FOR
LATIN AMERICA
AND THE CARIBBEAN

UNITED NATIONS

ECLAC

Nº 102
DECEMBER • 2010

CEPAL

REVIEW

ECONOMIC
COMMISSION FOR
LATIN AMERICA
AND THE CARIBBEAN

UNITED NATIONS

ECLAC

ISSN 0251-2920

The *CEPAL Review* was founded in 1976, along with the corresponding Spanish version, *Revista de la CEPAL*, and is published three times a year by the United Nations Economic Commission for Latin America and the Caribbean, which has its headquarters in Santiago, Chile. The *Review*, however, has full editorial independence and follows the usual academic procedures and criteria, including the review of articles by independent external referees. The purpose of the *Review* is to contribute to the discussion of socio-economic development issues in the region by offering analytical and policy approaches and articles by economists and other social scientists working both within and outside the United Nations. The *Review* is distributed to universities, research institutes and other international organizations, as well as to individual subscribers.

The opinions expressed in the signed articles are those of the authors and do not necessarily reflect the views of the organization. The designations employed and the way in which data are presented 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.

A subscription to the *CEPAL Review* in Spanish costs US$ 30 for one year (three issues) and US$ 50 for two years. A subscription to the English version costs US$ 35 or US$ 60, respectively. The price of a single issue in either Spanish or English is US$ 15, including postage and handling.

The complete text of the *Review* can also be downloaded free of charge from the ECLAC web site (**www.cepal.org**).

This publication, entitled the CEPAL Review, is covered in the Social Sciences Citation Index (SSCI), published by Thomson ISI, and in the Journal of Economic Literature (JEL), published by the American Economic Association.

To subscribe, please apply to ECLAC Publications, Casilla 179-D, Santiago, Chile, by fax to (562) 210-2069 or by e-mail to publications@eclac.cl. The subscription form may be requested by mail or e-mail or can be downloaded from the Review's Web page: http://www.cepal.org/revista/noticias/paginas/5/20365/suscripcion.pdf.

United Nations publication
ISSN 0251-2920
ISBN 978-92-1-121755-1
e-ISBN 978-92-1-054553-2
LC/G. 2468-P

Printed in Santiago, Chile

ARTICLES

Explanatory notes

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., 2006/2007.

(-) Use of a hyphen between years (e.g., 2006-2007) indicates reference to the complete period considered, including the beginning and end years.

The word "tons" means metric tons and the word "dollars" means United States dollars, 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.

KEYWORDS

Economic conditions
Macroeconomics
Capital movements
Business cycles
Economic development
Productivity
Fiscal policy
Monetary policy
Foreign exchange rates
Capital markets
Latin America

Macroeconomics for development: from "financierism" to "productivism"

Ricardo Ffrench-Davis

Contrary to the belief that the region has found its way to an efficient macroeconomic policy, this paper argues that macroeconomic failures have been partly responsible for its disappointing economic and social performance in recent decades. Producers of GDP have had to cope with extremely unstable demand, exchange rates and access to financing, which have discouraged productivity and investment. Financial capital flows have been a determinant of this macroeconomic instability. This paper examines their intrinsically procyclical behaviour and concludes that an environment friendly to production development requires countercyclical regulation of financial flows. It describes how regulation of aggregate demand needs to be reconciled with the evolution of potential GDP, the real exchange rate with the current account, and financial flows with a far-reaching reform of the capital market reforms, away from "financierism" and towards "productivism".

Ricardo Ffrench-Davis
Professor at the Faculty of Economics and Business, University of Chile
rffrench@fen.uchile.cl

I
Introduction

The countries of Latin America have introduced deep economic reforms in the context of what came to be known as the Washington Consensus, including far-reaching trade and financial liberalization, privatization and a new fiscal discipline, in the belief that this would be enough to ensure stability, economic growth and increased well-being. The reality is that price stability has been achieved (inflation generally under control) over the past two decades, but with low average gross domestic product (GDP) growth and a high degree of instability in the real economy, i.e., in output and employment.

It is often argued that the region has learned to cope effectively with the macroeconomic challenge and that its failures are microeconomic. However, the fact is that production and employment have been affected by large cyclical swings in economic activity, overall demand, credit access and exchange rates. These are key macroeconomic variables, forming the environment in which producers of goods and services operate. This article will analyse how the volatility of these macroeconomic variables has discouraged capital formation, employment and actual productivity. Financial capital flows have played a central role here.

Notwithstanding the diversity of the Latin American countries, the direction of fluctuations in economic activity, aggregate demand, real exchange rates, saving rates, investment and capital flows have coincided to a great extent. The synchronicity is particularly marked among the majority of large and medium-sized countries, with less-developed countries showing substantial differences.

The macroeconomic environment is mainly the outcome of the effects and interrelationships of fiscal, monetary and exchange-rate policies, domestic capital markets and the capital account. In turn, this environment influences the speed and stability of economic growth and the distribution of its benefits, chiefly through its effects on capital formation and employment. Contrary to what tends to be implied by the traditionally very high degree of compartmentalization between micro- and macroeconomic analyses, the fact is that poverty reduction efforts, the degree of equity in a society and economic growth are strongly affected by the quality of the macroeconomic environment.

The emergence of the global crisis, for all its severe recessionary and regressive effects, has had a salutary result insofar as it has buttressed arguments for the central importance of macroeconomic policymaking styles and the need to consider ways of correcting current practices (see, for example, Blanchard, Dell'Ariccia and Mauro, 2010). This is essential in a development strategy designed to achieve growth with equity.

The core argument of this article is that there is a need to move from the strong "financierist" and "short-termist" bias that prevails at present to an approach that explicitly prioritizes productive development and its effects on equity. This requires an integrated approach incorporating the interrelationships between the micro- and macroeconomy and taking account of the implications of profound structural heterogeneity in national markets and the intrinsically procyclical nature of international financial flows. One aspect of this is that the instabilities referred to have very different effects on large and small firms, on investment and consumption, and on skilled and unskilled workers. The gradualism of policies and the quality of coordination between their monetary, exchange-rate, financial and fiscal aspects, for example, make a great difference to economic growth and its distributional effects, and particularly the quality of employment.

Interrelationships encompass static and dynamic effects. An example of the former are the effects on the utilization rate of the productive capacity of labour and capital. Fluctuations in this rate have repeatedly opened up large gaps between installed capacity or potential GDP and the GDP actually generated. These gaps, and the volatility of variables such as the real exchange rate, have had far-reaching dynamic effects on, for example, the investment ratio and its influence

□ I have written a number of texts on macroeconomic policy styles in emerging economies since an article published in *CEPAL Review* No. 60 (1996). The subject is explored further in Ffrench-Davis (2006) and (2008), plus versions produced for different events and publications. Here I try to summarize what I have learned about the subject. I am grateful for the assistance of Rodrigo Heresi and Felipe Labrín.

on the trend of development; the amount of value added to exports and their interrelationship with other components of GDP; innovation; the development of small and medium-sized enterprises; and formality or lack of it in the labour market. Macroeconomic policymaking styles have a considerable influence on all these variables and have been a crucial factor in the very modest 3.2% annual growth rate of regional GDP between 1990 and 2008.[1]

Consequently, while safeguarding the progress made with inflation control and fiscal discipline, there is a need to progress towards the creation of a macroeconomic environment that is more "friendly" to the different agents generating GDP. Section II summarizes the achievements and failures of the countries of Latin America as a group since 1990. The variables on which the analysis is focused exhibit effects that to a great degree are common to the majority of the population, despite the manifest differences between countries. Successes with inflation control, fiscal discipline and export dynamism are highlighted. There follows an exposition of how these successes have been accompanied by weak economic growth and very low levels of capital formation. Section III documents the great instability of aggregate demand and exchange rates with which the different agents have had to cope and the way this has been associated with recurrent external shocks in financial capital flows and, more recently, to the terms of trade as well.

Section IV examines the characteristics of financial flows and addresses the issue of why these tend to be intrinsically procyclical in emerging economies. Section V considers the effects of instability, in particular with respect to the emergence of recessive gaps between potential output or the production frontier and actual GDP. This is followed by an analysis of how these gaps adversely affect firms' balance sheets and expectations, along with employment. It is shown that the main impact of instability in economic activity over these years has been on the underutilization of production capacity for the domestic market. This is the part of GDP (non-exported GDP) that is most dependent on the domestic macroeconomic environment, which is what this article is about. Then comes an examination of the dynamic consequences of the region's frequent recessions, manifested in falling ratios of productive investment and a deteriorating employment situation.

Section VI presents policy lessons for a development macroeconomics approach with a view to making the transition from "financierism" to "productivism", the aim being to contribute more effectively to growth with equity. It focuses on fiscal, monetary, exchange-rate, domestic finance and capital account policies. Section VII concludes.

II

Deep economic reforms and poor economic growth since the 1990s

During the gestation of the so-called Washington Consensus, inflation was an extremely serious problem in a number of the region's countries. Consequently, the reformers of the 1990s gave priority to combating it and imposing fiscal discipline. As one ingredient of this, they sought to insulate monetary management against pressure from governments running budget deficits. This entailed a tendency whereby central banks came to operate monetary and exchange-rate policies independently of other areas of macroeconomic policy, with their actions confined to controlling inflation as a "primary if not exclusive" goal (Blanchard, Dell'Ariccia and Mauro, 2010, p. 3).

By the mid-1990s, inflation was under control; since 1997, average annual rates have been in single digits. Control of inflation was naturally associated with substantial improvements to fiscal balances and their financing. With these two important achievements and the abandonment of public-sector intervention in the region's markets, the approach in fashion assumed that economic growth would arise spontaneously (see World Bank, 1997; IDB, 1997; Fischer, 1993). In parallel with macroeconomic achievements, far-reaching

[1] Ffrench-Davis (2006) examines the reforms and their effects. See also IDB (1997), World Bank (1997), ECLAC (1998 and 2000), Williamson (2003), World Bank (2005, in an interesting self-critical reaction), Rodrik (2006) and Ocampo (2008).

liberalization of imports (as one main stimulus for exports), domestic financial markets and the capital account was expected to play a strategic role as a driver of development.

It is clear that, with variations and to differing degrees, most of the countries of Latin America met these requirements of neoliberal macroeconomic balance as laid down by the Washington Consensus. Even the performance of export volumes was satisfactory as they expanded vigorously, growing at a rate one third faster than world trade.

Nonetheless, the results in terms of economic growth and equity have been poor. As table 1 shows, annual GDP growth (which includes production of exportables and non-exportables) averaged just 3.2% between 1990 and 2008, a far lower rate than East Asia's and similar to that of the United States, whose per capita income is four times as high as the region's. These averages include the catch-up of the post-2003 boom; GDP growth averaged 5.4% in the five years from 2004 to 2008, a figure not seen since the 1970s. As the boom came to an abrupt halt in 2009, with actual output and employment falling because of the global crisis, the calculation ends in 2008 so that the evaluation and the quantitative data underlying it can be focused on more structural aspects.[2]

The data on GDP variability and the negative effects caused by it show that a macroeconomic approach focusing on the two pillars referred to (low inflation and fiscal discipline) proved unsatisfactory from the perspective of stability in the real economy, which is where GDP is generated. In fact, sharp fluctuations are observed in the GDP growth rate. These fluctuations, which affected the great majority of Latin Americans, were due not to sudden structural or microeconomic changes but to major swings in aggregate demand and the exchange rate (which affects its composition), and in the expectations or mood of economic actors. All of these are macroeconomic variables.

The greatest determinant of these macroeconomic changes, which generated recessive gaps between

[2] To sustain the modest average of 3.2% since 1990 through to 2012, cumulative GDP growth of 15% would be required in the three years from 2010 to 2012.

TABLE 1

Latin America (19 countries): GDP growth rates, 1971-2009
(Annual percentage averages)

	1971-1980	1981-1989	1990-1997	1998-2003	2004-2008	1990-2008	2009
Argentina	2.8	-1.0	5.0	-1.3	8.4	4.0	0.9
Brazil	8.6	2.3	2.0	1.5	4.7	2.6	-0.2
Chile	2.5	2.8	7.0	2.7	4.9	5.4	-1.5
Colombia	5.4	3.7	3.9	1.1	5.5	3.6	0.8
Mexico	6.5	1.4	3.1	2.9	3.5	3.1	-6.5
Peru	3.9	-0.7	3.9	2.0	7.6	4.4	0.9
Uruguay	2.7	0.4	3.9	-2.1	8.3	3.2	2.9
Venezuela (Bolivarian Republic of)	1.8	-0.3	3.8	-2.7	10.3	3.5	-3.3
Latin America (19)							
Total GDP	5.6	1.3	3.3	1.4	5.4	3.2	-1.9
GDP per worker	1.7	-1.5	0.6	-1.1	3.0	0.6	-3.8
Per capita GDP							
Latin America (19)	3.0	-0.8	1.5	-0.2	4.0	1.7	-2.9
Asia (6)	4.9	5.0	5.6	2.0	3.5	3.9	-1.0
United States	2.2	2.4	1.6	2.1	1.4	1.7	-3.3
World	1.9	1.4	1.2	1.3	2.2	1.5	-3.2

Source: prepared on the basis of data from the Economic Commission for Latin America and the Caribbean (ECLAC), World Bank and International Monetary Fund (IMF). Figures for 2009 are provisional.
Note: Asia (6) includes the Republic of Korea, Indonesia, Malaysia, Philippines, Thailand and Taiwan province of China, excepting 1971-1980 for the last of these.

GDP: gross domestic product.

potential and actual GDP over much of the 1990-2009 period, have been cyclical variations in capital inflows and outflows.[3]

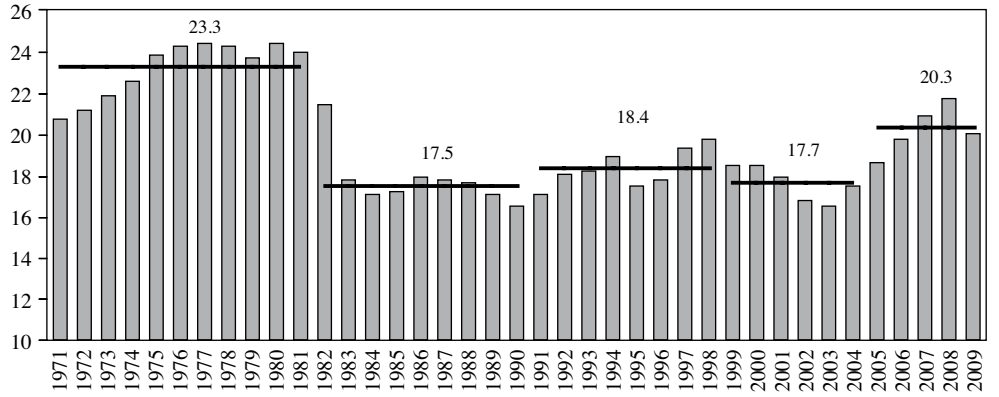
The low figures for growth indicate that the per capita GDP gap between the region and the developed countries has remained very high. The data available for 2008 show that per capita GDP in Latin America stood at just 27% of the level enjoyed by the inhabitants of the Group of Seven (G-7) and 23% of that of the United States. In addition, an enormous social gap remains, as the ratio between the tenth and first income deciles in the Latin American countries is almost treble that of the G-7 (a multiple of 34 as opposed to 12).

The dynamism of GDP depends on a number of factors, a very important one being the investment ratio. Spending on equipment and machinery, commercial and residential construction and infrastructure, which constitute gross fixed capital formation (GFCF), is closely associated with the macroeconomic environment that productive investors face and anticipate for the future. It transpires that the capital formation ratio has been remarkably low (see figure 1), compared both to that of successful emerging economies and to what the region itself achieved in the 1970s. In 1990-2008, the GFCF ratio averaged 18.5% of GDP, as against 23.3% in the 1970s.

[3] The resulting macroeconomic instability is not only recessionary and growth-depressing, but also has a markedly regressive bias (ECLAC, 2010; Ffrench-Davis, 2010b).

FIGURE 1

Latin America: gross fixed capital formation, 1971-2009
(Percentages of GDP)



Source: based on data from the Economic Commission for Latin America and the Caribbean (ECLAC).
Note: In 2000 prices. Figures for 2009 are preliminary. The figures above the horizontal lines are annual averages for the respective periods.

GDP: gross domestic product.

III

Price stability versus instability in the real economy

The performance of the Latin American countries has been shaped by a macroeconomic environment in which the main actors —businesses, workers, investors and the State— have had to cope with considerable fluctuations in aggregate (or domestic) demand, economic activity and macroeconomic prices (ECLAC, 2000, chapter 6; ECLAC, 2010, chapter II; Ffrench-Davis, 2006, chapters I and II). Figure 2 tellingly illustrates the "rollercoaster" behaviour of aggregate demand.

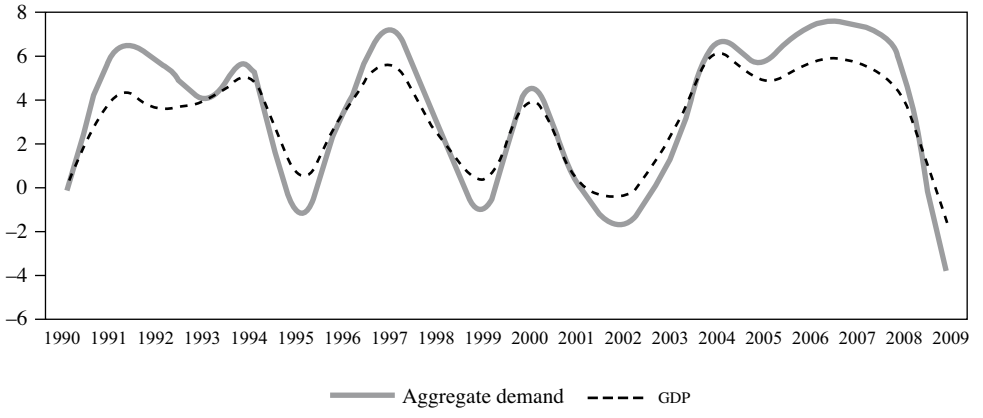
In these two decades of Washington Consensus reforms, the macroeconomy has been a determining factor in the volatile and unsatisfactory performance of regional output. It can be seen that fluctuations in demand are quickly followed by fluctuations in GDP; by definition, this involves fluctuations in the utilization rate of available capital and labour. If the economy were in macroeconomic balance, meaning that there was no substantial "recessive gap" between potential GDP (GDP*) and actual GDP (GDPa),[4] strong and persistent growth in domestic demand (like that recorded during the 1990-1994, 1996-1997 and 2004-2007 periods) would have been followed by a fairly steady evolution of the GDP trend and by upsurges in inflation and a deterioration in the external balance similar to the level of additional growth in aggregate demand, and this has not usually happened; instead, it has brought upsurges in GDPa, something that is possible only if there is a gap between the two measures of GDP.

The conclusion, which has major implications, is that since the 1980s the region has routinely been operating well below its production frontier, with

[4] This definition of a recessive or output gap differs from the one used modally (two consecutive quarters of falling GDP). The modal definition seems relevant in economies with small fluctuations in economic activity, but not in the countries of Latin America.

FIGURE 2

Latin America (19 countries): aggregate demand and GDP, 1990-2009
(Annual percentage growth rates)



Source: R. Ffrench-Davis, *Reforming Latin America's Economies after Market Fundamentalism*, New York, Palgrave Macmillan, 2006, and updated figures from Economic Commission for Latin America and the Caribbean (ECLAC), *Time for equality: closing gaps, opening trails* (LC/G.2432(SES.33/3)), Santiago, Chile, May 2010, figure II.5, for 19 countries.

GDP: gross domestic product.

fluctuations that have carried output closer to or further from potential GDP, but without this ever remaining there. When GDPa has come close to GDP*, it has usually been accompanied by large external deficits influenced by excessive currency appreciation. This represents a serious macroeconomic imbalance.

Whereas aggregate demand fluctuations prior to the 1990s were often due to fiscal deficits financed by printing money, it can be said as a rule that recent fluctuations have mainly been caused by external shocks in the capital account and terms of trade. Figure 3 presents an index of external shocks suffered by the region, including fluctuations in the terms of trade and net capital flows (net resource transfers) and their relationship with the evolution of aggregate demand.

One very important point is that, in many countries, fluctuations in the excess of spending over output (i.e., the external deficit) have largely been confined to the private sector.[5] The fact that the fiscal accounts have become more stable and balanced allows the conclusion that instability in aggregate demand and the external balance is mainly a private-sector problem, as documented by Marfán (2005). This does not mean that fiscal policy has been fully balanced, efficient or effective in serving development and combating inequality; it simply means that procyclical fluctuations in the external balance have been concentrated more in the private sector than in the public accounts.

This behaviour has generally been due to the signals arising from the combination of a large supply of external financing and permissive, procyclical domestic macroeconomic policies (Kaminsky, Reinhart and Vegh, 2004; Ocampo, 2007), many of them being lauded by the financial markets and risk rating agencies.

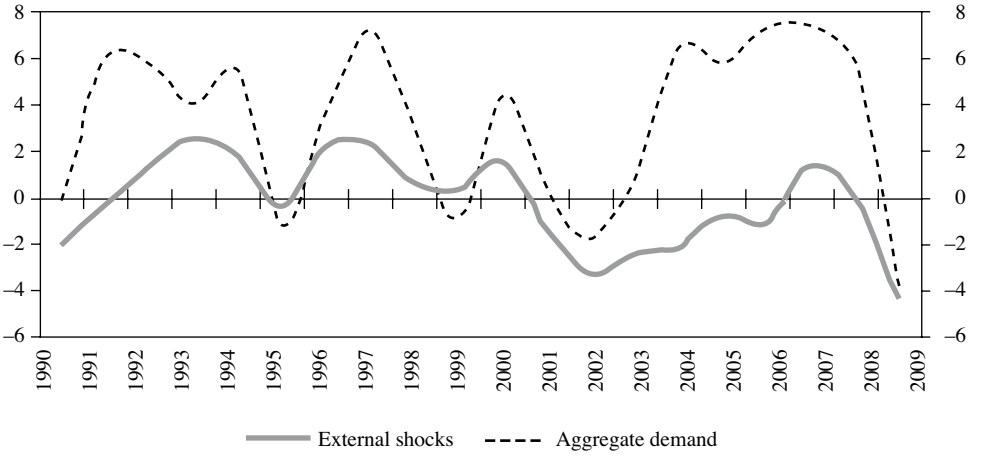
An adjustment process similar to the one seen in 1990-1994 and 1996-1997 took place again in 2004-2008, although with a boom essentially driven

[5] The transition towards imbalance has not started in national economies but has usually been caused by a positive external financial shock. The combination of growing external deficits with the build-up of international reserves during upturns is evidence of this. Following the initial external shock, the flow tends to become endogenous as a result of procyclical domestic monetary and exchange-rate policies. An examination of different national episodes reinforces these conclusions for the region as a whole. See Ffrench-Davis (2006), chapter VII.

FIGURE 3

Latin America (19 countries): external shocks and aggregate demand growth, 1990-2009

(Percentages of GDP, annual growth rates)



Source: R. Ffrench-Davis, *Reforming Latin America's Economies after Market Fundamentalism*, New York, Palgrave Macmillan, 2006, and updated figures, based on official data from the Economic Commission for Latin America and the Caribbean (ECLAC).
Note: External shocks include net transfers of resources from abroad plus the terms-of-trade effect, both measured as percentages of gross domestic product (GDP). Net resource transfers include net capital flows (including errors and omission) plus the net factor income balance plus the net current transfers balance, excluding emigrants' remittances.

on this occasion by improvements in the terms of trade. A substantial part of this improvement involved higher public revenues, reflected in a reduction of public liabilities and sometimes the establishment of stabilization funds, with a rising primary fiscal surplus between 2003 and 2007 (ECLAC, 2009a and 2010, chapter II). The scale of the terms-of-trade improvement meant that the region had a substantial current account surplus at that time. The region was better placed now than in the two previous cycles, thanks to the build-up in its own resources and the reduction of liabilities. Consequently, when the external balance abruptly reversed in 2008-2009 with the international financial crisis, many of the region's governments were able to implement countercyclical policies to mitigate the recessionary and regressive effects of contagion (ECLAC, 2009a).

In parallel with their repercussions for aggregate demand, each of the cycles of expansion in the supply of external financing tended to generate a process of currency appreciation in most of the Latin American countries. Expectations of persistent appreciation encouraged financial agents, operating within the time horizon of the appreciation outlook for local currencies, to channel additional funds into the region.[6] The experience of the Latin American countries has been that the real exchange rate, a macroprice that is vital to decisions about production and spending on tradables, has behaved in an extremely procyclical way. Exchange-rate movements have been strongly correlated with cyclical financial capital flows (ECLAC, 2010, figure II.8). Every upsurge in the supply of funding has routinely led to large currency appreciations, and these have repeatedly resulted in overshooting of the current account. The combination of an open capital account, large liquid liabilities and emerging expectations of depreciation have led to large-scale and usually sudden capital outflows, generally accompanied by traumatic devaluations, once the markets have become aware of these vulnerabilities.

An outstanding implication for policy design is that a development strategy that is supposed to be led by exporting success cannot entrust the setting of the exchange rate to the "short-termist" behaviour of some financial agents; opting for this approach denotes a severe policy inconsistency.

IV

Intrinsically procyclical financial flows

An outstanding feature of recent macroeconomic crises in East Asia and Latin America is that they have affected economies classified as "successful" by international financial institutions, financial agents and risk rating agencies.[7] As a consequence, emerging economies have been "rewarded" with large flows of private capital and diminishing spreads, in parallel with a build-up of increasing volumes of external liabilities.

The Latin American countries have thus moved into areas of vulnerability: varying combinations of growing and highly liquid external liabilities; domestic credit booms; currency and maturity mismatches; substantial external deficits; appreciated exchange rates; high stock market price/earnings ratios; high prices for luxury real estate; low rates of productive investment. At the same time, macroeconomic expectations have largely come to be dictated by the opinions of agents specializing in short-term segments of the financial market.

There is a very substantive literature on sources of financial instability: information asymmetries between lenders and borrowers and a failure to properly assimilate the negative externalities generated by each agent (in the form of growing vulnerability) have created the basis for cycles of abundance and scarcity of external financing (Krugman, 2000; Rodrik, 1998; Stiglitz, 2000; Harberger, 1985). As Heymann (2000) and Ocampo (2007) have emphasized, finance deals with the future, and concrete "information" about this is obviously not available. The tendency to equate opinions and expectations with "information"

[6] If appreciation is seen as lasting, this process will tend to discourage investment in the production of tradables that are intensive in local inputs. Consequently, it is very important to observe what happens to exchange rates during the expansionary phase or boom. This is when external imbalances and currency and maturity mismatches tend to arise.

[7] See Fanelli (2003), Frenkel (2003) and Reisen (2003) for complementary analyses.

contributes to a herd mentality and to multiple equilibria. And there have in fact been episodes of runaway contagion, first of excessive optimism and then of excessive pessimism, in the financial crises experienced over the last three decades, these imbalances often being encouraged by the risk rating agencies (Reisen, 2003).

An obvious contagion of overoptimism among lenders tends to be categorized as risk "appetite" among the agents following the "leaders", but what prevails is ignorance or underestimation of the underlying risks.[8] Meanwhile, as discussed below, the "leaders" tend not so much to have a particular appetite for risk as to believe that capital gains are assured. As regards borrowers, at times of overoptimism the evidence is that most of them do not borrow with the intention of not repaying or in the hope of being bailed out or benefiting from a moratorium. What usually prevail are rather expectations of large benefits —from continued currency appreciation, for example. Borrowers also fall victim to financial euphoria during booms.

Beyond these factors, two further characteristics of financial creditors are of vital relevance in explaining why they tend to exhibit intrinsically procyclical behaviour. One is the particular nature of the leaders acting on the supply side. There are natural asymmetries in the behaviour and goals of different economic agents. Agents oriented towards the financial markets are specialists in liquid investment, operate within short time horizons and thus are extremely sensitive to changes in the variables affecting short-term returns.

The second characteristic is the gradual spread of information about investment opportunities in emerging economies among agents who are in a position to expand supply. Agents in the different financial market segments are gradually attracted to new international markets as they learn of profitable opportunities in emerging economies that they had hitherto overlooked or been unaware of. This explains, on the supply side, why capital flows into the countries of Latin America (in 1977-1981, 1991-1994, 1995-1997 and 2004-2007) have followed a growth path over periods of several years rather than there being sudden one-off upward shifts in the supply of capital.

Feedback effects have been generated by the existence of installed capacity (potential GDP) that has been underused at the start of each of these processes and gradually brought back into operation during the upturn; this is something the authorities, markets and certain econometricians have often wrongly interpreted as a persistent structural increase in total factor productivity (TFP).[9] All this is self-reinforcing so that some variables – stock markets, exchange rates, risk ratings and real-estate prices – can move in a particular direction, first recovering and then overshooting so that they move away from sustainable equilibria for prolonged periods, offering economic agents the "assurance" that financial markets will move in only one direction and stimulating capital flows that pursue capital gains (rent-seeking flows).

This being so, it is important to highlight the significance for public policy design of the distinction between two different types of volatility in financial capital flows: short-term or random walk fluctuations and medium-term instability. The latter means that variables such as the exchange rate, stocks and shares and real-estate prices can move persistently in a particular direction, giving the market the false assurance already mentioned of asset prices and returns moving in a single direction. This stimulates further continuing flows that at some point become increasingly detrimental to macroeconomic fundamentals, but that still offer successive short-term windfall gains. These agents naturally specialize in the search for capital gains rather than productivity gains, until asset prices and the real exchange rate reach what are clearly outlying levels. Then someone sounds the alarm and there is a rush to reverse flows, with a strong and costly procyclical bias. Unlike fixed capital investment, which is to a large degree irreversible, this financial capital is wholly reversible.

Lenders' sensitivity to bad news will increase greatly at some point (and probably quite suddenly) once the country has entered "areas of vulnerability". Then lenders will take note of: (i) the volume of assets they hold in that market, (ii) the degree to which that market depends on additional net flows, something that is connected to the current account deficit, (iii) the level of exchange-rate appreciation, (iv) share price/earnings ratios and (v) the country's stock of short-term and liquid liabilities. It is therefore unsurprising

[8] Calvo and Mendoza (2000) examine how globalization can spur contagion by discouraging the collection of information as it creates stronger incentives to imitate the portfolio of the market. This introduces a new information asymmetry, this time between market "leaders" and "followers".

[9] A systematic distinction between potential GDP and actual GDP would allow this faulty interpretation to be avoided, being an essential component of a development-oriented macroeconomic policy.

that expectations become more and more likely to reverse as valuations move further into these areas of vulnerability.

The deeper and longer-lasting an economy's incursion into areas of vulnerability, the greater the likelihood of crises and the severer their effects. This highlights the vital need to apply effective regulations to ensure that capital flows strengthen productive investment and are consistent with a sustainable macroeconomic environment. The composition, volume and deviation from trend of the flows are crucial variables. Against a background of ubiquitous structural heterogeneity, the explanation lies in the differing capacity for action and reaction of the agents typically operating in the different domestic markets.

To sum up, the interaction between two factors —the nature of agents and that of the adjustment process— accounts for the dynamic of capital flows over time: the factors leading suppliers to continue providing funds even when the real macroeconomic fundamentals are deteriorating.

Consequently, both the accumulation of external assets by providers of finance until this expansionary stage of the cycle is far advanced and the sudden subsequent reversal of flows can be considered "rational" responses by individual agents, given their short-term horizons. This is because the question of whether the real fundamentals are improving or worsening is not relevant to these investors as long as they continue to make investments motivated by expectations of short-term returns. What does matter to them is whether the indicators which are critical from their standpoint —real-estate, bond and share prices and the exchange rate— can continue to yield short-term gains and, of course, whether markets are liquid enough for them to reverse their decisions in a timely fashion if necessary. They will thus continue to originate net flows until expectations of an imminent reversal emerge.

It needs to be stressed again that, for financial operators, the most relevant variables are not the long-term fundamentals of the country's economy but the short-term returns it yields. This explains why their view of a particular country can alter swiftly and radically even though its economic fundamentals, other than foreign-currency liquidity, may remain unaltered even as financiers' mood switches from overblown optimism to overblown pessimism.

Once debtor markets have made a "sufficient" downward adjustment, of course, the opposite process arises and can be sustained for some years, examples being 1991-1994 and 1995-1997, and probably the aftermath of the global crisis of 2008-2009. In conclusion, economic agents specializing in financial investments, who might be highly efficient in their field, operate with short-term planning horizons because of their training and the rewards they can thereby obtain, and they have largely dictated macroeconomic developments owing to the decisive influence they have had on policy design in the countries of Latin America. This means that a "financierist" attitude prevails over the "productivist" approach, and this enters into conflict with the twofold objective of growth with equity, which requires better incentives to increase productivity rather than giving priority to financial rent-seeking or capital gains. For growth with equity to be sustainably achieved, the views and priorities of the different economic and social actors need to be brought back into balance.

The heterogeneity characterizing the capital account in the recent era of financial globalization makes it essential to distinguish between the behaviour and effects of its different components. Greenfield direct foreign investment and long-term credits associated with imports of capital goods are relatively stable over the cycle, and are indissolubly linked to productive investment. By contrast, financial flows have shown great procyclical volatility, and this very property of theirs means that only a small share of them have gone into the financing of productive investment (Uthoff and Titelman, 1998); these flows usually end up financing purchases of existing assets and consumption, creating bubbles and crowding out national savings. Often, indeed, they have destabilized the macroeconomy instead of stabilizing it, and have not contributed to productive capital formation.[10]

[10] Opposing positions in two important papers published by the International Monetary Fund (IMF) appear in Prasad and others (2003) and Singh (2006).

V

The recessionary and growth-depressing effects of instability in the real macroeconomy[11]

Real macroeconomic instability has recessionary and regressive effects, associated with price inflexibility, incomplete factor markets and the deep structural heterogeneity of the region's economies. The consequences, over the cycle, are various. One of them, the most standard in the literature although somewhat overlooked during these years of neoliberal fashion, is an increased disparity between aggregate supply and demand, the result being a recurrent gap between potential production capacity and its utilization, particularly in the stop phases that follow go phases. Demand is restrained in some sectors by full capacity utilization, while in others it is markedly inadequate. Consequently, in a stop-and-go situation, the instability of overall demand inevitably means that average net utilization is lower than production capacity and that actual productivity is lower than it would be in a situation of stable proximity to the production frontier. The greater the instability, obviously, the larger the recessive output gap and the worse the effects on the labour market, with increasing informality (ECLAC, 2010, chapter V).

1. Structurally heterogeneous markets and instability

The connection between inequality and instability in the real macroeconomy stems from the great structural heterogeneity characterizing developing economies. This includes the differing capacity for action and reaction of the agents typically found in different market segments (large versus small businesses, high- versus low-skilled workers, productive or GDP-generating investors versus financial investors or buyers of existing assets, productive investors versus consumers) and the asymmetries between their respective responses to the instability of economic activity and macroprices. During upturns (as opposed to a relatively stable trend in economic activity), liquidity constraints tend to be relaxed faster for consumers than for productive investors, given the weakness of the long-term segments of capital markets. Again, consumers can react faster than productive investors because the latter need to identify, design and develop new projects, which is a lengthy process. Furthermore, the irreversibility of investment means that favourable expectations have to be perceived as sustainable by long-term investors before these begin new investment processes.

The production frontier obviously sets a bound to the recovery of actual GDP; only for short periods can this exceed potential GDP. During recessions, conversely, actual GDP can fall well short of potential GDP for long periods. Consequently, economic instability is intrinsically asymmetrical and, on average, inevitably entails underutilization of potential productivity and lower actual output. Recovery increases the flow of present output until the point where existing capacity is fully utilized, but output not generated in the past cannot be recovered. As long as the recessionary gap between the two levels persists, so will the depressive effects on productive investment, the labour market and the situation of small and medium-sized enterprises and informal sectors.

In consequence, the size of the gap between actual demand and the production frontier has major static and dynamic effects. First, it affects observed productivity (actual TFP) and the returns on projects implemented. Second, higher capital utilization rates generally mean that the average employment level is higher and the workforce interacts with a larger stock of physical capital in use. The consequent rise in observed productivity means that the welfare of workers and investors (wages and profits) can improve immediately, by virtue of the higher average capacity utilization index. Fiscal revenues also rise. The usual thing is for poverty to diminish in these situations, while the probability of an income distribution improvement rises during the recovery stage. The sign of the distributive effect depends on the micro- and mesoeconomic reforms accompanying recovery. Growth in itself may be either progressive or regressive. In the first case, it is usually sustainable and increasingly "endogenized", in the second it tends to be reversible and limited (Bourguignon and Walton, 2007).

[11] A report prepared for the International Labour Organization (ILO) examines the effects of this instability, which are likewise regressive (Ffrench-Davis, 2010b).

2. Recessive gaps and the dynamics of capital formation

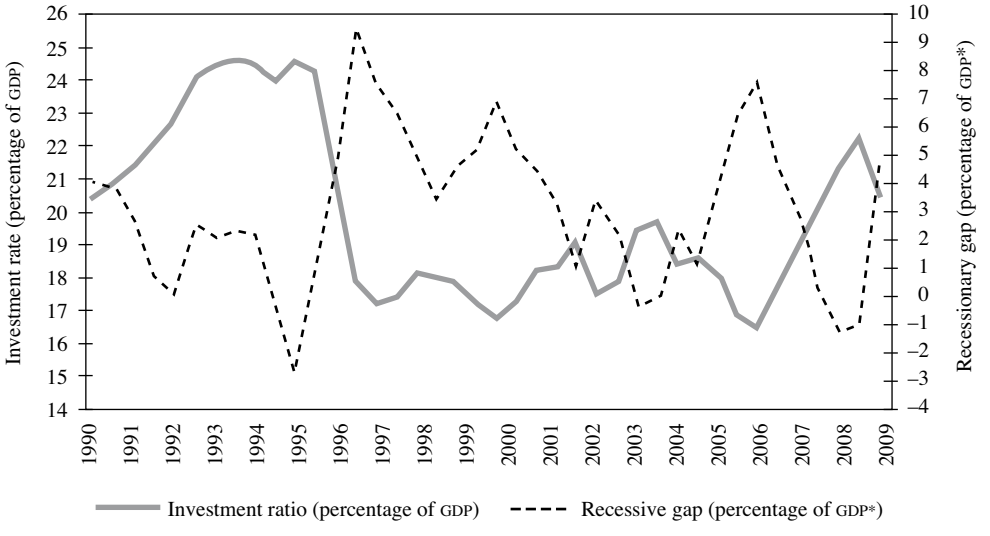
In the dynamic dimension, the degree of stability has a number of effects on the construction of the future. Higher utilization indices and the consequent increase in average actual productivity (in conventional econometrics this would appear as an increase in TFP) will tend to stimulate investment in new capacity. The dynamic effect on the investment ratio will be much more substantial if solid expectations are generated among economic actors regarding the ability of public policies to keep actual demand close to the production frontier, and if the authorities additionally undertake reforms to complete long-term capital markets while at the same time taking steps to improve labour force training and innovation.

Figure 4 shows the close relationship that has existed between the recessive gap and the investment ratio in Latin America, reflecting one of the main negative dynamic effects of the underutilization of production factors. This relationship is accounted for by a number of reasons (Ffrench-Davis, 2006, chapter III; Aizenman and Marion, 1999): (i) a large idle capacity naturally discourages investment in new productive assets; (ii) an environment in which economic activity and the exchange rate are volatile discourages irreversible investment; (iii) underutilization means lower profits and a lack of internally generated funding, usually coinciding too with a reluctance by the capital market to finance firms that have liquidity problems in recessions; (iv) the recessionary gap and its fluctuations tend to affect the quality of project evaluation; (v) the disincentives to the acquisition of new machinery and equipment dampen the technological innovation associated with these; and (vi) large recessionary fluctuations tend to depress public revenue, leading to cuts in the public investment needed to complement private investment (Easterly and Servén, 2003).

Figure 4 also shows the relevance of continuity in recovery processes and of the sustainability of

FIGURE 4

Latin America (9 countries): the recessive gap and gross investment ratio, 1990-2009



Source: R. Ffrench-Davis, *Reforming Latin America's Economies after Market Fundamentalism*, New York, Palgrave Macmillan, 2006, and updated figures from Economic Commission for Latin America and the Caribbean (ECLAC), *Time for equality: closing gaps, opening trails* (LC/G.2432(SES.33/3)), Santiago, Chile, May 2010, figure II.9, based on data from ECLAC and A. Hofman and H. Tapia, "Potential output in Latin America: a standard approach for the 1950-2000 period", *Estudios estadísticos y prospectivos* series, No. 25 (LC/L.2042-P), Santiago, Chile, Economic Commission for Latin America and the Caribbean (ECLAC), 2003. United Nations publication, Sales No. E.03.II.G.205.

Note: Includes Argentina, Bolivarian Republic of Venezuela, Brazil, Chile, Colombia, Costa Rica, Mexico, Peru and Plurinational State of Bolivia. The investment ratio measures the proportion between gross fixed capital formation and actual GDP. The recessive output gap measures the difference between actual gross domestic product (GDP) and potential gross domestic product (GDP*) as a percentage of the latter.

the equilibria that arise as the recessionary gap is progressively closed. It reveals that more prolonged periods of economic recovery have led to an ongoing escalation of the investment ratio. The two-year period 2007-2008, when investment ratios were at their highest since the 1970s, followed the prolonged recovery that had begun in 2003; thereafter, with the contagion from the crisis, the ratio shed two of the percentage points it had gained up to 2008.

The case of Chile well illustrates the effects of the macroeconomic environment on investment. After averaging about 15% of GDP in the 1970s and 1980s, and following a recessive adjustment in 1990, the ratio recovered steadily to gain about 8 points by 1998 (Ffrench-Davis, 2010c). The persistence of the process is a crucial factor in effectively stimulating investors to tie up their funds and credit for the long periods production activity requires.

Paradoxically, fluctuations in the capital formation ratio have responded far more to economic cycles than to the micro- and mesoeconomic reforms carried out in the region to raise productivity and reduce structural heterogeneity. Assuming the macroeconomy were performing well, micro- and mesoeconomic reforms would be expected to determine the evolution of the investment ratio.

Consequently, there is a clear link between volatility and long-run economic growth that operates via its effects on the volume of fixed capital investment. Capital formation, in fact, functions as a major variable in the evolution of potential GDP, feeding through to employment and TFP (Ffrench-Davis, 2006, figures III.1 and III.2; De Long and Summers, 1991). The unsatisfactory experience of the region by comparison with the Asian countries can be attributed not just to the crucial issue of productive development policies, but to macroeconomic failures and the nature of the Washington Consensus capital market reforms (see section VI).

3. The instability of a crucial macroprice: the exchange rate

Section III discussed the exchange-rate instability that has been generated by financial flows. This instability in the real exchange rate has been detrimental to the growth and diversification of exports and their integration into domestic economies (Agosin, 2007). Rates that fluctuate so much cannot be taken as effective reflections of shifting levels of "sustainable equilibrium"; "sustainable equilibrium" levels respond to the evolution of relative productivity between the domestic economy and trading partners (and to volumes of net capital inflows that are sustainable and can be absorbed efficiently). The changes these "structural" variables undergo are usually gradual rather than sudden. Consequently, the large swings in many countries' real exchange rates have generally been misalignments caused by procyclical capital flows.

Repeated cycles of currency appreciation, particularly after the substantial import liberalization that took place in the region (ECLAC, 1998, chapter V), meant that with each upturn, recovering aggregate demand from both individuals and firms has been increasingly import-intensive. Alongside a welcome increase in imports of capital goods, there have been large rises in other imports, many of them competing with savings and the output of local small and medium-sized enterprises (SMEs); thus, not only have the volume and quality of exports been crowded out, but so have production sectors that compete with imports.

4. Systemic competitiveness and the real macroeconomy

It is meaningful for the analysis to examine where fluctuations in economic activity have been located. Between the 1990-1997 and 1998-2003 periods, for example, 90% of the adjustment in the region's GDP growth (a fall of 1.9 points in the average growth rate) was concentrated in production for the domestic market, i.e., GDP that is not exported (Ffrench-Davis, 2006, chapter VI; ECLAC, 2010, chapter II). This reflects two facts, one micro, the other macro. The micro fact is the difficulty of repeatedly reallocating resources from the production of non-tradable goods to that of exportable goods and import substitutes and back. Switching policies in the region have been weakened by liberalization and changes in international trade institutions (ECLAC, 1998; Rodrik, 2008). Consequently, the main instrument actually available, namely the exchange rate, has become far more important. To decline to regulate it by permitting it to float freely without intervention by the economic authority is in stark contradiction with a strategy of export-led development.

The second point is macroeconomic. The GDP share which is not exported (about four fifths) depends on the local macroeconomic environment, while exports depend more on the global macroeconomy. The information available indicates that national

markets have been the main victims of instability (Ffrench-Davis, 2006, chapter V).[12] Here once again it is possible to appreciate the deficient quality of macroeconomic policy in the region, where it has been run very procyclically and has thus amplified rather than softened the transmission of external trade and financial shocks. The progress made over recent years remains insufficient.

[12] Vigorous growth means that non-exported GDP also expands fast. This has been the experience of emerging economies with an export model that has been successful in productive development, examples being the Republic of Korea for several decades and Chile between 1990 and 1998, when non-exported output grew by 6.5% a year (Ffrench-Davis, 2010c, table VII.6). In Latin America as a whole, on the other hand, non-exported output grew by a mere 2.7% (ECLAC, 2010, table II.2).

VI

Development macroeconomics: from "financierism" to "productivism"

Premature, indiscriminate and poorly sequenced liberalization of domestic financial markets and the capital account has become a source of costly destabilizing shocks. The high costs generated by economic cycles in the Latin American countries are related, as has been demonstrated, to the close links forged between domestic financial markets and procyclical segments of the international financial markets. As liberalization has taken place, there has been a major upsurge in financial saving without any increase in domestic saving, with a very low investment ratio and large fluctuations in economic activity and employment. The central cause is a financial market overly dominated by agents specializing in the short term rather than in productive investment. Consequently, only a small share of capital inflows have financed productive investment, a shortcoming aggravated by the financial and currency crises to which their volatility has given rise, and whose recessive effects have weighed down on capital formation by local companies and employment.

A consistent set of countercyclical fiscal, monetary, exchange-rate, domestic financial market and capital account policies is essential to foster a macroeconomic environment that allows potential GDP to be fully utilized and encourages the generation of new capacity.[13] For this to happen, such a development-friendly environment needs to be complemented by efforts to "complete" markets for capital, labour and innovation.

In economies that are highly vulnerable to external shocks, relying on just one particular policy instrument during adjustment processes can produce macroeconomic outcomes inferior to what can be achieved by the balanced implementation of the different macroeconomic policies. Distributing the adjustment across different policies usually yields superior macroeconomic outcomes in terms of macroprices that are more closely aligned with sustainable levels and actual GDP that is consistently closer to its potential.

Procyclical, volatile flows are a component of external funding, which includes the foreign savings required to supplement domestic savings if a substantial increase in the investment ratio is to be achieved. An "all or nothing" option is therefore not viable. Thus, a fundamental goal of macroeconomic policies (and of reforms to domestic financial markets) should be to reap the potential benefits of external savings in support of national development, while moderating the intensity of capital account cycles and their negative effects on domestic economic and social variables.

1. Fiscal policy

The international financial crisis has revealed the central importance of fiscal policy as a macroeconomic stabilization tool (Krugman, 2009; Griffith-Jones, Ocampo and Stiglitz, 2009; Blanchard, Dell'Ariccia

[13] ECLAC (2010, chapter II) examines these different macroeconomic policies. Ffrench-Davis (2008) provides a fuller analysis.

and Mauro, 2010). The concept of structural fiscal balance is an outstanding component of any countercyclical policy package. Its essential feature is budgetary measurement over the economic cycle to estimate what level of public spending would be consistent with trend public-sector revenue or full employment of productive capacity. Stabilizing spending in this way makes it more efficient and insulates it from cyclical fluctuations in fiscal revenue, while mitigating or removing the procyclical bias of an annually balanced budget policy. A number of countries in the region have been developing a form of budgetary planning that is not constrained by narrow annual limits. Chile is a particular example (Ffrench-Davis, 2010a).

Part of an approach of this kind is the creation of stabilization funds for fiscal revenues from exports whose prices are highly unstable. These funds can help to stabilize normal fiscal expenditure, provide supplementary financing for crisis situations like that of 2009, and additionally stabilize markets for foreign exchange by regulating its supply. For this, there is once again an essential need for close coordination between the fiscal authorities and those responsible for exchange-rate policy, which are usually based in different institutions (Martner and Tromben, 2004).

2. Monetary policy

Even if the countercyclical role of fiscal policy is successfully enhanced, however, this will not usually be enough. A crucial fact in the region is that fiscal spending accounts for only a fraction (around a fifth) of aggregate demand. Little will be achieved by operating an active fiscal policy over the cycle if other policies with a great influence on private-sector spending are dependent upon volatile flows and the opinions of procyclical financial agents.

Monetary policy, along with the independence achieved as regards financing of the fiscal balance, has been key to the large reduction in inflation rates. However, price stabilization can go together with large variations in the gap between potential and actual GDP (Blanchard, Dell'Ariccia and Mauro, 2010), as has been clearly demonstrated in practice. One critical feature of monetary policy is the weighting given to each macroeconomic variable in the work of the central bank; another is the coordination with other economic authorities already referred to. In a region where inflation is mainly in single digits, the tendency for central banks to ignore other macroeconomic goals has lost the justification it might have had in earlier contexts of fiscal irresponsibility and high inflation. Mediocre outcomes as regards growth, employment and capital formation can be partly explained by the way these have been left out in the cold by policies that have focused on inflation targets without adequately considering the effects on these other areas that have such a critical direct impact on development.

For monetary policy to be genuinely countercyclical and contribute to development, explicit account must be taken of its repercussions on other macroeconomic variables such as economic activity, the recessive gap, external equilibrium and employment, with a sustainable balance between different objectives and the pursuit of indispensable coordination with fiscal policy rather an exclusive concentration on maximizing anti-inflation effects. Real exchange rates are a macroeconomic variable that has brought severe conflict with anti-inflation policy. Inflation targets have frequently been met thanks to exchange-rate appreciation that has destabilized the economy.

3. Exchange-rate policy

The exchange rate is a macroeconomic variable that is essential for the sustainability of macroeconomic equilibria and resource allocation. Conventional approaches whereby the only exchange-rate options are a fixed nominal rate or a completely free float assume that the market will benignly set a sustainable equilibrium real exchange rate. Formally, a number of the region's countries have adopted a free exchange-rate regime. Although central banks have intervened on a number of occasions to dampen fluctuations, real exchange rates have responded very strongly to changes in the balance of payments, more so than to changes in the current account.

Although the predominance of free-floating regimes prevented the kind of currency crises characteristic of fixed-rate regimes, many of the region's currencies became extremely sensitive to procyclical changes in the supply of external funding. A severe contradiction therefore arose, with serious negative consequences for resource allocation and, especially, accumulation. Reforms to liberalize imports ushered in a leading role for tradable sectors, which meant that the exchange rate became crucial to international competitiveness (Williamson, 2000; Agosin, 2007; Rodrik, 2008; Eichengreen, 2008). Paradoxically, the authorities adopted a policy that

led to volatile exchange rates, dominated by short-term financial operators.

This obviously distorts project evaluation, encourages speculative rather than productive investment, artificially crowds out local production of importable tradables (many produced by SMEs) and discourages producers from adding value to exports.

This serious failure of exchange-rate policy is a severe constraint on export-led development strategies, particularly as regards non-traditional exports and those with greater value added that generate externalities and interact with SMEs. The management of exchange-rate policy is an essential component of the set of variables required for success in this area.

Intermediate regimes involving managed exchange-rate flexibility, such as different varieties of crawling pegs, dirty floats or both, represent a serious pragmatic attempt to correct this contradiction (Williamson, 2000). Neoliberal approaches tend to represent any exchange-rate intervention as going against "the market" and being doomed to failure. However, the idea behind the alternative approach we favour is to ensure that the real forces of the market —producers of exportables and importers and producers of importables, who are the major players in trade for production development and equity— are the ones that prevail in the setting of the exchange rate. This is the "market" that ought to set rates, rather than the market of short-term operators and rent-seekers imposing their interests over those of the drivers of innovation and productivity increases. Consistent, selective intervention by the economic authority is essential for this, even though it is obviously not infallible. It is always necessary to weigh the risk of mistakes when acting against the high likelihood of error when the exchange rate is left to float freely in a context of large flows of procyclical funds.

In summary, exchange-rate policy requires a far-reaching correction if it is to be consistent with a development strategy in which the production of tradable goods and services plays a central role. This would also contribute to systemic competitiveness, i.e., to the development of production capacity for both the domestic and external markets. The domestic market is home to the great majority of workers and firms. Improved systemic competitiveness achieved in this way helps to reduce domestic structural heterogeneity, a precondition for greater equality in the labour market and between the array of different-sized business.

4. Creating deeper capital markets to finance development

As this article has emphasized, the capital market has a major influence on macroeconomic equilibria, employment and capital formation. This is due to two features of the region's economies. A very prominent one, first, is the "incompleteness" of capital markets, with some segments weak or non-existent. The distributive and resource allocation effects of capital market failures are aggravated by the marked structural heterogeneity between different economic agents, to the detriment of SMEs, low-skilled workers, innovation and agents with limited assets. Heterogeneity in access to financing reinforces inequalities in productive capabilities and participation in broader markets, in a vicious circle that condemns less well-capitalized production units to vulnerability and makes it hard for them to grow.

The close relationship with more volatile international financial markets that has been a feature of recent decades has exacerbated these shortcomings and bears part of the responsibility for low levels of productive investment and the fragility of labour markets. Indeed, that link has contributed to an intensification of instability.

(a) *From the Washington Consensus to innovative development financing*

Where interest rates and maturities are concerned, high financial costs have been the rule. Instead of "deep markets" for investment financing, as the neoliberal approach expected, the result has been markets that are deeply segmented and excessively focused on the short term (Stallings and Studart, 2005). The Washington Consensus reforms to domestic capital markets have actually tended to weaken development banking and the long-term segment. Consequently, these reforms have not been characterized by consistency with the recommendations of the Monterrey Consensus (United Nations, 2007), whose goal was to increase the resources going to economic and social development and give an effectively inclusive and countercyclical character to the working of capital markets.

National financial systems were certainly quite imperfect and inadequate before the Washington Consensus reforms, notwithstanding which they financed a higher investment ratio in the 1970s than was achieved under the neoliberal reforms, and supported substantially higher GDP growth

(see table 1 and figure 1). It would be unwise in the extreme to ignore these two realities and the contrast between them.

Consequently, the reform to the reform of national financial systems should be aimed at channelling resources into savings and productive investment, which generates sustainable jobs. The institutional structure required includes a vigorous long-term intermediation segment to allocate savings to productive investment, and there need to be prudential and countercyclical regulations. This system needs to include an active role for public- and private-sector development banks (ECLAC, 2010).

In emerging economies like those of Latin America, domestic markets are extremely difficult to reform when the capital account is indiscriminately open. Effective and efficient countercyclical regulation of the capital account emerges as an unavoidable condition of progress towards a development macroeconomics, with space for monetary policy and exchange-rate sustainability (Ffrench-Davis, 2006, chapters II and V; Ocampo, 2008). Regulation of capital flows can create space for consistent and countercyclical exchange-rate and monetary policies simultaneously.

Extreme liberalization of external financing, like that introduced since the 1990s, entails integration into the most speculative segments of international financial markets. Consequently, the most dynamic segment of the capital market has been large-scale financial activity involving short-term inflows and outflows, characterized not just by its procyclical volatility but also by the tenuousness of its links with productive investment.

Regulation of the more volatile capital accounts can act as a countercyclical macroeconomic instrument, acting on boom and bust cycles right at their source. It can mitigate pressures for currency appreciation and make it possible to adopt contractionary monetary policies in periods of financial euphoria. Also relevant is that the use of precautionary regulations during booms subsequently creates space for expansionary monetary and fiscal policies in episodes like the global crisis.

At the same time, countercyclical regulation of capital account inflows and outflows provides room for a reorganization of the domestic financial system aimed at channelling resources into productive investment with a bias towards inclusiveness, helping to reduce the structural heterogeneity between different economic and social sectors.

(b) *National experiences with countercyclical measures*

On the whole, experience with the use of restrictions on short-term or liquid capital inflows has shown them to play a useful role in creating space for countercyclical macroeconomic policies, thereby contributing to growth and employment (Ocampo, 2008; Stiglitz, 2000; Williamson, 2003). These restrictions are designed to create a more stable macroeconomic environment during booms and minimize costly recessionary adjustments in the retreat from positions of disequilibrium due to domestic overheating or external imbalance.

The success of the Chilean experience in the first half of the 1990s is a robust proof of the effectiveness of countercyclical regulations. Having recently returned to democracy, in 1990 Chile was confronted with a larger supply of external financing (relative to GDP) than other nations in Latin America, owing to its better economic performance, smaller economy and renewed political stability. This supply of funding was perceived by the authorities as an excess that would destabilize the country's macroeconomy (particularly aggregate demand and its consistency with potential GDP and a sustainable external balance) and its export strategy.

Accordingly, the authorities regulated the amount and composition of capital inflows by adding to the cost of short-term flows of funds, whether in the form of credits or stock market investments. This was done by establishing a non-interest-bearing reserve requirement (*encaje*), calculated as a proportion of the gross flow and to be held at the central bank for a given period, the rate of the *encaje* and period varying with the supply of external funding. By regulating the composition and amount of inflows, the reserve requirement provided effective room for simultaneously implementing active countercyclical monetary and exchange-rate policies (Magud and Reinhart, 2006; Edwards and Rigobon, 2009). They allowed Chile to maintain a level of aggregate demand consistent with its productive capacity and a sustainable exchange rate. These equilibria led to a substantial increase in the investment ratio and in the potential and actual GDP growth rate, with average GDP growth exceeding 7% a year. In the second half of the 1990s, Chile went along with more fashionable policy thinking and allowed the regulatory power of the reserve requirement to weaken before liberalizing the capital account in 2001 (Ffrench-Davis, 2010c, chapters VIII and IX; Le Fort and Lehmann, 2003).

It thus fell victim to contagion from the Asian crisis in 1999 and saw its GDP growth rate drop from 7.1% in 1990-1998 to 3.8% in 1999-2008.

The region has had a great variety of experiences with capital account regulation. Controls on capital outflows, combined with large fiscal deficits and clearly overvalued exchange rates, are usually very inefficient and destabilizing. The purpose of the capital account regulation proposed here is to achieve sustainable equilibria in the real macroeconomy, which is the opposite of seeking to perpetuate imbalances. In pursuit of these equilibria, some interesting regulatory experiments have recently been implemented in the region, largely to prevent excessive currency appreciation. Mention may be made of the cases of Argentina, Brazil, Colombia and Peru (ECLAC, 2009b).

Other regions can show many positive recent or current experiences with capital account regulation. The Republic of Korea is an example of an economy that maintained strict capital controls during a remarkable spell of economic growth lasting a third of a century. Following liberalization of capital inflows in the early 1990s, which led to its 1998 crisis, the country has maintained controls on residents' transfers of funds abroad (Mahani, Shin and Wang, 2005). Meanwhile, more traditional controls like those applied in China and India (such as bans on short-term financial borrowing, stock market investment quotas and controls on capital outflows) have proved very effective in achieving the macroeconomic objective of a more drastic weakening of the link between the domestic economy and the volatility of international markets (Ocampo, 2008). Both nations have successfully withstood the current crisis and their controls created the conditions for the remarkable recovery programmes now in operation.

In summary, reforms of the Washington Consensus reforms are urgently needed and should give greater priority to the linkages of the financial system (both local financial markets and the capital account) with the domestic investment process and the domestic economy than to those with short-term external financial markets. They should also improve the stability of domestic demand and macroprices such as the exchange rate.

VII

The great macroeconomic challenge

These reflections come at a time when the world seems to have avoided what many analysts feared could have been a depression similar to that of the 1930s. That this did not occur was due not to good fortune or self-correcting markets, but to deliberate public action to correct the global macroeconomic situation. The correction ran counter to the dominant paradigm of a passive State and neutral economic policies. The response of the world's leading economies has been a vigorous countercyclical fiscal policy. With collapse avoided, fortunately for the great majority of the world's population, the task now is to deal with what is still a recessive situation and complete a set of corrections, both in global institutions and in each of our countries.

This article has documented how the choice of macroeconomic approach decisively affects the stability and speed of growth, and influences the degree of equity built into the structure of domestic markets. National financial systems have a crucial role to play in capturing savings and allocating them to investment. Foreign capital, meanwhile, can play a valuable role as a supplement to local savings; the composition and stability of flows are crucial here. Indiscriminate financial liberalization proved highly inefficient as a way of achieving economic development and real macroeconomic stability and of reducing inequality. The global crisis —an example of the risks of unregulated "financierism"— has created space for more pragmatic policies and countercyclical regulation of the capital account in the region.

The way reform of national capital markets and their relationship with international capital markets is undertaken represents a critical challenge on the road to a sustainable macroeconomics that is conducive to economic and social development.

(Original: Spanish)

Bibliography

Agosin, M. (2007), "Trade and growth: why Asia grows faster than Latin America", *Economic Growth with Equity: Challenges for Latin America*, R. Ffrench-Davis and J.L. Machinea (eds.), New York, Palgrave Macmillan.

Aizenman, J. and N. Marion (1999), "Volatility and investment: interpreting evidence from developing countries", *Economica*, vol. 66, No. 262, London, London School of Economics and Political Science.

Akerlof, G. (2007), "The missing motivation in macroeconomics", *American Economic Review*, vol. 97, No. 1, Nashville, Tennessee, American Economic Association, March.

Blanchard, O., G. Dell'Ariccia and P. Mauro (2010), "Rethinking macroeconomic policy", *IMF Staff Position Note*, SPN/10/03, Washington, D.C., International Monetary Fund, February.

Bourguignon, F. and M. Walton (2007), "Is greater equity necessary for higher long-term growth in Latin America?", *Economic Growth with Equity: Challenges for Latin America*, R. Ffrench-Davis and J.L. Machinea (eds.), New York, Palgrave Macmillan.

Calvo, G. and E. Mendoza (2000), "Rational contagion and the globalization of securities markets", *Journal of International Economics*, vol. 51, No. 1, Amsterdam, Elsevier.

De Long, J.B. and L. Summers (1991), "Equipment investment and economic growth", *Quarterly Journal of Economics*, vol. 106, No. 2, Cambridge, Massachusetts, The MIT Press.

Dodd, R. and S. Griffith-Jones (2007), "Report on derivatives markets: stabilizing or speculative impact on Chile and a comparison with Brazil" (LC/W.134), Santiago, Chile, Economic Commission for Latin America and the Caribbean (ECLAC).

Easterly, W. and L. Servén (2003), *The Limits of Stabilization: Infrastructure, Public Deficits and Growth in Latin America*, Stanford, California, Stanford University Press.

ECLAC (Economic Commission for Latin America and the Caribbean) (2010), *Time for Equality: Closing Gaps, Opening Trails* (LC/G.2432(SES.33/3)), Santiago, Chile, May.

_____ (2009a), *Economic Survey of Latin America and the Caribbean 2008-2009* (LC/G.2410-P), Santiago, Chile, July. United Nations publication, Sales No. E.09.II.G.2.

_____ (2009b), *Preliminary Overview of the Economies of Latin America and the Caribbean 2009* (LC/G.2424-P), Santiago, Chile. United Nations publication, Sales No. E.09.II.G.149.

_____ (2000), *Equity, Development and Citizenship* (LC/G.2071/Rev.1-P), Santiago, Chile. United Nations publication, Sales No. E.00.II.G.81.

_____ (1994), *Latin America and the Caribbean: Policies to Improve Linkages with the Global Economy* (LC/G.1800(SES.25/3), Santiago, Chile.

Edwards, S. and R. Rigobon (2009), "Capital controls on inflows, exchange rate volatility and external vulnerability", *Journal of International Economics* 78, Amsterdam, Elsevier.

Eichengreen, B. (2008), "The real exchange rate and economic growth", *Working Paper* No. 4, Washington, D.C., Commission on Growth and Development.

Fanelli, J.M. (2003), "Micro-macro interactions, competitiveness and sustainability", *Development Economics and Structuralist Macroeconomics*, A. Dutt and J. Ros (eds.), Aldershot, Edward Elgar.

Ffrench-Davis, R. (2010a), "Latin America: the structural fiscal balance policy in Chile", *Journal of Globalization and Development*, No. 1, Berkeley Electronic Press.

_____ (2010b), *Macroeconomía para el empleo decente en América Latina y el Caribe*, Santiago, Chile, International Labour Organization.

_____ (2010c), *Economic Reforms in Chile: From Dictatorship to Democracy*, New York, Palgrave Macmillan.

_____ (2008), "From financieristic to real macroeconomics: seeking development convergence in EEs", *Working Paper*, No. 44, Washington, D.C., Commission on Growth and Development.

_____ (2006), *Reforming Latin America's Economies after Market Fundamentalism*, New York, Palgrave Macmillan.

Fischer, S. (1993), "The role of macroeconomic factors in growth", *Journal of Monetary Economics*, vol. 32, No.3, Amsterdam, Elsevier.

Frenkel, R. (2003), "Globalization and financial crises in Latin America", *CEPAL Review*, No. 80 (LC/G.2204-P), Santiago, Chile, August.

Griffith-Jones, S., J.A. Ocampo and J. Stiglitz (eds.) (2009), *Time for a Visible Hand: Lessons from the 2008 World Financial Crisis*, New York, Oxford University Press.

Harberger, A. (1985), "Observations on the Chilean economy, 1973-1983", *Economic Development and Cultural Change*, vol. 33, No. 3, Chicago, University of Chicago Press, April.

Heymann, D. (2000), "Major macroeconomic disturbances, expectations and policy responses", *CEPAL Review*, No. 70 (LC/G.2095-P), Santiago, Chile, April.

Hofman, A. and H. Tapia (2003), "Potential output in Latin America: a standard approach for the 1950-2000 period", *Estudios estadísticos y prospectivos series*, No. 25 (LC/L.2042-P), Santiago, Chile, Economic Commission for Latin America and the Caribbean (ECLAC). United Nations publication, Sales No. E.03.II.G.205.

IDB (Inter-American Development Bank) (1997), *Latin America after a Decade of Reforms*, Washington, D.C.

Kaminsky, G., C. Reinhart and C. Vegh (2004), "When it rains, it pours: pro-cyclical capital flows and macroeconomic policies", *NBER Working Paper*, No. 10780, Cambridge, Massachusetts, National Bureau of Economic Research, January.

Krugman, P. (2009), *The Return of Depression Economics and the Crisis of 2008*, New York, W.W. Norton.

_____ (2000), "Crises: the price of globalization?", *Proceedings*, Kansas, Federal Reserve Bank of Kansas City.

Le Fort, G. and S. Lehmann (2003), "The unremunerated reserve requirement and net capital flows: Chile in the 1990s", *CEPAL Review*, No. 81 (LC/G.2216-P), Santiago, Chile, December.

Magud, N. and C. Reinhart (2006), "Capital controls: an evaluation", *NBER Working Paper*, No. 11973, Cambridge, Massachusetts, National Bureau of Economic Research, January.

Mahani, Z., K. Shin and Y. Wang (2005), "Ajustes macroeconómicos y la economía real en Corea y Malasia: la respuesta a la

crisis asiática", *Investigación económica*, vol. 64, No. 254, Mexico City, National Autonomous University of Mexico, October-December. Also published in R. Ffrench-Davis (ed.), *Seeking Growth under Financial Volatility*, London, Palgrave Macmillan/ECLAC, 2006.

Marfán, M. (2005), "La eficacia de la política fiscal y los déficit privados: un enfoque macroeconómico", *Los desafíos del desarrollo en América Latina*, J.A. Ocampo (ed.), Bogotá, D.C., ECLAC/Alfaomega.

Martner, R. and V. Tromben (2004), "Tax reforms and fiscal stabilization in Latin American countries", *Gestión pública series*, No. 45 (LC/L.2145-P), Santiago, Chile, Economic Commission for Latin America and the Caribbean (ECLAC), June. United Nations publication, Sales No. E.04.II.G.73.

Ocampo, J.A. (2008), "A broad view of macroeconomic stability", *The Washington Consensus Reconsidered: Towards a New Global Governance*, N. Serra and J.E. Stiglitz (eds.), New York, Oxford University Press.

_____ (2007), "The macroeconomics of the Latin American economic boom", *CEPAL Review*, No. 93 (LC/G.2347-P), Santiago, Chile, December.

Prasad, E., K. Rogoff and others (2003), "Effects of financial globalization on developing countries: some empirical evidence", *IMF Occasional Paper*, No. 220, Washington, D.C., International Monetary Fund, September.

Reisen, H. (2003), "Ratings since the Asian crisis", *From Capital Surges to Drought*, R. Ffrench-Davis and S. Griffith-Jones (eds.), London, Palgrave Macmillan.

Rodrik, D. (2008), "The real exchange rate and economic growth", *Brookings Papers on Economic Activity*, Washington, D.C., Brookings Institution Press.

_____ (2006), "Goodbye Washington Consensus, Hello Washington Confusion? A review of the World Bank's *Economic Growth in the 1990s: Learning from a Decade of Reform*", *Journal of Economic Literature*, vol. 44, No. 4, Nashville, Tennessee, American Economic Association, December.

_____ (1998), "Who needs capital account convertibility?", *Should the IMF Pursue Capital Account Convertibility?*, P. Kenen (ed.), *Princeton Essays in International Finance*, No. 207, Princeton, Princeton University.

Singh, A. (2006), "Macroeconomic volatility: the policy lessons from Latin America", *IMF Working Paper*, No. 166, Washington, D.C., International Monetary Fund, July.

Stallings, B. and R. Studart (2005), *Finance for Development: Latin America's Banks and Capital Markets after Liberalization*, Washington, D.C., Brookings Institution.

Stiglitz, J. (2000), "Capital market liberalization, economic growth and instability," *World Development*, vol. 28, No. 6, Amsterdam, Elsevier, June.

United Nations (2007), *Follow-up to and implementation of the outcome of the International Conference on Financing for Development. Report of the Secretary-General* (A/62/217), New York, August.

Uthoff, A. and D. Titelman (1998), "The relationship between foreign and national savings under financial liberalisation", *Capital Flows and Investment Performance: Lessons from Latin America*, R. Ffrench-Davis and H. Reisen (eds.), Paris, Organisation for Economic Co-operation and Development (OECD).

Williamson, J. (2003), "Overview: an agenda for restarting growth and reform", *After the Washington Consensus: Restarting Growth and Reform in Latin America*, P-P. Kuczynski and J. Williamson (eds.), Washington, D.C., Institute for International Economics.

_____ (2000), *Exchange Rate Regimes for Emerging Markets: Reviving the Intermediate Option*", Washington, D.C., Institute for International Economics.

World Bank (2005), *Economic Growth in the 1990s: Learning from a Decade of Reform*, Washington, D.C.

_____ (1997), *The Long March: a Reform Agenda for Latin America and the Caribbean in the Next Decade*, Washington, D.C.

Zahler, R. (2005), "Estabilidad macroeconómica e inversiones de los fondos de pensiones: el caso de Chile", *Crecimiento esquivo y volatilidad financiera*, R. Ffrench-Davis (ed.), Bogotá, D.C., Mayol Ediciones.

KEYWORDS

Economic crisis
Economic growth
Environmental degradation
Natural resources
Commodities
Commodity prices
Inflation
Wealth
Income distribution

Global economic crises, environmental-resource scarcity and wealth concentration

Ramón López

Three new structural factors underlie the most recent global crisis: (i) the fact that several high-population countries have joined the growth process; (ii) the increasing scarcity of environmental and certain natural resources; and (iii) the extraordinary concentration of income and wealth that has occurred in the advanced economies over the last two decades. These structural changes have significantly strengthened the links between global growth and commodity demand; they have made world commodity supply increasingly inelastic, and have rendered economic growth more dependent on easy monetary and financial policies. The combination of these factors could make the world economy highly crisis-prone and may hinder recovery from the current one.

Ramón López
Professor of Economics,
Department of Agricultural and
Resource Economics,
University of Maryland at College
Park, Maryland, USA
rlopez@arec.umd.edu

I
Introduction

This article focuses on new structural factors that have played an important but often neglected role in the genesis of the current global crisis, and may also have profound implications for long-term economic growth. It examines the synergetic interactions that exist between three structural factors: (i) the awakening of heavily populated countries, such as China and India, from centuries of economic lethargy, and their emergence as world growth leaders, and as large-scale providers of industrial goods and consumers of commodities; (ii) the increasing scarcity of environmental and certain natural resources which, for the first time ever, is starting to be felt in rich and poor countries alike; (iii) the dramatic concentration of wealth that has occurred over the last two decades especially in the advanced economies.

Structural factors (i) and (ii) have made commodity prices highly responsive to economic growth. The increased economic weight of population giants that are still at an early stage of development, such as China, India and others, has increased the commodity- and energy-intensity of global economic growth, as growth in these countries is largely based on a rapid expansion of commodity- and energy-intensive industries (Farrell and Grant, 2005). Moreover, while the advanced economies have largely dematerialized their production by focusing more and more on services and knowledge-intensive industries, their consumption has not dematerialized by nearly as much. This asymmetry has meant that the advanced countries increasingly rely on the rest of the world to meet their growing net demand for material goods, energy and other commodities.

As a result of the second factor —the increasing worldwide scarcity of natural resources— the response capacity of commodity supply is becoming less and less flexible, at a time when world growth is fuelling higher commodity demand. As a result, rapid world economic growth is now closely accompanied by rising commodity prices.

Central banks respond to these prices hikes by tightening monetary policy to prevent the corresponding pressures being validated in higher inflation indices, while hopefully still leaving some room for the economy to continue growing, albeit perhaps at a more modest pace. Nonetheless, for reasons to be made clear below, the third structural factor —the increasing concentration of wealt— has made the real economy much more sensitive to tight monetary and financial policies than in the past, and this has considerably reduced the space available for continued economic growth in periods of tight monetary policy.

Over the last two decades, the concentration of wealth in advanced economies, and in many middle-income countries too, has gone hand-in-hand with stagnation among the middle class; and both phenomena are at least partly the result of new policies implemented in this period (Krugman, 2006). This suffocation of the middle class —what Paul Krugman has dubbed "the great wealth transfer"— has made the vast majority of households increasingly reliant on new borrowing to finance their consumption, thereby causing debt-to-income ratios to explode (Taylor and others, 2008). As a large portion of that debt is often subject to periodic interest-rate adjustments, household disposable income (after debt service) has become much more sensitive to interest-rate hikes. Moreover, as global economic expansion depends heavily on rapid consumption growth, especially in the United States and other advanced countries, this has made economic growth ever more reliant on easier monetary and financial policies in those countries. Such policies are essential to allow consumers easy access to credit and support a relatively low debt-service burden despite their higher household debt/income ratios. As result, middle-class households were for long able to increase their consumption, financed by higher debt levels, even though their real income were being squeezed or remained flat.

In the United States, the share of total household income received by the richest 10% of households grew continuously from about 35% in the 1980s to 50% in 2007, its highest level ever (Saez, 2009). Ominously, the only other period in the last century in which the wealthiest 10% received close to 50% of total income

□ Research assistance was provided by Asif Islam and Amparo Palacios, graduate students at the University of Maryland at College Park.

was in the years leading up to the Great Depression. In fact, in 1917-28 the share of the top 10% grew almost exactly as fast as in 1996-2007, rising from 40% to 49% in the earlier period and from 40% to 50% in the latter (Picketty and Saez, 2003). Apparently, this is more than mere coincidence; when income concentration reaches such extreme levels it makes economies that depend on domestic demand increasingly vulnerable to deep and prolonged crises, because it becomes increasingly difficult to keep demand growing. While domestic income concentration may not directly make economies where growth is primarily export-led, such as those of Latin America, more crisis-prone, it certainly does in countries such as the United States that rely on domestic demand. Moreover, as this and a number of other rich countries provide a large proportion of the rest of the world's markets, a demand slump in these countries could easily drag the world into recession.[1]

As shown below, the greater sensitivity of commodity prices to world economic growth, combined with the greater reliance of economic growth on loose monetary policies in the advanced countries, are likely to make rapid economic growth with price stability much harder to sustain in the future; and they may make it more difficult to recover from the current crisis.

Section II of this article makes a detailed analysis of the emergence of the recent great recession, against the backdrop of a new economic order generated by the three structural factors mentioned above. The analysis starts by reviewing interactions between rich and poor countries, highlighting how the economic success achieved by many highly populated but historically poor countries has made global economic growth more widespread. Section III reviews environmental aspects and their relevance for global economic growth. Section IV considers the consequences of the radical new policies implemented by various conservative governments, based on the intellectual *laissez-faire* ideology reinvented in the advanced economies, to which many middle-income countries around the world have signed up. One result of these policies was greater concentration of wealth in most countries (Jerzmanowski and Nabar, 2008; Philippon and Reshef, 2009). Thus, while economic growth across countries has become more inclusive, with the growth club expanding to encompass a number of previously excluded countries, income growth within individual economies has increasingly become the privilege of a relatively small minority. Section V describes the relationship between energy and raw materials prices, inflation, and the outbreak of the crisis. Lastly, section VI concludes by showing how the unusual nature and intensity of the current crisis have been the corollary to these developments.

[1] In most countries with open markets, such as those of Latin America, China, India and many others, wealth concentration may not be a direct obstacle, since economic growth is mostly fuelled by foreign demand. This causes a social "tragedy of the commons" type of syndrome: governments that tend to be driven almost exclusively by gdp growth-maximization objectives have no incentive to prevent a concentration of wealth, since this will not affect an individual country's growth potential. Nonetheless, if most countries in a globalized world apply the same pro-growth policies with no concern for the distribution of wealth, wealth concentration will become contagious across the world. This, in turn, restricts global demand and thus provides a precarious foundation for sustained global economic growth.

II

The new economic order: retrospect

1. The dematerialization of production in the North

For much of the twentieth century, continuous economic growth was the privilege of an exclusive club comprising no more than a fifth of the world's population (henceforth referred to as the "North"). As the North grew ever richer, it underwent continuous structural change, and its production became increasingly "dematerialized" (López and Stocking, 2009). The structure of gross domestic product (GDP) in the North became increasingly focused on services and, generally, on human-capital- and knowledge-dependent activities. As a counterpart, the total-output share of natural-resource-based activities, and later most manufacturing sectors, gradually shrank. Figure 1 illustrates the intensity of this process over the last 50 years in the United States, as the shares

of both commodity production and manufacturing in total GDP have steadily declined. Manufacturing, agriculture, forestry, fishing, oil extraction and mining have seen their combined GDP share plummet from over 40% in the early 1950s to less than 20% in the early 2000s.

This dematerialization of the North's production has not been matched by a similar dematerialization of consumption. While there was some shift in the structure of consumption —which ceased to be centred on commodities such as foodstuffs— consumers in the North continued to expand their demand for energy and industrial products, especially durable goods, at a rate that often outpaced their per capita income (Ghertner and Fripp, 2007). The North's consumption dematerialized much more slowly than its domestic production, making it increasingly reliant on the rest of the world (the "South") to supply, firstly, commodities (including energy and other raw materials), and then manufacturing goods too, particularly over the last three decades.

An analysis of trade flows in the United States, for example, shows a rapid increase in net imports of commodities and industrial goods as a proportion of that country's total imports and GDP. Figure 2 illustrates this, showing a steep rise in the share of industrial goods in total imports and large increases in imports of manufactured and other industrial goods as a proportion of domestic production. The same is true for most other commodities including metals (see figure 3). Imports of industrial goods and commodities have grown particularly fast over the last two decades.

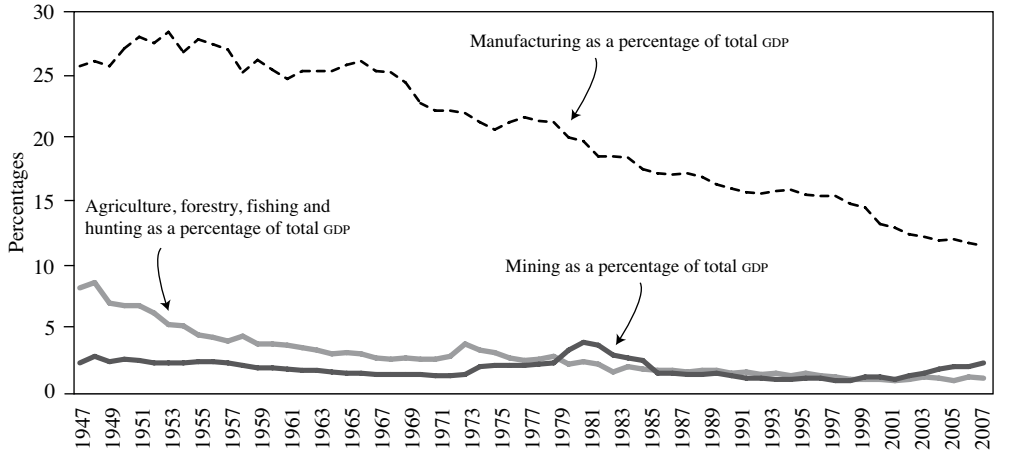
2. The roles played by the South

This article distinguishes two groups of economies in the South: those that are rich in natural resources (most of Latin America, Sub-Saharan Africa and parts of Asia); and those that are labour-abundant (nearly all Asian countries). The labour-abundant South includes population giants such as China and India and a few others, which are home to a very large proportion of the world's total population; in contrast, the resource-rich part of the South is relatively sparsely populated. While there have been periods in which the South as a whole has been able to achieve modest growth, until the last three or four decades few countries in the South had been unable to sustain growth for lengthy periods of time. The resource-rich South essentially became a passive supplier of energy and other commodities to the North.

Since colonial times the North has maintained efficient commodity-producing enclaves in the resource-rich part of the South, selling almost exclusively to northern markets. Burgeoning demand

FIGURE 1

United States: sector composition of GDP, 1947-2007
(Percentages of total GDP)

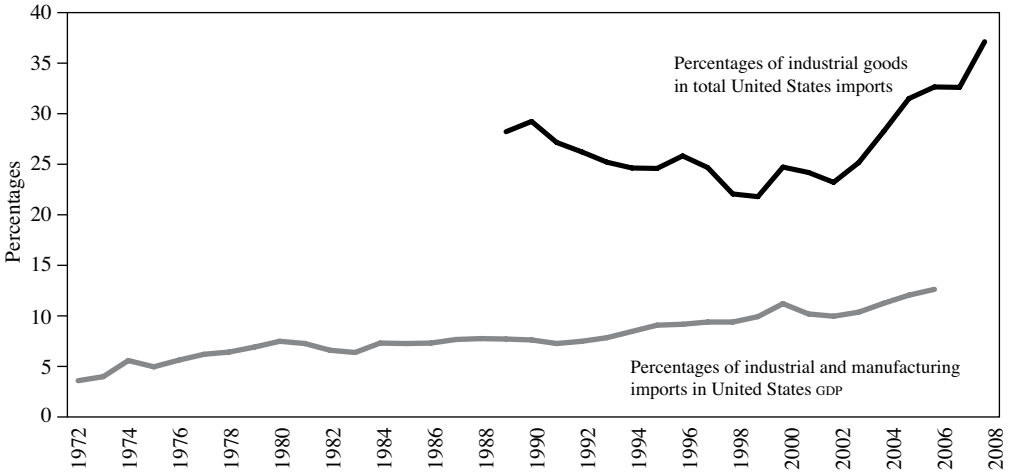


Source: United States Bureau of Economic Analysis.

GDP: Gross domestic product.

FIGURE 2

United States: share of industrial imports in total imports and GDP,[a] 1972-2008
(Percentages)



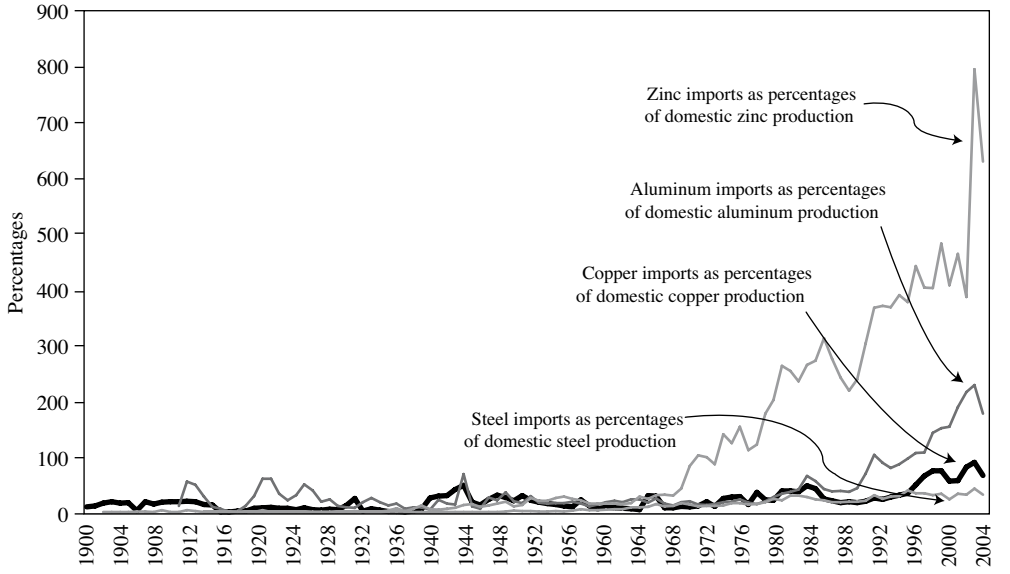
Source: United States Bureau of Economic Analysis; R.C. Feenstra, J. Romalis and P. Schott, "United States imports, exports and tariff data, 1989-2001", NBER Working Paper, No. 9387, Cambridge, Massachusetts, National Bureau of Economic Research, 2001), http://www.internationaldata.org/.

[a] The categories are: fuel and lubricants, paper and paper base stocks, materials associated with non-durables, selected construction materials, unfinished and finished metals associated with durable goods, non-metals associated with durables.

GDP: Gross domestic product.

FIGURE 3

United States: Metal imports as a proportion of domestic production, 1900-2004
(Percentages)



Source: United States Geological Survey.

for commodities in these markets was matched by continuous exploration and new investments made by the North to expand these enclave economies. The presence of abundant resources in the South and the absence of effective regulations limiting the environmental damage caused by their extraction, together with continuous Northern investments in resource extraction in the (resource-rich) South, kept the global commodity supply curve relatively flat for a long time (in other words, supply was highly elastic), which helped to keep real commodity prices constant (López and Stocking, 2009). As has been documented in several studies, productive enclaves of this type have had few linkages with the rest of the South's economies, and generally contributed little to their growth (de Janvry, 1975).

Both the resource-rich and the labour-abundant segments of the South remained essentially stagnant, exerting little pressure on commodity demand; and this, in turn, helped to keep commodity prices stable. The fact that this process lasted well into the second half of the twentieth century (Sokoloff and Engerman, 2000; Acemoglu, Johnson and Robinson, 2001; Khor, 2000) allowed the North to grow while enjoying the luxury of commodity prices that stayed constant and sometimes even fell in real terms, despite continuously expanding demand from the North throughout much of the century (see figure 4).

3. The labour-abundant South emerges in the global economy

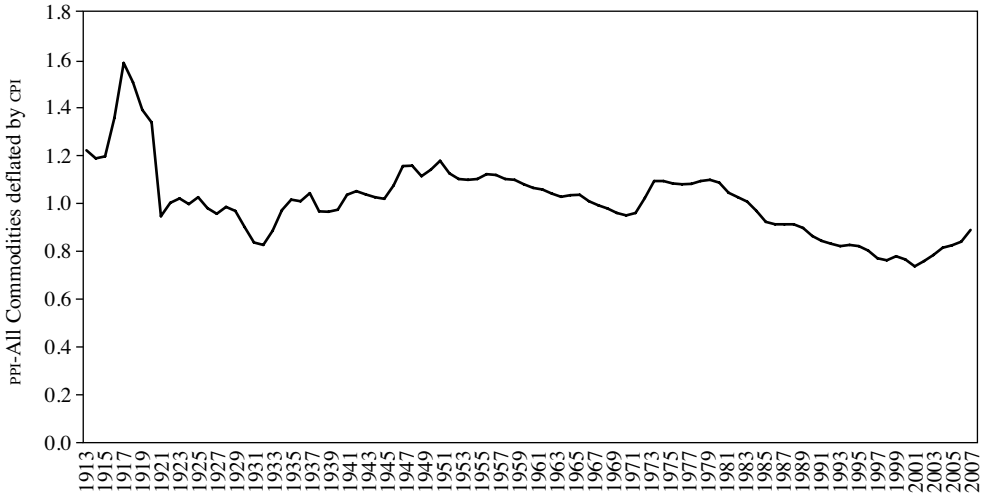
As documented by the so-called "resource-curse" literature, the constant expansion of natural resource extraction in enclaves of resource-rich countries of the South has in most cases been insufficient to promote their sustained economic growth (Barbier, 2005). In contrast, the labour-abundant South was better placed to benefit from the growing demand for industrial goods arising from the dematerialization of production in the North, especially over the last three decades.

(a) *First, the "Little Giants"*

The labour-abundant South started to shake off its lethargy in the 1970s with the emergence of a few relatively small countries in South East Asia (Korea, Taiwan, Hong-Kong (Special Administrative Region of China), Singapore and a number of others), which succeeded in maintaining rapid growth over long periods

FIGURE 4

Real commodity price trends: producer price index (PPI),[a] all commodities, 1913-2007



Source: United States Bureau of Labor Statistics.

[a] Producer price index for commodities as defined by the United States Bureau of Labor Statistics.

CPI: Consumer price index.

of time, driven by manufacturing exports. Although small in size, these newly industrializing countries (NICs) became major exporters of manufactured goods to the North. The emergence of these little manufacturing-export giants allowed the North to intensify its relative specialization in environmentally clean non-material products, while relying increasingly on the NICs as efficient suppliers of industrial goods to satisfy growing demand among consumers in the North (Krugman, 1994; Lall and Albaladejo, 2004; Noland, 1997).

Although population density in the NICs was high, their overall population was too small to have a significant impact on global demand for commodities, despite their rapid growth. This meant that NIC expansion did not exert greater demand pressure on world commodity prices, which, as shown in figure 4, remained broadly stable during the NIC boom period in the 1970s and 1980s.

Thus, by the late 1980s the world had achieved a remarkable equilibrium: the North plus a few NICs achieved rapid growth —the North on the basis of environmentally clean service-oriented production which greatly facilitated its low-cost, ecologically "sustainable" development; while the NICs supplied the North with an increasing portion of its growing net demand for industrial goods at low market prices, albeit produced at high domestic environmental cost; whereas the still slow-growing resource-rich part of the South supplied raw materials, also at low prices but at the cost of a continuous erosion of its natural resources and environment.[2]

(b) *Then, the "Real Giants"*

The late 1980s brought even more dramatic changes, as other large labour-abundant countries of the South implemented major pro-growth policy reforms. The emergence of the new industrial giants (NIGs), mainly China, India and a few other large initially poor countries that succeeded in growing very rapidly, was partly a consequence of the drastic policy reforms implemented in those countries. The new policies included pro-market reforms, privatization of State-owned enterprises, export promotion through exchange-rate policies and other incentives, and weak regulation of pollution, which effectively gave a green light to expand manufacturing production with few environmental constraints. The success of these new policies was assured by the burgeoning demand for industrial goods among consumers in the North.

The NIGs were as efficient as the NICs in supplying industrial goods, but on a much larger scale (Bosworth and Collins, 2008; Panagariya, 2006; López, 2008; Lall and Albaladejo, 2004); and industrial-export-led growth has afforded them unprecedented economic growth for over two decades. Growth in these countries was based on the rapid expansion of industrial exports, facilitated by undervalued exchange rates.[3] This, in turn, meant a huge accumulation of foreign exchange (see figure 5) which was recycled to the North, especially the United Sates and parts of Europe, thus creating large current-account deficits in the respective economies.[4] Copious financial flows into the North helped keep interest rates low and massively expanded the supply of credit in those countries, while the capital flow from the NIGs and oil-exporting countries to the North fuelled a continuous rise in equity and real-estate prices, which prolonged the economic boom and helped form a financial bubble. This, in turn, further fed the North's huge appetite for industrial imports from the emerging giants.

Another remarkable (temporary) equilibrium was thus created: the massive financial flows from the new giants, created by their industrial-export success, fuelled the boom in the North which, in turn, fed consumer demand for industrial imports and thus reinforced the NIGs' continuous expansion. Real GDP growth rates of China and India over the last two decades have been consistently above 8% per year, more than three times the rates achieved by the advanced economies

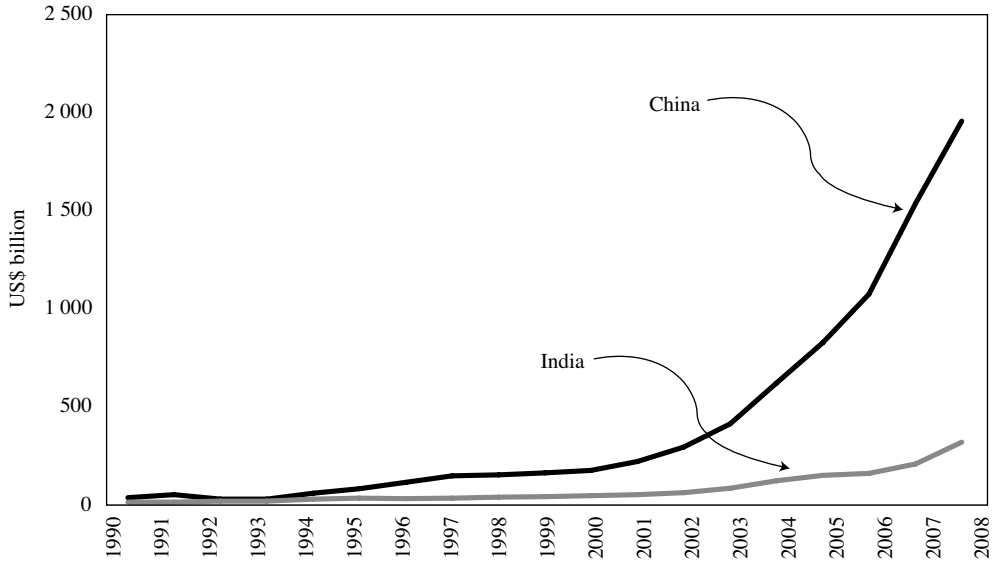
[2] It is interesting to note that the North started to dramatically reverse its environmental degradation in the mid-1970s, just as the NICs were emerging as suppliers of low-cost goods produced by polluting industries. This process also coincides with the time when most of the modern environmental regulation in the North begun to be implemented. Increasingly stringent environmental regulation in the North may have been rendered politically acceptable precisely because of the emergence of foreign suppliers of goods produced by polluting industries. López (2008) shows that one of the reasons why the North is able to enforce significant environmental regulation at very low cost (estimated at less than 2% of GDP) is the emergence of the NICs, and later of other large industrial suppliers, which allowed it to shift its production away from most polluting industries. This view is consistent with the econometric evidence provided by Levinson and Taylor (2008).

[3] According to Rodrik (2007), not only have China and India kept their exchange rates undervalued over the last two decades, but the extent of undervaluation has increased steadily during the period. Most of the NICs also based their industrial export take-off on undervalued exchange rates, although, unlike the NIGs, some of them recently have allowed their real exchange rates to become less undervalued —even overvalued at times.

[4] The exceptions were Germany and Japan, both of which are large exporters of technologically advanced goods and services.

FIGURE 5

China and India: foreign exchange reserves, 1990-2008
(Billions of dollars)



Source: Bank of India and State Administration of Foreign Exchange, People's Republic of China.

(IMF World Economic Outlook). More importantly, the NIGs contributed over US$ 350 billion per year to global growth in the early 2000s —that is, over a third of the total annual growth of the world economy, estimated at about US$ 1.1 trillion. As shown in table 1, the contribution made by China and India, of about US$ 200 billion, represented nearly 20% of total annual world growth in the 2000-07 period, compared to just 5% in the 1980s.

4. Finally, commodity-demand pressures

The new North-NIG boom equilibrium of the second half of the 1990s and early 2000s differed in one important respect from the North-NIC equilibrium of earlier decades: nearly 50% of all humanity lives in the "new giants", compared to at most 5% who live in the "little giants". In other words, the emergence of the NIGs dramatically expanded the growth club,

TABLE 1

Distribution of real global GDP growth by region and period
(Millions of dollars)

Period	China and India	Advanced Economies	Rest of the World	World
1961-1969	5 154	366 333	120 100	491 587
1970-1979	12 198	514 740	47 150	574 088
1980-1989	36 740	472 160	94 735	603 635
1990-1999	86 310	519 200	116 040	721 550
2000-2007	201 375	618 875	274 525	1 094 775

Source: World Development Indicators, World Bank.

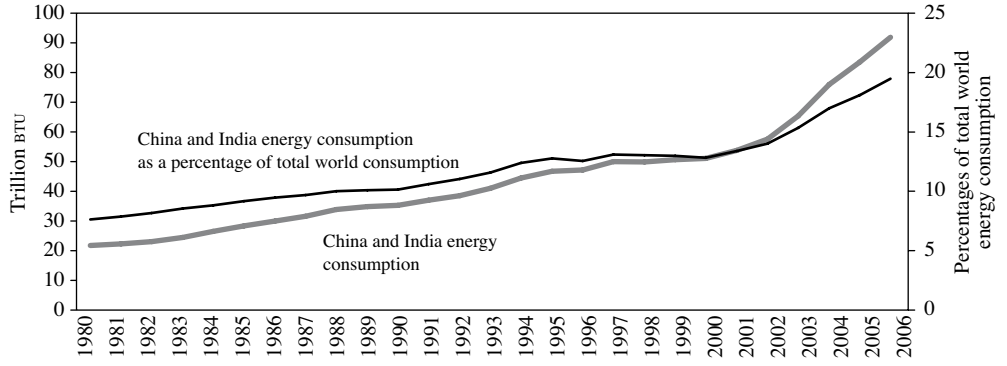
and, for the first time ever, a large part of the South (most of the labour-abundant portion) became an increasingly important and rapidly growing net consumer of energy, raw materials and other commodities. Constant economic growth ceased to be the exclusive preserve of just a few countries; now, for the first time in history, most the world's population were living in countries that were capable of growing. Only the resource-rich South remained more or less at a standstill.

The NIGs' recent leadership of world economic expansion has made growth much more commodity- and energy- intensive than in previous decades, when it was mainly confined to the North. This is a consequence of the large increase in the absolute scale of economic growth caused by the incorporation of 50% of the world's population, formerly dormant, into the growth process. Moreover, the fact that the massive population of the now expanding NIGs still has a low per capita income means that their income-elasticity for food, energy and other commodities is much higher than in the advanced economies. In other words, the continuous upward shift of the world commodity demand curve caused by income growth is now occurring much faster than before.

The rapid growth of the NIGs thus generated not only a drastic expansion of the supply of industrial goods but also a dramatic increase in the NIGs' demand for energy, food, and other commodities. As the NIGs started from very low levels of consumption, their increased demand for such commodities initially had little repercussion on world markets. By the mid-1990s, however, they had become significant net importers of commodities, energy, and other raw materials. After more than a decade of annual growth of 8-10%, the huge populations of the NIGs expanded their demand for commodities to account for a sizable portion of total world demand, and so it remains today. Figures 6 and 7 illustrate the rapid rise of consumption levels in China and India and their shares of total world consumption of energy and certain other commodities over the last two decades.

FIGURE 6

China and India: Energy consumption and share in world consumption, 1980-2006
(Trillions of BTU)

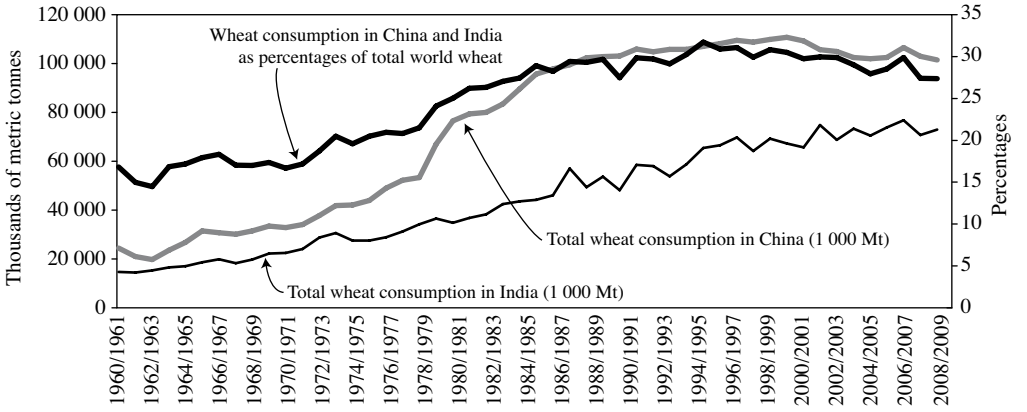


Source: United States Energy Information Administration (EIA).

BTU: British thermal unit (unit of energy).

FIGURE 7

India and China: wheat consumption and share in total world consumption, by period
(Percentages and thousands of metric tonnes)



Source: United States Department of Agriculture (USDA).

III

Environmental-resource scarcity and commodity supply

As a result of the incorporation of the NIGs into the growth process, the relation between commodity demand and economic growth is becoming ever closer. At the same time, the natural resources that the South used to possess in large quantities, are becoming less abundant, and governments are finally starting to take seriously some of the dire environmental consequences of the frenetic expansion of natural resource extraction. While most underground raw materials may still be plentiful, there are signs that their supply will have to tap increasingly more expensive sources; moreover, resource extraction has led to massive environmental costs affecting extremely valuable ecosystems, water quality, forests and other increasingly scarce environmental assets.[5]

Under increasing pressure from international non-governmental organizations (NGOs), local communities that are becoming aware of their ancestral rights over natural resources, and other parts of domestic civil society, governments in the resource-rich South are finally starting to implement policies aimed at limiting some of the large environmental costs that commodity extraction entails. More countries are now enforcing environmental regulations, albeit very timidly, on the use of many ecosystems that tend to be destroyed by unscrupulous resource extraction.[6] In other words,

[5] With some important exceptions, the real limits on commodity supply are not so much the scarcity of in-ground raw materials, but the high and rising environmental costs that their production entails (Simpson, Toman and Ayres, 2005). Resource extraction (mining, oil drilling and other activities) seriously affects water quality, soils and forests (for example, mountain top removal for coal extraction). The United States could significantly increase its oil production by expanding off-shore or Alaskan production, at the cost of ever higher risk of environmental destruction.

[6] Brazil, for example, has passed an environmental crime act with stiff sanctions including imprisonment; and NGOs frequently lobby for the enforcement of laws to protect the Amazon (Da Motta, 2003). In 1991 Colombia's constitutional law completely reorganized environmental management, resulting in stricter enforcement. Priority has been given to the use of economic instruments, and especially to punitive and compensatory pollution charges or taxes (Huber, Ruitenbeck and da Motta, 1998). More recently, China, Taiwan Province of China, and Korea have implemented

the countries of the South have started doing what the few natural resource-rich countries in the North (the United States, Australia and Canada) did several decades ago —impose heavy restrictions on resource extraction to mitigate the damage caused to fragile ecosystems and other renewable natural resources. These restrictions ultimately make commodity extraction more costly.

This means that now, perhaps for the first time ever, the long-run supply curve of natural-resource-based commodities has become relatively inelastic. The finite capacity of the natural world is finally being reflected in an increasingly vertical commodity supply curve. Ironically this is due not so much to the scarcity of non-renewable underground materials but mainly to the increasing degradation of ecosystems and other renewable natural resources that are crucial to life, resulting from the massive exploitation of non-renewable resources.

In short, economic growth is now shifting the world commodity demand curve upwards more rapidly at a time when the long-run commodity supply curve has become relatively inelastic. Hence, it appears that world economic growth and commodity prices are now more intimately related than in earlier decades. This may explain the large hike in commodity prices that took place in 2003-2007 and, as shown below, helped trigger the current recession. Since the increased sensitivity of commodity prices to economic growth is symmetric, it also explains the rapid fall in commodity prices in the depths of the crisis and their rapid recovery in the second half of 2009 and 2010, when world economic growth began to stage a tentative recovery.

Climate change

The capacity of the atmosphere to absorb greenhouse gases without major climatic disturbance is another prominent example of how renewable resources act as a constraint on the expansion of commodity- and raw-materials-based industries. Once environmental policies incorporate the costs of climate change, this will eventually have to be reflected in higher prices for these products. The emergence of the new industrial giants and ever-increasing demand for material goods by consumers in the North have kept greenhouse-gas emissions at high levels, resulting in increasing concentrations of such gases in the upper atmosphere (Bohringer and Loschel, 2003; Loschel and Zhang, 2002). At the same time, the scientific evidence shows that an impending climate disaster can only be avoided if greenhouse-gas emissions are drastically cut. The demands made on the atmosphere have for too long outweighed its capacity for self-renewal, leading to an accumulation of climate-changing gases with the potential to cause major setbacks for human development.

The NIGs are responsible for a growing share of total global emissions, which had surpassed 25% in 2006 (see figure 8). More importantly, the elasticity of emissions with respect to economic growth in the NIGs is very high, because production in those countries makes intensive use of fossil fuels, including oil, natural gas and coal. This means that continued economic growth in those countries will entail a large increase in their emissions.

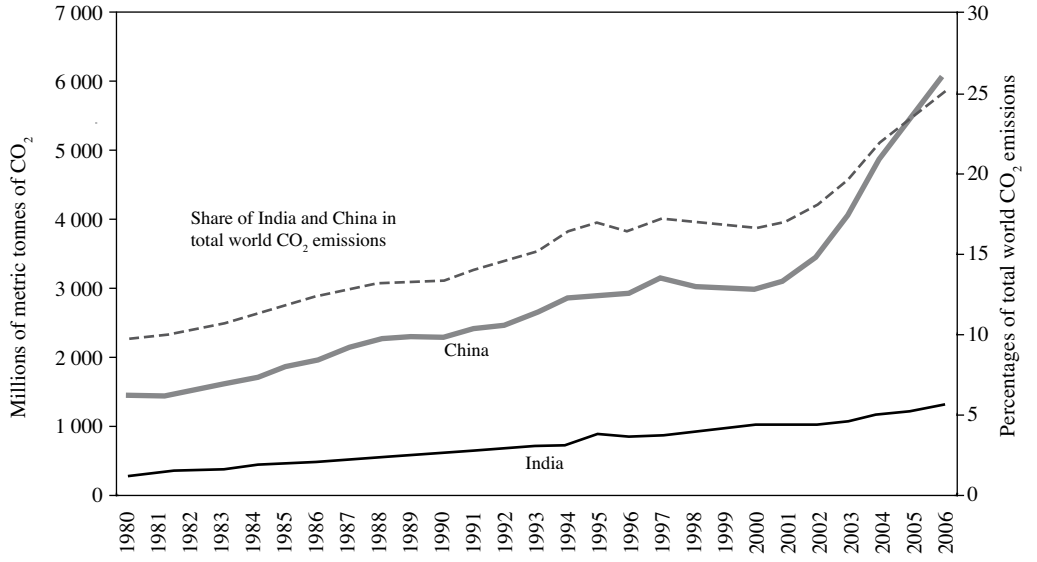
Global economic growth is thus even more closely tied to carbon emissions than it was during the late twentieth century; and even merely stabilizing world emissions at their current unsustainable levels will require much more intensive and costly efforts by the North to compensate for increasing emissions from the NIGs. Reducing the magnitude of climate change constitutes a new constraint on economic growth that has been largely ignored over the last few decades. In other words, an important yet hitherto ignored limit to growth in the North has been aggravated by the rise of the NIGs as powerhouses of global economic growth and by the constantly expanding consumption of material goods in the North..

The Kyoto Protocol of the United Nations Framework Convention on Climate Change, originally envisaged as a mechanism to deal with the growing emissions of climate-changing gases, has thus far proven ineffective owing to a lack of commitment by the largest carbon emitters, such as the United States, China, India and other major economies. Notwithstanding the, hitherto mostly failed, attempts to develop an international institution that is effective in limiting carbon emissions, sooner or later the world's economies will have to recognize this constraint; but imposing heavy restrictions on carbon emissions would imply new constraints on the growth of the NIGs and could even threaten their role as large-scale providers of (highly carbon-intensive) industrial goods to the North.

policies that are broadly in line with international standards for targeting and reducing greenhouse gases (Shapiro, 2009). Engel and López (2008) provide a detailed account of how organized local communities have recently emerged around the world to uphold their rights and restrict access to natural resources.

FIGURE 8

China and India: carbon dioxide (CO_2)emissions from the consumption and burning of fossil fuels
(Millions of metric tons)



Source: United States Energy Information Administration.

IV
The new economic policies

1. The North "sets its economy free" from the "shackles of government"

By the early 1980s the North had embarked upon its own "structural-change" process. The view that the "omnipotent state" and excessively high taxes were suffocating the private economy became widely accepted among policy-makers and economists; and the belief was that tackling these issues would make it possible to increase economic efficiency and thus allow for faster economic growth. Most countries in the North embarked on a far-reaching policy experiment that was justified as a means to enhance market incentives for investment and innovation.

The new structural-change policies included several pro-market measures, together with a massive withdrawal of government from the economy. The process included large tax cuts and fiscal retrenchment affecting a broad spectrum of social spending, together with large-scale financial deregulation. Even more important than the financial deregulation, however, was a process of "de-supervision" involving deliberate government actions to weaken the powers of the regulatory agencies by cutting their budgets, staffing, and powers (Caprio, Demirguc-Kunt and Kane, 2008). Moreover, structural change, especially in the United States and other advanced countries, also included policies and legislation (some subtle and others less so) aimed at weakening labour unions; and successive governments allowed the real minimum wage to gradually erode (Autor, Katz, and Kearney, 2006). These policies significantly weakened workers' bargaining power, which ultimately may have led to lower real labour earnings for all except top executives, financial traders and a few others. Although some of these policies may have encouraged investments, especially financial ones, they also appear to have had a number of serious, yet presumably unintended, consequences for wealth distribution.

Against a backdrop of increasing globalization, and in some cases under pressure from international financial institutions (IFIs), these policies in the North

were imitated by governments in many countries of the South. Indeed the same ideological consensus that led the North to adopt such radical reforms took root among most policy-makers and economists in developing countries. Of course, this extreme ideology of more market and less government (and less taxation) was partly created and supported by institutions directly or indirectly funded by élite groups, including think-tanks, universities and most of the media. The ever greater ideological influence gained by the élites through these institutions has been replicated in many countries in the South and has become their main and most effective lobbying mechanism. This paved the way for IFIs to serve as vehicles for instituting similar policy reforms in developing countries. Policy-makers in developing countries, many of whom had received their academic training in United States universities, did not need much persuasion —particularly since following IFI advice would significantly improve access to international credit and other types of international support.

2. The state does not regulate… wealth becomes more concentrated

These policies, which were intended to increase economic efficiency and maintain economic growth in the North, may also have been partly responsible for the massive concentration of income that has occurred over the last two and a half decades.[7] While not necessarily the sole cause of growing inequality, the evidence suggests that these policies contributed significantly to the trend (Sloan, 1997). Figures 9 and 10 give an idea of the massive concentration of income that has taken place in the United States and United Kingdom, respectively, especially since the early 1980s. In the United States, the middle class (defined here as the second, third and fourth quintiles of the household income distribution, representing about 200 million people) saw its share of national income shrink continuously from nearly 54% in 1980 to 47% in 2006, while the share of the poor (the lowest 20%) shrank from 5% to 4%. Most of these reductions benefited the wealthiest 5% of the population, whose share of national household income grew from 17% to 22.5% over the same period (in 2007 the richest 400 households in the United States, with an average annual income of US$ 350 million per household, received more than 1.5% national income). Similarly in the United Kingdom, the share of the middle class fell from 55% to 50% between 1980 and 2006, while that of the poorest quintile also dropped from 10% to 8%. Concomitantly, the share of the top 10% of households grew from 20% to 27% of total household income.

As a result of this concentration of income, the average real household income of the middle class in the United States has hardly increased over the last three decades, rising from US$ 48,000 in the early 1980s to just US$ 52,000 in 2007 (see figure 11). In other words, while annual per capita GDP grew by more than 2% in this period, income growth in middle-class families was about 0.3% per year. In contrast, the mean real household income of the wealthiest 5% of the population rose dramatically over the same period, from US$ 155,000 to US$ 290,000, far outpacing per capita GDP.

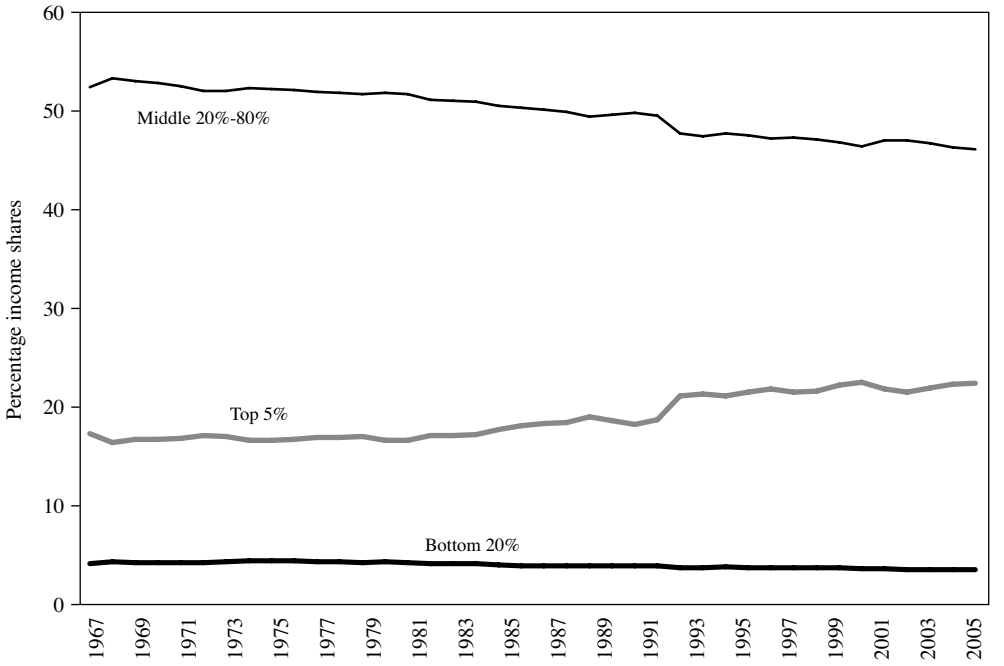
3. Dilemmas and paradoxes

The new regulatory environment provided incentives for business profit that were highly positive for economic expansion. Nonetheless, income concentration made it harder to match those incentives with the expansion of domestic demand needed to sustain such profits and fuel economic growth over time. When income becomes more concentrated, the middle class —the backbone of the consumer economy— is squeezed; and

[7] The rapid conversion of the economies of the North towards services and away from traditional industrial activities was another factor that may have contributed to the erosion of the middle-class and to greater income concentration. A large part of the industrial labour force which had achieved middle-class status lost their jobs and suffered significant income losses over the adjustment period (this is sometimes referred to as the "lost generation"). There are many studies that try to explain the concentration of income by focusing mainly on the distribution of wages among broad groups within the labour force (Autor, Katz and Kearney, 2006; Autor, Levy and Murnane, 2003; Card and di Nardo, 2002). Many of these emphasize the implications for the wage distribution of new technologies and higher levels of human capital. A study by Gordon and Dew-Becker (2007) analyses the consequences for wage polarization of the decreasing unionization of the labour force —a process that has accelerated precisely since the time of structural change. Few studies have analysed the upper end of the wage distribution, in which a high portion of wage incomes seems to have become concentrated. Kaplan and Rauh (2009) do this by showing that the amazing pay increases received by executives of investment banks, hedge fund, private equity and mutual fund managers and traders, as well as other categories of top executives in large corporations, largely explain the observed polarization of wage income. This increase in executive pay, benefiting perhaps no more than 1% of the labour force, is mainly due to the extreme financial deregulation and other structural adjustment policies implemented in the early 1908s.

FIGURE 9

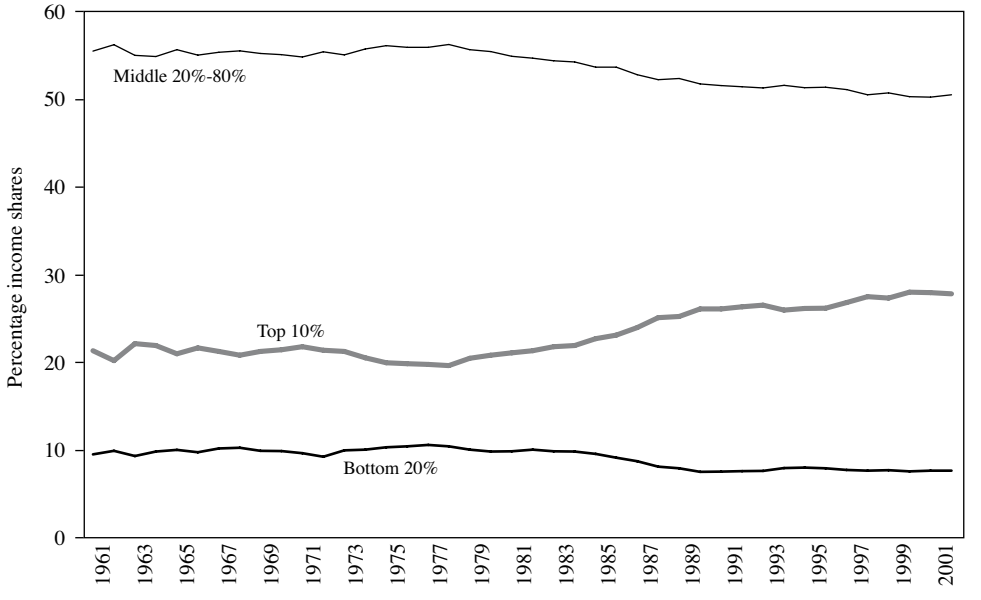
United States: income distribution: income shares by group, 1967-2005
(Percentages)



Source: United States Census Bureau.

FIGURE 10

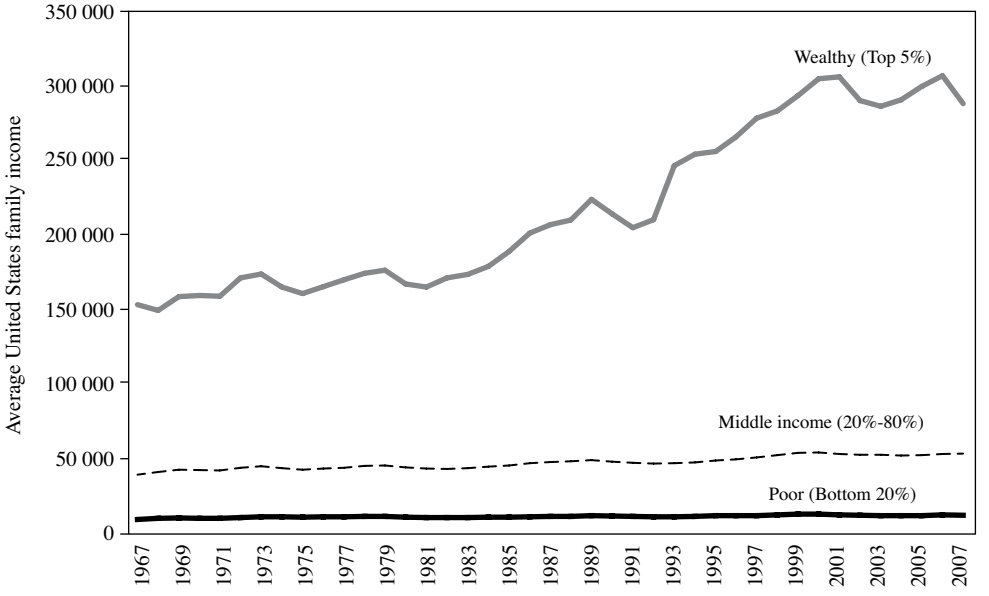
United Kingdom: income distribution: income shares by group, 1961-2001
(Percentages)



Source: World Income Inequality Database (WIID), United Nations University – World Institute for Development Economics Research (UNU-WIDER).

FIGURE 11

United States: real family income, by income group, 1967-2007
(Dollars at 2007 prices)



Source: United States Census Bureau.

this makes it much harder to keep domestic demand growing at the rates needed to sustain high rates of business profits and economic growth.

The maintenance of a high rate of economic expansion in the North depends largely on matching rates of consumption growth (Boone, Girouard, and Wanner, 2001; McConnell, Mosser, and Perez-Quiros, 1999).[8] A continuous and rapid expansion of consumption, in turn, needs a prosperous middle class; but middle-class incomes grew by less than one eighth of the per capita GDP growth rate. So, in principle, the rate of growth of middle-class household income was insufficient to support the consumption expansion that the economies of the North needed to sustain their historical growth rates.[9] The dilemma was how to persuade the middle class to expand its consumption sufficiently to keep per capita GDP growing by 2-2.5% per year, when its real income was only growing at 0.3%.

In these conditions, the required consumption growth could only occur if credit was plentiful and cheap, and if the middle class could be induced to accept ever higher debt levels. Why did households fall so readily into the temptation of easy credit? The massive inflow of capital from the NIGs and oil-exporting countries was one contributory factor; and the almost complete financial deregulation and de-supervision by government agencies, which fuelled the proliferation of exotic and highly risky financial instruments, was another. The final ingredient was monetary policy. The Federal Reserve implemented a generally permissive monetary policy in the face of obvious asset-price bubbles; and it also failed in its financial regulation function.

These three factors generated an unprecedented flood of credit and other financial resources into the consumer economy, thereby creating the conditions for the last component of the miracle (or, to be more precise, mirage), namely an extraordinary rise in the price of assets of all kinds, but particularly real-estate and equities, driven by the burgeoning demand for

[8] Another possible source of growth was exports, but given the historically low weight of export products in GDP, even rapid export expansion cannot provide enough support to domestic growth. Investment, at less than 20% of GDP, is also too small to make a significant difference for growth. Moreover, investment is largely driven by consumer demand itself.

[9] While the phenomenal increase in the after-tax income of the wealthy meant that their consumption expanded rapidly, the fact that their propensity to consume is far lower than that of the middle class means that the net effect on total consumption of redistributing income from the middle class to the wealthy is, *ceteris paribus*, negative.

such assets that was being fuelled by the massive and continuous inflow of financial resources into the economy. The resulting capital gains made ordinary citizens feel wealthier, even though their income flows were hardly growing at all —since most of the income growth accrued to a very small minority of the population. The confusion between (necessarily) short-term capital gains and higher permanent income induced middle-class families to cut savings and rapidly expand their consumption, financed by easy credit with inflated assets often used as collateral.

Although the dilemma of sustaining economic growth with stagnant middle-class incomes was "resolved" through easy credit and low interest rates, there were ominous implications. This growth model requires ever-rising household debt levels, backed by ever-increasing equity and real-estate prices. The debt burden of the average household almost tripled in two decades, from 45% of its annual income in the mid-1980s to 120% in 2004 (Taylor and others, 2008). Obviously this could not be considered a permanent solution.

4. The monetary and financial regulation syndrome: "kicking the can down the road"

Almost two decades of monetary and financial profligacy, driven by increasing income concentration, created the conditions for a potential economic tsunami. The middle class' continuing and ever-greater reliance on new borrowing to finance a large proportion of its consumption, made the consequences of restricting monetary policy and tightening financial regulation much more costly than previously. In fact, for every percentage-point hike in interest rates, median household income, after debt service, falls nearly three times as fast when the debt/income ratio is 1.2, as it is today, than when it is 0.45 as it was in the early 1980s. Moreover, with middle-class income growing so slowly, consumption growth depends almost exclusively on new borrowing. Hence, monetary tightening, or stricter enforcement of financial regulations to curtail the manifestly risky practices of the financial sector, have much more depressing effects on consumption. Consequently, the effects of financial and monetary tightening on the real economy are likely to be much more pronounced than previously, when household consumption was more closely linked to permanent income than to new borrowing.

Awareness of the serious consequences of tightening monetary and financial policies in a heavily indebted economy makes the monetary and financial authorities more reluctant to respond promptly to monetary and financial imbalances; and it may make them less likely to implement effective regulations even when there are clear signs of serious financial "anomalies", such as prolonged asset bubbles. This has helped generate a serious regulatory syndrome, namely addiction. The longer the implementation of corrective monetary and financial measures is delayed, the greater are the short-run adjustment costs of doing so, and hence the greater the incentives for policy-makers to "kick the can down the road" to gain time and leave the financial and monetary authorities of the future to solve the problem. This makes the monetary and financial authorities progressively more complacent, which fosters even deeper imbalances which, by the time the authorities are finally forced to adopt the necessary measures, could potentially trigger a much more serious crisis.

V

Commodity prices and inflation: the crisis

Given the relatively small weight of commodities in consumer demand in the North, the sharp commodity-price hikes associated with the rapid growth of the world economy initially had little impact on inflation in the North. Nonetheless, the magnitude of the commodity-demand implied by 8-10% annual growth rates in the NIGs forced their prices relentlessly upwards (Bosworth and Collins, 2008). As commodity prices skyrocketed in 2003-2007 they started to have an effect on the overall consumer price index (CPI) in the United States and many other countries, which were not always immediately internalized by the monetary authorities provided they did not affect so-called "core" inflation. However, core inflation was eventually affected, and by the end of 2004 the CPI had more than doubled. This ultimately forced the Federal Reserve to belatedly tighten monetary policy, causing a steady rise in the effective Federal

Funds Rate from 1.25% in the last quarter of 2004 to over 5% by the end of 2007 (see figure 12).

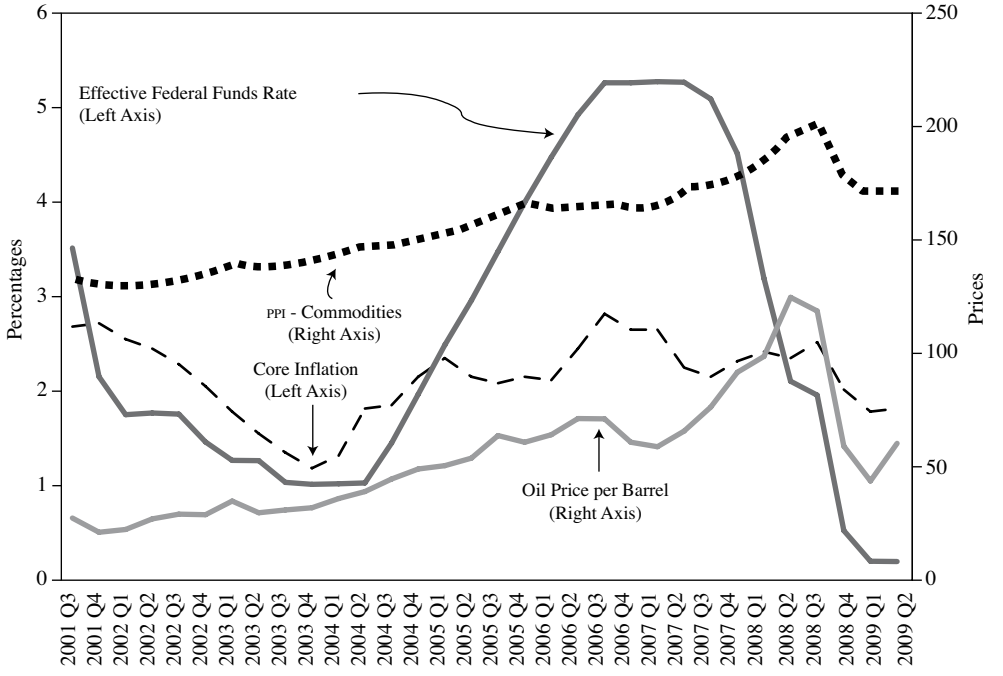
As economic growth had become so dependent on ever-expanding household debt and easy credit, which the income-squeezed middle class needed to sustain growth, the monetary tightening of 2005-2007 had much more traumatic consequences than previous episodes of this type.[10] In fact, it brought the economic boom to an end and triggered the start of the great global recession. Firstly, the interest-rate hike caused a concomitant rise in mortgage rates, leading to the sudden collapse of the housing market, which —at the peak of a prolonged price bubble— was highly sensitive to mortgage rates. With falling real-estate values and the consequent evaporation of a significant part of earlier capital gains, consumers suddenly realized that they were not nearly as wealthy as they had been led to believe.

Moreover, the drying up of credit made it increasingly difficult for the middle class to continue borrowing; and, more importantly, the cost to consumers of servicing their accumulated debt went up significantly. The high debt-to-income ratio has made consumer incomes highly vulnerable to interest-rate hikes, possibly more so than ever before. All of this prompted an initial round of adjustment in household consumption and caused widespread defaults in the housing sector, which, in turn, led to financial collapse inflicting capital losses on equity markets. Large-scale capital losses meant the evaporation of the collateral asset values that underpinned continued borrowing, thereby causing a second-round impact on consumption which ultimately allowed the crisis to further impact the real sector, causing high unemployment. This now appears to be inducing a third-round effect on middle-class income and consumption levels, which we are experiencing today. The subsequent collapse

[10] The financial and real-estate bubble that was fuelled by cheap credit was part of the same phenomenon.

FIGURE 12

United States: commodity prices, inflation, and monetary policy, 2001-2009



Source: United States Bureau of Labor Statistics, Federal Reserve Bank of St. Louis.
Note: The letter "Q" placed after the year indicates quarter.

PPI: Producer price index.

of the world economy also led to a steep fall in world commodity demand and prices.

Policy responses to the crisis

Governments in the North responded to the crisis in the following way: (i) a widespread bail-out process involving massive transfers of State resources to large financial and non-financial corporations; (ii) a drastic relaxation of monetary policy bringing short-term nominal interest rates to near zero levels, together with quantitative monetary easing —whereby the central banks have effectively become direct lenders, in particular buying medium- and long-term government bonds and lending directly to large financial institutions on an unprecedented scale; (iii) massive fiscal expansion that has created deficits seldom seen in the history of the advanced economies.

The new deficit-financed fiscal spending is expected to partly make up for the sharp retrenchment of private expenditure and, in the process, stimulate an economic recovery that may entice consumers once again to revive their pace of consumption. This view sees private consumption being stimulated by the economic recovery generated by the fiscal expansion and by the near-zero interest rates permitted by the more expansionary monetary policy.

Nonetheless, given the obstinate persistence of the new dominant ideology that rejects any structural change that would shift the tax burden towards the wealthy, fiscal deficits effectively imply large new financial liabilities for ordinary citizens, which sooner or later will have to be paid for through higher taxes and/or inflation. In other words, the capital losses suffered by consumers as a result of the real-estate and equity-market collapse, are now compounded by potential new obligations in the future generated by the fiscal deficits and increased public debt. This negative wealth effect is likely to become a serious impediment to any significant revival of private consumption. As a result, the multiplier effect of large but temporary fiscal and monetary stimulus is likely to be small, as impoverished middle-income households will be wary when increasing their spending. Consequently, once the bulk of the various temporary fiscal and monetary stimulus measures are withdrawn, there is a risk that the world economy may slip back into recession.

Moreover, the generous assistance given by governments to financial enterprises and other parts of the large corporate sector effectively means a new large-scale transfer of resources from the average citizen to the wealthiest segments of society. A second "great wealth transfer" may be underway, thus further exacerbating one of the key factors making economic growth vulnerable to deep crises. The fact that little has been done to tighten financial regulation of the major financial companies means that "business as usual" has been resumed in the financial sector. This has resulted in huge profits for that sector and, once again, massive bonuses to top financial executives and traders; while most workers, and the non-financial part of the economy generally, continue to suffer the dramatic consequences of the crisis.

In sum, policy responses to the crisis have been dominated by the prevailing pro-élite ideology. Some of the structural fundamentals that caused the crisis have thus been exacerbated rather than mitigated, which means that the world may remain highly susceptible to massive crises, and recovery from the current one is likely to be protracted and painful for the middle classes.

VI
Conclusion

The confluence of three major factors has made the world more vulnerable to deep and long-lasting crises, as never before since the Great Depression. One of these factors —extreme income inequality both in advanced and in many developing countries— is as glaring as in the years prior to the 1929 crash. Unlike the 1920s, however, the world is now experiencing an increasing scarcity of environmental resources, which is making commodity supply less responsive at a time when the elasticity of the global demand for commodities with respect to world growth has increased. As a result, economic growth is now more closely tied to rising commodity prices than in the twentieth century. The combination of greater wealth inequality and the concomitant impoverishment of the middle class and the high responsiveness of

commodity prices to global expansion, have potentially lethal economic consequences.

Although inflation in the advanced economies is now less directly affected by commodity prices, the persistent rise in commodity prices fuelled by rapid world growth eventually feeds into core inflation, thereby forcing governments to adopt restrictive monetary policies. As a consequence of the highly concentrated distribution of wealth, monetary and financial tightening now have much more dramatic effects on the real sector of the economy than they used to have. As a result, the world is now more crisis-prone, and global economic growth may become much more difficult to sustain than in the past.

In response to the current crisis, the world has seen an unprecedented combination of massive bail-outs for large firms, together with fiscal stimulus measures and extraordinarily lax monetary policies in the advanced countries. Public lending has taken up part of the slack left by the reduced private lending caused by the financial debacle. A new period of easy credit seems to have arrived, with the sole difference that the source is now public rather than private. Moreover, fiscal stimulus has spawned gigantic fiscal deficits, as governments have generally been reluctant to raise taxes on the rich —apart from a few isolated token gestures— to at least partly finance the deficits. While probably preventing the great recession from becoming a new great depression, the anti-crisis policies seem to have done little to address the structural problems discussed in this article. Moreover, the bail-out of large firms may have exacerbated one of those factors by inducing a second-round wealth transfer from the middle class to the super-rich.

The crisis brought about a temporary reduction of commodity prices and a concomitant fall in inflation. Nonetheless, with the structural factors discussed in this paper remaining in place, and given the huge fiscal and monetary imbalances created in confronting the crisis, even the modest recovery of the world economy has revived commodity-demand pressures, and the prices of several commodities have more than doubled over the last eighteen months. If the global economy continues its recovery, commodities may again fuel inflationary pressures, which is likely to prompt governments to reverse monetary policy and reduce fiscal deficits by raising taxes. Furthermore, the large fiscal imbalances mean that any economic recovery would put additional upward pressure on long-term interest rates. Lastly, unless governments are able to divorce themselves from the persistent dominant economic ideology, the necessary future tax increases will almost certainly fall mostly on the middle class, thus exacerbating the third structural factor even further.[11] The net effect of all this is that the incipient growth recovery is likely to be suffocated.

Restoring sustained economic growth requires addressing the underlying structural factors: make aggregate demand less dependent on easy credit, by reversing the concentration of wealth and income; and promote rapid, environmentally and commodity-saving technological change to make global economic growth less commodity- and environmentally-intensive.

The new policies must start from an understanding that the economic orthodoxy based on the ideology of the élites has failed. Excessive wealth concentration needs to be corrected, to enable the middle-class incomes to expand again. The major threat of looming middle-class tax hikes should be removed by drastically raising taxes on the super rich, the only group that has benefited and continues to benefit from the crisis. The burden of overcoming those huge fiscal imbalances must be borne by the wealthy rather than the middle class —not merely out of fairness but as a prerequisite for sustained economic recovery. Restoring fiscal balance by raising taxes on the highest-income groups would do much to boost confidence among ordinary citizens, enabling them to start expanding consumption again, thus increasing the fiscal and monetary multipliers and paving the way for a more permanent economic recovery.

Would large tax hikes on the super-rich undermine economic growth, as the prevailing ideology would have us believe? The top marginal income tax rate in the United States is currently 35%, compared to 91% in the 1950s and early 1960s, for people earning more than the equivalent of US$ 2 million per year in 2009 dollars; and it was above 60% for many years until the early 1980s. Yet this did not prevent the United States economy from growing rapidly in those decades. There are many other examples around the world where extremely high marginal tax rates affecting the super rich have not impeded rapid growth.

Many of the other policies introduced in the 1980s need to be reversed. The financial sector needs to be tightly regulated, and the so-called "Tobin tax",

[11] As this article was being written, the idea of a national sales tax in the USA was gaining increasing support among economists and policy-makers. As is well known, however, such taxes are highly regressive and would inflict yet another blow on the middle class.

consisting on a small levy on financial transactions should be implemented worldwide. Apart from reducing incentives for unproductive financial speculation, this could provide a major source of tax revenue that could help to correct the alarming fiscal imbalances prevailing in most of the advanced economies.

The advanced countries have made significant progress in controlling the destruction of environmental and natural resources caused by their production systems, thanks to the dematerialization of their production and appropriate environmental regulation. These policies made it possible to internalize many environmental and natural-resource costs generated by the productive sector. Nonetheless, the externalities concept used was too narrow, since it focused only on the domestic environmental effects of economic activity. The fact that much of the natural-resource and environmental destruction taking place worldwide is linked to highly material consumption structures in the developed economies does not seem to bother governments in the North. The North's failure to recognize its consumption as a source of natural resource degradation in the South has exacerbated global commodity shortages. An obvious implication is that consumption needs to be regulated in the North to internalize environmental impacts worldwide and not just the domestic environmental impacts in the North. This would help the South prevent further environmental destruction, which in the long run would make the commodity supply curve even more inelastic.

(Original: English)

Bibliography

Acemoglu, D., S. Johnson and J.A. Robinson (2001), "The colonial origins of comparative development: an empirical investigation", *American Economic Review*, vol. 91, No. 5, Nashville, Tennessee, American Economic Association.

Autor, D., L. Katz and M. Kearney (2006), "The polarization of the U.S. labor market", *American Economic Review*, vol. 96, No. 2, Nashville, Tennessee, American Economic Association.

Autor, D., F. Levy and R.J. Murnane (2003), "The skill content of recent technological change: an empirical exploration", *The Quarterly Journal of Economics*, vol. 118, No. 4, Cambridge, Massachusetts, The MIT Press.

Barbier, E. (2005), *Natural Resources and Economic Development*, Cambridge, Cambridge University Press.

Böhringer, C. and A. Löschel (2003), "Market power and hot air in international emissions trading: the impacts of US withdrawal from the Kyoto Protocol", *Applied Economics*, vol. 35, No. 6, London, Taylor & Francis.

Boone, L., N. Girouard e I. Wanner (2001), "Financial market liberalisation, wealth and consumption", *OECD Economics Department Working Papers*, No. 308, Paris, Organisation for Economic Co-operation and Development (OECD).

Bosworth, B. and S.M. Collins (2008), "Accounting for growth: comparing China and India", *Journal of Economic Perspectives*, vol. 22, No. 1, Nashville, Tennessee, American Economic Association.

Caprio, G., A. Demirguc-Kunt and E. Kane (2008), "The 2007 meltdown in structured securitization: searching for lessons, not scapegoats", *Policy Research Working Paper Series*, No. 4756, Washington, D.C., World Bank.

Card, D. and J.E. DiNardo (2002), "Skill biased technological change and rising wage inequality: some problems and puzzles", *Journal of Labor Economics*, vol. 20, No. 4, Chicago, University of Chicago Press.

Da Motta, R.S. (2003), "Determinants of environmental performance in the Brazilian industrial sector", *Working Paper Series*, Rio de Janeiro, Institute of Applied Economic Research (IPEA).

De Janvry, A. (1975), "The political economy of rural development in Latin America: an interpretation", *American Journal of Agricultural Economics*, vol. 57, No. 3, Oxford, Oxford University Press.

Engel, S. and R. López (2008), "Exploiting common resources with capital-intensive technologies: the role of external forces", *Environment and Development Economics*, vol. 13, Cambridge, Cambridge University Press.

Farrell, D. and A. Grant (2005), "China's looming talent shortage", *McKinsey Quarterly*, McKinsey & Company.

Feenstra, R.C., J. Romalis and P. Schott (2001), "U.S. imports, exports and tariff data, 1989-2001", *NBER Working Papers*, No. 9387, Cambridge, Massachusetts, National Bureau of Economic Research.

Ghertner, D.A. and M. Fripp (2007), "Trading away damage: quantifying environmental leakage through consumption-based, life-cycle analysis", *Ecological Economics*, vol. 63, No. 2-3, Amsterdam, Elsevier.

Gordon, R. and I. Dew-Becker (2007), "Selected issues in the rise of income inequality", *Brookings Papers on Economic Activity*, vol. 38, No. 2, Washington, D.C., The Brookings Institution.

Huber, R.M, J. Ruitenbeek and R.S. da Motta (1998), "Market-based instruments for environmental policymaking in Latin America and the Caribbean: lessons from 11 countries", *World Bank Discussion Paper*, No. 381, Washington, D.C., World Bank.

Jerzmanowski, M. and M. Nabar (2008), "Financial Development and Wage Inequality: Theory and Evidence" [online] http://ssrn.com/abstract=1280978

Kaplan, S.N. and J. Rauh (2009), "Wall Street and Main Street: what contributes to the rise in the highest incomes?", *Review of Financial Studies*, vol. 23, No. 3, Oxford, Oxford University Press.

Khor, M. (2000), "Globalization and the South: some critical issues", *Discussion Paper*, No. 147 (UNCTAD/OSG/DP/147), New York, United Nations Conference on Trade and Development (UNCTAD), April.

Krugman, P. (2006), "The great wealth transfer", *Rolling Stone*, November.

_____ (1994), "The myth of Asia's miracle", *Foreign Affairs*, vol. 73, No. 6, New York, Council on Foreign Relations.

Lall, S. and M. Albaladejo (2004), "China's competitive performance: a threat to East Asian manufactured exports?", *World Development*, vol. 32, No. 9, Amsterdam, Elsevier.

Levison, A. and M. Scott Taylor (2008), "Unmasking the pollution haven effect", *International Economic Review*, vol. 49, No. 1, Philadelphia, University of Pennsylvania/Osaka University Institute of Social and Economic Research Association.

López, R. (2008), "'Sustainable' economic growth: the ominous potency of structural change", document presented at the Sustainable Resource Use and Dynamics Conference (Switzerland, June 2008), College Park, University of Maryland.

López, R. and A. Stocking (2009), "Bringing growth theory 'down to earth'", *Working Papers*, No. 09-01, College Park, University of Maryland.

Löschel, A. and Z. Zhang (2002), "The economic and environmental implications of the US repudiation of the Kyoto Protocol and the subsequent deals in Bonn and Marrakech", *Review of World Economics*, vol. 138, No. 4, New York, Springer.

McConnell, M.M., P.C. Mosser and G. Perez-Quiros (1999), "A decomposition of the increased stability of GDP growth", *Current Issues in Economics and Finance*, New York, Federal Reserve Bank of New York.

Noland, M. (1997), "Has Asian export performance been unique?", *Journal of International Economics*, vol. 43, No. 1-2, Amsterdam, Elsevier.

Panagariya, Arvind (2006), "India and China: trade and foreign investment", *Working Paper*, No. 302, Stanford, Stanford Center for International Development.

Philippon, T. and A. Reshef (2009), "Wages and human capital in the US financial industry: 1909-2006", *NBER Working Paper*, No. 14644, Cambridge, Massachusetts, National Bureau of Economic Research.

Piketty, T. and E. Saez (2003), "Income inequality in the United States, 1913-98", *Quarterly Journal of Economics*, vol. 118, No. 1, Cambridge, Massachusetts, The MIT Press, February.

Rodrik, D. (2007), *The Real Exchange Rate and Economic Growth: Theory and Evidence*, Cambridge, Massachusetts, John F. Kennedy School of Government, Harvard University, July.

Saez, E. (2009), *Striking it Richer: the Evolution of Top Incomes in the United States (Updated with 2007 estimates)*, Berkeley, University of California.

Shapiro, M.A. (2009), *Environmental Sustainability in East Asia: Policies and Technological Output*, document prepared for the APPAM-KDI Conference, Seoul.

Simpson, D.R., M. Toman and R. Ayres (2005), *Scarcity and Growth Revisited: Natural Resources and the Environment in the New Millennium*, Washington, D.C., Resources for the Future.

Sloan, J. (1997), "The Reagan presidency, growing inequality, and the American dream", *Policy Studies Journal*, vol. 25, No. 3, New York, John Wiley & Sons.

Sokoloff, K.L. and S.L. Engerman (2000), "History lessons: institutions, factor endowments, and paths of development in the new world", *Journal of Economic Perspectives*, vol. 14, No. 3, Nashville, Tennessee, American Economic Association.

Taylor, P. and others (2008), *Inside the Middle Class: Bad Times Hit the Good Life*, Washington, D.C., Pew Research Center.

KEYWORDS

Economic development
Local development
Development strategies
Regional economy
Public goods
Social capital
Decentralization in government
Developing countries

From national to local economic development: theoretical issues

Mario D. Tello

Based upon a brief and selective survey of the literature on local economic development (LED), this paper analyses four theoretical aspects that distinguish "local" economic development theories from their "national" counterparts. These are: location factors, local public goods, active participation by a variety of private agents, and the multidisciplinary approach of LED theories. This analysis could be used to design an academic discipline of LED, which seldom exists in developing countries, and shed light on the objectives and roles of agents involved in ongoing decentralization and LED processes in those countries.

Mario D. Tello
Associate Professor and
Research Fellow
CENTRUM Católica and
Economics Department,
Catholic University of Peru
mtello@pucp.edu.pe

I
Introduction

Local economic development (LED) theory and national economic development (NED) theory are well established disciplines that are taught in most universities in industrialized countries (Blakely, 2003). Although developing countries, particularly in Latin America, have a long tradition of implementing regional policies, such as river-basin planning, tax incentives, regional development agencies, growth hubs and integrated rural development, which spawned an initial wave of graduate programmes on local economic development back in the early 1960s, Latin American universities rediscovered an interest in this field as recently as the early 2000s. The decentralization processes implemented in many developing countries over the last two decades (Montero and Samuels, 2004; Stren and others, 2002; Aghón, Alburquerque and Cortés, 2001; Oxhorn, Tulchin and Selee, 2004; and Rondinelli and Cheema, 1983) have contributed to this renewed interest. International development institutions are also paying attention to LED, and providing funding for it, as shown on their specific Internet pages.[1] Generally speaking, LED issues in geographic areas (regions, counties, provinces and/or departments) in developing countries are analysed through the NED framework without including distinctive local-development features. Moreover, government agencies operating at the local level, along with economic, social and political agents residing in specific areas, perceive that economic development in those areas depends heavily on the interventions and economic policies implemented by central (or federal) government.[2] This article seeks to form a bridge between current LED and NED theories, by briefly surveying four additional theoretical aspects that distinguish local economic development theories from their national counterparts in the analysis of LED issues.

The first of these consists of the location factors associated with specific geographic areas in an economy; most *national* economic development theories implicitly assume an economy's territory to be homogenous. The second aspect is the nature of the goods and services provided by different levels of government, where NED theories implicitly assume that the influence of the public goods and services provided by government covers all geographic areas of the economy. The third aspect is the role and participation of specific agents living in the local areas. Traditional NED theories do not incorporate this explicitly. Over the last two decades, however, modern NED theories have given a more important role to the behavior of agents, although not necessarily linked to the local areas in which they live. The fourth aspect is the multidisciplinary approach to the analysis of LED issues. In NED theories, the focus is on economic fundamentals and the workings of markets, institutions and organizations, irrespective of the geographic areas in which they operate. A brief and selective survey of these four aspects is presented in sections III to VI of this article. As a starting point, however, section II discusses some of the definitions of LED found in the literature; and, lastly, section VII makes a number of concluding comments.

☐ The author thanks Esteban Hnyilicza and an anonymous referee for their helpful comments, and particularly for drawing his attention to early applications of local economic development policies in Latin American countries.

[1] For example, the World Bank (WB), Economic Commission of Latin America and the Caribbean (ECLAC), United Nations (UN), the Inter-American Development Bank (IDB), the Organisation for Economic Co-operation and Development (OECD), and others.

[2] See the study of two regions of Peru (Tello, 2008).

II

Definitions of local economic development (LED)

The World Bank website states that: *"Local economic development (LED) offers local government, the private and not-for-profit sectors, and local communities the opportunity to work together to improve the local economy. It focuses on enhancing competitiveness, increasing sustainable growth, employment generation and ensuring that growth is inclusive. LED encompasses a range of disciplines including physical planning, economics and marketing. It also incorporates many local government and private sector functions including environmental planning, business development, infrastructure provision, real estate development and finance."*

In various contributions from ECLAC (Aghón, Alburquerque and Cortés, 2001; Finot, 2001) and IDB (Albuquerque, Llorens and Del Castillo, 2002; Llisteri, 2000), LED is defined as: *"the structural and growth process which, by making full use of local resources, leads to a continuous increase of the welfare of the people living in a local area or region within a country. The process includes three dimensions: economic (encompassing the means of production that allow local firms to make efficient use of local resources, generate scale economies and increase their productivity and market competitiveness); sociocultural (characterized by the social and economic network, in which local values and institutions support the LED process); and the administrative and political dimension (involving local initiatives that create an appropriate local and business environment to foster local economic development)."*

In the economics literature and from the industrialized-economy standpoint (Blair, 1995; Bartik, 1995; Bingham and Mier, 1993; and Malizia, 1985) LED is traditionally defined as changes that affect a local economy's capacity to *"increase economic growth, generate employment and create new wealth for local residents."* A modern definition is given by Blakely (2003) and Blakely and Bradshaw (2002), who state that the LED field is a combination of disciplines and an amalgamation of policies and practices; and that today LED scholarship is a small and growing industry in its own right. The LED concept is based on four factors: (i) indigenous resources and local control; (ii) new wealth formation; (iii) new capacity building, and (iv) resource expansion.

These definitions of LED share several aspects that tend not to be explicitly included in definitions of national economic development. The first is the geographic location of the LED process. Countries are usually divided, geographically, politically or administratively (see OECD, 2002) into different territorial levels or spatial units (such as states, regions, departments, provinces, districts, municipalities and so forth); and residents (economic, political and social agents) are identified with those territories. Greffe (2004) postulates at least three justifications for the local approach to the economic development process: one concerns the specific features of certain territories, which may affect the adequate functioning of spontaneous market mechanisms, or the policies formulated for an area without taking local features into account. Another involves the multidimensional nature of employment problems, which are usually presented as a mismatch between supply and demand. Thus, factors such as training, housing or mobility, health care, minimum-wage constraints and others can in fact only be identified and managed in a precise manner and close to the stakeholders involved, which means that initiatives must be planned, executed and coordinated at the local level. The last concerns the global-economy context of the countries, under which the local approach is justified, because it allows for greater synergy between the economic and social foundations of markets.

The second aspect of LED processes relates to the provision of public goods and services at the local level.[3] Governments use a variety of instruments to achieve the goals of efficient and equitable resource allocation in an economy, which entails providing public goods and services such as infrastructure (roads, bridges, etc.) and social services (education, health, etc.). Most of these "goods and services" are supplied locally, however, (Tiebout, 1956a) and mainly benefit the residents of the areas in which they are supplied. Thus, the difference between the public goods and services that affect residents in a country's territories

[3] McGuire and others (1994), conceptualize local-development capacity in terms of three broad factors: citizen participation, community or local structure, and development instruments. The latter relate to local government policies and the supply of goods and services that support the LED process.

or spatial units and what the literature calls "local public goods and services" is another aspect that is not explicitly considered in definitions and theories of national economic development.

The third specific feature of LED (in contrast to NED) is its explicit consideration of the role and participation of local residents (economic, political and social agents and citizens) in the local development process. Firstly, local residents demand goods and services from government (at all levels), to support business activities and reduce income inequalities in local areas. Secondly, through the activities and participation of specific agents, together with alliances and partnerships among agents in local areas, they can also help increase the supply of public goods and services, boost local economic growth and also influence economic policies at the local and national levels (OECD, 2007).

The fourth aspect that distinguishes LED from the NED is the former's multidisciplinary approach. To address the geographic-location or spatial dimension of the LED process, the key tools for analysing the LED process are drawn from the fields of regional, urban, rural and geographical economics. In contrast, the public-finance approach is used to analyse the provision of local public goods; and tools drawn from political economy, sociology and psychology are needed to understand the actions and interventions of agents in the LED process. The following sections deal with each of these four aspects in turn.

III

Location, geography and regional economic theories as inputs for LED theories

The starting point for understanding the relevance of economic theories of geographic and regional location for LED is regional economic-base theory borrowed from regional economics. Andrews (1953) defines the "economic base" as the set of activities in a "region" (defined as a local geographic area or a specific spatial unit) which "exports" goods and services to points outside its economic boundaries, or sells its goods and services to outsiders. Based on this definition, regional economic-base theory postulates that the "region's" economic growth is driven by growth in the "export activities" of the economic base (Sirkin, 1959, Tiebout, 1956b, and North, 1955).[4] In LED theories, the set of factors that determine location, activities and economic growth of the economic base of a region[5] come from the areas of the economics of location and economic geography (both traditional and new). These factors include: external factors not localized in the "region", local resource endowments (human, natural and capital)

[4] The economic base activities or *basic industries* of a "region" are the set of "export commodities or industries or staples". The non-base activities or *non-basic industries* of a "region" consist of "subsidiary industries" developed and/or derived from the economic-base activities, demand for which is determined locally by the residents of the "region" (North, 1955).

[5] Taking market demand, its geographic distribution, and local resource endowments as exogenous, location theories (Weber, 1957; Isard, 1956; von Thünen, 1826) and central place theories (Christaller, 1966) introduced transport costs and distance from markets as factors explaining the fixed and spatial distribution of the region's (base and non-base) activities. Marshall (1890) introduced the concept of "external or spatial economies" associated with the proximity of economic actors within a given location. Such economies arise from three types of cost and market-location advantages: job creation and the capacity to absorb workers with specialized skills, who are attracted to local areas and form a supply pool of workers; creation of demand for (specialized and complementary) inputs which are profitable to produce, given the proximity of production markets; and the generation of technological spillovers via the exchange of information and production methods among firms located in the same spatial unit. Weber (1957) introduced the concept of "agglomeration economies" which arise from transaction cost savings stemming from the proximity of firms within a specific area; and "external economies of scale", also introduced by Marshall (1890), which are defined as the cost savings accruing to a firm because of the size or growth of output of the industry as a whole. Such economies contrast directly with internal scale economies, which are a source of increasing returns based on larger plant size. These external economies are essentially spatial externalities, which can generally be defined as economic side-effects of the proximity between economic actors. They can be negative or positive, static or dynamic, pecuniary or technological. The static variety is reversible, whereas dynamic externalities are associated with the technological progress, greater specialization, and division of labour that accompanies and/or drives growth and development (Young 1928). Pecuniary externalities are internalized thorough market mechanisms whereas technological ones are not, although they may be internalized through non-market mechanisms.

and physical infrastructure, together with factors such as distance from markets (where the economic base's export demand comes from); transport costs (which affect production features of the goods and services produced in the economic base and the spatial distribution of the production of those goods and services); spatial (or external) agglomerations; and external economies of scale.

In a series of papers (Fujita and Krugman, 1995, Fujita and Mori, 1997; Fujit, Krugman and Mori, 1999; and Stahl, 1987), the new economic geography theories put forward by Krugman (1991) and Fujita (1988) have introduced all these concepts in a formal way (using rational optimizing decisions by agents, interactions among agents, skilled labour and capital mobility in a general equilibrium framework), allowing for the endogenous location of manufacturing and agriculture activities and explaining the agglomeration of activities around cities and the economic growth of the regions. The agglomeration of consumer and producer activities in a given spatial unit is formally shown as the outcome of two forces: centripetal or push forces, and centrifugal, dispersion or expulsion forces. The first of these stems from spatial, agglomeration, and external scale economies and the creation and development of backward and forward linkages, or market-size effects. The second group of forces is generated by the immobility of factors such as land and workers (factor rewards decrease as the distance from the agglomerated activities increases), fierce competition, and pure external diseconomies (Krugman, 1999 and Fujita and Thisse, 1996).

The LED literature has formulated various mechanisms through which the agglomeration of economic (base and non-base) activities generates economic growth and development for the local region or spatial unit as a whole. The first is the income/employment-multiplier mechanism (Sirkin, 1959), whereby higher income and employment in economic-base activities will increase the demand for goods and labour in non-base activities. A second mechanism, drawing on "staples and vent-for-surplus models" (Findlay and Lundahl, 1994), operates through the backward and forward inter-sectoral linkages generated by the economic-base activities when there are idle resources in the region. Multi-sector and multi-regional growth models have been based on these two mechanisms of economic base LED models (Loveridge, 2004; Nijkamp, Rietveld and Snickars, 1987).

The third mechanism involves "external and agglomeration economies", as sources of the centripetal and centrifugal forces of the agglomerated activities in a region. Growth and development poles (Perroux, 1950, 1955, and 1988) and product-cycle development theories (Vernon, 1966) formulated the basic ideas, which were then formally modelled by the new economic geography approach (Fujita and Thisse, 2003; Walz, 1996; Baldwin and Forslid, 2000, Black and Henderson, 1999; Martin, Gianmarco and Ottaviano, 1999 and 2001).

According to Perroux (1950), an economic space, conceptualized as a field of forces, consists of centres (or poles) from which centrifugal forces expel and to which centripetal forces attract. As a reason for such agglomeration, Perroux argued that dominant (leading) firms are comparatively efficient, and they can make effective use of innovations and thus expand their output by more than other firms. This effect would be propagated and perceived throughout society through a multiplier process. Hence, a polarization process is needed for the population at large to benefit. Perroux (1955) also postulated that economic growth does not occur everywhere at once, but manifests itself in points or "poles" of growth of varying intensities; and it spreads through different channels with variable final effects on the economy as a whole. Consequently, a growth pole is an "aggregation of propulsive industries" connected to the surrounding (or peripheral) environment. It is a "set with capacity to induce growth [defined as a lasting increase of a dimensional indicator] in another set". Perroux (1988) adds that the development pole is a "set with capacity to generate economic and social structures, whose effect is to increase the complexity of the whole and expand its multidimensional performance". Based on the dynamic stages of products or "product cycles" (Levitt, 1965),[6] and evoking the leading role of specific and efficient firms in the creation of the

[6] Levitt (1965) distinguish four stages: (i) the stage of market development, or creation of a product before there is a proven demand for it and often before it has been fully tested technically in all respects. Sales volumes are low and progress slowly; (ii) the market-growth stage, when a product has survived its introduction, demand starts to pick up, and the size of the total market expands rapidly; this is the peak stage for any product; (iii) the market-maturity stage, when sales growth has started to slow and is approaching the point where the inevitable decline will begin. In the two latter stages, products that are considered economic-base activities generate demand in subsidiary industries or non-base activities, and are "exported" to markets outside the spatial units in which they are produced; and (iv) the market-decline phase, in which the product starts to lose consumer appeal and a downward slide in sales sets in. Vernon (1966) called this the standardization phase, and it starts at the product-maturity stage.

growth poles, Vernon (1966) argues that regions which are capable of producing at the market-development and growth stages of the product cycle grow more rapidly. The region's capacity to produce in those two stages depends, among other things, on the degree of technological innovation among the firms located in the region, the region's endowment of innovative firms and the income generated in the region.

These ideas have been formalized by the new economic geography (NEG) approach, in which the sources of agglomeration, the rate of technological innovation (associated with investment in R&D activities) and technological spillovers are modelled as the key mechanisms driving local economic growth. Furthermore, and as a result of the local growth models of NEG, regions are divided in two groups: central regions, which are the more developed; and peripheral regions which are less developed. Central regions produce goods included in the first three stages of a product cycle, while the peripheral regions produce at the standardized stage. These concepts of "centre and periphery" first appeared in the seminal work of Prebisch (1959).

A fourth and related mechanism shared by NEG models is the "circular and cumulative causation" mechanism generated by the "lock-in" effects of agglomeration (Fujita and Thisse, 1996; Arthur, 1989). Under this mechanism, the set of (usually differentiated final and/or intermediate) goods at the first two stages of the product cycle will be produced by innovative firms in locations where there is a relatively large and attractive market (measured the number of workers or consumers). But the market will be relative large and attractive if a relatively large number of producers locate their production there. Thus, the concentration of the leading sector (usually manufacturing) in a given location is generated and reinforced through this circular and cumulative causation mechanism.[7] The initial activity (usually with increasing-returns technology) and its location, which generate this mechanism, stem from lock-in effects caused by accidental or historical conditions (or events).[8]

While local economic development depends on location factors that fuel development in the economic base and region through mechanisms of transmission between the economic base and non-base activities, local public goods and services and economic policies are also key ingredients in local development capacity.

[7] A relatively large market is attractive for firms because of the potential demand that may exist for their goods (a relatively large number of consumers) and the availability of (particularly skilled) labour (many consumers also means many workers). Thus, firms will demand inputs and labour through backward linkages. Moreover, large market size results in lower prices and higher real wages, which may induce workers to migrate to locations where the leading sector is concentrated. Thus, firms will increase the supply and number of (differentiated) goods through forward linkages, and they will lower their prices in the locations where their goods are produced. According to Fujita and Thisse (1996), the backward and forward linkages of the circular cumulative causation mechanism turn increasing returns to scale at the firm level into increasing returns to scale for the region as a whole.

[8] Increasing-returns technology may lead to multiple equilibriums; with economic conditions and random events determining which equilibrium actually occurs.

IV

Local public goods, local government and LED policies

Several issues arise when local public goods and services (LPGs) are introduced into the analysis of the local economic development process, three of which are analysed here.[9] The first is the efficiency with which the central (federal) government provides LPGs. The second is the level of local government that can supply LPGs most efficiently. The third is the economic and social role of local government in the LED process.

On the first of these topics, the pioneering studies by Oates (1972) and Olson (1969) provided a starting point for the analysis of decentralization theory or fiscal federalism. Recent surveys of this strand of literature are contained in Oates (1999, 2005) and Bardhan (2002) among others. According to the Oates (1972) decentralization theorem, in the absence of heterogeneous consumer preferences, and if LPGs overflow the local jurisdictions in which they are supplied, the most efficient arrangement is for central government to provide a common level of public goods and services to all localities. In contrast, when preferences are heterogeneous and there are no spillovers across jurisdictions, local governments are more efficient in providing LPGs to their respective localities. In the first-generation theory of fiscal federalism, Oates (2005) envisioned a setting in which governments at different levels provide public goods, whose spatial patterns of benefits are encompassed by the geographical scope of their jurisdictions. This allocation of LPGs is called a "perfect mapping" or "fiscal equivalence" to use Mancur Olson's (1969) terminology.

Based upon (i) public-choice and political-economy studies focusing on political processes and the behavior of political agents, and (ii) the extensive literature on information problems, the modern theory of fiscal federalism (or second-generation fiscal theory) summarized by Oates (1999, 2005) analyses the workings of different political and fiscal institutions in an imperfect-information and control setting, focusing basically on the incentives that those institutions embody and the behaviour they induce from utility-maximizing participants. In this context, the first issue of whether to centralize or decentralize public activities is analysed from this new perspective; and the trade-offs between the inefficiencies under centralized provision of public services stemming from more uniform outputs that fail to reflect divergences in local tastes and conditions, versus the inefficiencies of local supply resulting from the failure to internalize inter-jurisdictional externalities, are analysed from a somewhat (but not altogether) different perspective.

The pioneering work of Tiebout (1956b) is the starting point for the second topic, which is closely related to the first one. Tiebout argues that government levels are directly related to the set of LPGs that governments supply within their jurisdictions. He showed that, under conditions of high household mobility, households can optimally (and efficiently) choose the jurisdiction of residence that offers the LPG package that best suits their preferences. In contrast to this non-spatial allocation theory of LPGs, the Hochman, Pines and Thysse (1995) geographic or locational approach to the provision of LPGs postulates that their consumption entails transport costs. These costs increase with the distance between residential locations and the public facilities where the goods and services in question are available. Consequently, decentralization does not have to be based on the types of LPGs supplied by local governments, but could be based on territories instead. It has been shown that optimal provision of LPGs can be decentralized only through metropolitan governments supplying the whole range of LPGs over one or more appropriate territory.[10]

[9] A fourth issue, which emerged in the 1990s and is partially discussed in this paper, is local governance, for which recent surveys can be found in Liou (2007) and Shah and Shah (2006). This issue concerns the various types of institutional framework that best enable governments to fulfil their economic role. It deals primarily with the failings of government institutions when intervening in markets. A fifth and related issue, not discussed here, is decentralization (in other words the transfer of specific functions from central government to local governments). Surveys on this topic are presented by Rondinelli and Cheema (1983), Litvack, Ahmad and Bird(1999), and Bardhan (2002) among others.

[10] The geographic jurisdiction of the metropolitan local government is a territory where the user charge collected from its residents, plus the corresponding aggregate land rent, is just equal to the

The third issue is what most LED practitioners have concentrated on. The starting point is the theory of the economic role of the government as formulated by Musgrave (1959) and Samuelson (1954). According these authors, efficiency, equity and (macroeconomic) stability are the three basic market principles on which the economic role of the government needs to be based. Consequently, market failures or distortions, such as the existence of public and merit goods, externalities and natural monopolies, have traditionally been viewed as market inefficiencies to be corrected by government. Income-distribution inequalities arising from the market allocation of resources is another area that requires government intervention in the economy. Consequently, market distortions and inequalities arising from the jurisdictional distribution of resources may be territorial features of the market; and central (or federal) and local governments share the role of intervening in the economy at the national and local level; while the goal of macroeconomic stability is left as an exclusive preserve of central government.[11]

Recently, Shah and Shah (2006) have summarized the evolving economic role and responsibilities of local governments since the Musgrave and Samuelson contributions. Using the traditional fiscal federalism approach, which is based on the market-failure and LPG- provision approach to the role of government, they classify the different expenditures, public goods and service provision and taxation responsibilities of government and assign them to three levels: central (or federal), regional (states or provinces) and local (municipalities and metropolitan areas).

Under the same market-failure and LPG-provision basis for the role of government, the new public management approach focuses on what local governments should do and how they should do it better. From this standpoint, (central and local) government is viewed as the "agent", and the population represents the "principal"; so its responsibility is to serve the public interest and create public value (defined by Moore (1996) and measured as improvements in social outcomes or quality of life). This approach also suggests a change in the way local government should fulfill its responsibility: from the top-down approach of fiscal federalism to a bottom-up approach in which local governments behave as managers serving the people living in their local jurisdictions (Shah, 2005 and Caulfield 2003).

In contrast to these two perspectives, public choice theory and the new institutional economics focus on government failings rather than market failures. These approaches propose different ways of organizing governance to avoid the inefficiencies caused by government failures. The public-choice literature endorses the self-interest doctrine of government and argues that the various stakeholders involved in policy formulation and implementation can be expected to use opportunities and resources to advance their self-interest. Consequently, for local governments to serve the people's interests, they need full local taxing and spending autonomy, and they must be subject to competition within and beyond government. In the absence of these prerequisites, local governments are likely to be inefficient and unresponsive to citizen preferences (Boyne 1998). In contrast, the new institutional economics postulates various orders of government (as agents) to serve the interests of citizens (as principals). The jurisdictional design should ensure that these agents serve the public interest while minimizing transaction costs for the principals (Williamson, 1985; Horn, 1997 and Shah, 2005).

Similar to the previous two approaches, the network form of governance is also concerned with the institutional arrangements of government, while focusing on both market and government failures. It provides specific guidance in dealing with government failures in a hierarchical form of public governance, and local government involvement in a partnership with multiple organizations. Under this perspective, a network mechanism of governance has been advanced for local governments, based on trust, loyalty and reciprocity between partners with no formal institutional safeguards. Networks formed on the basis of shared interests (interest-based networks) can provide a stable form of governance provided their membership is limited to partners that can make significant resource contributions and there is a balance of powers between members. Local government may thus have an opportunity to serve as a catalyst in facilitating the roles of both interest-based and hope-based networks to improve social outcomes for local residents (Dollery and Wallis 2001).

A more proactive approach to the role of the local government is formulated in the LED literature summarized by Liou (2007), Bartik (1995 and 2003),

cost of supplying all the LPGs provided by the metropolitan government.

[11] Watt (2006) and King (1984), among others, argue that stabilization and redistribution (usually through transfers) are the main roles of central government, whereas the local-government role is to allocate local public goods efficiently.

Bachtler and Yuill (2001), Blair (1999) and Blackely and Bradshaw (2002), among others. In addition to the traditional and modern theories of the economic role and responsibilities of local government summarized in Shah and Shah (2006), LED practitioners (from the United States and Europe) propose another role for local government, namely to implement policies that foster local economic development. The rational for this role is based on specific aspects of theoretical models of LED, which may be related to, or are claimed to be consistent with, the efficiency and equity roles of the traditional and modern theories of local government.

According to Bartik (2003), local-government economic-development policy is defined as the special activities undertaken by local government to promote economic development. Activities referred to as "economic development programmes" fall into two categories: (i) providing incentives and customized assistance for individual businesses from which greater economic development benefits are expected; and (ii) strategic initiatives to alter more general tax, spending, and government regulatory policies to promote local economic development.

Blackely and Bradshaw (1999), Blair (1999) and Bachtler and Yuill (2001) distinguish up to three "waves" of LED programmes and policies implemented by LED practitioners in developed countries: the first wave, before the 1980s and based on location theories of LED, dominated by (incentive and subsidy) programmes designed specifically to attract footloose firms from old industrial areas into growing regions; the second wave, during the 1980s, based upon traditional and neoclassical regional development theories, dominated by local growth programmes (such as creating new businesses, increasing investment capital, developing incubators, and providing technical assistance); and the third wave, from the 1990s onwards, based on LED competitiveness and cluster theories, dominated by policies aimed at providing an appropriate regional business environment, emphasizing public-private partnership, collaboration and coordination.

From the institutional standpoint of local governance models, however, local governments are not the "principals" of the LED process; instead, the citizens' residents in local areas, in their diversity of composition and roles, are the "principals", and their active participation is also a distinctive feature of LED theories. The next section considers the role and participation of the citizen or private agent in the local economic development process.

V

The role and participation of local private agents in the LED process

Various groups of citizens or private agents (entrepreneurs, women, groups representing social capital, etc.) play multiple roles in the LED literature,[12] and they affect the local development process through a variety of mechanisms. Entrepreneurship (Bates, 1993 and Malecki, 1994), "intrapreneurship"(Pinchot III, 1985) or entrepreneurship capital (Audretsch and Keilbach, 2004a) are the terms normally used in the LED literature to identify one of the oldest mechanisms used by entrepreneurs and managers to generate knowledge creation and innovation (Schumpeter, 1934) leading to regional/local economic growth (Audretsch and Keilbach, 2007, 2005, 2004b; and Lawton, Glasson and Chadwick, 2005). Innovation, however, is not the only business activity that affects the LED process. An OECD (2003) report summarizes activities and interactions among entrepreneurs in local areas that affect the area's economic development and growth process. Entrepreneurs are sources of investment, savings, job creation, networks and agent-coordination, which may enhance the development capacity of the local areas in question.

A second way a group of citizens can affect the LED process is through social capital (Trigilia, 2001; Putnam, 1993). Although social capital (defined as

[12] For example: (i) governors (such as owner-authorizers, voters, taxpayers, community members); (ii) activist-producers (such as providers of services, co-producers, self-helpers obliging others to act); and (iii) consumers (clients and beneficiaries) (Moore 1996).

the level of interpersonal trust, civic engagement and organizational capability prevailing in a community or among group of citizens) is a feature of specific local geographic areas, the concept was originally conceived as having economic-development consequences at the national level (Woolcock and Narayan, 2000; and Zabojnik and Francois, 2005).

Moreover, as pointed out by Durlauf (2002), social capital also has an impact on issues relating to political participation (DiPasquale and Glaeser, 1999), development traps (Woolcock, 1998), human-capital formation (Coleman, 1988) and the efficiency of the judicial system (La Porta and others, 1997). In the first case, citizens' investment in social capital might include membership of a social organization leading to better coordination and political actions within a community. In the second case, a lack of trust among citizens (or social capital) in a community could contribute to the persistence of development traps. In the third case, the trust and coordination dimensions of social capital may improve information channels, communications skills, and knowledge creation, transfer and flows, leading to human capital formation. In the last case, the efficiency of the judicial system may affect the level of trust among people and thereby encourage or discourage the formation of social capital.

Women are another group of citizens whose participation has recently been discussed in the LED literature (Blumenberg, 1998). Beyond the gender-inequality and social-exclusion issues in the development process (Weinberger and Jütting, 2001; Blumenberg, 1998; and Elson, 1998), gender issues and women's role in the family also have been linked to other aspects of social development (as defined in Mokate, 2004), such as: poverty; fertility rates; human-capital formation; household nutrition; infant, child and maternal mortality rates (Elson, 1998). In terms of women's role in economic development and growth at the local level, the LED literature emphasizes three roles: as entrepreneurs, as innovators (particularly in retail and service industries), and in forming social capital (Forsyth, 2000; and Molyneux, 2002).

Another way citizen participation can influence the LED process is through local "partnerships" (or cooperation, collaboration, coordination or association) between two or more group of agents (including institutions, community or private organizations, and government entities) sharing common development objectives that are location-based and operate within spatially defined social, cultural, economic and political relationships. Local partnerships are territorial entities by definition (OECD, 2007). Aside from issues relating to the definitions and forms of partnerships (such as vertical supplier or purchaser associations, horizontal and lateral governmental partnerships, and private and public partnerships, as listed in Camarero, Hernández and San Martín (2008), it is through their social-capital and local-governance features that partnerships can influence local economic development process (OECD, 2007).

The "trust" element in social capital may improve knowledge flows among the members of local partnerships (Jones, Kashlak and Jones, 2004) and overcome the market failures arising from market activities based on the partnerships' development objectives and programmes (for the supply of public infrastructure); and the local-governance feature may make it possible to overcome government failures arising from the process of achieving development objectives (by providing stability in a turbulent environment caused by economic, social and political changes; and by improving market efficiency through adequate control and allocation of resources and responsibilities among partnership participants, as suggested in Walsh and Meldon (2004) and OECD (2001)).

Active participation by these and many other groups of agents and entities —such as the economic or interest groups analysed by Gray and Lowery (1988); and the local developers analysed by Laukkanen and Niittykangas (2003)— is considered by LED practitioners as a part of any planning strategy to foster local economic development (Walsh and Meldon, 2004).

VI

The multi-disciplinary approach of modern LED theories

The fourth distinctive feature of LED theories is their multidisciplinary approach to the LED process. As shown in previous sections, LED theories draw on a variety of disciplines (such as spatial and location theories, public finance theory and the theory of governance, among others). Nonetheless, these are considered separately to emphasize different aspects of the local development dynamic. Until the 1980s there was a consensus among LED practitioners in terms of the various factors taken into account in the LED process (Thompson, 1968). In the early 1990s, however, a set of LED theories emerged with a multi-disciplinary, multi-dimensional or multi-factorial approach to the LED process within a unified framework. Those theories have been associated with the third wave of LED policies and local government programmes, and they highlight simultaneous interactions between various factors to achieve local development goals. Two of the most widely studied "modern" theories in the LED literature are the regional/local competitiveness or "cluster" approach developed by Porter (1990) (surveyed, by Budd and Hirmis, 2004 among others), and the cluster-development approach (Rainess, 2003; Rocha, 2004; Enright, 1996).

Leaving aside the problems involved in defining the concept of competitiveness (Lall, 2001) and clusters (Martin and Sunley, 2003), Porter's approach to LED is based on his competitiveness diamond which contains four sources of the competitive or productivity advantages of national or regional economies. These are: demand conditions; factor (or input) conditions; firm strategy, structure and rivalry; and related and supporting industries, According to Porter (2000), "demand conditions at home have much to do with whether firms can and will move from imitative, low-quality products and services to competing on differentiation. In low-productivity economies, the focus is heavily on foreign markets. Advancement requires the development of more demanding local markets. The presence or emergence of sophisticated and demanding home customers presses firms to improve and provides insights into existing and future needs that are hard to gain in foreign markets. Local demand also can reveal segments of the market where firms can differentiate themselves. In a global economy, the quality of local demand matters far more than does its size."

"Factor inputs", argues Porter (2000), "range from tangible assets such as physical infrastructure, to information, the legal system, and university research institutes that all firms draw on in competition. To increase productivity, factor inputs must improve in efficiency, quality, and (ultimately) specialization to particular cluster areas. Specialized factors, especially those integral to innovation and upgrading (a specialized university research institute), not only are necessary to attain high levels of productivity but also tend to be less tradable or available from elsewhere."

On the other hand, the context for firm strategy and rivalry concerns the rules, incentives, and norms governing the type and intensity of local rivalry. Low-productivity economies with are characterized by little local rivalry. Most competition, if present at all, comes from imports; local rivalry, if any, involves imitation. Price is the sole competition variable, and firms hold down wages to compete in local and foreign markets. Competition entails minimal investment.

Moving to an advanced economy requires the development of vigorous local rivalry. Rivalry must shift from low wages to low total cost, and this requires upgrading the efficiency of manufacturing and service delivery. Ultimately, rivalry must also evolve from cost alone to include differentiation. Competition must shift from imitation to innovation and from low investment to high investment —not only in physical assets but also in intangibles (skills and technology, for example). Clusters play a key role in these transitions.

While the nature of rivalry in a given location is heavily influenced by many aspects of the business environment (factor availability, local demand conditions), the investment climate and competition policies set the context. Issues such as macroeconomic and political stability, the tax system, labour-market policies affecting the incentives for workforce development, and intellectual property rules and their enforcement, affect the willingness of companies to invest in upgrading capital equipment, skills and

technology. Antitrust policy; government ownership and licensing rules; and policies toward trade, foreign investment, and corruption play a vital role in defining the intensity of local rivalry.

Lastly, related and supporting industries refer to the local presence or absence of suppliers of materials, components and machinery and equipment, together with related industries that support enterprise productivity and competitiveness (Porter, 1998). The level and rate of growth of productivity in a particular location, according to Porter (1998), depend less on what industries and firms compete on and more on how they compete. The sources of competition define the factors that influence how firms compete, thus affecting productivity and the LED process in local areas.

A structural variant of Porter's approach is the systemic competitiveness approach outlined in Meyer-Stamer, Altenburg and Hillebrand (1998) and Meyer-Stamer (2005), in which the concept of systemic competitiveness seeks to capture the political and economic determinants of successful industrial development. It refers to a pattern in which State and social actors deliberately create the conditions for successful industrial development. The concept distinguishes four levels: the "micro-level" of the firm and inter-firm networks; the "meso-level" of specific policies and institutions; the "macro-level" of generic economic conditions; and the "meta-level" of "qualitative" variables such as sociocultural structures, the basic economic order and orientation, and the capacity of social actors to formulate strategies.

At the local level, the vehicle through which specific geographic areas may become more competitive and achieve systemic competitiveness for successful economic and industrial development is a "geographic cluster" (Porter, 1996 and 1998). This is defined as a "geographically proximate group of interconnected companies and associated institutions in a particular field, linked by commonalities and complementarities. The geographic scope of clusters ranges from a region, a state, or even a single city, to span nearby or neighbouring countries" (Porter, 2000).

In contrast to the competitiveness approach to LED, the cluster approach focuses on the way specific features of clusters affect the LED process, which are intrinsically associated with the "economies" and properties generated in a geographic location. Under the former approach, Porter (1990) and Meyer-Stammer, Altenburg and Hillebrand (1998) claim that competitiveness (and its cluster vehicle) can be also applied at the national level, and is not necessarily linked to the development properties of specific geographic areas.[13] In addition to the factors that determine competitiveness locally, the cluster approach focuses on the following LED features of clusters localized in specific geographic areas: linkages and interdependency among firms and activities within a given space (Feser, 1998b); externalities (including technological spillovers) and agglomeration economies arising from location (Feser, 1998a); the formation of non-market social networks among agents within the geographic cluster (Jones, Hesterly and Borgatti, 1997; Powell, 1990); the innovation environment (Audretsch, 1998; Audretsch and Feldman, 1996); and path-dependency and lock-in effects (Kenney and von Burg, 1999; and Antonelli, 2000).

Location-factor and cluster-development features have been also associated with the emerging literature on (national and regional) innovation systems, learning and knowledge-based economies (Lundvall and Johnson, 1994; Morosini, 2004; Maskell, 2001 and Cooke, 2001). In a knowledge-based economy —defined by OECD (1995) as an economy which is directly based on the production, distribution and use of knowledge and information— location and cluster features can serve as vehicles for knowledge creation and economic growth in local areas. In this regard, Cappellin (2003) postulates that the knowledge-creation process is interactive and combinatorial; and that closer geographical proximity and greater cognitive proximity makes it easier to combine complementary pieces of knowledge and facilitates interaction between various complementary actors. Maskell (2001) adds that the cluster is considered the territorial configuration most likely to enhance learning processes. Lastly, Leydesdorff (2006) states that the "dynamic of a knowledge-based economy has important consequences for the function of regions. The locales may serve as the incubators where production, innovation, and diffusion processes are closely coupled. The density of the local interactions increases the chances for "lock-in" and therefore the (co-)shaping of trajectories within the system. The density of the interactions within clusters and regions determines this capacity. Therefore, one can expect metropolitan regions to hold an advantageous position in the knowledge-based economy."

[13] Enright (1998), Raines (2001) and Camagni (2002) summarize the geographic or territorial aspect of the competitiveness concept.

VII
Conclusions

The decentralization process pursued by developing countries over the last three decades has generated a demand for conceptual frameworks to define the appropriate objectives and roles of private and public agents in LED processes. Unlike the situation in the industrialized world, LED is practically non-existent as an academic discipline in most developing countries. This discipline provides four theoretical approaches to the analysis of the LED process in developing economies, going far beyond the economic fundamentals, institutions and the market-failure approach of national economic development theories. Thus, location factors, local public goods, active participation by diverse private agents, and the multidisciplinary framework of the LED theories can help elucidate the necessary objectives and roles of agents in decentralization and LED processes in developing countries.

(Original: English)

Bibliography

Aghón, G., F. Alburquerque and P. Cortés (eds.) (2001), *Desarrollo económico local y descentralización en América Latina: un análisis comparativo* (LC/L.1549), Santiago, Chile, Economic Commission for Latin America and the Caribbean (ECLAC).

Alburquerque, F., J.L. Llorens and J. del Castillo (2002), *Estudio de casos de desarrollo económico local en América Latina*, Washington, D.C., Inter-American Development Bank.

Andrews, R. (1953), "Mechanics of the urban economic base: historical development of the base concept", *Land Economics*, vol. 29, No. 2, Wisconsin, The University of Wisconsin Press.

Antonelli, C. (2000), "Path Dependence, Localized Technological Change and the Quest for Dynamic Efficiency", Turin, Italy, unpublished.

Arthur, W.B. (1989), "Competing technologies, increasing returns, and lock-in by historical events", *The Economic Journal*, vol. 99, No. 394, Oxford, Royal Economic Society.

Audretsch, D. (1998), "Agglomeration and the location of innovative activity", *Oxford Review of Economic Policy*, vol. 14, No. 2, Oxford, Oxford University Press.

Audretsch, D. and M. Feldman (1996), "R&D spillovers and the geography of innovation and production", *American Economic Review*, vol. 86, No. 3, Nashville, Tennessee, American Economic Association.

Audretsch, D. and M. Keilbach (2007), "The theory of knowledge spillover entrepreneurship", *Journal of Management Studies*, vol. 44, No. 7, Oxford, Blackwell Publishing.

_____ (2005), "Entrepreneurship capital and regional growth", *The Annals of Regional Science*, vol. 39, No. 3, Heidelberg, Springer.

_____ (2004a), "¿Does entrepreneurship capital matter?", *Entrepreneurship Theory and Practice*, vol. 28, No. 5, Texas, Baylor University.

_____ (2004b), "Entrepreneurship and regional growth: an evolutionary interpretation", *Journal of Evolutionary Economics*, vol. 14, No. 5, Heidelberg, Springer, December.

_____ (2004c), "Entrepreneurship capital and economic performance", *Regional Studies*, vol. 38, No. 8, London, Taylor & Francis, November.

Bachtler, J. and D. Yuill (2001), "Policies and strategies for regional development: a shift in paradigm?", *Regional and Industrial Policy Research Paper*, No. 16, Glasgow, European Policies Research Centre.

Baldwin, R. and R. Forslid (2000), "The core-periphery model and endogenous growth: stabilizing and destabilizing integration", *Economica*, vol. 67, No. 267, London, London School of Economics and Political Science.

Bardhan, P. (2002), "Decentralization of governance and development", *Journal of Economic Perspectives*, vol. 16, No. 4, Nashville, Tennessee, American Economic Association.

Bartik, T. (2003), "Local economic development policies", *Upjohn Institute Staff Working Paper*, No. 03-91, Kalamazoo, Michigan, The W.E. Upjohn Institute for Employment Research, January.

_____ (1995), "Economic development strategies", *Upjohn Institute Staff Working Paper*, No. 95-33, Kalamazoo, Michigan, The W.E. Upjohn Institute for Employment Research, January.

Bates, T. (1993), "Theories of entrepreneurship", *Theories of Local Economic Development: Perspectives from across the disciplines*, R. Bingham and R. Mier (eds.), London, Sage Publications.

Bingham, R. and R. Mier (1993), *Theories of Local Economic Development: Perspectives from Across the Disciplines*, London, Sage Publications.

Black, D. and V. Henderson (1999), "A theory of urban growth", *The Journal of Political Economy*, vol. 107, No. 2, Chicago, University of Chicago Press.

Blair, J. (1999), "Local economic development and national growth", *Economic Development Review*, vol. 16, No. 3.

_____ (1995), *Local Economic Development: Analysis and Practice*, Thousand Oaks, Sage Publications.

Blakely, E. (2003), "Conceptualizing local economic development: Part 1", *Public Finance and Management*, vol. 3, No. 2,

Elizabethtown, Southern Public Administration Education Foundation.

Blakely, E. and T. Bradshaw (2002), *Planning Local Economic Development: Theory and Practice*, London, Sage Publications.

_____ (1999), "What are "third-wave" state economic development efforts?: from incentives to industrial policy", *Economic Development Quarterly*, vol. 13, No. 3, London, Sage Publilcations.

Blumenberg, E. (1998), "Gender equity planning: inserting women into local economic development", *Journal of Planning Literature*, vol. 13, No. 2, London, Sage Publications, November.

Boyne, G. (1998), *Public Choice Theory and Local Government*, Houndmills, Macmillan.

Budd, L. and A. Hirmis (2004), "Conceptual framework for regional competitiveness", *Regional Studies*, vol. 38, No. 9, London, Taylor & Francis.

Camarero, C., C. Hernández and S. San Martín (2008), "Developing relationships within the framework of local economic development in Spain", *Entrepreneurship & Regional Development*, vol. 20, No. 1, London, Taylor & Francis, January.

Camagni, R. (2002), "On the concept of territorial competitiveness: sound or misleading", Dortmund, Germany, European Regional Science Association.

Cappellin, R. (2003), "Networks and technological change in regional clusters", *Innovation Clusters and Interregional Competition*, J. Bröcker, D. Dohse and R. Soltwedel (eds.), Heidelberg, Springer-Verlag.

Caulfield, J. (2003), "Local government reform in comparative perspective", *Reshaping Australian Local Government*, B. Dollery, N. Marshall and A. Worthington (eds.), Sydney, University of New South Wales Press.

Coleman, J. (1988), "Social capital in the creation of human capital", *American Journal of Sociology*, vol. 94, Chicago, The University of Chicago Press.

Cooke, P. (2001), "Regional innovations system, clusters and the knowledge economy", *Industrial and Corporate Change*, vol. 10, No. 4, Oxford, Oxford University Press.

Christaller, W. (1966), *The Central Places of Southern Germany*, Englewood Cliffs, Prentice-Hall.

DiPasquale, D. and E. Glaeser (1999), "Incentives and social capital: do homeowners make better citizens", *Journal of Urban Economics*, vol. 45, Amsterdam, Elsevier.

Dollery, B. and J. Wallis (2001), *The Political Economy of Local Government*, Cheltenham, United Kingdom, Edward Elgar.

Durlauf, S. (2002), "On the empirics of social capital", *The Economic Journal*, vol. 112, No. 483, Oxford, Royal Economic Society.

Elson, D. (1998), *Gender and Economic Development*, Rotterdam, Directorate General for International Cooperation (DGIS)/NEI.

Enright, M. (1998), "The globalization of competition and the localization of competitive advantage: policies toward regional clustering", *Working Paper*, Hong Kong, University of Hong Kong.

_____ (1996), "Regional clusters and economic development: a research agenda, *Business Networks: Prospects for Regional Development*, U. Staber, N. Schaefer and B. Sharma (eds.), Berlin, Walter de Gruyter.

Feser, E. (1998a), "Enterprises, external economies and economic development", *Journal of Planning Literature*, vol. 12, No. 3, London, Sage Publications.

_____ (1998b), "Old and new theories of industrial clusters", *European Research in Regional Science*, vol. 8, Regional Science Association International.

Findlay, R. and M. Lundahl (1994), "Natural resources, 'vent-for-surplus,' and the staples theory", *From Classical Economics to Development Economics*, G.M. Meier, New York, St. Martin's Press.

Finot, I. (2001), "Descentralización en América Latina: teoría y práctica", *Gestión pública series*, No. 12 (LC/L.1521-P), Santiago, Chile, Economic Commission for Latin America and the Caribbean (ECLAC). United Nations publication, Sales No. S.01.II.G.64.

Forsyth, F. (2000), "Women's enterprise and business development", *Local Economy*, vol. 15, No. 1, London, Taylor & Francis.

Fujita, M. (1988), "A monopolistic competition model of spatial agglomeration: a differentiated product approach", Regional Science and Urban Economics, vol. 18, No. 1, Amsterdam, Elsevier.

Fujita, M. and J. Thisse (2003), "Does geographical agglomeration foster economic growth? And who gains and who loses from it?", *The Japanese Economic Review*, vol. 54, Hoboken, John Wiley & Sons.

_____ (1996), "Economics of agglomeration", *Journal of the Japanese and International Economies*, vol. 10, No. 4, Amsterdam, Elsevier.

Fujita, M. and P. Krugman (1995), "When is the economy monocentric?: von Thünen and Chamberlin unified", *Regional Science and Urban Economics*, vol. 25, No. 4, Amsterdam, Elsevier.

Fujita, M., P. Krugman and T. Mori (1999), "On the evolution of hierarchical urban systems", *European Economic Review*, vol. 43, No. 2, Amsterdam, Elsevier.

Fujita, M. and T. Mori (1997), "Structural stability and evolution of urban systems", *Regional Science and Urban Economics*, vol. 27, No. 4-5, Amsterdam, Elsevier.

Gray, V. and D. Lowery (1988), "Interest group politics and economic growth in the US States", *American Political Science Review*, vol. 82, Washington, D.C., American Political Science Association.

Greffe, X. (2004), *Local Governance and Partnerships*, Paris, Université Paris I - Sorbonne.

Hochman, O., D. Pines and J. Thisse (1995), "On the optimal structure of local governments", *American Economic Review*, vol. 85, No. 5, Nashville, Tennessee, American Economic Association.

Horn, M. (1997), *The Political Economy of Public Administration*, Cambridge, Cambridge University Press.

Isard, W. (1956), *Location and Space Economy*, New York, John Wiley & Sons.

Jones, R., R. Kashlak and A. Jones (2004), "Knowledge flows and economic development through microenterprise collaboration", *Journal of Entrepreneurship*, vol. 7, No. 1, London, Sage Publications.

Jones, C., W. Hesterly and S. Borgatti (1997), "A general theory of network governance: exchange conditions and social

mechanism", *The Academy of Management Review*, vol. 22, No. 4, Briarcliff Manor, Academy of Management.

Kenney, M. and U. von Burg (1999), "Technology, entrepreneurship and path dependence: industrial clustering in Silicon Valley and route 128", *Industrial and Corporate Change*, vol. 8, No. 1, Oxford, Oxford University Press.

King, D.N. (1984), *Fiscal Tiers*, London, Allen & Unwin.

Krugman, P. (1999), "The role of geography in development", *Proceedings of the World Bank Annual Conference on Development Economics*, Washington, D.C., World Bank.

_____ (1991), "Increasing returns and economic geography", *Journal of Political Economy*, vol. 99, No. 3, Chicago, University of Chicago Press.

La Porta, R. and others (1997), "Trust in large organizations", *American Economic Review*, vol. 87, No. 2, Nashville, Tennessee, American Economic Association.

Lall, S. (2001), "Competitiveness indices and developing countries: an economic evaluation of the global competitiveness report", *World Development*, vol. 29, No. 9, Amsterdam, Elsevier.

Laukkanen, M. and H. Niittykangas (2003), "Local developers as virtual entrepreneurs. Do difficult surroundings need initiating interventions?", *Entrepreneurship & Regional Development*, vol. 15, No. 4, London, Taylor & Francis, October-December.

Lawton, H., J. Glasson and A. Chadwick (2005), "The geography of talent: entrepreneurship and local economic development in Oxfordshire", *Entrepreneurship & Regional Development*, vol. 17, No. 6, London, Taylor & Francis, November.

Levitt, T. (1965), "Exploit the product life cycle", *Harvard Business Review*, vol. 43, Boston, Harvard Business Publishing, November-December.

Leydesdorff, L. (2006), "While a storm is raging on the open sea: regional development in a knowledge-based economy", *The Journal of Technology Transfer*, vol. 31, No. 1, New York, Springer.

Liou, K. (2007), "Applying good governance concept to promote local economic development: contribution and challenge", *International Journal of Economic Development*, vol. 9, No. 1-2, Elizabethtown, Southern Public Administration Education Foundation, Inc..

Litvack, J., J. Ahmad and R. Bird (1999), *Rethinking Decentralization at the World Bank*, Washington, D.C., World Bank.

Llisterri, J. (2000), "Competitividad y desarrollo económico local", *Working Paper*, Washington, D.C., Inter-American Development Bank, March.

Loveridge, S. (2004), "A typology and assessment of multi-sector regional economic impact models", *Regional Studies*, vol. 30, No. 3, London, Taylor & Francis.

Lundvall, B. and B. Johnson (1994), "The learning economy", *Journal of Industry Studies*, vol. 1, No. 2, London, Routledge.

Malecki, E. (1994), "Entrepreneurship in regional and local development", *International Regional Science Review*, vol. 16, No. 1-2, London, Sage Publications.

Malizia, E. (1985), *Local Economic Development: A Guide to Practice*, New York, Praeger.

Marshall, A. (1890), *Principles of Economics*, London, Macmillan and Co. Ltd.

Martin, R. and P. Sunley (2003), "Deconstructing clusters: chaotic concept or policy panacea?", *Journal of Economic Geography*, vol. 3, No. 1, January, Oxford, Oxford University Press.

Martin, P., I. Gianmarco and P. Ottaviano (2001), "Growth and agglomeration", *International Economic Review*, vol. 42, No. 4, Philadelphia, University of Pennsylvania/Osaka University Institute of Social and Economic Research Association.

_____ (1999), "Growing locations: industry location in a model of endogenous growth", *European Economic Review*, vol. 43, No. 2, Amsterdam, Elsevier.

Maskell, P. (2001), "Towards a knowledge based theory of geographical clusters", *Industrial and Corporate Change*, vol. 10, No. 4, Oxford, Oxford University Press.

McGuire, M. and others (1994), "Building development capacity in non-metropolitan communities", *Public Administration Review*, vol. 54, No. 5, New York, John Wiley & Sons.

Meyer-Stamer, J. (2005), "Systemic competitiveness revisited", Mesopartner WP, Duisburg, Germany.

Meyer-Stamer, J., T. Altenburg and W. Hillebrand (1998), *Building Systemic Competitiveness. Concept and Case Studies from Mexico, Brazil, Paraguay, Korea and Thailand*, Berlin, German Development Institute.

Mokate, K, (2004), *Women's Participation in Social Development: Experiences from Latin America and the Caribbean*, Washington, D.C., Inter-American Development Bank.

Molyneux, M. (2002), "Gender and the silences of social capital, lessons from Latin America", *Development and Change*, vol. 33, No. 2, New York, John Wiley & Sons.

Montero, A. and D. Samuels (2004), *Decentralization and Democracy in Latin America*, Notre Dame, University of Notre Dame Press.

Moore, M. (1996), *Creating Public Value*, Cambridge, Massachusetts, Harvard University Press.

Morosoni, P. (2004), "Industrial clusters, knowledge integration and performance", *World Development*, vol. 32, No. 2, Amsterdam, Elsevier.

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

Nijkamp, P., P. Rietveld and F. Snickars (1987), "Regional and multiregional economic models: a survey", *Handbook of Regional and Urban Economics*, P. Nijkamp (ed.), vol., No. 1, Amsterdam, North Holland.

North, D. (1955), "Location theory and regional economic growth", *The Journal of Political Economy*, vol. 63, No. 3, Chicago, University of Chicago Press.

Oates, W. (2005), "Toward a second-generation theory of fiscal federalism", *International Tax and Public Finance*, vol. 12, No. 4, New York, Springer.

_____ (1999), "An essay on fiscal federalism", *Journal of Economic Literature*, vol. 37, No. 3, Nashville, Tennessee, American Economic Association.

_____ (1972), *Fiscal Federalism*, New York, Harcourt Brace Jovanovich.

OECD (Organisation for Economic Co-operation and Development) (2007), *Enhancing the Capacity of Partnerships to Influence Policy*, Paris.

_____ (2003), *Entrepreneurship and Local Economic Development: Programme and Policy Recommendations*, Paris.
_____ (2002), *Redefining Territories: Functional Regions*, Paris, Working Party on Territorial Indicators.
_____ (2001), *Local Partnership for Better Governance*, Paris.
_____ (1995), *The Knowledge Based-Economy*, Paris.
Olson, M. (1969), "The principle of 'fiscal equivalence': the division of responsibilities among different levels of government", *American Economic Review*, vol. 59, No. 2, Nashville, Tennessee, American Economic Association.
Oxhorn, P., J. Tulchin and A. Selee (eds.) (2004), *Decentralization, Democratic Governance, and Civil Society in Comparative Perspective: Africa, Asia and Latin America*, Baltimore, The Johns Hopkins University Press.
Perroux, F. (1988), "The pole of development's new place in a general theory of economic activity", *Regional Economic Development: Essays in Honor of Francois Perroux*, B. Higgins and D. Savoie (eds.), Boston, Unwin Hyman.
_____ (1955), "Note sur la notion de pole de croissance", *Economie appliquée*, No. 8, Paris, Les Presses de l'ISMEA. Published also in I. Livingstone, *Development Economics and Policy: Selected Readings*, London, George Allen & Unwin, 1979.
_____ (1950), "Economic space: theory and applications", *Quarterly Journal of Economics*, vol. 64, Cambridge, Massachusetts, The MIT Press.
Pinchot III, G. (1985), *Intrapreneuring: Why You Don't Have to Leave the Corporation to Become an Entrepreneur*, New York, Harper & Row.
Porter, M. (2000), "Location, competition and economic development: local clusters in the global economy", *Economic Development Quarterly*, vol. 14, No. 1, London, Sage Publications.
_____ (1998), "The Adam Smith address: location, clusters and the new microeconomics of competition", *Business Economics*, vol. 33, No. 1, Washington, D.C., National Association for Business Economics.
_____ (1996), "Competitive advantage, agglomeration economies and regional policy", *International Regional Science Review*, vol. 19, No. 1-2, London, Sage Publications.
_____ (1991), *La ventaja competitiva de las naciones*, Buenos Aires, Vergara.
Powell, W. (1990), "Neither markets nor hierarchy: network forms of organizations", *Research in Organizational Behavior*, B. Staw and L. Cummings (eds.), Greenwich, JAI Press.
Prebisch, R. (1959), "Commercial policy in the underdeveloped countries", *American Economic Review*, vol. 49, No. 2, Nashville, Tennessee, American Economic Association.
Putnam, R. (1993), "The prosperous community: social capital and economic growth", *The American Prospect*, No. 13.
Raines, P. (2003), *Cluster Development and Policy*, Burlington, Ashgate Publishing.
_____ (2001), "Local or national competitive advantage?: the tensions in cluster development policy", *Regional and Industrial Policy Research Paper*, No. 43, Glasgow, European Policy Research Center.
Rocha, H. (2004), "Entrepreneurship and development: the role of clusters", *Small Business Economics*, vol. 23, No. 5, New York, Springer.
Rondinelli, D.A. and G.S. Cheema (1983), *Decentralization and Development: Policy Implementation in Developing Countries*, London, Sage Publications.
Samuelson, P. (1954), "The pure theory of public expenditure", *The Review of Economics and Statistics*, vol. 36, No. 4, Cambridge, Massachusetts, The MIT Press.
Schumpeter, J. (1934), *The Theory of Economic Development: An Inquiry, Profits, Capital, Interest and the Business Cycle*, Cambridge, Massachusetts, Harvard University Press.
Shah, A. (2005), "A framework for evaluating alternate institutional arrangements for fiscal equalization transfers", *Policy Research Working Paper Series*, No. 3785, Washington, D.C., World Bank.
Shah, A. and S. Shah (2006), "The new vision of local governance and the evolving roles of local governments", *Local Governance in Developing Countries*, A. Shah (ed.), Washington, D.C., World Bank.
Sirkin, G. (1959), "The theory of the regional economic base", *The Review of Economics and Statistics*, vol. 41, No. 4, Cambridge, Massachusetts, The MIT Press.
Stahl, K. (1987), "Theories of urban business location", *Handbook of Regional and Urban Economics*, vol. 2, E. Mills (ed.), Amsterdam, North Holland.
Stren, R. and others (2002), *Decentralization in Global Perspective: A Review of Twenty-Eight Country Experiences*, report prepared for UN-Habitat, Nairobi, Kenya, October.
Tello, M. (2008), *Desarrollo económico local, descentralización y clusters: teorías, evidencias y aplicaciones*, Lima, Centrum Católica, Catholic University of Peru.
Thompson, W. (1968), "Internal and external factors in the development of urban economics", *Issues in Urban Economics*, H. Perloff and L. Wingo Jr. (eds.), Baltimore, Johns Hopkins Press.
Tiebout, C. (1956a), "A pure theory of local expenditures", *The Journal of Political Economy*, vol. 64, No. 5, Chicago, The University of Chicago Press.
_____ (1956b), "Exports and regional economic growth", *The Journal of Political Economy*, No. 64, Chicago, University of Chicago Press.
Trigilia, C. (2001), "Social capital and local development", *European Journal of Social Theory*, vol. 4, No. 4, London, Sage Publications.
Vernon, R. (1966), "International investment and international trade in the product cycle", *The Quarterly Journal of Economics*, vol. 80, No. 2, Cambridge, Massachusetts, The MIT Press.
Von Thünen, J. (1826), *Isolated State*, New York, Pergamon Press.
Walsh, J. and J. Meldon (2004), *Partnerships for Effective Local Development*, Charleroi, Université Libre de Bruxelles.
Walz, U. (1996), "Transport costs, intermediate goods and localized growth", *Regional Science and Urban Economics*, vol. 26, No. 1, Amsterdam, Elsevier.
Watt, P. (2006), "Principles and theories of local government", *Economic Affairs*, vol. 26, No. 1, Oxford, Blackwell Publishing.
Weber, A. (1957), *The Theory of Location of Industries*, Chicago, The University of Chicago Press.

Weinberger, K. and J. Jütting (2001), "Women's participation in local organizations: conditions and constrains", *World Development*, vol. 29, No. 8, Amsterdam, Elsevier.

Williamson, O. (1985), *The Economic Institutions of Capitalism*, New York, Free Press.

Woolcock, M. (1998), "Social capital and economic development: towards a theoretical synthesis and policy framework", *Theory and Society*, vol. 27, No. 2, New York, Springer.

Woolcock, M. and D. Narayan (2000), "Social capital: implications for development theory, research, and policy", *The World Bank Research Observer*, vol. 15, No. 2, Oxford, Oxford University Press, August.

World Bank (2010), "Making Local Economic Development Strategies: a Trainer´s Manual" [online] www.worldbank.org/urban/local/toolkit/.../defining-led.htm

Young, A. (1928), "Increasing returns and economic progress", *The Economic Journal*, vol. 38, Oxford, Royal Economic Society.

Zabojnik, F. and P. Francois (2005), "Trust, social capital, and economic development", *Journal of the European Economic Association*, vol. 3, No. 1, Cambridge, Massachusetts, The MIT Press.

KEYWORDS

Competitiveness
Regional development
Measurement
Governance
Economic development
Productivity
Human resources
Business management

A competitiveness index for the regions of a country

Jorge Benzaquen, Luis Alfonso del Carpio, Luis Alberto Zegarra and Christian Alberto Valdivia

This article, which forms part of a comprehensive research project on measuring the level of competitiveness, sets out a proposal for developing an index to measure the competitiveness of the regions of a country. The aim is to develop a new approach to competitiveness, by measuring how resources and capacities are managed in a given region of a country, to generate a sustained increase in business productivity and the well-being of its population. The following pillars of the competitiveness of regions are identified: (i) government and institutions; (ii) economic development; (iii) productive infrastructure; (iv) human capital; and (v) business efficiency. For each of these pillars, five factors and their variables are identified to measure the various aspects of regional competitiveness. These represent a second and third level of disaggregation that enhance the analysis that can be made using the results obtained.

Jorge Benzaquen
Professor at CENTRUM Católica, Pontificia Universidad Católica del Perú
jbenzaq@pucp.edu.pe

Luis Alfonso del Carpio
Professor at CENTRUM Católica, Pontificia Universidad Católica del Perú

Luis Alberto Zegarra
Professor at CENTRUM Católica, Pontificia Universidad Católica del Perú
zegarra.la@pucp.edu.pe

Christian Alberto Valdivia
Professor at CENTRUM Católica, Pontificia Universidad Católica del Perú
cvaldiviao@pucp.edu.pe

I
Introduction

The competitiveness of the regions comprising a country play a major role in ensuring that efforts are targeted on achieving sustained development that improves the well-being of the population. How to measure the competitiveness of regions has been discussed in various studies, which identify some of the components needed to develop a methodological measurement proposal.

This article sets forth a proposal for developing a regional competitiveness index (RCI) for a country, as part of a comprehensive research project on measuring competitiveness.

With this index, the aim is to highlight a new approach to competitiveness, measuring how resources and capacities are managed in a given region of a country, to generate a sustained increase in business productivity and in the well-being of its population.

II
Competitiveness

Although the origins of the concept of competitiveness date back to trade theories that are over three centuries old, today there is no consensus on its definition, nor therefore on how to measure it.

The concept of competitive advantage is based on "productivity" and on the factors that determine this within an enterprise. Productivity is the relation between the output obtained from a system for producing goods and services and the resources used to obtain that output —in other words their efficient use, or the relation between the results obtained and the resources used and time taken to obtain them (D'Alessio, 2004, p. 223). This approach, which is broadly accepted from the standpoint of its constituent factors, was described by Prokopenko in 1972, in *Productivity Management, A Practical Handbook* (Prokopenko, 1972).

This vision of productivity was enhanced in the 1980s with the widespread dissemination of *Competitive Strategy: Techniques for Analysing Industries and Competition* (Porter, 2000), which develops a model for understanding industries and competition, and for formulating a global strategy. The model presents the five competitiveness factors that determine the attractiveness of an industrial sector, along with their causes, as well as those that change through time and can be modified through the strategy.

Porter (1999) proposes the competitive advantage approach as the value that a firm creates for its customers, over and above its costs. This value corresponds to what individuals are willing to pay, and is best represented by the extent that supply prices are lower than those offered by the competition. This requires individuals to obtain equivalent benefits. The optimal strategy should reflect an adequate understanding of the business environment.

Although the origin of the concept of a nation's competitiveness dates back to trade theories several centuries old, it was Porter (1991), in *The Competitive Advantage of Nations*, who laid the foundations and recognized changes in the environment and the instability of generic strategies, pointing out the need for more dynamic models for thinking about the competitive advantage of nations.

Krugman (1994) points out that competitiveness becomes irrelevant at the national level, since the leading countries are not competing with each other, so it is more a domestic issue of the nation in question than an external one. On this point, Porter (1991) argues that a nation's competitiveness depends on the capacity of its industries to innovate and improve, and that certain firms are capable of doing so consistently, tirelessly seeking improvements and an ever better source of competitive advantage.

In his book *Economía urbana*, Camagni (2005) highlights the debate arising from the position adopted on international competitiveness by Krugman (1998, p. 5), which casts doubt on the idea that the prosperity of a country depends on its commercial success.

For Camagni (2005), the principle of comparative advantage cannot be applied when analysing competitiveness between local economies and inter-regional trade, since this theory is based on the concept of relative costs and prices that adjust because of a lack of mobility among productive factors, currency devaluation, and the downward rigidity of prices and wages in a situation of autarchy or isolation. Nonetheless, in the case of regional economies, there are factors that divert the principle of comparative advantage from its foundations:

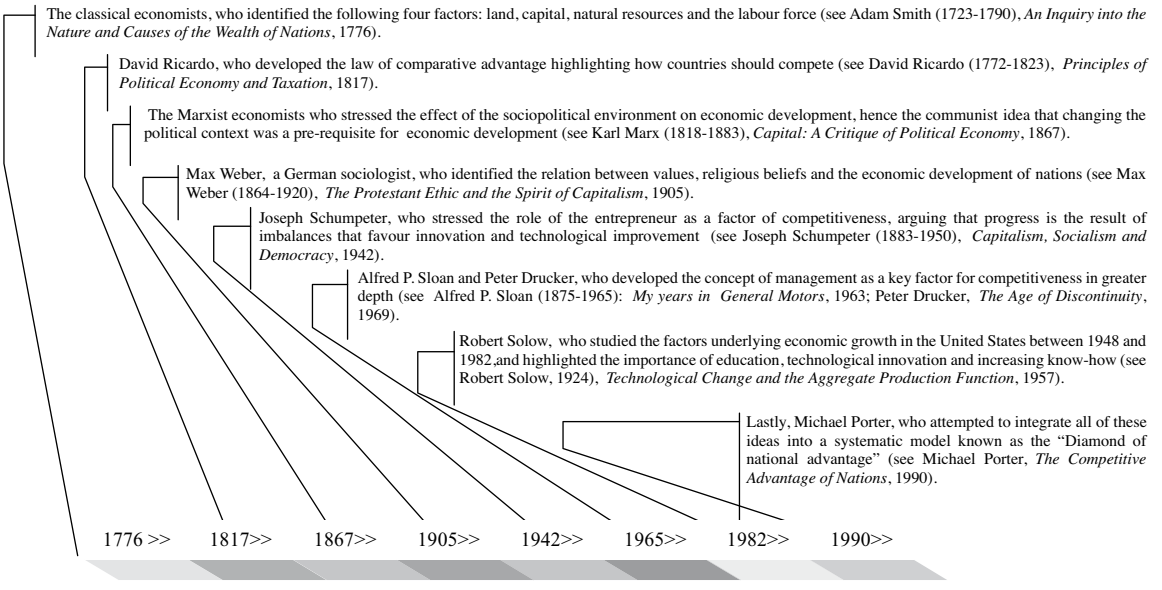
i) one cannot speak of autarchy or isolation: the link between average productivity and real wages is lost;
ii) productive factors move between regions: a region that has an absolute disadvantage in all goods will display imbalances in the labour market owing to factor mobility. This region adjusts quicker through emigration and depopulation than through price adjustments; and
iii) there is no regional currency or specific exchange-rate for each territory. Starting from a situation of equilibrium in which each region has an absolute advantage in some product, if a region sees that its productivity is growing by less than that of other regions and its output is becoming less competitive, it cannot devalue its currency as a country could.

In brief, the evolution of the theory of competitiveness starts from two fundamental theories: traditional economic theory and modern economic theory. The first is represented by the international-trade model (Smith, 1776), and the second by the competitive-advantage-of-nations model (Porter, 1991), also known as the "diamond of national advantage", which gives rise to the determinants of national competitive advantage, and has produced the two most widely recognized studies worldwide measuring the competitiveness of nations based on its theoretical models. These studies are discussed in the next subsection.

The evolution of the concept of competitiveness is summarized in figure 1, which traces its development, through the thinkers of the last three centuries, as a process of aggregation culminating in the current proposal of Michael Porter.

FIGURE 1

Evolution of the concept of competitiveness



Source: S. Garelli, "The competitiveness of nations: the fundamentals", IMD World Competitiveness Yearbook 2006, 2006 [on line] http://www.imd.ch/documents/wcc/content/Fundamentals.pdf

III

Global and regional competitiveness indices

The two most important world-level studies measuring the determinants of the competitiveness of nations are the Global Competitiveness Report (GCR) produced by the World Economic Forum (WEF), and the World Competitiveness Yearbook (WCY) produced by the International Institute for Management Development (IMD). Both are published annually, based on information obtained from statistical data and surveys conducted in each participating country. Partner institutions are responsible for compiling the data and conducting the survey. The two studies use similar competitiveness factors, drawn from the literature and empirical analyses.

The IMD classifies the data in terms of four factors: (i) economic performance; (ii) government efficiency; (iii) business efficiency; (iv) infrastructure; and each of these is divided into five subfactors. In contrast, the WEF classifies the data in 12 factors without additional classifications. It should noted that the WEF classification corresponds to the 2008-2009 edition of the GCR, because major methodological changes were made in the three preceding editions, causing variations in the factors used to measure competitiveness.

Production of the global competitiveness index is led by Professor Xavier Sala-i-Martin, Chief Adviser to the WEF Global Competitiveness Network (2008, p. 3). The index is based on 12 competitiveness pillars and offers a detailed overview of the competitiveness scenario of the world's countries at all levels of development. It has been produced and published annually since 1979; and its 2008-2009 edition evaluated 134 developed and developing economies.

The ranking of the variables of the Growth Competitiveness Index (GCI) 2008-2009, using statistical data or the survey used to measure the competitiveness of the nations evaluated in this study, consists of 110 variables, of which 79 (72%) are obtained from the survey, and the rest (31) come from statistical data obtained from secondary sources.

The World Competitiveness Yearbook, meanwhile, has been published since 1989 by the IMD to provide government and business leaders worldwide with information on the status and main trends in the competitiveness of the participating countries. In 2008, the IMD published its index on 55 countries from different regions of the world, for which it is assisted by strategic partners in each country involved.

The WCY is based on two types of information:

i) statistical indicators (hard data) compiled specially from international organizations such as the World Bank, the Organisation for Economic Co-operation and Development (OECD), the World Trade Organization (WTO) the Inter-American Development Bank (IDB) and the United Nations, among others; and
ii) annual surveys of entrepreneurs around the world (soft data).

Two thirds of the Yearbook is based on statistical information and one third on opinions and perceptions drawn from the business world. This distinguishes it from the Growth Competitiveness Index, which mostly compiles business opinions. Another difference with respect to the GCI is that whereas the WEF theory is based on the Porter (1990) diamond model, the IMD applies its own theory, which sees countries as managing their environments according to four fundamental forces that form the competitive environment.

Table 1 provides a summary of the indices and subindices published by the following organizations: (i) IMD 2008; (ii) Heritage Foundation (2008) with its Index of Economic Freedom; and (iii) the last three versions of the WEF Global Competitiveness Report. The latter include the Growth Competitiveness and Business Competitiveness Indices, which appeared in WEF publications of earlier years, and the Global Competitiveness Index, which is the current version of the index produced by this Swiss institution.

The regional competitiveness indices calculated in Latin America are derived mainly from research undertaken by Michael Porter and the Forum, and also from the IMD Global Competitiveness Yearbook. Although several of these regional indices have been discontinued, in all cases their development provided crucial support for decentralization processes. The Latin American indices include those of Mexico, Chile, Colombia and Peru.

Table 2 presents a summary of the determinants of regional competitiveness according to a sample of regional competitiveness indices produced in Latin America, showing the existence of a number of constant factors irrespective of the theoretical framework used.

TABLE 1

Comparison of competitiveness indices

IMD[a] 2008	World Economic Forum 2005-2006		World Economic Forum 2006-2007		World Economic Forum 2008-2009	Heritage Foundation 2009	World Bank 2009
Main index							
World Competitiveness Yearbook	Growth Competitiveness Index	Business Competitiveness Index	Global Competitiveness Index	Business Competitiveness Index	Global Competitiveness Index	Index of Freedom Economic	*Doing Business*
Subindices							
Economic performance	Technology	Ranking of company operations and strategy	Institutions	Ranking of company operations and strategy	Institutions	Business freedom	Overview
Government efficiency	Public institutions	Ranking of the quality of the national business environment	Infrastructure	Ranking of the quality of the national business environment	Infrastructure	Trade freedom	Starting a business
Business efficiency	Macroeconomy		Macroeconomy		Macroeconomic stability	Fiscal freedom	Dealing with construction permits
Infrastructure			Health and primary education		Health and primary education	Government size	Employing workers
			Higher education and training		Higher education and training	Monetary freedom	
			Market efficiency		Goods market efficiency	Investment freedom	Registering property
			Technological readiness		Labour market efficiency	Financial freedom	Getting credit
			Business sophistication		Financial market sophistication	Property rights	Protecting investors
			Innovation		Technological readiness	Freedom from corruption	Paying taxes
					Market size	Labor freedom	Trading across borders
					Business sophistication		
					Innovation		

Sources: Prepared on the basis of International Institute for Management Development, *IMD World Competitiveness Yearbook, 2008*, Lausanne, Switzerland, 2008; World Economic Forum, *Global Competitiveness Report 2008-2009*, Geneva, 2008; *Global Competitiveness Report 2006-2007*,Geneva, 2006; *Global Competitiveness Report 2005-2006*,Geneva, 2005; *Global Competitiveness Report 2004-2005*,Geneva, 2005; Heritage Foundation, *2009 Index of Economic Freedom* [on line] http://www.heritage.org/Index/, 2008; World Bank, *Doing Business*, 2008 [on line] http://www.doingbusiness.org/

[a] International Institute for Management Development.

TABLE 2

Comparison of regional competitiveness indices for a sample of countries

Country	Peru	Chile	Colombia	Colombia	Mexico	Mexico
Author	National Competitiveness Council (CNC)	Office of the Under-Secretary for Regional and Administrative Development (SUBDERE)	National University of Colombia	National University of Colombia	Mexican Institute of Competitiveness (IMCO)	Centre for Economic Research and Education (CIDE)
Main index (latest available)	Regional Competitiveness Index 2008	Regional Competitiveness Index 2003	Structural Departmental Competitiveness Index 2002	Revealed Departmental Competitiveness Index 2002	State Competitiveness 2008	Competitiveness of Mexican cities 2007
			Subindices			
1.	Institutional framework	Economic results factor	Infrastructure and location	Economic growth	Reliable and impartial legal system	Economic
2.	Infrastructure	Enterprises factor	Natural resources	External competitiveness	Sustainable management of the environment	Sociodemographic
3.	Macroeconomy	Persons factor	Human capital and employment	Quality of life	Inclusive, skilled and healthy society	Urban-environmental
4.	Health	Innovation factor	Firms		Stable and dynamic economy	Institutional
5.	Education	Science and technology	Innovation and technology		Stable and functional political system	
6.	Labour market	Infrastructure factor	Institutions		Efficient factors markets	
7.	Financial market	Government factor	Government administration		World-class leading sectors	
8.	Technological readiness	Natural resources factor	Integration into the global economy		Efficient and effective governments	
9.	Market		Economic growth		Exploitation of international relations	
10.	Business sophistication		External competitiveness		Vigorously competing economic sectors	
11.	Innovation					
12.	Natural resources					

Sources: National Competitiveness Council (CNC), *Índice de Competitividad Regional*, Lima, 2008; Office of the Under-Secretary for Regional and Administrative Development (SUBDERE), *Informe de competitividad regional*, Santiago, Chile, LOM Ediciones Ltda., 2003; National University of Colombia, *Sistema de indicadores de competitividad departamental*, Bogotá, D.C., Centre for Development Research (CID), 2002; Mexican Competitiveness Institute (IMCO), *Competitividad estatal*, Veracruz, 2008; Centre for Economic Research and Education (CIDE), *Competitividad de las ciudades mexicanas*, Mexico City, 2007.

IV
Definition of regional competitiveness

In this article, the word "regional" refers to the geographic division of a country, which can be defined in terms of various factors, including demography, history, culture, economics and climate, among others.

An analysis of earlier literature shows that competitiveness can be approached from two perspectives: firstly, as a set of factors that determine the level of productivity; and secondly, as a determinant of the sustained increase in the population's well-being. Based on these perspectives, regional competitiveness can be defined as the management of resources and capacities to obtain a sustained increase in business productivity and in the well-being of the region's population.

The definition proposed is tested by applying measures of competitiveness, which are statistical approximations, to evaluate the consistency of the definition proposed in the light of empirical evidence.

This evidence is obtained by comparing indicators of the country's economic development, such as gross domestic product (GDP) per capita, total factor productivity (TFP), or indicators of recognized prestige similar to the regional competitiveness index (RCI), such as the Human Development Index (UNDP, 2006) and the global competitiveness indices published by the IMD and WEF using specific parameters.

This analysis takes as its reference point the work done by Tello (2004) in his report on competitiveness factors in Peru.

The first measure of competitiveness compares the final result of the IMD Global Competitiveness Yearbook with per-capita GDP for the 55 economies covered by the IMD Yearbook in 2008. The analysis shows that countries ranked higher in terms of competitiveness have higher per-capita GDP.

The second measure of competitiveness is obtained by comparing the WEF growth competitiveness index (2005a, the latest year in which this index was published) and the per- capita GDP growth rate for the period 2007-2003. The analysis shows that countries with a higher GCI also had a higher per-capita GDP growth rate; so an increase in competitiveness is correlated with economic growth.

The third measure of competitiveness is obtained by relating the GCI to the annual average growth rate of total factor productivity (TFP) for the period 2004-2000. This shows a positive relation that suggests that improving competitiveness is also related to TFP growth.

V
Proposal for a regional competitiveness index

The proposal for the regional competitiveness index (RCI) includes a frame of reference for constructing indices, supported by both a general and a specific model. Lastly, the statistical calculation method is analysed and defined, and all of its components are presented.

The RCI belongs to the category of social indicators, generally linked to social research and the design and management of social projects. As such, it is a summary statistic, relating to the quantity or magnitude of a set of parameters or attributes of a given society.

Among indicators normally used for project management, social indicators can be classified as either simple or complex (Bobadilla, Del Águila and Morgan, 1998). The RCI is in the complex category, since it requires a theoretical framework and there is no simple way to corroborate its results.

Indicators can also be classified according to their measurement purpose —for example impact, effect and compliance (Bobadilla, Del Águila and Morgan, 1998). The RCI can be classified as an impact indicator, because it measures competitiveness through the results obtained by each region at the end of a given period, generally a year, in which the actions of economic agents have increased or reduced competitiveness.

Social research deal with phenomena that differ in terms of their greater or lesser complexity and abstraction. The RCI and the concept of

competitiveness that underlies it can be categorized as an abstract concept which cannot feasibly be observed empirically and thus cannot be measured. It is therefore necessary to undertake a process of decomposition and transformation, what Lazarsfeld (1958) calls "operationalization", to convert the notion and concept of competitiveness into a set of indicators susceptible to empirical observation, and, in the specific case of this study, the construction of an index.

According to Blalock (1970), this operationalization process must consider that, firstly, the index is conceptualized using theoretical reflections based on a literature review and the author's own thinking; and secondly, that the measurement makes it possible to assign values to social phenomena according to specific criteria.

Lazarsfeld (1958) argues that the operationalization process makes it possible to express concepts in terms of empirical indices and consists of the following stages: descriptive representation of the concept; specification of the concept, identifying the dimensions of its components or semantic subdivisions; and the choice indicators for each dimension. Once the dimensional indicators have been chosen, they are synthesized by constructing indices.

Figure 2 illustrates the process of ICR operationalization, based on the framework proposed by Lazarsfeld (1958). As there is no universally accepted definition of competitiveness, the process starts by proposing a definition, which is broken down into its initial components, namely dimensions, and which also give form to the pillars comprising the definition of competitiveness (in this case regional competitiveness). Lastly, the variables of the factors contained in the pillars are weighted, and possible combinations are defined. The assignment of weights attempts to express differences in relative importance in the RCI.

The methodology proposed is based on a review of similar experiences in constructing competitiveness indices and also on the literature review. The factors used, the statistical techniques applied and the information required mean that the RCI can be generalized and replicated in other countries that have a similar human development index (UNDP, 2006), to minimize the repercussions of the inherent economic and social differences in each country. The adaptations needed to apply the RCI in specific countries should not invalidate comparisons between their regions, to provide a base for expanding the scope of comparison, together with increasingly advanced reference points representing national objectives to be achieved.

VI

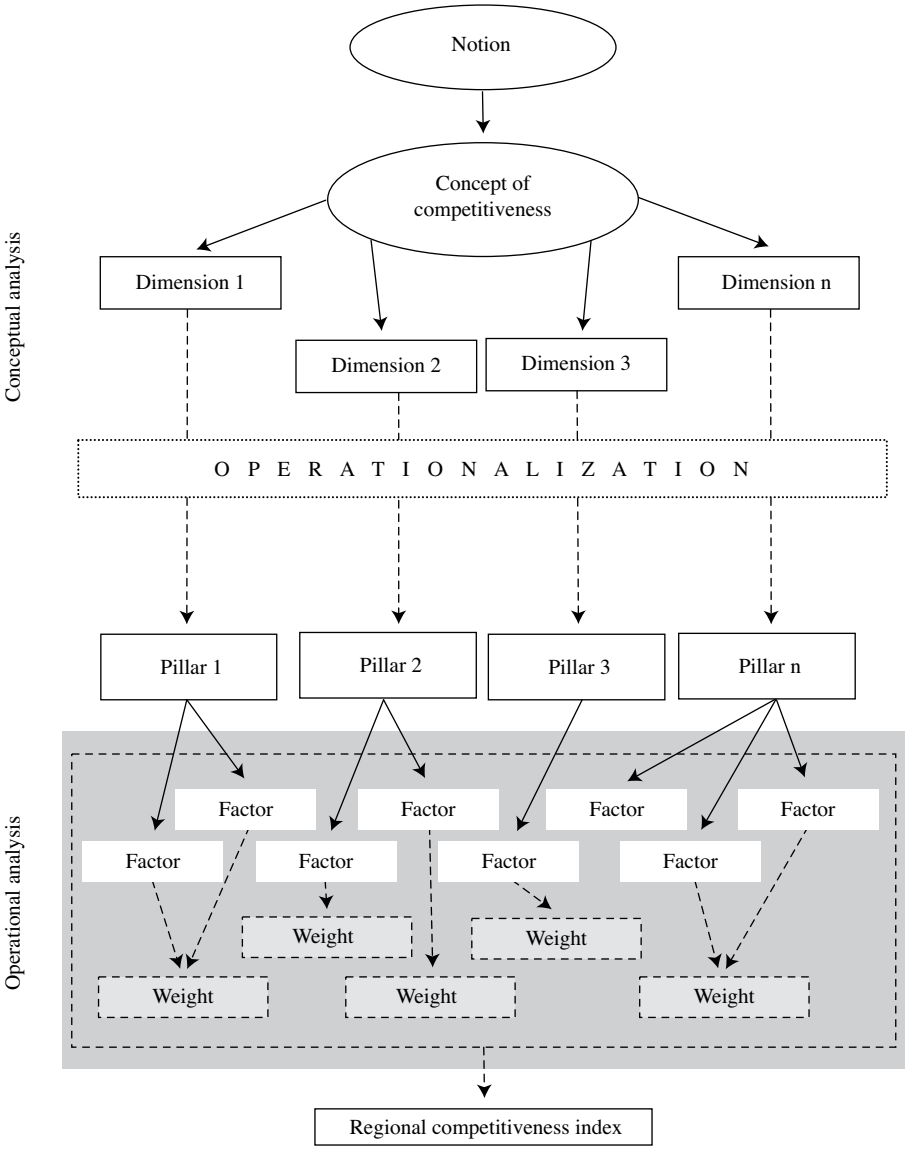
Determinants of regional competitiveness

The determinants of regional competitiveness, which in this study are referred to as pillars, have been defined on the basis of previous literature and the analysis of international experiences. There is no explicit consensus on what determines competitiveness; on the contrary, defining the pillars of competitiveness is in practice a process of choosing between different criteria. These can relate to availability, frequency and consistency with the concept (Joy Way, 2004); or to (i) consistency with the definition and conceptual framework; (ii) statistical support, in the sense that the factor is statistically related to an economy's performance indicators; (iii) the fact that it can be measured in some form (qualitatively or quantitatively), and be easily distinguishable from other factors (Tello, 2004).

This process of selecting pillars is a common denominator of competitiveness indices, particularly regional ones. Different methodologies are used in the process, ranging from the holding of workshops or interviews with experts, through to the application of *sui-generis* models. To identify the pillars of regional competitiveness proposed, the bases of the regional competitive advantages identified in Kitson, Martin and Tyler (2004) have been taken as a reference model (see figure 3). The concept of regional competitiveness related to these bases captures the notion that, although there are competitive and uncompetitive firms in each region, there are common elements in a region that affect the competitiveness of all firms. The approach followed by the authors is that of regional externalities, in other words resources that are external to the firm

FIGURE 2

Process of "operationalization" of the regional competitive index (RCI)



Source: P.F. Lazarsfeld, "Evidence and inference in social research", *American Academy of Arts & Sciences*, vol. 87, No. 4, Cambridge, Massachusetts, American Academy of Arts & Sciences.

but which are used directly or indirectly and have repercussions on its efficiency, innovation, flexibility and dynamism —in other words, on its productivity and competitive advantage.

The procedure followed in this study to determine the pillars of regional competitiveness based on regional competitive advantages (Kitson, Martin and Tyler, 2004) includes two stages: generalization of the definitions of the six bases (see table 3) and an analysis of their applicability, comparing the bases with the existing global and regional competitiveness indices (see table 4).

The pillars are obtained mainly from the bases of regional competitive advantages (Kitson, Martin and Tyler, 2004) —except for cultural capital, since no empirical evidence has been found that is consistent with the definition of those bases in international experiences relating to the inclusion of cultural capital. Moreover, neither the global nor the regional indices include cultural capital as an individual factor or variable. In several cases, as in the pillars used in this study, it is considered in relation to education and hence as human capital.

The proposed regional competitiveness pillars are now described, along with their foundations.

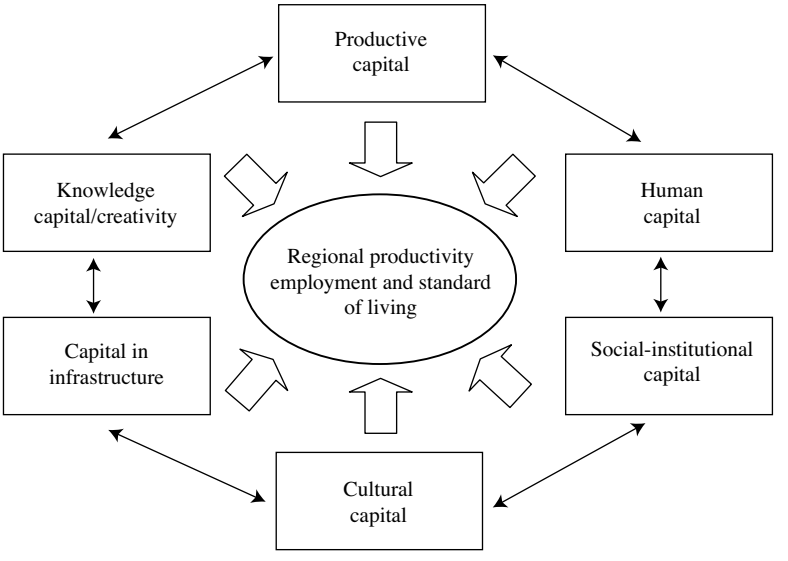
1. Government and institutions (P1)

The first pillar of regional competitiveness relates to the government and its institutions. In the case of the government, it is important to identify its main functions to be able to lay the foundations of the components with respect to its competitiveness. Although government is not the same as State, the government exercises the power of the State; in other words it undertakes State activities that are commonly referred to as powers of State. Accordingly, the State role can be analysed on three fronts: microeconomic, macroeconomic and institutional (León, 2003).

The government and institutions pillar arises from the microeconomic aspect of the role of the State; in other words, according to microeconomic theory, in a situation of perfect competition, the free market makes it possible to attain equilibrium. Nonetheless, this scenario is utopian because market failures or distortions prevent it from operating in such a situation of efficiency. The most common of these market failures is the existence of public goods that need to be managed with economic and social criteria, taking account of externalities, or generating monopolies and oligopolies. It is in these cases that

FIGURE 3

Bases of regional competitive advantages



Source: M. Kitson, R. Martin and P. Tyler, "Regional competitiveness: an elusive yet key concept?", *Regional Studies*, vol. 38, No. 9, London, Taylor & Francis, 2004.

TABLE 3

Classification of the pillars of global competitiveness indices according to the bases of regional competitive advantage

Basis of regional competitive advantage	World Competitiveness Yearbook 2008	Global Competitiveness Index 2008-2009
1. Productive capital	— Economic Development	— Macroeconomic stability — Financial market sophistication — Market size
2. Human capital	— Economic Development	— Labour market efficiency
3. Cultural capital		
4. Social-institutional capital	— Government efficiency	— Institutions
5. Capital in infrastructure	— Infrastructure	— Infrastructure — Health and primary education — Higher education and training
6. Knowledge capital/ creativity	— Business efficiency	— Innovation — Goods market efficiency — Technological readiness — Business sophistication

Source: Prepared on the basis of M. Kitson, R. Martin and P. Tyler, "Regional competitiveness: an elusive yet key concept?", *Regional Studies*, vol. 38, No. 9, London, Taylor & Francis; International Institute for Management Development, *IMD World Competitiveness Yearbook, 2008*, Lausanne, Switzerland, 2008; World Economic Forum, *Global Competitiveness Index*, Geneva, 2008.

TABLE 4

Classification of the factors of regional competitiveness indices according to the bases of regional competitive advantage

Bases of regional competitive advantage	Regional competitiveness index 2003 (Chile)	Structural departmental competitiveness index 2002 (Colombia)	State competitiveness 2008 (Mexico)	Competitiveness of Mexican cities 2007
1. Productive capital	– Economics results factor	– Integration into the global economy	– Stable and dynamic economy	– Economic
	– Natural resources factor	– Economic growth	– Vigorously competing economic sectors	
		– External competitiveness	– Exploitation of international Relations	
		– Natural resources	– Efficient factors markets	
			– World-class leading sectors	
			– Sustainable management of the environment	
2. Human capital	– Persons factor	– Human capital and employment		
3. Cultural capital				
4. Social-institutional capital	– Government factor	– Institutions	– Stable and functional political system	– Institutional
		– Government administration	– Efficient and effective governments	
			– Reliable and impartial legal system	
			– Inclusive, skilled and healthy society	
5. Capital in infrastructure	– Infrastructure factor	– Infrastructure and location		– Urban-environmental
				– Sociodemographic
6. Knowledge capital/ creativity	– Enterprises factor	– Firms		
	– Innovation factor	– Innovation and technology		
	– Science and technology			

Source: Prepared on the basis of M. Kitson, R. Martin and P. Tyler, "Regional competitiveness: an elusive yet key concept?", *Regional Studies*, vol. 38, No. 9, London, Taylor & Francis, 2004 ; Office of the Under-Secretary for Regional and Administrative Development (SUBDERE), *Informe de competitividad regional*, Santiago, Chile, LOM Ediciones Ltda., 2003; National University of Colombia, *Sistema de indicadores de competitividad departamental*, Bogotá, D.C., Centre for Development Research (CID), 2002; Mexican Institute of Competitiveness (IMCO), *Competitividad estatal*, Veracruz, 2008; Centre for Economic Research and Education (CIDE), *Competitividad de las ciudades mexicanas*, Mexico City, 2007.

the theoretical foundations for State intervention arise and, hence, the foundation for government action through its institutions, based on Keynes (1936), who envisioned the role of the State.

2. Economic development (P2)

The second pillar of regional competitiveness involves, firstly, the performance of the regional economy; secondly the process of internationalization; and, thirdly, job creation capacity. The economic development pillar is linked to economic growth theory, the most representative model of which in modern growth economics is that of Solow (1956), who explained growth in terms of capital accumulation, expansion of the labour force and technical change. Nonetheless, unlike the Solow (1956) growth accounting framework, and with the goal of describing it, the economic development pillar aims to measure the effect of economic growth, which also needs to respond to criteria of sustainability, social responsibility and environmental stewardship.

According to Sachs and Larraín (1994), economic growth is necessary to improve the standard of living of a growing population. Malthus (1798), in contrast, argued that population growth would be limited by the amount of resources that the earth could provide. He claimed that population growth exceeded the possibilities of the planet (resources) to provide subsistence to man; and death, in the form of wars or widespread famine, would inevitably bring the relation into balance.

3. Productive infrastructure (P3)

Infrastructure is the primary intervention of human beings in a territory to gain access to it and activate its development potential (ADC, 2008). A region's infrastructure level is closely related to its degree of development, and, in the event of backwardness, severely constrains the possibility of achieving significant progress in the material well-being of its population.

One of the most important specialized indices is the infrastructure ranking published by *América Economía* (2008). The productive infrastructure pillar uses the methodology of that ranking and is seen as the current capacity of each region to sustain productivity and business competitiveness. It can therefore be said that the effects of infrastructure are seen in job creation and improvements in the region's competitiveness and the quality of life of its citizens.

4. Human capital (P4)

Human capital is the value of the income-earning potential of individuals. Although it has a natural-resource component, it mostly stems from investment in education, skill development and health. These investments make human capital more productive (Larroulet and Mochon, 1995).

Economic science began to value the role of human capital following the contributions by Schultz (1961), who was the first to argue that skills and knowledge are a form of capital. A few years later, Becker (1964) considered human capital as a primary economic factor in his study on knowledge. Nonetheless, the most significant contribution, namely introducing human skills into the production function, appears in the work of Uzawa (1965) and Lucas (1988). According to these authors, an economy's production function of can be represented through a *Cobb-Douglas* function ($Y = A \cdot K^{\alpha} \cdot H^{1-\alpha}$), which considers production (Y) to be determined by human capital (H) and by physical capital (K).

5. Business efficiency (P5)

National prosperity is created, not inherited; and it depends on the capacity of its industry to innovate and improve (Porter, 2001). In other words, competitiveness is synonymous with productivity; and this is achieved by promoting innovation, which is driven by the four attributes of a nation, which in turn depend on the capacity of its firms to innovate and improve.

Although the basis of the approach with regard to industry, firms and productivity is a solid foundation for the business efficiency pillar, there is also the systemic competitiveness approach (Esser and others, 1996, pp. 39-52) which, without contradicting the approach taken by Porter (2001), strengthens the view of the role of the firm. At the micro level, the systemic approach is directly related to the development of the firm and its immediate environment. This approach proposes that, to successfully cope with new demands, firms need to reorganize themselves both internally and in their surrounding environment.

VII

Factors and variables of regional competitiveness

The factors correspond to the elements comprising the pillars. Overall, the pillar is given a certain significance according to the variables involved and the weightings they receive. The determination of the factors and variables is the result of an individual process in each country, which is established through the fulfilment of criteria that need to be applied to achieve a selection in accordance with the country's characteristics. The criteria for determining factors in variables of the regional competitiveness index are:

— That the variables comprising the factors be compiled from an official source. There is an initial group of institutions that generate information, either obtaining it through fieldwork or making back-office calculations, and a second group that compiles information from the first to summarize it and produce statistical publications.
— That it has a continuous historical record lasting over three years. This criterion makes sure the information is continuous and that it is not a variable calculated to meet a particular need; this reduces the chances that it will not be available in the following year.
— That it has a breakdown by region.
— That the methodology used in the sources is rigorous and stable through time, so as not to affect the potential for comparative analysis.

These criteria are applied in the process of selecting factors and variables in the country's official sources of statistical information, and are considered in the framework formed by the pillars.

The "government and institutions" pillar consists of factors that quantify the capacity of regional and local governments to fulfil their State role, to provide services to their inhabitants (resources, expenditure, investment, security, presence of the State) and thus promote the sustained development of their region.

The factors comprising the "economic development" pillar include growth of production, employment, and others, which are not only the result of the previous year, but also of the latest representative period or business cycle. Similarly, the variables are expressed not only in current - value terms but also in real (deflated) terms to avoid bias arising as a result of price volatility. Foreign trade not only includes the value exported, but also the volume, and to a greater extent, the region's internationalization and diversification process.

With regard to the factors comprising the "productive infrastructure" pillar, these take as a reference the concept of physical base (Joy Way, 2004), which implies the set of physical factors on which a region's competitiveness is founded. Accordingly, the pillar includes infrastructure support, such as the road network, energy and transport, and in particular the way in which factors of production in each region are organized, such as connectivity and tourism.

The "human capital" pillar includes factors that are directly related to the theories discussed above. It takes account of school and higher education, distinguishing between private and public; it also considers non-university job training and, finally, health.

Lastly, the "business efficiency" pillar encompasses factors directly related to the firm, such as productivity, business skills and innovation, and those pertaining to its immediate environment, such as the business climate and job creation.

The variables form the general definition of the indicator used and jointly comprise the factor. Each of the variables is linked to an indicator, either a simple indicator or a relative one. The simple indicator shows the absolute value of the variable, whereas a relative one calculates it in relation to another variable, such as population or GDP, among others. Tables 5 to 9 show the factors and variables classified by each factor for each of the pillars.

TABLE 5

Components of the government and institutions pillar

Factor	Measurement Variable
1. Resources obtained	Relative direct revenue collection Direct revenue collection
2. Resources transferred	Relative transfer income Transfer income
3. Investment expenditure	Investment as proportion of total expenditure Investment expenditure
4. Public security	Crimes Misdemeanors Terrorism
5. Presence of the State	Presence of primary and secondary schools Presence of health establishments Presence of a police station or post Presence of the municipality

Source: Prepared by the authors.

TABLE 6

Components of the economic development pillar

Factor	Measurement variable
1. Economic size	Real GDP Per-capita GDP
2. Economic growth	Real and current gdp growth
3. Internationalization	Export value Exports as a percentage of GDP and export volume Export growth
4. Diversification	Destination countries Outputs
5. Employment	Employed EAP Relative employed EAP Remuneration of executives, employees and manual workers

Source: Prepared by the authors.

EAP: Economically active population.

TABLE 7

Components of the productive infrastructure pillar

Factor	Measurement variable
1. Energy	Electric power Unregulated customers and consumption by unregulated customers Regulated customers and consumption by regulated customers
2. Road network	National road network and density of national network Departmental road network and density of departmental network Neighbourhood road network and density
3. Transport	Land transport and density of land transport Air transport and density of air transport International freight traffic in airports Export freight traffic in ports
4. Tourism	1, 2, 3, 4 and 5- star hotels Hostels Other establishments
5. Connectivity	Fixed telephony and density of fixed telephony Cellular telephony and density of cellular telephony

Source: Prepared by the authors.

TABLE 8

Components of the human capital pillar

Factor	Measurement variable
1. School education	Reading comprehension and understanding of mathematics in primary school Reading comprehension and understanding mathematics in secondary school
2. Public higher education	Graduates from public universities and density Professional qualification from private university and density
3. Private higher education	Graduates from private university and density Professional qualification from private university and density
4. Private higher education	Density of graduates from private university
5. Job training	Graduates from private university Density of persons with professional qualifications from private university Persons with professional qualification from private university Density of technical training centres
6. Health	Technical training centres Job training centres and density Infant mortality Life expectancy Morbidity Medical coverage

Source: Prepared by the authors.

TABLE 9

Components of the business efficiency pillar

Factor	Measurement variable
1. Productivity	Average labour productivity Variation in average labour productivity Employed EAP
2. Business climate	Number of firms Penetration and coverage of the financial system Enterprise start-ups Effort to develop firms Presence of successful firms
3. Entrepreneurial skills	Management capacity Long-term vision Capacity for adaptation and internationalization
4. Innovation	Existence of innovative products/services Cases of innovative firms or persons Creation of new products or services Improvement of techniques and processes
5. Innovation	Existence of innovative products/services
6. Job creation	Cases of innovative firms or persons Creation of new products or services Improvement of techniques and processes Access to well-paid jobs Opportunities for self-employed workers Stable employment Wage level Non-wage labour costs

Source: Prepared by the authors.

EAP: Economically active population.

VIII

Calculation of the regional competitiveness index

The regional competitive index (RCI) can be represented as follows:

$$ICR = \frac{\sum_{k=1}^{l} Pillar_k}{l}$$

where the RCI is the average of the l pillars comprising it, and in which each pillar ($Pillar_k$) is represented by the average of the m factors comprising it.

$$Pillar_k = \frac{\sum_{k=1}^{m} F_j}{m}$$

Lastly, the factor (F_j) is the sum total of the n variables comprising it, weighted by

$$F_j \sum_{i=1}^{n} V_i P_i$$

In other words, a factor (P_i)

where (V_i) = *Variable i*

and P_i = *Weight of the variable i*

Calculating the RCI faces two types of problem that are common to any effort to construct indices: (i) How to "standardize" different criteria?; and (ii) How to integrate the "standardized" criteria in the index? Both problems can be summarized in a single question: How to transform the variables so as to integrate them into a single index?

With regard to the first problem, there are three options that are the most widely accepted statistics

for standardizing criteria that stem from dissimilar variables or indicators: (i) conversion of scale; (ii) percentile rank; and (iii) standard result. Table 10 provides a summary of the capacities of each method with respect to the analytical criteria described.

It can also be seen that all methods fail to comply with at least one of the established criteria. The first criterion could be the most subjective of the three, and the negative observation for the standard result is due to the relative comparison with other methods that are simpler to understand because they do not have negative values and a restricted scale.

The relative difficulty of interpreting the standard result can be overcome by applying the scale conversion, after calculating the standard result. Moreover, the standard result is the only method that fulfils the third criterion of being able to provide an objective reference of the relative distance between the results of each region. For the reasons described, the standard result method complemented by scale conversion, which we will simply refer to as standardization, is the best method for calculating the RCI.

It is not feasible to define a uniform criterion to assign weightings to each of the variables comprising the RCI; nonetheless, some principles can be established to make the process less arbitrary. These basically relate to the unit of measurement of the variables, giving rise to the following two cases:

i) Variables with an original measurement unit, in other words the unit derived from the main source from which it was compiled.

ii) Variables with a transformed measurement unit, in other words the unit derived from a simple process of expressing the variable in relation to another variable that has some element of transcendence for the region, known as the transformer variable.

The first criterion in assigning weightings for each of the variables comprising each pillar of the RCI is to consider the two types of variables described; and, as the variable with a transformed measurement unit better represents the relative position of a given region or department, it is given a relatively higher weight than the variable with an original measurement unit. This is only true of cases in which both types of variable were used.

The second criterion for assigning weights relates to consistency with the identification of competitive advantages in the region. These are variables directly related to the fundamental aspects derived from the definition of regional competitiveness, such as (i) productivity; (ii) creativity; (iii) internationalization; and (iv) social welfare, among others. All of the variables related to these concepts must be assigned a relatively larger weight than the other accompanying variables.

The availability and quality of information at the subnational level in Latin America is a significant obstacle, owing to the different territorial divisions, degree of disaggregation of the information, and the relative importance of the factors or pillars in each country, which could affect the weighting criteria. All of this stems from the fact that the methodology is mainly based on secondary information.

TABLE 10

Comparison of methods for calculating the regional competitiveness index (RCI)

"Standardization" method	Easy to interpret?	Allows ranking?	Allows relative distances to be calculated?
Scale conversion	Yes	Yes	No
Percentile range/rank	Yes	Yes	No
Standard result	No	Yes	Yes

Source: Prepared by the authors.

IX
Conclusions

The definition and scope of competitiveness will remain a work in progress, possibly until partial consensus is reached, either in its definition or in its spheres of application. Moreover there is no proven and disseminated methodology for judging the quality of the results of a given competitiveness index, either at the world level or, as in this case, at the regional level. The difficulty stems from its status as a relative indicator, in other words, it does not determine which region is competitive but provides a relative view of the competitiveness of a given region compared to its peers. A country's RCI is therefore a tool or a guide for business or State policies. The role of global competitiveness indices, as a guide for countries' development, has also been sharply criticized when compared to the development results of certain countries or regions. Such is the case of the annual global competitive indices published every year by the WEF and IMD, the WEF index since 1979 and the IMD index since 1988.

Developing a regional competitive index for a country is a real effort to construct a tool to support its development. Every aspect of its preparation needs to be the best, taking into account the literature review, critical analysis of other experiences, but above all, clearly understanding the concept of regional competitiveness that it is intended to measure, namely to prepare a tool that contributes to development by efficient management of the region's resources to the benefit of its inhabitants and increasing business productivity.

The determinants of the competitiveness of regions were identified and referred to as pillars: (i) government and institutions; (ii) economic development; (iii) productive infrastructure; (iv) human capital; and (v) business efficiency. For each of these pillars, five factors and their variables were identified, which measure various aspects of regional competitiveness. These represent a second and third level of disaggregation which contributes to the analysis that can be performed with the results obtained.

A country's RCI is a structurally specific model, consisting of statistical information obtained from secondary sources, and, to a lesser extent, primary data obtained from a survey of entrepreneurs in all of the country's regions. From these two large data sources, indicators are selected to form the five pillars of the RCI. The selected indicators go through a standardization process to consolidate units of measurement and ultimately obtain three types of results: (i) the global RCI results; (ii) the partial RCI results for each of the region; and (iii) the partial RCI results for each of the five pillars.

(Original: Spanish)

Bibliography

ADC (Andean Development Corporation) (2008), "Infraestructura" [online] http://www.caf.com/view/index.asp?pageMS=34366&ms=17

AméricaEconomía (2008), *Ranking infraestructura AE* [online] http://beta.americaeconomia.com/Multimedios/Otros/5503.pdf

Becker, G. (1964), *Human Capital*, Chicago, University of Chicago Press.

Blalock, H. (1970), *Introducción a la investigación social*, Buenos Aires, Amorrortu.

Bobadilla, P., L. del Águila and M. de la L. Morgan (1998), *Diseño y evaluación de proyectos de desarrollo*, Lima, Pact-USAID.

Camagni, R. (2005), "El concepto de competitividad territorial", Economía urbana, R. Camagni, Barcelona, Antoni Bosch Editor.

CIDE (Economic Research and Teaching Centre) (2007), *Competitividad de las ciudades mexicanas*, Mexico City.

CNC (Consejo Nacional de Competitividad) (2008), *Índice de Competitividad Regional*, Lima.

D'Alessio, F.(2004), *Administración y dirección de la producción. Enfoque estratégico y de calidad*, Mexico City, Pearson.

Drucker, P. (1969), *The Age of Discontinuity*, New York, Harper & Row.

Esser, K. and others (1996), "Systemic competitiveness: a new challenge for firms and for government", *CEPAL Review*, No. 59 (LC/G.1931-P), Santigo, Chile.

Garelli, S. (2006), "The competitiveness of nations: the fundamentals", *IMD World Competitiveness Yearbook 2006* [online] http://www.imd.ch/documents/wcc/content/Fundamentals.pdf

Heritage Fundation (2008), *2009 Index of Economic Freedom* [online] http://www.heritage.org/Index/

IMCO (Mexican Institute of Competitiveness) (2008), *Competitividad estatal*, Veracruz, Mexico.

IMD (International Institute for Management Development) (2008), *IMD World Competitiveness Yearbook 2008*, Lausanne, Switzerland.

Joy Way, R.L. (2004), *Cómo establecer prioridades en las regiones del Perú: una propuesta de índice de competitividad regional sostenible*, Piura, Peru, Consorcio de Investigación Económca y Social (CIES).

Keynes, J.M. (1936), The *General Theory of Employment, Interest and Money*, Cambridge, Cambridge University Press.

Kitson, M., R. Martin and P. Tyler (2004), "Regional competitiveness: an elusive yet key concept?", *Regional Studies*, vol. 38, No. 9, London, Taylor & Francis.

Krugman, P. (1998), *Pop Internationalism*, Cambridge, Massachusetts, The MIT Press.

_____ (1994), "Competitiveness: a dangerous obsession", *Foreign Affairs*, vol. 73, No. 2, New York, Council on Foreign Relations.

Larroulet, C. and F. Mochon (1995), *Economía*, Madrid, McGraw Hill.

Lazarsfeld, P.F. (1958), "Evidence and inference in social research", *American Academy of Arts & Sciences*, vol. 87, No. 4, Cambridge, Massachusetts, American Academy of Arts & Sciences.

León, J. (2003), ¿Cuál es el rol del Estado?, *Revista de la Facultad de Ciencias Económicas de la UNMSM*, year 5, No. 15, Lima, Universidad Nacional Mayor de San Marcos.

Lucas, R. (1988), "On the mechanics of development planning", *Journal of Monetary Economics*, vol. 22, No. 1, Amsterdam, Elsevier.

Malthus, T. (1798), *First Essay on Population 1796*, Englewood-Cliffs, Prentice Hall.

Mark, K. (1867), *El capital: crítica de la economía política*, Mexico City, Siglo XXI Editores.

National University of Colombia (2002), *Sistema de indicadores de competitividad departamental*, Bogotá, D.C., Centre for Development Research (CID) .

Porter, M.E. (2001), ¿Dónde radica la ventaja competitiva de las naciones?, *Harvard Deusto Business Review*, No. 44, Barcelona, Planeta.

_____ (1999), *Ser competitivo: nuevas aportaciones y conclusiones*, Bilbao, Deusto.

_____ (1990), *The Competitive Advantage of Nations*, New York, The Free Press.

_____ (1980), *Competitive Strategy: Techniques for Analyzing Industries and Competitors*, New York, The Free Press.

Prokopenko, J. (1972), *Productivity Management. A Practical Handbook*, Geneva, International Labour Organization.

Ricardo, D. (1817), *Principios de economía política y tributación*, Madrid, Pirámide.

Sachs, J.D. and F.B. Larraín (1994), *Macroeconomía en la economía global*, Mexico City, Prentice Hall.

Schultz, T. (1961), "Investment in human capital", *American Economic Review*, vol. 51, No. 1, Nashville, Tennessee, American Economic Review.

Schumpeter, J. (1942), *Capitalism, Socialism and Democracy*, New York, Harper & Brothers.

Sloan, A. (1963), *My Years with General Motors*.

Smith, A. (1776), "An inquiry into the nature and causes of the wealth of nations", *The Harvard Classics*, E.W. Charles, New York, P.F. Colliers & Son Corporation.

Solow, R. (1957), "El cambio técnico y la funcion de producción agregada", *Economía del cambio tecnológico*, N. Rosenberg, Mexico City, Fondo de Cultura Económica.

_____ (1956), "A contribution to the theory of economic growth", *Quarterly Journal of Economics*, vol. 70, No. 1, Cambridge, Massachusetts, The MIT Press.

SUBDERE (Office of the Under-Secretary for Regional and Administrative Development) (2003), *Informe de competitividad regional*, Santiago, Chile, LOM Ediciones Ltda.

Tello, M.D. (2004), *Los factores de competitividad en el Perú*, Lima, Catholic University of Peru.

UNDP (United Nations Development Programme) (2006), *Informe sobre desarrollo humano del Perú*, Lima.

Uzawa, H. (1965), "Optimal technical change in aggregative model of economic growth", *International Economic Review*, vol. 6, No. 1, Oxford, Blackwell Publishing.

Weber, M. (1905), *The Protestant Ethic and the Spirit of Capitalism*, London.

World Bank (2008), "Doing business" [online] http://www.doingbusiness.org/

World Economic Forum (2008), *Global Competitiveness Report 2008-2009*, Geneva.

_____ (2006), *Global Competitiveness Report 2006-2007*, Geneva.

_____ (2005a), *Global Competitiveness Report 2004-2005*, Geneva.

_____ (2005b), *Global Competitiveness Report 2005-2006*, Geneva.

KEYWORDS

Taxation
Income tax
Tax reform
Income distribution
Measurement
Case studies
Guatemala

The paradox of progressivity in low-tax countries: income tax in Guatemala

Santiago Díaz de Sarralde, Carlos Garcimartín and Jesús Ruiz-Huerta

The Kakwani and Reynolds-Smolensky indices are used to analyse the consequences of tax reforms in terms of a tax's progressivity and redistributive capacity. Nonetheless, these indices can only serve as a basis for normative judgments in reforms where revenue remains constant. As this is generally not the case with tax reforms in practice, the Reynolds-Smolensky index is usually broken down into changes in the average tax rate and changes in the Kakwani index. This article argues that this procedure has serious disadvantages, especially in countries with low levels of tax revenue. To help overcome these problems, a number of alternative indicators are proposed based on the traditional indices, to make it possible to analyse the redistributive and progressivity effects of reforms that generate changes in revenue. These indicators are then used to analyse hypothetical reforms to income tax in Guatemala.

Santiago Díaz de Sarralde
Professor at the Department of Applied Economics II
Universidad Rey Juan Carlos. Vicálvaro Campus
santiago.diazsarralde@correo.aeat.es

Carlos Garcimartín
Professor at the Department of Applied Economics II
Universidad Rey Juan Carlos. Vicálvaro Campus
carlos.garcimartin@urjc.es

Jesús Ruiz-Huerta
Chair Professor at the Department of Applied Economics II
Universidad Rey Juan Carlos. Vicálvaro Campus
jesus.ruizhuerta@urjc.es

I
Introduction

The literature generally analyses a tax reform's effects on progressivity and redistributive capacity through variations in the Kakwani (1977) and Reynolds-Smolensky (1977) indices. Nonetheless, these indices may not be appropriate as a basis for normative judgments in the case of tax reforms that involve significant changes in revenue. This problem has traditionally been overcome in two ways. The first consists of comparing the after-tax income distribution using generalized concentration curves (generalized Lorenz curves); while the second method uses a decomposition of the Reynolds-Smolensky index, which splits the variation of the tax's redistributive capacity into a portion caused by changes in the average effective rate and another part caused by changes in its progressivity.

Nonetheless, as this article will attempt to show, this net separation of the effects of a reform on the average rate and progressivity is questionable, because a reform that not only increases the redistributive capacity of the tax, but also widens the differences between the amounts of tax paid by high- and low-income taxpayers, might actually appear to be regressive. This problem, if it exists, is more serious in low-tax countries, where levels of evasion are also usually high and tax systems are inequitable. As a result, the reforms needed to increase public-sector funding may not be implemented because they are considered regressive, when in fact they could produce both a greater income-redistribution capacity and a significant difference between the amounts paid by high- and low-income taxpayers. In fact, in the specific case of Guatemala, one of the lowest-tax countries in Latin America, it will be shown how two measures that significantly increase income-tax revenue capacity may seem regressive if the traditional indicators are used. Nonetheless, this progressivity is merely apparent, and possibly contrary to the subjective perception that citizens have of progressivity.

Consequently, other mechanisms need to be developed to evaluate tax reforms that generate changes in revenue (as they usually do) and to complement the information provided by the traditional indicators. To that end, two concepts will be used which are indeed separable: the tax level and the differences between net incomes or tax liabilities. This separation, which will be used to designate level and distance effects, aims to provide an additional analytical tool for evaluating the progressivity and redistributive capacity of tax structures that produce different revenue outcomes, in each case estimating the individual contributions made by the two effects.

Sections II and III of this article describe the main weaknesses which, in the authors' opinion, the indices normally used to evaluate this type of reform suffer from. Section IV formulates a proposal; and, lastly, section V uses the indicators developed to evaluate hypothetical reforms to income taxes in Guatemala.

II
Measurement instruments: inequality, progressivity and redistribution

Any analysis of the redistributive effects of a tax reform firstly requires an instrument that summarizes the income distribution in the pre-and post-reform situations. A widely used tool for this purpose is the Lorenz curve (Lx), which provides a standardized system for measuring the percentage shares of total income received by different proportions of the population. The Gini coefficient (G_x), derived from the Lorenz curve, is generally used as a single synthetic indicator of relative inequality. Graphical speaking, this coefficient compares the area between the Lorenz curve and the diagonal and the total area under the diagonal, expressed mathematically for discrete income distributions as:

$$G_x = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left|x_{ij} - x\right|}{2N^2\mu} \quad (1)$$

where μ represents average income, x_i and x_j are the incomes of individuals i and j respectively, and N represents the population. The Gini coefficient thus expresses the average difference between income pairs divided by twice the average income, and can take values between zero (absolute equality) and 1 (a single person receives the whole of the population's income). As is true of the Lorenz curve, this coefficient shows the relative inequality of a set of incomes, but not absolute inequality; so it is difficult to interpret in welfare terms when two populations have different average-income levels.

Just as these indicators can be used to compare different distributions in time and space, they can also be reformulated to compare changes in the income distribution caused by the tax system. To simplify, if all units of the population with the same income are assumed to have exactly the same tax pressure[1] —in other words, the amount of tax paid depends only on income— then the distribution of those incomes can be represented using a technique similar to that represented in the Lorenz curves, to obtain the income concentration curve (*Lt*) —and related concentration coefficient *Ct*, which is analogous to the Gini coefficient. We could also obtain the after-tax income concentration curve (*Lx-t*) and its corresponding concentration coefficient (*Cx-t*) (assuming for simplicity that no reordering occurs, then *Cx-t* = *Gx-t*) (see figure 1).

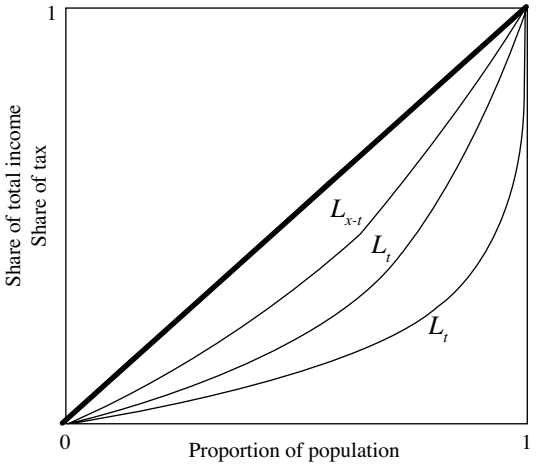
In particular, the tax liability concentration index would be

$$C_t = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left|t(x_i) - t(x_j)\right|}{2N^2\mu t};\ 0 \le t \le 1 \quad (2)$$

where $t(x_{i,j})$ represents the tax payable by taxpayers i and j, and t is the effective average rate. The after-tax income concentration coefficient would be:

FIGURE 1

Lorenz income curve and tax-liability and net-income concentration curves



Source: Prepared by the authors.

L_x: Pre-tax income concentration curve.
L_{x-t}: After-tax income concentration curve.
L_t: Tax liability concentration curve

$$C_{x-t} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left|(x_i - t(x_i)) - (x_j - t(x_j))\right|}{2N^2\mu(1-t)} \quad (3)$$

In the case of a progressive tax, the amounts of tax payable depart systematically from proportionality with respect to pre-tax income. The fact that tax liabilities are more unequally distributed than incomes means that the tax liability concentration curve lies further from the diagonal than the pre-tax Lorenz income curve; in other words, using the normal notation, *Lx*>*Lt*. As *Lx* represents not only the Lorenz curve for pre-tax income, but also the tax liability concentration curve for a proportional tax of equal revenue capacity, the gap between the two curves (*Lx*-*Lt*) can be used to measure the departure of the tax from proportionality. This is precisely the purpose of the Kakwani index of departure from proportionality (*K*), which is widely used in literature, to measure twice the area between the Lorenz curve of pre-tax income and the tax liability concentration curve. In other words, it represents the difference between the tax liability concentration coefficient and the Gini coefficient of the distribution of pre-tax income:

[1] Following the ECLAC definition, "tax pressure" is understood to mean the ratio of total taxes paid to total household income.

$$K = Ct - Gx \quad (4)$$

A progressive tax would also generate changes in the income distribution before and after its payment. This redistributive effect is normally quantified by measuring the distance between the Lorenz curves before and after the tax (*Lx-t - Lx*), since the two curves would be the same in the case of a proportional tax. The gap between the two can be synthesized using the Reynolds-Smolensky index (*RS*):

$$RS = Gx - Cx\text{-}t \quad (5)$$

Clearly, a tax's departure from proportionality and its redistributive effect are two closely related phenomena; the corresponding indices are related by the following equality:[2]

$$RS = \frac{t}{1-t} K \quad (6)$$

This shows that the redistributive effect is determined by the departure from proportionality and the level of the tax. In other words, it depends not only on progressivity but also on level.

[2] This expression is explained in Lambert (2001, p. 206 and ff).

III

Progressivity, redistribution and tax reforms

The indices described above are those most commonly used to analyse the progressivity and income-redistribution consequences of tax reforms.[3] Nonetheless, they may not be suitable when analysing reforms that involve significant changes in revenue. The specific purpose of this article is to develop indicators that help overcome some of the disadvantages of these indices.

As noted above, the Lorenz curve compares the income distribution with respect to proportionality, whereas the Gini coefficient, derived from it, measures the sum of the differences between income pairs in relation to average income. Both are therefore relative comparisons in which proportions matter rather than levels. When comparing these indicators in situations where levels vary significantly in time or space, most studies explicitly admit the shortcomings of these instruments when making welfare judgments.[4] To overcome these problems, developments based on the work of Atkinson (1970) and Shorrocks (1983) are often used, through the generalized Lorenz curve —namely the ordinary Lorenz curve multiplied by average income. This makes it possible to compare not just distributions, but also levels, which is more appropriate for making a normative assessment of changes or differences in the income distribution across a wide range of situations, although several cases persist in which it is hard to make a welfare judgment.

Nonetheless, there seems to be less reticence when evaluating a tax reform, in which Lorenz and concentration curves (and their related indices of inequality, progressivity and redistribution) are widely used to compare pre- and post-reform values, and thus infer "normative" consequences of the design of the reform based on the differences observed. These comparisons and normative judgments are correct if total revenue remains unchanged. Otherwise, the fact that the reform produces a more progressive outcome or greater redistribution, for example, does not in itself have any normative content, because the "superiority" of a tax's progressivity or redistributive effect can only be justified by comparing it with a proportional tax that generates the same revenue.

To overcome this problem, two approaches have been used in studies that evaluate tax reforms. The first consists of comparing after-tax income distributions using generalized Lorenz curves. In our opinion, this approach is questionable. Imagine a tax reform that leaves the after-tax Lorenz curve unaltered when

[3] Many empirical studies have been undertaken, both nationally and internationally, using those indices.

[4] The same is true when the Lorenz curves intersect each other (see Lambert, 2001, p. 44 and ff).

taxes are reduced. In that case, using generalized curves would indicate a welfare improvement. Yet, this conclusion is highly debatable, since it assumes that public expenditure *per se* generates less utility than private expenditure; and that would be the sole reason for the apparent increase in welfare. The effect would be similar to that produced by constructing Lorenz curves after spending on a specific good. If expenditure on this good increases proportionately owing to a change in the population's preferences, the generalized curves would indicate that welfare had decreased, because there is less income available for other goods. Yet, in reality all that has happened is that preferences have changed. Moreover, it is possible for changes in revenue to be offset by other taxes, in which case their effects also need to be taken into consideration. Ultimately, a comparison using generalized Lorenz curves is justified because the income distribution is not the only thing that matters; the absolute level of average income is also relevant. Nonetheless, this does not seem appropriate in a tax reform, since the average pre-tax income of the country in question will not necessarily affected by the reform, at least in the short run; but if it is, the effect should be made explicit.[5]

The second approach to evaluating the effects of a tax reform that changes revenue uses the decomposition of the Reynolds-Smolensky index (*RS*), mentioned in the previous section, separating the variation in the tax's redistributive capacity caused by changes in its effective average rate (*t/1-t*) from the variation caused by changes in its progressivity (*K*). A decrease (increase) in the level of the tax owing to a decrease/increase in *t* would always have a negative (positive) effect on *RS* when the tax is progressive; and a decrease (increase) in progressivity would have the same effect, measured by *K*. Thus, with a reform that lowers *t*, one should only expect that the increase in progressivity is sufficient to compensate for the change in the level of the tax.

This procedure is useful because it seems to permit a separate evaluation of what happens in terms of revenue level and progressivity, making the trade-off "explicit". It would thus be possible, for example, to positively value the increase in progressivity, measured through *K* and attributable to the design of reform, while attributing the decrease (or smaller increase) in redistribution exclusively to the amount of the tax reduction.

Nonetheless, in the usual case of tax reforms that alter revenue, this separation of responsibilities between the design of the reform and the magnitude of its revenue effect is incorrect, since the change in revenue not only changes the measurement scale but also the distribution of tax pressure. Only a tax reform that altered all tax liabilities equi-proportionately would leave the progressivity indicator (*K*) invariant, so that all of the change in redistributive capacity (*RS*) could be attributed to changes in the tax's revenue capacity. Thus, even though the measurement of progressivity is scale-invariant in itself, a change in progressivity resulting from a specific reform will ultimately depend on the same decisions that change the level of the tax. Thus, since *RS* can be expressed as:

$$RS = \frac{t}{1-t}K = \frac{t}{1-t}(C_t - G_x) = \frac{t}{1-t}\left(\frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left|t(x_i)-t(x_j)\right|}{2N^2\mu t} - G_x\right) \quad (7)$$

it is clear that progressivity is not separable from the level of the tax, since the latter affects the former, measured by *K*, through changes in the denominator of expression (7), whereas the individual distribution of the change in level affects *K* by altering its numerator. Only in the case of an equi-proportional change in all tax liabilities would these variations leave the quotient unaltered. In other words, level and progressivity are not separable concepts when evaluating the design of a tax reform and its effects; and using indicators derived from them to evaluate tax reforms could lead to errors of interpretation.

[5] A tax cut does not necessarily lead to an immediate economic expansion. Moreover is not easy to quantify its repercussion in economic contexts in which other variables play a significant role.

IV

Evaluating tax reforms: an alternative proposal based on level and distance effects

1. Tax reforms and redistribution

Although tax level and progressivity are not separable concepts in the sense discussed in the previous section, it is possible to separate changes in the tax level and in the differences between net incomes or tax liabilities, thus making it possible to more appropriately analyse tax reforms that cause changes in revenue. This is the basic idea underlying the proposal developed in this article; and, at this point, it is worth recalling that a reform will increase the redistributive effect of a tax if the Reynolds-Smolensky index after the change (*RS'*) is greater than it was before. Otherwise, the reform will reduce the redistributive effect, if $RS' - RS < 0$, or be neutral if $RS' = RS$.

The proposal of this article consists of decomposing the change in the Reynolds-Smolensky index, to separate changes in the differences between net incomes from changes in the average tax rate. In other words, starting from the expression:

$$RS' - RS = (G'x - C'_{x-t}) - (Gx - C_{x-t}) \tag{8}$$

and, to simplify, assuming that the Gini coefficient before and after the reform does not change,[6] in other words, $G'x = Gx$, then

$$RS' - RS = C_{x-t} - C'_{x-t} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left|(x_i - t(x_i)) - (x_j - t(x_j))\right|}{2N^2\mu(1-t)} - \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left|(x_i - t'(x_i)) - (x_j - t'(x_j))\right|}{2N^2\mu(1-t')} \tag{9}$$

Where the superscript (') represents the value of the variable in question after the reform. This expression can be broken down as follows:

$$RS' - RS = C_{x-t} - C'_{x-t} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left|(x_i - t(x_i)) - (x_j - t(x_j))\right|}{2N^2\mu(1-t)} - \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left|(x_i - t'(x_i)) - (x_j - t'(x_j))\right|}{2N^2\mu(1-t')} + \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left|(x_i - t(x_i)) - (x_j - t(x_j))\right|}{2N^2\mu(1-t')} - \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left|(x_i - t(x_i)) - (x_j - t(x_j))\right|}{2N^2\mu(1-t')} \tag{10}$$

So,

$$RS' - RS = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left|(x_i - t(x_i)) - (x_j - t(x_j))\right|}{2N^2\mu(1-t)}\left(1 - \frac{1-t}{1-t'}\right) + \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left|(x_i - t(x_i)) - (x_j - t(x_j))\right| - \sum_{i=1}^{N}\sum_{j=1}^{N}\left|(x_i - t'(x_i)) - (x_j - t'(x_j))\right|}{2N^2\mu(1-t')} \tag{11}$$

To further clarify the meaning of this expression, let β represent the rate of change of average net income after tax; D the sum of the distances between net incomes before the reform; and D' the sum of distances between net incomes after the reform, in other words:

$$\beta = \frac{(1-t') - (1-t)}{(1-t)} \tag{12}$$

$$D = \sum_{i=1}^{N}\sum_{j=1}^{N}\left|(x_i - t(x_i)) - (x_j - t(x_j))\right| \tag{13}$$

and

$$D' = \sum_{i=1}^{N}\sum_{j=1}^{N}\left|(x_i - t'(x_i)) - (x_{jj} - t'(x_{\ }))\right| \tag{14}$$

Thus, equation (11) can be written as

[6] In empirical exercises to evaluate tax reforms, the different fiscal structures resulting from the reform are generally applied to the same income distribution.

$$RS' - RS = C_{x-t}\left(1 - \frac{1}{1+\beta}\right) + \frac{D - D'}{2N^2\mu(1-t')} \qquad (15)$$

Consequently, the change in the Reynolds-Smolensky index would be the sum of what may be called a level effect (*LE*) and a distance effect (*DE*):

$$LE = C_{x-t}\left(1 - \frac{1}{1+\beta}\right) \qquad (16)$$

$$DE = \frac{D - D'}{2N^2\mu(1-t')} \qquad (17)$$

The two effects can be interpreted as follows. The level effect represents the difference between the net income concentration curve before the reform, and what it would have been if the reform had been implemented via an equal (positive or negative) transfer for all individuals, thus keeping distances between net incomes constant. It is important to note that, in our analysis, the "level effect" does not refer to the redistributive effect of any possible change in the tax that alters revenue by a given amount, but to that resulting from a change specifically made by transferring a lump sum to all taxpayers, for which reason it differs from the "average-type effect" present in the decomposition of the Reynolds-Smolensky index. Thus, for example, for a tax reduction, the graphic representation of this effect is shown in figure 2, where *Lx-tn* represents the concentration curve following that hypothetical reform.

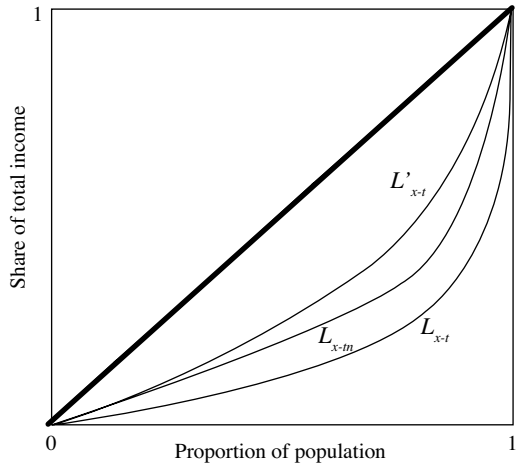
The distance effect, meanwhile, expresses the difference between the concentration curve that would exist if the reform had been implemented through equal transfers (positive or negative) for all individuals, holding constant the distances between net incomes and the concentration curve after the real reform (*L'x-t*) (which, strictly speaking, means evaluating the effects of a purely redistributive reform that alters the distances between net incomes by the same amount as the original reform analysed, but without changing revenue). For the case of a reduction in distances, this effect is also shown in figure 2.

The advantage of splitting the distributive effect into distance and level components, compared to the traditional division between level and progressivity, is that it makes it possible to clearly identify whether each of the factors —changes in the average rate and

FIGURE 2

Tax reform with an increase in redistribution

Level and distance effects



Source: Prepared by the authors.

L_{x-t}: After-tax income concentration curve (pre-reform)
L'_{x-t}: After-tax income concentration curve (post-reform)
L_{x-tn}: Income concentration curve following a reform of equal revenue effect made via a lump sum tax (in this case, transfer).

in distances— contributes positively or negatively to the change in the redistributive capacity of the tax, because it compares distances under a hypothetical equal-revenue scenario. Thus, both the level effect and the distance effect may be positive (greater redistribution) or negative (less redistribution).

In particular, for the level effect (*LE*):

If $\nabla t \Rightarrow \beta > 0 \Rightarrow LE > 0$
If $\Delta t \Rightarrow -1 < \beta < 0 \Rightarrow LE < 0$
If $\beta = 0 \Rightarrow LN = 0$

where ∇t represents a tax reduction, and Δt represents an increase.

Whereas for the distance effect (*ED*):

If $D > D' \Rightarrow DE > 0$
If $D < D' \Rightarrow DE < 0$
If $D = D' \Rightarrow DE = 0$

With this decomposition, the effect of changes in the tax level is perfectly isolated, which does not happen with the traditional breakdown. For example, if one imagines a tax reform that only lowers the average rate without altering the distances between individual tax

liabilities, the traditional breakdown might show that: (i) the redistributive capacity of the tax has increased; and (ii) the reduction in the average rate has contributed negatively to that increase; for which reason (iii) the increase in progressivity is exclusively responsible for the tax's greater redistributive capacity.

Clearly, however, any increase in progressivity is entirely due to the decrease in the average rate; so this has made a positive net contribution to redistributive capacity, and is in fact the only factor causing it to increase. In contrast, the breakdown presented here would show that the reduction in the average rate has a positive effect on the redistributive capacity of the tax and, moreover, is the only factor responsible for its increase, whereas the reform would be neutral in distance terms.

2. Tax reforms and progressivity

The decomposition applied to redistribution can also be done for progressivity, using the Kakwani index (K).

$$K'-K=(C't-G'x)-(Ct-Gx) \tag{18}$$

If, again to simplify, it is assumed that the Gini coefficient before and after the reform does not change, in other words $G'x=Gx$, then

$$K'-K=C'_t-C_t=\frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left|t'(x_i)-t(x'_j)\right|}{2N^2\mu t'}-\frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left|t(x_i)-t(x_j)\right|}{2N^2\mu t} \tag{19}$$

This expression can be rewritten as follows:

$$K'-K=C'_t-C_t=\frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left|t(x_i)-t(x_j)\right|}{2N^2\mu t}\left(\frac{t}{t'}-1\right)+\frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left|t'(x_i)-t'(x_j)\right|-\sum_{i=1}^{N}\sum_{j=1}^{N}\left|t(x_i)-t(x_j)\right|}{2N^2\mu t'} \tag{20}$$

Defining β as the rate of change in the average tax rate, and D and D' as the sum of the distances between the pre-and post-reform tax liabilities, then

$$\beta=\frac{t'}{t}-1 \tag{21}$$

$$D=\sum_{i=1}^{N}\sum_{j=1}^{N}\left|t(x_i)-t(x_j)\right| \tag{22}$$

and

$$D'=\sum_{i=1}^{N}\sum_{j=1}^{N}\left|t'(x_i)-t'(x_j)\right| \tag{23}$$

This would give the following expression:

$$K'-K=C_t\left(\frac{1}{1+\beta}-1\right)+\frac{D-D'}{2N^2\mu t'} \tag{24}$$

Thus, the variation in the Kakwani index would be the sum of the level effect (*LE*) and the distance effect (*DE*), now defined as

$$LE=C_t\left(\frac{1}{1+\beta}-1\right) \tag{25}$$

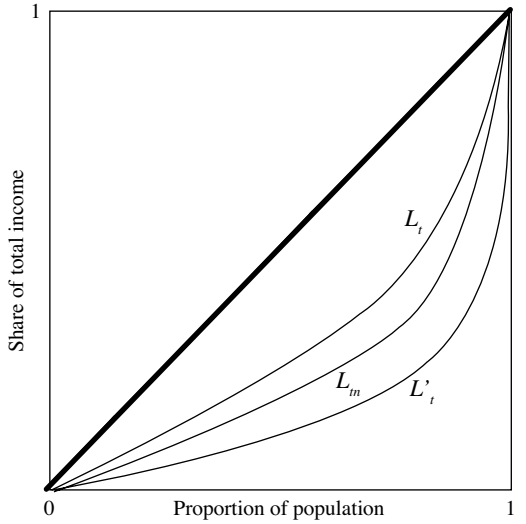
$$DE=\frac{D'-D}{2N^2\mu t'} \tag{26}$$

In this case, the level effect would represent the difference between the tax-liabilities concentration curve before the reform, and what it would be if the reform had been implemented through an equal (positive or negative) transfer for all individuals, keeping the distances constant. Figure 3 illustrates this effect, for the case of a hypothetical tax cut, where *Ltn* represents the post-reform tax-liabilities concentration curve. The distance effect, in contrast, measures the difference between the concentration curve that would exist if the reform had been made through equal (positive or negative) transfers for all individuals, holding constant the distances between tax liabilities and the concentration curve after the real reform (*L't*). For the case of an increase in distances, this effect would also be as shown in figure 3.

Once again, the advantage of this breakdown is that it makes it possible to separate the effect caused by the average tax level from that corresponding to

FIGURE 3

Level and distance effects in progressivity



Source: Prepared by the authors.

L_t: After-tax income concentration curve (pre-reform)
L'_t: Tax-liabilities concentration curve after the real reform
L_{tn}: Tax-liabilities concentration curve following a reform of equal revenue effect implemented with a lump sum tax (in this case, a transfer)

the differences between tax libilities. In other words, whereas the traditional analysis only shows whether progressivity has changed, but not whether this is caused by the change in the average tax rate or by real changes in the differences between individual tax liabilities, the decomposition proposed here does allow that distinction to be made. Thus, both the level effect and the distance affect can take a positive sign (positive contribution to progressivity) or a negative sign (negative contribution).

In particular, for the level effect (*LE*):

If $\Delta t \Rightarrow \beta > 0 \Rightarrow LE < 0$
If $\nabla t \Rightarrow <\beta<0 \Rightarrow LE > 0$
If $\beta = 0 \Rightarrow LE = 0$

where ∇t represents a tax reduction and Δt an increase.

Whereas, for the distance effect (*ED*):

If $D > D' \Rightarrow LE < 0$
If $D < D' \Rightarrow LE > 0$
If $D = D' \Rightarrow LE = 0$

3. Typology of tax reforms

Once the distance and level effects produced by a tax reform have been defined, the different tax-reform modalities can be classified in terms of those effects. In the case of the redistributive capacity of the tax, the classification proposed is shown in figure 4.

FIGURE 4

Redistribution. Types of tax reform

Types of reform
- ΔRS
  - $DE > 0$ Strong redistributive reform
  - $DE < 0$ Weak redistributive reform
- ∇RS
  - $DE > 0$ Weak non-redistributive reform
  - $DE < 0$ Strong non-redistributive reform

Source: Prepared by the authors.

RS: Reynolds-Smolensky index.
EE: Distance effect

Using this typology, tax reforms would be classified not only in terms of the tax's redistributive capacity, but also by what happens to the distances between taxpayers' net incomes. The following synthetic indicator (distance-level redistribution index) can be used to evaluate the different reforms:

$$I_R = \frac{\Delta RS}{|\Delta RS|}\left(1+\frac{DE}{|DE|+|LE|}\right)$$
$$\frac{\Delta RS}{|\Delta RS|} = \pm 1;\ 0 \le \left(1+\frac{DE}{|DE|+|LE|}\right) \le 2 \qquad (27)$$

The $\frac{\Delta RS}{|\Delta RS|}$ component would thus contribute the sign of the indicator, and $\left(1+\frac{DE}{|DE|+|LE|}\right)$ its absolute value, which would reflect the relative importance of the distance effect in the reform. Consequently, according to this indicator, reforms would be classified as:

(i) strong redistributive reform, if $1<I_R\le 2$ (ΔRS, $DE>0$)

(ii) weak redistributive reform, if $0< I_R \leq 1$ (ΔRS, $DE<0$).
(iii) weak non-redistributive reform, if $-2\leq I_R <-1$ (∇RS, $DE>0$)
(iv) strong non-redistributive reform, if $-1\leq I_R \leq 0$ (∇RS, $DE<0$)

The analysis in the case of progressivity is similar. Here again there would be four possible reform types based on the change in the Kakwani index and the sign of the distance effect; and the corresponding indicator (distance-level progressivity index) would be:

$$I_K = \frac{\Delta K}{|\Delta K|}\left(1+\frac{DE}{|DE|+|LE|}\right) \qquad (28)$$

$$\frac{\Delta K}{|\Delta K|} = \pm 1;\ 0 \leq \left(1+\frac{DE}{|DE|+|LE|}\right) \leq 2$$

The meaning and interpretation follow the same patterns, but in this case applied to progressivity:
(i) strong progressive reform, if $1<I_K\leq 2$ (ΔK, $DE>0$)
(ii) weak progressive reform, if $0< I_K \leq 1$ (ΔK, $DE<0$)
(iii) weak regressive reform, if $-2\leq I_K <-1$ (∇K, $DE>0$)
(iv) strong regressive reform, if $-1\leq I_K \leq 0$ (∇K, $DE<0$)

These indicators enhance the classification of tax reforms, augmenting the traditional classifications "redistributive" or "progressive" (based on the positive or negative value of *RS* and *K*) —which are consistent with the classical indicators— with the modifiers "strong" or "weak" obtained from the contribution made by the distance effect in each case. Moreover, provided the value of the indices is standardized to the magnitude of the revenue effects, it will be possible to compare tax reforms that have different quantitative effects.

4. Tax reform and elements of the tax: effects on level and distance

The decomposition of variations in the redistribution and progressivity indicators made in the foregoing subsections makes it possible to evaluate the effects of tax reforms, distinguishing between the amount of a tax cut and the effects of the various elements of the tax used to implement the reform. Thus, the level effect makes it possible to isolate the pure tax increase or decrease component, which is the same for all tax reforms that affect revenue equally. This procedure thus makes it possible to observe the differences between the various possible reforms that produce the same revenue change —differences that will be reflected in the distance effect. By way of example, table 1 summarizes the effects of three possible income-tax reform measures, where the previous tax rate was progressive.

TABLE 1
Distance and level effects in alternative reforms

1. Deductions from tax liability		
Introduction or increase in an equal deduction from tax liability for all taxpayers		
Progressivity	LE > 0 DE = 0	Increase in K
Redistribution	LE > 0 DE = 0	Increase in RS
2. Base reductions		
Introduction or increase in an equal deduction for all tax payers		
Progressivity	LE > 0 DE < 0	K
Redistribution	LE > 0 DE< 0	RS
3. Lowering of tax rates		
Lowering of marginal rates in the tax schedule (irrespective of which rates are reduced)		
Progressivity	LE > 0 DE < 0	K
Redistribution	DE > 0 LE < 0	RS

Source: Prepared by the authors.

LE: Level effect.
DE: Distance effect.
RS: Reynolds-Smolensky index.
K: Kakwani index.

The distance effect separately measures a variation in the distribution (or progressivity) and has normative significance, because it compares two income distributions (or two tax structures) in terms of their departure from proportionality, holding average net

income (or tax revenue) constant. As the level effect is constant for a given amount of tax reduction, the design of the reform (the instruments used) can be evaluated normatively. In relation to the examples given in table 1, the design of reform 1 (deductions from tax liability) would be neutral in terms of its contribution to progressivity and redistribution ($ED = 0$), whereas the design of reforms 2 and 3 (reductions in the base and lower tax rates) would be negative (a decrease) in terms of the progressivity and redistributive capacity of the tax ($ED < 0$). The valuation of other reform alternatives (deductions or reduction in amounts that vary according to income levels; alterations to the brackets of the tax schedule; a combination of increases and decreases in deductions, reductions or rates; changes in the calculation of taxable income, among others), and the joint effects of combining the different measures, is more complex, and the corresponding indicators would have to be calculated as appropriate.

V

An application to personal income tax in Guatemala

Despite efforts made over the last decade to improve the administration and design of Guatemala's fiscal policy, the country's tax take remains one of the lowest in Latin America.[7] Although nominal tax rates are not very different from the Latin American average, a high degree of informality, compounded by unequal income distributions, narrow tax bases and high levels of fraud explain these revenue shortfalls. Personal income tax plays a very small part in the Guatemalan tax structure,[8] whereas corporate and consumption taxes are relatively important (see table 2).

A more detailed analysis of taxation by income sources shows that wages (personal income tax on wages) contribute revenue equivalent to just 0.13% of GDP, or 3.92% of the total revenue raised through income tax. This contrasts with the share of wages in GDP, which was 30% in 2006 according to the national accounts. In other words, only 0.34% of total gross wages paid goes in tax, which is an excessively low proportion. Explanations for this low level of revenue go beyond the high level of informality and extreme inequality prevailing in Guatemala and include the legislation defining taxable income and deductible expenses. Firstly, incomes received in the form of gratifications, bonuses (*Bono 14*), length-of-service payments or pensions (retirement or other) are not subject to income tax. Secondly, deductions are allowed for life-insurance policies, certain donations, medical expenses and, also, a tax credit in respect of VAT paid on purchases for up to 12% of net

TABLE 2

Guatemala: tax revenue in 2006

(Percentages of GDP)

Direct taxes	2.9
Income tax	2.1
Corporate	1.7
Personal	0.3
Tax on financial products (corporations)	0.1
Tax on financial products (individuals)	0.0
IETAAP[a]	0.8
Wealth taxes	0.0
Indirect taxes	7.7
Domestic value added tax	1.9
Value added tax on imports	3.4
Customs duties	1.0
Tax on tobacco and tobacco products	0.1
Taxes on the distribution of beverages	0.2
Tax on vehicle circulation	0.1
Tax on the distribution of oil and petroleum products	0.7
Total tax revenue	11.1

Source: Superintendency of Tax Administration (SAT).

[a] Special temporary tax to support the peace accords.
GDP: Gross domestic product.

[7] Between 2003 and 2008, the tax burden in Latin America and the Caribbean as a whole increased from 15.5% to 17.8% of GDP, but in Guatemala it dropped from 11.9% to 11.6% of GDP (ECLAC, 2009).

[8] Nonetheless, this is in keeping with most tax systems in Latin America, where direct taxation remains weak despite partial reforms.

income. Lastly, there is a tax-free allowance of 36,000 quetzales, representing the minimum living wage and tax threshold, which, while not particularly high in the Central American context (2.03 times per capita GDP in 2006), nonetheless contributes to the loss of revenue and progressivity.

The taxation of personal capital is broadly based despite a clearly dual structure: capital gains pay income tax at a 10% rate, interest and other similar income pays financial products tax also at 10%, whereas dividends are exempt provided they have been subjected to tax withholding at source.

As there is no disaggregated information on taxable income and the taxes paid in Guatemala, personal income tax in the case of wage earners was analysed using data from the National Employment and Income Survey (ENEI) 2004. The database in question was adapted to the structure of the tax, making assumptions about the basic variables not included in the survey, to approximate the baseline scenario (legislation of 2006) to actual revenue outcomes.[9]

This replica of the 2006 baseline scenario reveals a low-revenue tax concentrated on a very small number of actual taxpayers, with little redistributive capacity despite its high formal progressivity. The exemption threshold is the first factor responsible for this low average revenue. Although there are four marginal rates: 15%, 20%, 25% and 31% (see table 3), 73% of wage earners fall outside the income tax net; and the top rate is applied only to very high incomes (16.7 times per capita GDP). All of this results in a scale of rates that is excessively complex in relation to its very narrow scope in terms of the income and individuals subject to it.

The existence of the VAT credit aggravates this situation, because after applying it, only 9.65% of formal workers pay the tax, all of whom are in the top decile of the distribution. Moreover, the VAT credit reduces the amount paid by 62%. A simulation shows that the joint effect of the minimum threshold and the VAT credit is such that eliminating the 31% tax rate would make no difference to revenue outcomes (no taxpayer would be affected). As a result of all of this, progressivity and redistribution indicators reveal a highly progressive tax (Kakwani = 0.6136), but one that is poorly redistributive (Reynolds-Smolensky = 0.0072), owing to its low revenue capacity (the effective average rate is 1.34%).

TABLE 3

Guatemala: personal income-tax rates, 2006

Tax brackets *(in quetzales)*	Marginal rate *(Percentages)*	Percentiles
up to 36 000[a]	0	1-73 (73%)
36 000 - 65 000	15	74-92 (19%)
65 000 - 180 000	20	93-99 (7%)
180 000 - 295 000	25	100 (1%)
over 295 000	31	

Source: Prepared by the authors.

[a] This amount is deducted from taxable income.

In view of this, and to highlight the shortcomings of traditional redistribution and progressivity indicators, this article has proposed two hypothetical reforms acting on these two elements of the tax. This first involves altering the minimum threshold. Here, it should be noted that the existence of the threshold pursues three basic objectives:[10] (i) to set the income threshold needed for a minimum living wage, exempting all incomes below this level; (ii) to reduce the tax burden in line with taxpayers' economic capacity, bearing in mind their family and personal circumstances; and (iii) to simplify the tax, both for the administration and for citizens, exempting large numbers of very low-income taxpayers.

When setting this minimum tax-exempt income threshold, reductions in the tax base (such as that in force in Guatemala), zero-rated income brackets, or tax credits may be equivalent; but the revenue cost is much higher in the case of reductions, and credits are less visible. Moreover, as the tax saving for each taxpayer occurs at his or her top marginal rate, the saving rises with income. Accordingly, at the international level, there is a trend to replace

[9] Incomes declared in the survey have been grouped in the following categories: taxable wage income, exempt wage income, pensions, interest, dividends, rental, capital gains, and incomes from agricultural and nonagricultural activities. These incomes have been adjusted to 2006 prices using the consumer price index, except for interest and dividends, where the recipients of such income have been imputed their proportional share of the total dividends and interest received by households, as reported in the national accounts. To calculate the baseline scenario of revenue equal to the actual figure, net income spent is estimated from the VAT credit at 45%. In addition, individuals below the minimum working age have been excluded from the database, and all persons lacking an employment contract have been classified as "informal". Lastly, the fact that the data were processed at the individual level (more appropriate for tax purposes), rather than by households, needs to be kept in mind when interpreting the inequality indicators.

[10] See, for example, Zee (2005).

reductions in the tax base by zero-rated income brackets or credits (OECD, 2006).

In view of the above, this article argues in favour of transforming the tax-exempt threshold into an equal-sized zero-rated bracket. In that case, revenue would increase by 21%, and there would be an increase in the Reynolds-Smolensky index and a fall in the Kakwani index. The interpretation of the reform based on these indicators would show higher redistributive capacity, but a reduction in progressivity, so the increase in redistributive capacity would merely be the consequence of a higher average rate (see table 4). Accordingly, a reform of this type could be criticized because the gain in redistributive capacity is merely the outcome of the higher average tax rate, whereas the tax is actually becoming less progressive. Nonetheless, as shown by the indices proposed here, what is really happening is that this reduction in progressivity, as measured by the Kakwani index, is also an outcome of the higher average rate. The level effect, related to the latter, contributes negatively to both progressivity and redistributive capacity. In contrast, the distance effect is positive, which means that not only are the distances between net incomes significantly reduced, but the differences between the amounts of tax paid by taxpayers grow. For that reason, it cannot really be said that the reform is regressive; and, if it was, the subjective perception that citizens have of progressivity would be lost. In contrast, the proposal developed in this article simulates a specific reform, separating what happens to the average rate from what happens to the differences between incomes and tax liabilities. In this example, where the difference between taxes paid by high-income and low-income individuals have widened, the indicators proposed would show that the reform was strongly redistributive (I_R=1.74) but weakly regressive (I_K=-1.50).

A second simulation involved a reform abolishing the VAT credit, which would generate a significant increase in revenue (164%). Its consequences for the redistributive capacity and progressivity of the tax would be the same as in the previous measure, namely an apparent loss of progressivity, resulting exclusively from the higher average rate (negative level effect) (see table 5). In contrast, the differences between taxes paid by low- and high-income individuals would increase (positive distance effect). This reform is again strongly redistributive (I_R=1,72) and weakly regressive (I_K=-1,47).

TABLE 4

Transformation of the minimum threshold to a zero-rated bracket

	Post reform	Pre reform	Variation
Revenue *(quetzales)*	354 348 097	291 670 901	21%
RS	0.0090	0.0072	0.0018
K	0.6119	0.6136	-0.0017
	RS	K	
Level effect	-0.0010	-0.1694	
Distance effect	0.0028	0.1677	
I_R	1.7381		
I_K		-1.4975	

Source: Prepared by the authors.

RS: Reynolds-Smolensky index.
K: Kakwani index.
I_R: Distance-level redistribution index
I_K: Distance-level progressivity index.

TABLE 5

Abolition of the VAT credit

	Post reform	Pre reform	Variation
Revenue *(quetzales)*	769 134 255	291 670 901	164%
RS	0.0189	0.0072	0.0118
K	0.5502	0.6136	-0.0634
	RS	*K*	
Level effect	-0.0076	-0.5944	
Distance effect	0.0194	0.5310	
I_R	1.7178		
I_K		-1.4718	

Source: Prepared by the authors.

RS: Reynolds-Smolensky index.
K: Kakwani index.
I_R: Distance-level redistribution index.
I_K: Distance-level progressivity index.
VAT: Value added tax.

VI
Conclusions

This article has attempted to highlight the shortcomings of the Kakwani and Reynolds-Smolensky indices for analysing the tax-progressivity and redistributive-capacity effects of tax reforms that produce changes in revenue. It has also shown that the traditional ways of overcoming these shortcomings by using generalized curves and the decomposition of the Reynolds-Smolensky index into effects caused by the average rate and progressivity, are questionable: in the first case, because they unjustifiably bias the results in favour of tax reductions; and, in the second case, because decomposing changes in the Reynolds-Smolensky index into effects caused by the average rate and by progressivity cannot be used to evaluate the design of the reforms, since the measurement of progressivity is altered in most real reforms by the same factors that alter the average tax rate. Consequently, indicators have been proposed that make it possible to quantify the effects of a tax reform involving changes in revenue based on concepts that are separable, namely the tax level and differences between net incomes or tax liabilities. These indicators make it possible to distinguish the effects on a tax's redistributive capacity and progressivity caused by changes in the average rate and variations in the differences in individual tax liabilities.

The level and distance effects thus developed also make it possible to partly recover the intuitive meaning of the concept of progressivity and redistribution. Deciding "who benefits most" from a tax reform is complicated and subject to value judgments. The traditional indicators (K, RS, and their respective breakdowns) provide a view based on relative differences in incomes or tax liabilities, which is very useful when making comparisons in a static context for reforms that do not cause changes in tax revenue. In other situations, the conclusions obtained may be counterintuitive. For example, how can a tax be made more progressive by a reform that lowers the tax liabilities of high-income individuals by much more, when they receive most of the tax reduction both in absolute and in relative terms? If this is true, then is it "good" to increase progressivity? Put another way, if they were well-informed, would a majority of citizens vote for a reform of this type?

As this article has attempted to show, interpreting the indicators used to evaluate tax reforms that change revenue causes confusion. This article proposes a different alternative. The level effect isolates the repercussions that a reform would have on income shares or tax liabilities between taxpayers (in particular, progressivity and redistribution indicators) if the distances between tax liabilities and incomes remain constant. The distance effect reflects the consequences of a specific reform design (in other words the elements of the tax that are altered) on progressivity and redistribution, when the level of revenue and total income remains constant. This separation of the effects, and the indicators constructed from it, make it possible to nuance and enhance the conclusions reached using the classical indicators.

In addition, this alternative breakdown of the effects of the tax reform makes it possible to highlight the effect of the change in distances between net incomes or between taxes payable, without ceasing to use traditional instruments that are based on a relative concept of inequality, which is of interest for two reasons. Firstly, given the explanation of the repercussions of a tax reform, public-sector managers and citizens may wish to know its consequences in absolute terms, and how it affects the distances between individual incomes (or tax liabilities). Secondly, the decomposition allows for an approach to "relative-income hypotheses". Although, strictly speaking, it cannot be claimed that the indicators described in this paper capture the effect on the relative position of taxpayers in the tax-reform scenario, they do reveal its aggregate effect on the set of relative positions through the calculation of the variation in distances between incomes or tax liabilities.

Lastly, it is worth noting that the problems with the traditional indicators discussed throughout this article are more serious in the case of low-tax countries. A reform that generates the necessary increase in the tax system's revenue capacity will often appear to be regressive, according to these indicators; and this may be used as an argument to reject the reform. Nonetheless, this apparent regressivity is merely the consequence of the increase in revenue capacity itself and not of a narrowing of the differences in the amount

of tax paid by high- and low-income taxpayers. As shown in the tax measures analysed for Guatemala, the traditional indicators would suggest that the tax has become less progressive. Nonetheless, in both cases the differences between taxes paid by higher- and lower-income individuals have actually widened, which seems to contradict the previous conclusion. In contrast, this proposal shows clearly that this reduction in progressivity is merely the consequence of the higher average rate following the reforms, but that the cost of the reforms will affect higher-income taxpayers more. In our opinion, this information is of considerable social interest.

(Original: Spanish)

Bibliography

Atkinson, A.B. (1970), "On the measurement of inequality", *Journal of Economic Theory*, vol. 2, No. 3, Amsterdam, Elsevier.

ECLAC (Economic Commission for Latin America and the Caribbean) (2009), *Economic Survey of Latin America and the Caribbean, 2008-2009* (LC/G.2410-P), table A-37, Santiago, Chile. United Nations publication, Sales No. E.09.II.G.2.

Kakwani, N.C. (1977), "Measurement of tax progressivity: an international comparison", *The Economic Journal*, vol. 87, No. 345, London, Royal Economic Society.

Lambert, P. (2001), *The Distribution and Redistribution of Income*, Manchester, Manchester University Press.

OECD (Organisation for Economic Co-operation and Development) (2006), "Fundamental reform of personal income tax", *Tax Policy Studies*, No. 13, Paris.

Reynolds, M. and E. Smolensky (1977), *Public Expenditures, Taxes and the Distribution of Income: The United States, 1950, 1961, 1970*, New York, Academic Press.

Shorrocks, A.F. (1983), "Ranking income distributions", *Economica*, vol. 50, No. 197, London, London School of Economics and Political Science.

Zee, H.H. (2005), "Personal income tax reform: concepts, issues, and comparative country developments", *IMF Working Papers*, No. 87, Washington, D.C, International Monetary Fund.

CEPAL REVIEW 102 • DECEMBER 2010

KEYWORDS

Fuels
Taxation
Gasoline
Diesel fuels
Automobiles
Supply and demand
Chile

Differential fuel taxes and their effects on automobile demand

Claudio A. Agostini

Fuel taxation policy in Chile has always been to keep taxes on diesel lower than those on gasoline. The proportion of automobiles with diesel engines has grown considerably as a result. Just 20% of diesel engine emissions are equivalent to 80% of gasoline engine emissions, and this affects the level of externalities associated with automobile use, especially in cities such as Santiago where pollution levels are high. This study estimates the effect of the fuel tax differential on automobile demand. The findings indicate diesel automobile demand elasticities of -3.4 and 2.1 with respect to the price of the vehicle and the tax differential. The scale of these effects means there is scope for substantial emissions cuts by way of tax changes to equalize gasoline and diesel tax rates and create a specific tax on diesel automobiles.

Claudio A. Agostini
Associate Professor,
ILADES-Alberto Hurtado University
agostini@uahurtado.cl

I
Introduction

In Chile, the government has greatly increased the specific tax on gasoline over the past 20 years while keeping the specific tax on diesel unchanged. The basic motive has simply been to increase or recoup tax revenues; because demand for fuel is assumed to be fairly inelastic, the thinking is that the distortions caused by taxation should be minimized by applying it in these markets. Proportionally equivalent increases in all fuel markets would be the natural outcome of this logic, but gasoline taxes have risen by proportionally more than the tax on diesel, thus changing the relative prices of the two. In 1990, the gasoline tax was increased from 3 UTM per cubic metre to 3.6186 UTM.[1] It was then raised to 4.4084 UTM in 1995, 5.2 UTM in 2000 and 6 UTM in 2001. In 2008 it was cut first to 4.5 UTM in March and then to 3.5 UTM in September. Throughout, the tax on diesel has been held at 1.5 UTM per cubic metre.

This relative price change deriving from tax policy in fuel markets has had two effects. The first has been an increase in the consumption of diesel relative to gasoline in all uses for which the two fuels are substitutes. The second has been an incentive to purchase diesel rather than gasoline cars. Imports of diesel-powered light passenger vehicles rose from 13,646 units in 1997 to 61,433 in 2007. Similarly, the total fleet of diesel automobiles rose from 267,341 units in 2002 to 566,122 in 2008, i.e., there was an increase of 112% in six years, while the fleet of gasoline automobiles grew by 30% in the same period. The diesel and gasoline automobile fleets thus grew by 13.3% and 4.5% a year between 2002 and 2008, respectively. The result, as figure 1 shows, is that the share of diesel automobiles in the total has grown steadily over recent years, rising from 12.7% in 2002 to 19.2% in 2008 in the country as a whole, and from 9.1% to 14.4% in the Metropolitan Region over the same period.

This change in the composition of the automobile fleet is important because of the externalities automobile use produces. There are externalities such as road congestion and traffic accidents that are unconnected with the type of engine in the automobile, so that different tax treatment for different fuel types ought not to affect their level directly.[2] However, there are externalities such as air pollution from emissions that are related to the type of engine.

Given that pollution from an automobile creates an external cost for society which automobile owners and users do not take into account in their decision-making, tax policy can play an important part in internalizing this cost. Three decisions are relevant here. The first is the number of kilometres driven since it is this, in combination with the automobile's emissions per kilometre travelled, that determines the total volume of pollution produced. The second is the type of fuel used by the automobile and the emissions associated with it. The third is the choice of vehicle characteristics that directly or indirectly affect emissions per kilometre travelled, such as the age of the car, its fuel efficiency (kilometres per litre), engine size, etc.

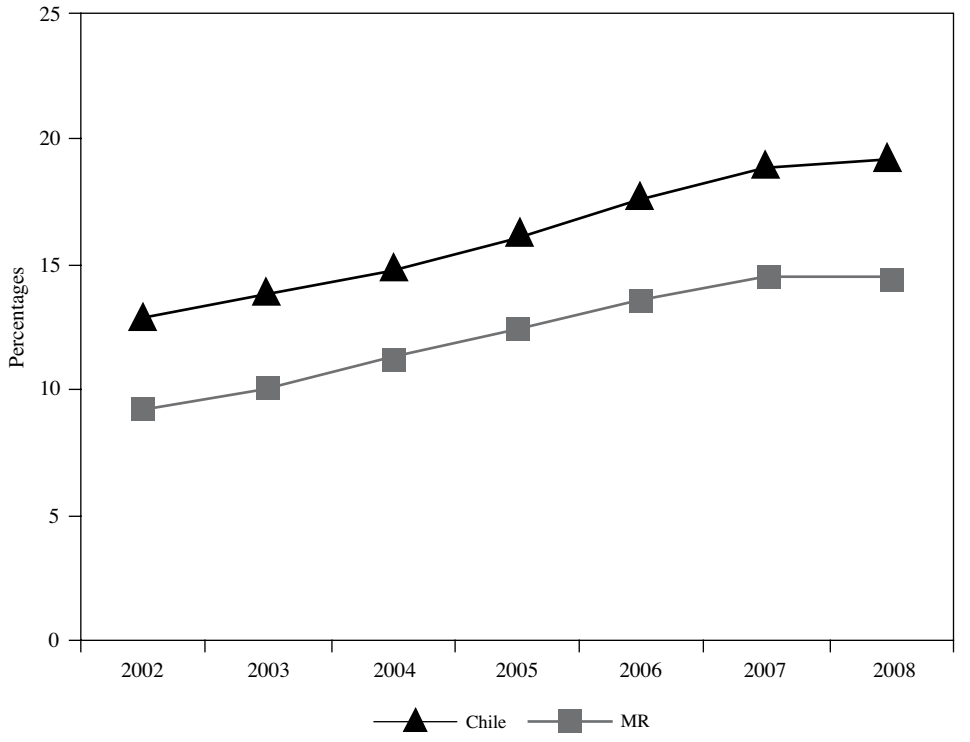
An optimal tax policy would be one that led to consumers incorporating external costs into their demand for automobiles, including fuel type. If a vehicle's emissions are proportional to its fuel consumption and do not depend on its design and engine type, as is the case with carbon dioxide emissions from automobiles, fuel use alone will determine the level of emissions and a tax per litre will allow consumers to internalize the social costs of their emissions. Thus, a gasoline tax encourages consumers to purchase smaller or more fuel-efficient vehicles, maintain them better and change their usage patterns (fewer kilometres travelled) to lower consumption (Portney and others, 2003; West and Williams, 2004; Bento and others, 2005). Furthermore, given that gasoline and leisure are relative complements, increasing the gasoline tax will increase the supply of labour, generating additional

[1] The monthly tax unit (UTM) is an index used to maintain the value of taxes in constant money. In June 2010, one UTM was worth 37,083 Chilean pesos.

[2] Tax rates obviously have a direct effect on levels of externalities, and potentially a considerable one. Thus, in the case of the United States, it has been estimated that an increase of US$ 1 per gallon (3.7854 litres) would cut gasoline consumption by between 15% and 20%, mileage by between 11% and 12% and road accident deaths by between 16% and 18%, while an extra US$ 100 billion in taxes would be raised over 10 years (Haughton and Sarkar, 1996).

FIGURE 1

Chile: proportion of diesel-powered automobiles
(Percentages)



Source: "Anuarios parque de vehículos en circulación", National Institute of Statistics (INE) [online] http://www.ine.cl/canales/chile_estadistico/estadisticas_economicas/transporte_y_comunicaciones/parquevehiculos.php.
MR: Metropolitan Region.

efficiency gains (West and Williams, 2004). Conversely, increasing fuel economy standards encourages greater automobile use by reducing the cost per kilometre travelled (Thorpe, 1997).

Nitrogen oxide and carbon monoxide emissions, on the other hand, are not proportional to fuel use but depend directly on a vehicle's engine size and type. A catalytic diesel automobile emits between 0.6 and 0.8 grams of nitrogen oxide (NO_x) and between 0.04 and 0.09 grams of particulate matter per kilometre, while a gasoline car with a catalytic converter emits about 0.1 grams of nitrogen oxide per kilometre and no particulate matter. What this works out at is that, on average, 20% of diesel automobile emissions are equivalent to 80% of gasoline automobile emissions, making the pollution externalities of the two types of vehicle very different.

Given the environmental situation in some Chilean cities, the externalities generated by diesel vehicles are greater than those of gasoline vehicles, which means that the difference in specific taxes favouring diesel is not consistent with a Pigovian tax approach designed to correct externalities. Broadly speaking, then, a tax treatment that favours diesel does not seem justifiable on environmental grounds (see, for example, Michaelis, 1995). In particular, emissions from diesel engines have major environmental disadvantages from the standpoint of urban air quality (Crawford and Smith, 1995). Even under the more demanding new standards for particulate matter established in Europe, there is evidence of negative effects on climate change (Jacobson, 2002) and on urban mortality (Mazzi and Dowlatabadi, 2007).

Other elements of taxation policy undoubtedly need to be considered for optimum gasoline taxes to be determined, particularly the role of diesel as a production input in some industries, but it is important to determine the effect of the current tax differential

on the composition of the automobile fleet in Chile in order to be able to identify its repercussions on certain cities with high levels of environmental pollution, such as Santiago and Temuco. For this, the important question is how sensitive diesel automobile sales are to the relative specific taxes and vehicle prices. These two elements are complementary from the standpoint of optimal tax policy, since besides the fuel tax it is possible to use a tax on diesel automobiles to correct externalities. The reason is that diesel automobiles, while more expensive, use less fuel per kilometre, and willingness to pay for this characteristic varies by consumer type, depending on how the automobile is to be used (distance driven per year). A tax on diesel automobiles makes it possible to discriminate by consumer type, raising prices by more for consumers who generate greater externalities owing to their more intensive vehicle use.[3]

The objective of this study is to estimate the elasticities of automobile demand by engine type in relation to price and specific fuel taxes. This is a first input of relevance for evaluating the effects of current tax policy on fuels in Chile and proposing tax changes to internalize external costs associated with automobile use, particularly as regards air pollution.

The remainder of the article continues as follows. Section II presents an automobile demand model together with methodological considerations for estimating it. Section III describes the data used in the empirical analysis. Section IV presents the findings of the estimation and section V concludes with some future research and policy recommendations.

II
An automobile demand model

Demand functions are generally estimated using uniform prices for an industry's products, and a single estimate for the price elasticity of demand is obtained as a result. Such estimates are valid for homogeneous products, but they omit information that is important for an understanding of consumers' substitution patterns in the case of differentiated or heterogeneous products. The reason is that differences between products translate into different demand elasticities, making it important when estimating them to take explicit account of the characteristics differentiating products from one another. This is true of automobile demand, since a consumer can choose between different vehicle makes and models. When deciding which automobile to buy, consumers compare different characteristics across the different makes and models available. Price is undoubtedly among the most important characteristics, but consumers will also consider the type of engine (gasoline or diesel), as this is a major determinant of the automobile's annual running costs. Because consumers have to choose between different makes and models with engines of differing types and sizes, demand for automobiles is a demand for differentiated products.

It is also important to realize that consumers have the option of not buying an automobile (outside option). It is important for this option to be considered in the estimation of demand, because if it did not exist then it would be possible to raise the prices of all automobiles uniformly without demand being affected, as relative prices would remain constant.

Estimating demand for differentiated products is complicated, chiefly because of the large number of parameters that have to be estimated. Where N different products are involved, N price elasticities and N(N-1) cross-price elasticities need to be estimated. This means that for most cases the econometric model is "overparameterized" and impossible to estimate. There are two ways of solving this problem. The first is to aggregate similar products until there are just a few product groups. The largest cost of this strategy is that some parameters which might be of interest are lost.

[3] There is evidence that this type of price discrimination by engine type is employed by vehicle manufacturers (Verboven, 2002). Between 75% and 90% of the price differential between diesel and gasoline automobiles appears to be due to price discrimination between heavy and light vehicle users.

The second way is to model product choice explicitly. This approach is based on the work of McFadden (1974), who developed discrete choice models to characterize the choice of products by a consumer, and it is the one used in the present study.[4]

A second difficulty in estimating demand for differentiated products is the heterogeneity of consumers. If consumers did not have different preferences they would all buy the same automobile, conditional upon income. This does not happen, though, because consumers have individual characteristics which lead them to prefer one model or make of automobile out of all those available, and these characteristics somehow need to be incorporated into the demand model.

Discrete choice

By contrast with a standard consumption model, where the quantity consumed is a continuous variable, what is analysed in a discrete choice model is a situation where the relevant consumption decision is discrete. In the case of the decision to buy an automobile, for instance, the relevant thing is what automobile a consumer buys, rather than how many. To be able to estimate demand for goods or services that are consumed discretely, a discrete choice model relates each consumer's choice statistically with his or her personal characteristics and the characteristics of the different products or services available to be chosen. In this way, the model can be used to estimate the likelihood of a consumer picking a specific alternative.[5]

To estimate models of this type it is necessary to specify, first, the indirect utility function of a consumer *i* purchasing an automobile with an engine of type *m* (cylinder capacity and gasoline or diesel) in year *t*. Following Berry (1994) and Nevo (2000), this is defined as:

$$U^*_{imt}(X_{mt}, P_{mt}, \xi_{mt}, \tau_t, v_i; \theta) \qquad (1)$$

where X is a vector of size k of an automobile's observable characteristics, ξ are an automobile's unobservable characteristics, τ is the specific tax on the fuel used by the automobile, P is the automobile's price, v are consumers' individual characteristics and θ are the unknown parameters to be estimated.

A simple functional form used in the literature for the utility function (1) is:[6]

$$U^*_{imt} = \alpha P_{mt} + \gamma \tau_t + X_{mt}\beta + \xi_{mt} + v_{imt} \qquad (2)$$

The functional form used in (2) assumes that the unobserved heterogeneity of consumers (the individual characteristics that determine their preferences) is captured by a random shock vimt.[7] The choice of a specific functional form along with the assumptions made about the distribution of v have a considerable effect on patterns of substitution between products (Berry, 1994).

As noted earlier, it is important for the analysis to consider the possibility that the consumer may decide not to purchase an automobile (possibly deciding to purchase a motorcycle, use only public transport, etc.). The utility function of this outside alternative is defined as:

$$U^*_{iot}(X_{ot}, \xi_{ot}, \tau_{ot}, v_i, \theta) \qquad (3)$$

and the specific functional form taken for this is:

$$\pi^*_{iot} = \xi_{ot} + v_{iot} \qquad (4)$$

The assumption implicit in this model is that consumers purchase just one automobile, namely the one that gives them the highest level of utility. Thus, consumer *i* will choose to purchase an automobile of engine type *m* if and only if that purchase provides him or her with the greatest utility of all the alternatives. In other words, if the following condition is met:

[4] Using this approach, Berry, Levinsohn and Pakes (1995 and 1998) study strategic price interactions between automobile manufacturers in the United States; Nevo (2001) estimates demand for cereals in the United States; and Agostini (2007) and Agostini and Jalile (2009) estimate the effect of taxes on foreign investment in the United States and Latin America, respectively.

[5] While the estimation may be carried out with data on individual consumer decisions, this is not a requirement for using these models and it can be done with aggregated market-level data for each alternative.

[6] See Berry (1994), Berry, Levinsohn and Pakes (1995) and Nevo (2000). The model is fairly general, however, and can be used with different specifications subject to minor adjustments.

[7] A consumer is thus defined as a vector of specific shocks by automobile engine type: (v_{iot}, v_{i1t}, …, v_{iMt}).

$$\alpha P_{mt} + \gamma\tau_t + X_{mt}\beta + \xi_{mt} + v_{imt} \geq \alpha P_{kt} + \gamma\tau_t + X_{kt}\beta + \xi_{kt} + v_{ikt} \quad (5)$$

for any k ≥ 0 and k ≠ m ($k = 0$ represents the outside option: an alternative that entails not buying an automobile).

Condition (5) implicitly defines a set of unobserved individual characteristics v_{imt} leading to the choice of automobile model/make m. This set is defined by:

$$A_{mt}(\delta_{.t}) = \left\{ v_{i.t} \mid \delta_{mt} + v_{imt} \geq \delta_{kt} + v_{ikt} \ \forall\ k = 0...M \right\} \quad (6)$$

where $\delta_{.t} = \alpha P_{.t} + \gamma\tau_t + X_{.t}\beta + \xi_{.t}$ is the mean utility of each make/model.

In the case of the outside option, the mean utility (δ_{ot}) cannot be identified without making more assumptions or normalizing one of the makes of automobile. The simplest way of identifying it, and the standard alternative used in the literature, is to normalize ξ_{ot} at zero.

It is then possible, for a given set of parameters, to predict the market share of each engine type as a function of the characteristics of the automobile, its price, taxes and unknown parameters. The market share of model/make m in period t, S_{mt}, is determined by the proportion of consumers for whom condition (5) is met, i.e., for whom consuming model m provides greater utility than any of the other alternatives, including the outside one. This proportion is determined by the likelihood of v_{it} being in set A_{mt}. Given a distribution function and a density for v, $F(v)$ and $f(v)$, respectively, the market share of model m in period t is:

$$s_{mt}(\delta_{.t}) = \int_{A_{mt}} f(v)dv \quad (7)$$

In order to calculate the integral of equation (7), it is necessary to assume a distribution for vimt. An assumption traditionally used in the literature is that vimt are independent and equally distributed in accordance with an Extreme Value Type I distribution function.[8] In this case, the market share of engine type m is:

$$s_{mt} = \frac{\exp\left(\alpha P_{mt} + \gamma\tau_{mt} + X_{mt}\beta + \xi_{mt}\right)}{1 + \sum_{k=1}^{M} \exp\left(\alpha P_{mt} + \gamma\tau_{kt} + X_{kt}\beta + \xi_{kt}\right)} \quad (8)$$

The estimation strategy consists in choosing the parameters that minimize the distance between the market shares predicted by the model and those observed, which means solving the following implicit equation system:[9]

$$s_{.t}\left(X_{.t}, P_{.t}, \tau_{.t}, \delta_{.t}; \theta\right) = S_{.t} \quad (9)$$

where $s_{.t}(\cdot)$ are the market shares defined by equation (7) and $S_{.t}$ are the observed market shares.

Equation (9) can now be resolved analytically to get:

$$\delta_{mt} = \ln\left(S_{mt}\right) - \ln\left(S_{ot}\right) \quad (10)$$

where S_{mt} and S_{ot} are the market shares of engine type m and the outside option, respectively.

Thus, the aggregate demand equation to be estimated is:

$$\ln\left(S_{mt}\right) - \ln\left(S_{ot}\right) = \alpha P_{mt} + \gamma\tau_{mt} + X_{mt}\beta + \xi_{mt} \quad (11)$$

The price elasticities of the market shares, as defined by equation (8), are:

$$\eta^{P}_{mt} = \frac{\partial s_{mt}}{\partial P_{kt}} \frac{P_{kt}}{s_{mt}} = \begin{cases} \alpha P_{mt}\left(1 - s_{mt}\right) & \text{if } m = k \\ \alpha \boldsymbol{P}_{kt} s_{kt} & \text{if } m \neq k \end{cases} \quad (12)$$

Similarly, the tax rate elasticity of the market shares is:

$$\eta^{\tau}_{mt} = \frac{\partial s_{mt}}{\partial \tau_{mt}} \frac{\tau_{mt}}{s_{mt}} = \begin{cases} \gamma\tau_{mt}\left(1 - s_{mt}\right) & \text{if } m = k \\ \gamma\tau_{kt} s_{kt} & \text{if } m \neq k \end{cases} \quad (13)$$

In summary, this model can be used to solve the problem of "overparameterization" of demand

[8] See, for example, McFadden (1974), Cardell and Dumbar (1980), Boyd and Mellman (1980) and Tardiff (1980).

[9] The integral of equation (7) can be calculated analytically using assumptions about the distribution of v.

for differentiated products and makes it possible to obtain consistent estimators of the elasticities concerned.[10] Lastly, it is important to consider that the price variable might be endogenous (with P_{mt}, ξ_{mt}), in which case it is necessary to estimate equation (11) using instrumental variables.

[10] This model has been used in the literature by studies dealing with automobile and fuel markets. For example, Agras and Chapman (1999) estimate the impact of emissions standards and fuel taxation on total greenhouse gas emissions in the United States transport sector; Levinsohn (1996) estimates the effect of different trade policies on automobile sales; and Greene (1986) analyses the evolving market share of diesel automobiles in the United States.

III

Data used in the empirical analysis

The empirical analysis uses monthly data on automobile imports registered by the National Customs Service for the period from 2002 to 2008. The breakdown for each entry identifies the type of engine (gasoline or diesel), cylinder capacity in four categories (under 1,000 cm^3, between 1,000 cm^3 and 1,500 cm^3, between 1,500 cm^3 and 3,000 cm^3, and over 3,000 cm^3), whether the automobile is an off-road or four-wheel drive (4x4) vehicle, and the country of origin. For each of these groups of automobiles, the customs register lists the number of units imported, the country of origin, the average unit price and the standard deviation of the average price.[11] Unfortunately, no breakdown of data by unit is available, nor was it possible to obtain a classification by model or even by make of automobile.

Table 1 gives a statistical summary of the data. The Quantity variable is the number of vehicles imported per month in each of the 304 categories.[12] It provides the basis for calculating the logarithm of the monthly share of vehicle imports in each category, which is used as a dependent variable in the regression ($\log S_{mjt}$). Although no data are available to calculate the proportion of consumers who prefer the option of not having a car each month, there are data on motorcycle imports, making it possible to calculate the market share of this outside option ($\log S_{ot}$), which is accordingly used in the empirical analysis.

The price variable is the average dollar unit price of automobiles in each category. The variables generated by the tariff categories of the different automobile import entries were used to determine seven dummy variables capturing these characteristics: Engine 1,000 for engine sizes under 1,000 cm^3, Engine 1,500 for engine sizes between 1,000 cm^3 and 1,500 cm^3, Engine 3,000 for engine sizes between 1,500 cm^3 and 3,000 cm^3, Engine 3,000+ for engine sizes over 3,000 cm^3, 4x4 for off-road or four-wheel drive vehicles, Diesel for diesel-engined vehicles and, lastly, Free Zone for all vehicles entering by one of the three free zones that exist in Chile (Arica, Iquique and Punta Arenas). The Tax variable is the monthly value, in dollars per litre, of the specific tax on gasoline. The IMACEC variable measures the monthly change in economic activity in the country, being based on the performance of 90% of the goods and services that make up gross domestic product (GDP); the monthly indicator of economic activity (IMACEC) is published by the Central Bank of Chile. Lastly, the Steel Price variable is the price of the type of stainless steel used by automobile producers, as published in the MEPS *Stainless Steel Review* on the basis of information from contracts between steel producers and consumers. Because steel is a production input in automobile manufacturing that ought not to be correlated with unobserved variables determining the demand for automobiles, its price is used as an instrumental variable for the price of automobiles when estimating demand.

[11] There are records of imports from 45 countries, but 18 countries account for over 95% of all vehicles imported. For this reason, the empirical analysis specifies the proportion of automobiles imported from each of these 18 and groups the rest into an "Other" category.

[12] The 304 categories of automobile in the data arise from the interaction of: 4 engine sizes, 2 engine types, 2 vehicle types (automobile and off-roader or 4x4) and 19 countries of origin.

TABLE 1

Statistical summary

	Average	Standard deviation	Minimum	Maximum
Quantity	151.22	343.63	1	3 544
S_{mjt}	0.0101	0.0218	0.004	0.197
Price	12 379	11 663.8	2 000	152 352
Engine 1,000	0.034	0.165	0	1
Engine 1,500	0.099	0.278	0	1
Engine 3,000	0.214	0.376	0	1
Engine 3,000+	0.131	0.295	0	1
4x4	0.087	0.126	0	1
Free Zone	0.203	0.327	0	1
Diesel	0.248	0.291	0	1
Tax	0.254	0.064	0.143	0.411
IMACEC	0.305	0.821	-1.8	1.7
Steel Price	3 421.25	1 328.71	2 091	4 732

Source: National Customs Department, Central Bank of Chile and MEPS *Stainless Steel Review*.

Note: Average unit prices in dollars.
Engine sizes are stated in cm^3.
IMACEC: Monthly indicator of economic activity.
Free Zone: Arica, Iquique and Punta Arenas.

IV

Results of the estimation

Table 2 shows the results of estimating equation (11). Model (1) is the ordinary least squares estimation, and therefore does not take account of the possible endogeneity of automobile prices. In Chile, 100% of light passenger automobiles are imported, meaning that supply is determined simply by the international price, additionally bearing in mind that Chile is a small country. For this reason, the problem of simultaneity in determining the equilibrium price in the automobile market does not generate problems of identification. It is possible that there may be measurement errors in the price variable, however, as it is an average for each category of automobile. To take this possibility into account and remove a possible endogeneity bias when estimating price elasticity, equation (11) is estimated once again using the price of steel as an instrumental variable for the price of automobiles.

The empirical results are generally quite satisfactory. The regressions are able to explain about 56% of the variation in the data, the signs of the coefficients are the expected ones for automobile demand and they are all statistically significant at a 5% confidence interval. The demand for automobiles is greater for engine sizes larger than the omitted category of engines smaller than 1,000 cm^3, but the most considerable effect is seen with engines of between 1,500 cm^3 and 3,000 cm^3, which is where most demand is concentrated. Similarly, the findings reveal a positive effect on aggregate automobile demand of diesel-powered as compared to gasoline-powered vehicles, and of four-wheel drive vehicles as compared to those with two-wheel drive. Although the size of the effects is fairly small, with elasticities of between 0.2 and 0.4, this is consistent with the growing market share of both diesel and 4x4 automobiles over time. The positive repercussions of the Free Zone reflect the opportunity to import automobiles free of customs duties in two regions of the country, increasing demand there relative to the other regions of Chile. On average, everything else being constant, having a free zone increases relative demand for automobiles by 3.6%. The positive effect of the IMACEC on automobile demand is consistent with a positive income elasticity, but also reflects the

TABLE 2

Estimates of the automobile demand equation

	(1)	(2)
Price	-0.00028*	-0.00028*
	(0.000091)	(0.00012)
Engine 1,500	0.0032*	0.0034*
	(0.0011)	(0.0013)
Engine 3,000	0.0026*	0.0028*
	(0.0011)	(0.0013)
Engine 3,000 +	0.0011*	0.001*
	(0.0004)	(0.0005)
4x4	0.0184*	0.0193*
	(0.0032)	(0.0043)
Free Zone	0.0462*	0.0484*
	(0.0152)	(0.0167)
Diesel	0.002*	0.002*
	(0.0002)	(0.0006)
Tax	-0.0301	-0.0318
	(0.0527)	(0.0507)
IMACEC	0.4011*	0.4036*
	(0.1785)	(0.1918)
Constant	-13.82*	-9.74*
	(1.56)	(1.67)
Country of origin dummies	Yes	Yes
R^2	0.557	0.563
F	141.7	152.1
N	8 307	8 307

Source: author's estimates based on data detailed in table 1.

* Significant at 5%.
Note: Model (1) estimated by ordinary least squares with robust errors; model (2) estimated by instrumental variables and bootstrapping with 1,000 repetitions for errors.
Engine sizes are stated in cm^3.
Free Zone: Arica, Iquique and Punta Arenas.
IMACEC: Monthly indicator of economic activity.
R^2: Goodness of fit. F: Fischer test. N: Number of observations.

role played by the economic cycle in vehicle imports. An increase of 1% in monthly economic activity is associated with an average increase of about 1.4% in automobile demand.

The coefficients of greatest interest in the findings are those related to price and the fuel tax. Price elasticity for aggregate automobile demand in Chile, as evaluated at the mean of the sample, is -3.4. This value is not very different from the results obtained by other studies in the economic literature. In fact, while somewhat greater than the value of -2.4 estimated by Trandel (1991), it is very close to the -3.28 estimated by Goldberg (1995) and within the range of -3.0 to -4.5 estimated by Berry, Levinsohn and Pakes (1995).[13] Price elasticities vary both by engine type and by engine size, reflecting consumers' patterns of substitution between different types of automobiles. Automobiles with engines larger than 3,000 cm^3 have the greatest price elasticity (-3.71, evaluated at the mean of the sample), while automobiles with engines of between 1,500 cm^3 and 3,000 cm^3 have the lowest (-2.6, also evaluated at the mean).[14] This range denotes a greater elasticity of demand between engine sizes than is found in developed countries. For the United States, for example, Bento and others (2009) estimate elasticities ranging from -1.44 for compact cars to -2.3 for minivans. The price elasticity of 4x4 vehicles, lastly, is estimated at -3.83, making them the most price-elastic vehicles in Chile.

The fuel tax elasticity of demand for diesel automobiles is estimated at -2.1, evaluated at the mean of the sample.[15] Given that the diesel tax has been kept unchanged throughout the period, the effect of the tax differential between gasoline and diesel is identified in this elasticity. The magnitude of this elasticity, while smaller than the price elasticity, is not small from the standpoint of the role played by tax policy. Fuel tax elasticity partly captures the annual cost of using the automobile, with greater use requiring higher gasoline consumption, so that narrowing the tax rate gap between gasoline and diesel and levying a tax on diesel automobiles would considerably reduce the air pollution associated with diesel.

In particular, increasing the diesel tax from its current rate of 1.5 UTM per cubic metre to 4.5 UTM and thereby equalizing it with the gasoline tax would reduce the proportion of diesel automobiles by almost 3 percentage points. In the case of the Metropolitan Region, assuming the current growth in diesel automobiles remained constant, this would imply about 11,300 fewer imports of diesel automobiles in the next five years. Assuming each vehicle covered

[13] The first studies conducted showed much lower elasticities, however, with values in the range of -1 to -1.5 (see, for example, Chow, 1957; Suits, 1958; Wykoff, 1973).

[14] Elasticities are calculated by the formula set out in (12) using market shares and price for the specific m group. Thus, for example, price elasticity for aggregate demand is calculated as -0.00028*12,379 (1-0.0101)=-3.4. For the different types of automobile, elasticity is calculated in the same way but using the respective average price and market share for each group of automobiles: engine size over 3,000 cm^3, engine size between 1,500 cm^3 and 3,000 cm^3, and 4x4. See Nevo (2000) for details on the derivation and calculation of elasticities.

[15] Tax elasticity is calculated by the formula set out in (13).

12,000 kilometres a year, this would mean 950 tons less nitrogen oxide and 88 tons less particulate matter being emitted annually. If a tax were additionally levied on diesel automobiles at a rate of 1% of vehicle value, the number of diesel vehicle imports would be cut by about 16,000 units over the next five years, leaving annual nitrogen oxide emissions 1,083 tons lower and emissions of particulate matter 106 tons lower.

V
Conclusions

In Chile, one of the important effects of the current difference in tax rates between gasoline and diesel is that it incentivizes the purchase of diesel automobiles. Thus, the proportion of the automobile fleet with diesel engines in the country rose from about 12% in 2002 to 19% in 2008. Since diesel automobiles emit more nitrogen oxide than gasoline automobiles, as well as particulate matter, the externalities associated with air pollution from automobile use have been increasing. This is particularly significant for cities with high levels of pollution in winter, such as Santiago and Temuco. To be able to quantify the scale of this problem and consider tax policy alternatives that would allow externalities to be internalized by the owners of diesel automobiles, it is necessary to know the price and fuel tax elasticities of automobile demand.

The elasticities estimated in this study show that the demand for diesel automobiles is quite sensitive to vehicle price and the tax differential between gasoline and diesel. This provides a basis for implementing a tax policy involving equalization of the gasoline and diesel tax rates and the introduction of a tax on diesel automobiles that would be quite effective in reducing externalities, and most particularly in reducing emissions of nitrogen oxide and particulate matter. To obtain a more precise estimate of the effects of a tax policy of this type, it would be worthwhile for future research to carry out an even more detailed disaggregation of automobile demand than this study has. It is particularly important to consider the different makes and models of automobile so that consumers' patterns of substitution between different automobiles in response to tax rises can be studied with greater accuracy.

(Original: Spanish)

Bibliography

Agostini, C. (2007), "The impact of corporate state taxes on FDI location", *Public Finance Review*, vol. 35, No. 3, London, Sage Publications.

Agostini, C. and I. Jalile (2009), "Efectos de los impuestos corporativos en la inversión extranjera en América Latina", *Latin American Research Review*, vol. 44, No. 2, Pittsburgh, Latin American Studies Association.

Agras, J. and D. Chapman (1999), "The Kyoto Protocol, CAFE standards, and gasoline taxes", *Contemporary Economic Policy*, vol. 17, No. 3, Huntington Beach, Western Economic Association International.

Bento, A.M. and others (2009), "Distributional and efficiency impacts of gasoline taxes", *American Economic Review*, vol. 99, No. 3, Nashville, Tennessee, American Economic Association.

_____ (2005), "Distributional and efficiency impacts of increased U.S. gasoline taxes: an econometrically based multi-market study", *American Economic Review*, vol. 95, No. 2, Nashville, Tennessee, American Economic Association.

Berry, S. (1994), "Estimating discrete choice models of product differentiation", *Rand Journal of Economics*, vol. 25, No. 2, Santa Monica, California, The Rand Corporation.

Berry, S., J. Levinsohn and A. Pakes (1998), "Differentiated products demand systems from a combination of micro and macro data: the new car market", *NBER Working Paper*, No. 6481, Cambridge, Massachusetts, National Bureau of Economic Research.

_____ (1995), "Automobile prices in market equilibrium", *Econometrica*, vol. 63, No. 4, New York, Econometric Society.

Boyd, H.J. and R.E. Mellman (1980), "The effect of fuel economy standards on the U.S. automotive market: an hedonic demand analysis", *Transportation Research*, vol. 14, No. 5-6, Amsterdam, Elsevier.

Cardell, N.S. and F. Dunbar (1980), "Measuring the societal impacts of automobile downsizing", *Transportation Research*, vol. 14, No. 5-6, Amsterdam, Elsevier.

Chow, G. (1957), *The Demand for Automobiles in the United States*, Amsterdam, North-Holland.

Crawford, I. and S. Smith (1995), "Fiscal instruments for air pollution abatements in road transport", *Journal of Transport Economics and Policy*, vol. 29, No. 1, Bath, University of Bath.

Goldberg, P. (1995), "Product differentiation and oligopoly in international markets: the case of the U.S. automobile industry", *Econometrica*, vol. 63, No. 4, New York, Econometric Society.

Greene, D. (1986), "The market share of diesel cars in the USA, 1979-83", *Energy Economics*, vol. 8, No. 1, Amsterdam, Elsevier.

Haughton, J. and S. Sarkar (1996), "Gasoline tax as a corrective tax: estimates for the United States, 1970-1991", *The Energy Journal*, vol. 17, No. 2, Cleveland, International Association for Energy Economics.

Jacobson, M. (2002), "Control of fossil-fuel particular black carbon and organic matter, possibly the most effective method of slowing global warming", *Journal of Geophysical Research*, vol. 107, No. 19, Washington, D.C., American Geophysical Union.

Levinsohn, J. (1996), "Empirics of taxes on differentiated products: the case of tariffs in the U.S. automobile industry", *Trade Policy Issues and Empirical Analysis*, R.E. Baldwin (ed.), Chicago, University of Chicago Press.

Mazzi, E. and H. Dowlatabadi (2007), "Air quality impacts of climate mitigation: UK policy and passenger vehicles choice", *Environment Science Technology*, vol. 41, No. 2, Washington, D.C., American Chemical Society.

McFadden, D. (1974), "Conditional logit analysis of qualitative choice behavior", *Frontiers in Econometrics*, Paul Zarembka (ed.), New York, New York Academic Press.

Michaelis, L. (1995), "The abatement of air pollution from motor vehicles: the role of alternative fuels", *Journal of Transport Economics and Policy*, vol. 29, Bath, University of Bath.

Nevo, A. (2001), "Measuring market power in the ready-to-eat cereal industry", *Econometrica*, vol. 69, No. 2, New York, Econometric Society.

_____ (2000), "Mergers with differentiated products: the case of the ready-to-eat cereal industry", *Rand Journal of Economics*, vol. 31, No. 3, Santa Monica, California, The Rand Corporation.

Portney, P.R. and others (2003), "The economics of fuel economy standards", *Journal of Economic Perspectives*, vol. 17, No. 4, Nashville, Tennessee, American Economic Association.

Suits, D. (1958), "The demand for new automobiles in the United States, 1929-1956", *Review of Economics and Statistics*, vol. 40, No. 3, Cambridge, Massachuetts, The MIT Press.

Tardiff, T. (1980), "Vehicle choice models: review of previous studies and directions for further research", *Transportation Research*, vol. 14, No. 5-6, Amsterdam, Elsevier.

Trandel, G. (1991), "The bias to omitting quality when estimating automobile demand", *Review of Economics and Statistics*, vol. 73, No. 3, Cambridge, Massachusetts, The MIT Press.

Thorpe, S.G. (1997), "Fuel economy standards, new vehicles sales and average fuel efficiency", *Journal of Regulatory Economics*, vol. 11, No. 3, Springer.

Verboven, F. (2002), "Quality-based price discrimination and tax incidence: evidence from gasoline and diesel cars", *Rand Journal of Economics*, vol. 33, No. 2, Santa Monica, California, The Rand Corporation.

West, S.E. and R.C. Williams (2004), "Empirical estimates for environmental policy making in a second-best setting", *NBER Working Paper*, No. 10330, Cambridge, Massachusetts, National Bureau of Economic Research.

Wykoff, F. (1973), "A user cost approach to new automobile purchases", *Review of Economic Studies*, vol. 40, No. 3, Oxford, Institute of Economics and Statistics, University of Oxford.

KEYWORDS

Education
Income
Income distribution
Wages
Equality
Labour market
Ethnicity
Social classes
Peru

Is education income-equalizing?

Evidence from Peru

Adolfo Figueroa

Is the education system income-equalizing? The evidence from developing countries shows that while education has expanded tremendously in recent decades, income distribution has not become more equal. This article seeks to resolve this seeming paradox. A theoretical model is constructed for the relationship between education and income in which ethnicity plays a key role in the distribution process. The model predicts that the education system is not income-equalizing. A broader set of predictions are derived from the theoretical model and then compared to Peruvian data. Statistical tests produce results that are consistent with the model's predictions.

Adolfo Figueroa
Centrum Business School
Catholic University of Peru
afiguer@pucp.edu.pe

I
Introduction

Education has expanded tremendously in developing countries in recent decades. World Bank data show that the net primary school enrolment rate rose from 78% to 89% between 1980 and 1997, while the secondary school rate went up from 53% to 63% (World Bank, 2001, table 6, p. 285). Yet, paradoxically, the degree of income inequality has not declined (Deininger and Squire, 1996, table 5; Li, Squire and Zou, 1998). Why is education not an income-equalizing system?

The literature has not provided a satisfactory scientific explanation for this paradox; and the most popular theoretical models on intergenerational mobility (Becker and Tomes, 1979; Durlauf, 1996) have not been subjected to statistical testing in a developing-country setting.

In contrast, the empirical literature on the relationship between education and income inequality in the developing world is vast but lacks theoretical underpinnings. In the particular case of Latin America, compared to the situation in more advanced countries, three empirical insights can be gleaned: (i) the distribution of education, measured in terms of years of schooling, is more unequal; (ii) the quality gap between schools attended by rich and poor students is greater; and (iii) income differentials between workers with high and low education levels are much wider, which some authors attribute to the relatively limited supply of educated workers (Bourguignon, Ferreira and Menendez, 2007; Birdsall, De la Torre and Menezes, 2008; Blom and Vélez, 2004).

This article explores the role of class and ethnicity in explaining the observed relationships between education and income, including the paradox. A simple theoretical model is developed for that purpose, and its predictions are tested against Peruvian data. The role of ethnicity in the relationship between education and incomes in developing countries is not widely discussed in the international literature. For Peru, Ñopo, Saavedra and Torero (2004) found that ethnicity played an important role in the statistical breakdown of urban wage differences. For Brazil, the study by Bourguignon, Ferreira and Menéndez (2007) cited above also found that race affected earnings differentials. Neither study has theoretical underpinnings, however.

The article is organized as follows. Section II presents a theoretical model in which class and ethnicity play a significant role in the education process. Section III discusses the role of initial inequality in family asset endowments in the process of human capital accumulation. The static and dynamic models are presented in sections IV and V. In section VI, the model's empirical predictions are tested statistically against Peruvian data, utilizing parametric and non-parametric tests (described in full in the Appendix). Section VII presents the conclusions.

II
A simple theoretical model

Consider a hypothetical capitalist society in which the distribution of economic and social assets among individuals is highly unequal. Individuals are assumed to participate in the economic process endowed not only with economic assets but also with social ones, thus introducing social factors into the economic process. Social assets are special goods, for they belong to the realm of rights and entitlements granted to individuals in a society. They are not physical goods, nor are they marketable.

In this study, social assets will basically refer to political and cultural assets. The former are defined as the capacity of individuals to exercise individual and collective rights, including the right to have rights. Unequal individual endowments of political assets generate a hierarchy of citizens in society—first-class

and second-class. As a result, not all individuals are equal before the law; moreover, not all individuals have the same degree of access to the public goods supplied by the State.[1]

Cultural assets are defined as the right of social groups to exercise cultural diversity in a multicultural and multi-ethnic society. Unequal endowments of cultural rights generate ethnic groups with a hierarchy of ethnic markers in society: first- and second-class races, languages, religions and customs. These markers are referred to as cultural because their hierarchies are socially constructed and they are transmitted from generation to generation. Inequality in cultural assets leads to social practices of segregation, exclusion and discrimination against certain ethnic groups.

Unequal individual endowments of political and cultural assets are assumed to be highly correlated in society, so political assets alone can be included in the theory. For simplicity, this fictitious society can be referred to as the "sigma society".

The unequal distribution of economic and political assets is one of the features of the sigma society. The other feature concerns factor endowments, which implies the existence of overpopulation. The marginal productivity of the whole labour force is too low for wage rates to clear labour markets.

In order to derive empirically refutable predictions from sigma theory, a specific sigma society, or a particular model of sigma theory, must now be constructed. A set of auxiliary assumptions are then introduced.

Race, class and citizenship constitute the social structure of sigma society. The total population can be divided into: (i) two social classes: capitalists and workers; (ii) two types of citizens: first-class and second-class; (iii) three ethnic groups: the Blues, the Reds and the Purples, the latter resulting from interbreeding between the other two races.[2]

Table 1 shows the social structure of sigma in matrix form. The Blues constitute the capitalist class, owning most physical capital. The Purples and Reds belong to the working class. Purples are endowed with skilled labour and Reds with unskilled labour only. In terms of citizenship endowments, the Blues and the Purples are first-class citizens, while the Reds are second-class, thus forming the initial inequality in individual asset endowments. The social matrix shows a society that is highly correlated in terms of asset endowment. Three social groups are thus identifiable; for easy reference and for reasons that will become apparent later on, they will be called by the letters A, Y and Z.

Sigma society can now be distinguished analytically from a socially homogeneous capitalist society, which could be referred to as "epsilon society". If epsilon society were represented by table 1, there would still be two social classes (capitalists and workers) but only one citizenship class (C). There would still be three ethnic groups, but only one level of citizenship for all (C). In epsilon society, therefore, racial differences would become unimportant and the social matrix would collapse into two social groups only, A (capitalists) and Y (workers); social group Z would not exist. While sigma society is intended to resemble the developing world, epsilon society corresponds to the first world.

The question considered by this paper is whether or not the process of human capital accumulation through education can reduce inequality in the initial distribution of assets (naturally including human capital) and thus make income flows less unequal. The answer is developed in the rest of the paper.

TABLE 1

Social structure of sigma society: race, class and citizenship

Ethnic group	Physical capital	Human capital	Citizenship	Social group
Blues	K_b	K_{h1}	C_1	A
Purples	0	K_{h1}	C_1	Y
Reds	0	K_{h0}	C_0	Z

Source: prepared by the author.

Symbols: K_b physical capital endowment; K_{h1} high-level human capital endowment; K_{h0} low-level human capital endowment; C_1 first-class citizenship endowment; C_0 second-class citizenship endowment.

[1] The concept of citizenship used in this study corresponds to that proposed by the sociologist T.H. Marshall in his classic 1950 essay: "Citizenship is a status bestowed on those who are full members of a community. All who possess the status are equal with respect to the rights and duties with which the status is endowed" (Marshall and Bottomore, 1992, p. 18).

[2] In a paper that analyses theoretical relationships between consumer preferences and culture, Akerlof and Kranton (2000) construct an abstract world of two ethnic groups, the Greens and Reds, in which the Greens are the dominant group. To use primary colours, call them Blues and Reds here; then introduce a third ethnic group: the Purples, resulting from miscegenation between the two races. As in that paper, it is assumed here that people cannot choose their ethnic identity, which is exogenous.

III

Human capital accumulation: the role of initial endowments

Human capital is usually defined as the stock of knowledge and productive skills embodied in workers. As individuals do not acquire this stock at birth, they need to invest in acquiring it through the education process.

The initial conditions that each individual brings to the educational system at each level are essential for learning. In the literature of related sciences (psychology, biology and neuroscience), the standard view is that initial talent endowments matter, and talents are multiple —the so-called "multiple intelligence theory" (Gardner, 1999).

Brain plasticity theory is another important component of the standard view. While the individual's initial talent endowments —his or her genetic inheritance— are exogenous (nature), talents become endogenous over time because brain development also depends on the social environment (nurture) and on the interaction between the two. Brain plasticity theory is usually stated as follows: "The brain is not a computer that simply executes predetermined programs. Nor is it a passive gray cabbage, victim to the environmental influences that bear upon it. Genes and environment interact to continually change the brain, from the time we are conceived until the moment we die" (Ratey, 2002, p. 17).

On aggregate, individual endowments of talents, genetically-based gifts, can be assumed to be normally distributed among the population (the result of a random mechanism). The distribution of learning skills, in contrast, will be generated by the social environment. The important distinction made by Rousseau (1755) refers precisely to these two factors. He identified two types of inequalities among individuals: the natural, determined by randomly assigned natural endowments; and the artificial, determined by the functioning of society.

The sigma model will assume that nutrition, health and language are the main channels through which wealthy families can develop higher levels of learning skills in their children as compared to poorer families. Personal access to nutrition and health, as either private or public goods, is assumed to be differentiated by social group.

Language is another factor of cognitive skill inequality associated with a family's socio-economic level. There are language differences between individuals, which in sigma society lead to language inequality. This inequality is seen in various aspects of language use in the dominant tongue in society, such as vocabulary, syntax, ways of speaking, and reading and writing skills. Sociolinguistic theory views language inequality as mostly due to personal experiences (the social environment) rather than genetic factors (Hudson, 1996, p. 204).

Inequalities in language skills between social groups imply unequal cognitive skills among their children. Abstract and complex thoughts are not just language-dependent, but complex language-dependent. As philosopher John Searle has stated: "Some thoughts are of such complexity that it would be empirically impossible to think them without being in possession of symbols. Mathematical thoughts, for example, require a system of symbols... Complex abstract thoughts require words and symbols" (Searle, 1995, p. 64). The implication seems to be that written language allows the individual to work with more abstract and complex thoughts than does oral language alone.

Consider a situation in which the Y workers live in a written culture and most of them are literate, but the Z workers live in an oral social environment, most of them are illiterate in the dominant language, and their aboriginal language is not a written one. Assume further that Z workers live in an oral culture in segregated communities. In a setting of this type, language skills in the dominant language will be very unequally distributed through society.

Due to their illiteracy, Z populations will be limited in the use of abstract and complex thoughts. To be illiterate in a written culture is different from being illiterate in an oral culture. The handicap will be greater in the former. Z populations will then show lower levels of language skills in the dominant language, and thus their children will display lower cognitive skill levels than the children of the A and Y populations.

The characteristic of multicultural, multilingual and hierarchical society makes sigma a heteroglossic society. This term comes from sociolinguistic theory and refers to the existence of various forms of, or variations in, the use of the dominant language, with a hierarchy separating those who are socially superior (with a good grasp of the dominant language) from those considered inferior.

The sigma model then assumes that language inequality plays a crucial role in generating unequal cognitive skills. Sociolinguistic theory goes even further: "Linguistic inequality can be seen as a *cause* of social inequality, but also as a *consequence* of it, because language is one of the most important means by which social inequality is perpetuated from generation to generation" (Hudson, 1996, p. 205).

To conclude, in a sigma society with pronounced initial inequality in economic and political asset endowments, and given the random distribution of talents in the population, people will start the human capital accumulation process endowed with unequal cognitive or learning skills. This initial inequality plays a major role in the process of human capital accumulation among social groups. It should be noted that this proposition is not empirically refutable or falsifiable, because learning capacity is unobservable. It will therefore be used as a primary assumption of the theory of human capital accumulation, presented below.

IV

The transformation of education into income: a static model

The static sigma model can be represented by the following system of equations:

$$k_h = f(E, X), f1 > 0, f_2 > 0, X = Z, Y, A \quad (1)$$

$$y = g(k_h, X, p), g_1 > 0, g_2 > 0, g_3 > 0 \quad (2)$$

Hence,

$$\begin{aligned} y &= g(f(E, X), X, p) \\ &= G(E, X, p), G_1 > 0, G_2 > 0, G_3 > 0 \end{aligned} \quad (3)$$

Equation (1) states that the average human capital endowment of individuals (k_h) depends upon their average education level measured as years of schooling (E) and the social group to which they belong (Z, Y, A). For a given social group, an exogenous increase in the average education level will result in a higher average human capital level; for a given education level, the average human capital level will be higher for individuals belonging to the higher-ranked social groups (where the hierarchy in ascending order is Z, Y, A).

Equation (2) says that the average income of individuals (y) depends upon their average endowments of human capital and the social group to which they belong, for given market conditions (p). In a given social group, the higher the average human capital endowment, the higher the average income will be; for a given endowment of human capital, average income will be higher for individuals belonging to higher-ranked social groups (Z, Y, A, in ascending order).

To simplify, assume that market conditions depend on the international terms of trade (p), which are the sole exogenous variable determining the aggregate level of output, employment and mean income in the short run. A higher p means a higher relative price for exported goods compared to imports, which in turn implies higher labour productivity and higher demand for labour by domestic firms. Thus, for given educational levels, the higher the value of p, the higher the average income of the social groups.

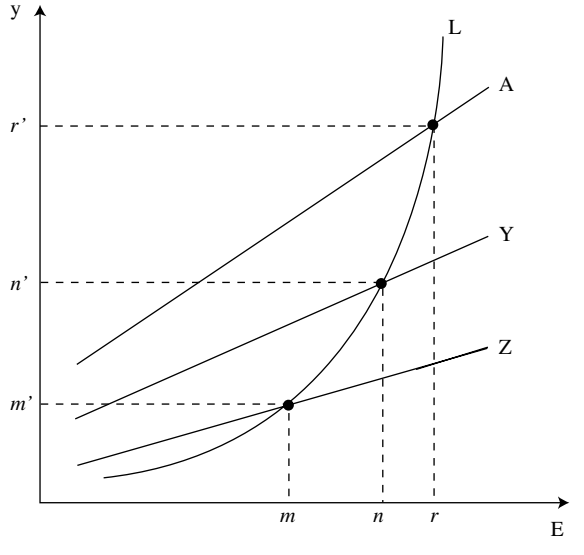
Sigma theory assumes that the social structure stems from the initial inequality in the distribution of economic and political assets. This inequality, the variable δ (delta), underlies the X term in equation (3). So income differences between social groups depend upon differences in their average education levels and asset endowments. As long as the values of the exogenous variables E and δ remain constant, average income in each social group will remain constant, and so will the income distribution across social groups.

Figure 1 depicts the static model. Points *m*, *n* and *r* are the average values of years of schooling in the three social groups A, Y and Z, and points *m'*, *n'* and *r'* are the corresponding average incomes. The initial inequality in the distribution of assets gives rise to a situation in which the capitalist class has the most years of schooling, human capital and income, followed by the other two social groups; moreover, additional years of schooling in group Z will result in a higher income level along curve Z. The same process applies to group Y and group A. The equilibrium situation is represented by curve L, which is empirically observable.

Under the equilibrium situation shown in figure 1, three cases can be considered to analyse the effect of exogenous changes in education on income inequality. Firstly, an additional year of education in group Z alone will increase the mean income of that group and reduce the initial inequality; but the effect will be small. Secondly, an additional year of schooling in all social groups will have a small effect on inequality; certainly, average income will rise in all groups, but it is not clear that relative incomes would change substantially. Thirdly, additional years of schooling in all three groups, with more years in group Z, will have an ambiguous effect or (at best) a small one.

FIGURE 1

Hypotheses regarding the relationship between education (*E*), mean income (*y*) and ethnic group (A, Y, Z)



Source: prepared by the author.

In short, the static model predicts that a reduction of inequality in years of schooling will not imply a significant reduction in income inequality; moreover, the same number of schooling years does not generate the same average income for all social groups. The reason is that the difference in years of schooling is only one of the factors generating income differentials between social groups.

The conclusion would be different if the relationship between years of schooling and incomes were not separable; that is, if just one curve existed for all groups. If the three curves were reduced to curve A in figure 1, an additional year of schooling would raise income by the same amount in all three social groups. Then, if differences in years of schooling were reduced, income inequality would decrease as well; moreover, the same number of years of schooling would generate the same average income for all social groups. But according to the sigma model, this is not how the capitalist system operates. Sigma society has three separate curves (class and race relationships), while epsilon society would only have two separate curves (class relationships).

The relationships presented in figure 1 refer to the average values of the variables for each social group. Successful individual cases may occur (people jumping from curve Z to Y or even to A), but the theory predicts that such cases will be the exception rather than the rule.

Changes in the exogenous variables will change the average incomes of social groups in specific ways, as indicated by the sign of their effects in equation (3). The model's empirical predictions are thus represented in equation (3), which is empirically refutable and can be used to test the model's validity statistically. The L curve in figure 1 relates the average values of years of schooling and income for each social group. The static equilibrium situation shown in figure 1 is thus represented by the L curve.

The sigma model departs from the standard analysis in several ways. Standard empirical studies usually measure the relationship between education and income along line L. Under the sigma model, however, the L curve cannot be used to compare the effect of exogenous changes in years of schooling on incomes, or rates of return on education (as is done incorrectly in the international literature), because the effect of education on income operates along different paths (lines A, Y, Z) for different social groups. Measuring the relationship along the L curve would certainly overestimate the effect of education on income.

It should also be noted that this model refers to group analysis, in which the aim is to explain differences in income and education across social groups. By contrast, in the standard literature the empirical relationship between income and education is studied using Mincerian earnings regressions (Mincer, 1974), which refer to individual analysis and thus seek to study differences in income and education between individuals (including regression variables such as experience, age, gender and other individual characteristics).

V
A dynamic model

Assume now that the quantity of education is determined endogenously in society and in each social group in the long run. Assume also the initial conditions of the economic process are given; in other words, individual endowments of economic and political assets are unequal at the outset, and these determine the social structure A, Y and Z. Likewise, ignore the short-run effects of the international terms of trade in this long-run model.

Dynamic equilibrium will now be defined as a sequence of static equilibrium situations over time. Therefore, the static equilibrium situation will now be considered as the solution of the initial period or generation, which is given by the L curve in figure 1. This curve shows the initial values of education and income for each social group, implying mean education and income values for society in the aggregate.

The dynamic model assumes that part of the total output produced in this period or generation will be allocated to human capital investment in the form of years of schooling. This investment is financed by families and by the State through fiscal policies, depending on the social group. As a result, there will be more education and human capital for the second period, because the average income in society increases. For simplicity, the model assumes that the growth of average income depends essentially upon the growth of human capital and hence also on the expansion of education. The other factors usually considered as affecting economic growth, such as physical capital accumulation and technological change, are assumed to be induced by higher levels of education.

Now consider the dynamic process by social group. In the next period or generation, education levels and average income will be higher in each social group (and on aggregate); there will be new investment, resulting in even higher education levels and average incomes for each social group (and for society as a whole) in the subsequent period, and so on. This dynamic process can be modelled through the L curve in figure 1. Given the initial equilibrium, the investment in education will imply an outward shift of the curve, say, to L1, which will in turn imply a new equilibrium with higher income for each social group. Further investment in education will take place in this period and the curve will shift outward again, to L2 say, and so on. In sum, the L curve will shift outward endogenously over time.

In the process of accumulating human capital through education, each social group will move along its specific path, given by lines Z, Y and A in figure 1. It was shown above that the process of education expansion along their corresponding paths does not generate significant reductions in income inequality between social groups. Differences in years of schooling can be reduced, but not income differences. Given the initial inequality in asset endowments, there are two reasons why incomes do not converge: (i) in the transformation of education into human capital, more equality in years of schooling does not imply more equality in human capital, since social groups accumulate human capital along different paths; (ii) in the transformation of human capital into incomes, the latter depend on access to basic markets, which is differentiated by social group, and this relationship is not altered by human capital accumulation alone. This means that education is transformed into income along different paths, depending on the social group. The relative incomes of the three social groups over time are path-dependent. In other words, the initial conditions of society matter; social history counts.

Initial inequality in economic and political asset endowments is thus the key to explaining the non-equalizing nature of the education system. Differences in years of schooling may eventually be eliminated, because there is a ceiling to years of education; but

this does not significantly reduce income inequality, for social groups move along different paths over time. Educational mobility is possible, but socio-economic mobility is much more problematic.

In intergenerational terms, the dynamic sigma model predicts that the "children" of a given social group will tend to inherit the economic status of their "parents". Long ago, the British biologist Francis Galton established the "law of regression towards the mean" (equalizing tendency) in relation to height differences between children and their biological parents. The sigma model presented here predicts that there will be "no regression towards the mean" in incomes between "children" and "parents" in the same social group.

The dynamic sigma model can explain the paradox as follows: developing countries are socially heterogeneous and hierarchical societies, i.e., they resemble sigma society. Inequality in the initial distribution of economic and political assets makes the society operate with exclusions; in particular, the processes of transforming education into human capital and this into incomes are differentiated by class and ethnicity relationships. In developed countries, the initial inequality is such that these processes are differentiated by class relations alone; but in the developing world, ethnicity also plays a significant role in the reproduction of inequality. As long as the initial inequality in asset endowments remains unchanged, the education system will not be income-equalizing.

VI

Empirical hypotheses tested against Peruvian data

A set of empirical predictions about the relationships between education and income can be derived from the static sigma model. These predictions will then be tested against data from Peru, a multi-ethnic and multicultural society, to see whether they confirm or refute the model. The data set comes from the National Household Survey (ENAHO) of 2003.

Empirical estimation of the size of ethnic groups is a complex task. In the case of Peru, the Z population corresponds to the descendants of indigenous populations. Up to four criteria were considered to measure the size of such groups, all of which have disadvantages:

— The mother tongue criterion: this criterion reduces ethnicity to the language marker, thereby underestimating its size. Indigenous people who do not speak an aboriginal language —either because they belong to a group that does not have a legacy of an aboriginal language or belong to a generation that has lost the aboriginal language— would formally not be counted as indigenous.
— The self-identification criterion: this measure understates the size of indigenous populations because individuals tend to conceal their ethnicity in a hierarchical society.
— The rural residence criterion: this also underestimates the size of the indigenous population because people migrating to the cities would not be counted as indigenous.
— The rural district birthplace criterion: this criterion again understates the size of the indigenous population because children born in a large city (following migration) would not be counted as indigenous.

This study chose the last of these to estimate the size of the indigenous population, since it seems to be the most reliable. Moreover, the analysis will refer to the adult population only, aged 25 or over; so people born in the rural districts of Peru in the three natural regions (the coast, the Andes and the Amazon), irrespective of where they currently live, are considered mostly indigenous. This represents the Z social group of the theoretical model.

People born in the most affluent residential districts of Lima, regardless of where they now reside, are considered to belong mostly to the A social group. The mestizo or Y group was calculated by the difference between the two. The estimates are as follows: social group A accounts for 3.5% of the total adult population, group Y for 27.5% and group Z for 69%.[3]

[3] The 2001 Household National Survey (ENAHO) contained a question on ethnic self-identification. The result was that

H1. Hypothesis of quantitative differences in years of schooling between social groups

As shown in figure 1, the static sigma model predicts that the average number of years of schooling varies by social group, increasing in the order Z, Y, A. This relationship is shown in table 2. Mean years of schooling vary between social groups: 14 years for group A, 11 years for group Y and 7 for group Z. The medians show the same ranking: 14, 11 and 5 years. Parametric and non-parametric statistical tests (see Appendix) show that the observed differences in mean values are statistically significant. The empirical data do not refute the hypothesis, but reveal pronounced inequalities in education.

Using education levels (basic, secondary, technical and university) instead of years of schooling also reveals major inequalities. Nearly 70% of group A have attained post-secondary education, compared to 36% of group Y and just 15% of group Z; the national average is 21%. The exclusion of group Z from post-secondary education is remarkable.

Few studies have been carried out on the conversion of education into human capital, by social group. Those that do exist have compared test scores, for a given grade, between two categories of schools: private-public and urban-rural. The results show that the average level of knowledge, measured by test scores, is higher among students attending private schools than those in public schools, and also higher in urban schools than rural ones (Rivera, 1979; Cueto, Jacoby and Pollit, 1997; Peru, Ministry of Education, 2005).

There is also some evidence of the negative role of malnutrition in the learning process at school. A study carried out as part of the Joint Programme on Latin American Economic Integration (ECIEL) shows a large and statistically significant negative correlation between school performance and degree of malnutrition, based on a sample of the student population in Lima and in urban and rural Puno (Rivera, 1979).

As social group Z is predominant in rural areas, it follows that the transformation of education into human capital in Peru seems to operate as shown in the structural equation (1). These results are consistent with the hypothesis. After 180 years as a republic,

among family heads, 43% declared themselves to be indigenous descendants, while 34% claimed to speak an aboriginal language (Hall and Patrinos, 2006, tables 2.1, 2.7 and 7.1). These estimates of the proportions of indigenous people are clearly smaller than the 69% calculated in this study, as would be expected from the definitions used. Given the order of magnitude of these estimates, the claim made by many Peruvian writers that the predominant ethnic group is Y ("we are a mestizo country") is not empirically supported.

TABLE 2

Peru: education level by social group, 2003
(Thousands of people and percentages)

Education level	Social group						Total	
	Z		Y		A			
	N[a]	(percentage)	N	(percentage)	N	(percentage)	N	(percentage)
None	1 283	14.1	110[b]	3.3	0.7[b]	0.2	1 394	10.9
Basic	3 903	43	556	16.5	10.1[b]	2.8	4 470	34.9
Secondary	2 611	28.7	1 475	43.8	94.9	26.9	4 180	32.6
Technical	725[b]	8	530	15.7	73.4	20.8	1 328	10.4
University	562[b]	6.2	697	20.7	174.4	49.3	1 433	11.2
Total N	9 083	100	3 368	100	353.4	100	12 805	100
Horizontal percentage	70.9		26.3		2.8		100	
Years of education								
Mean	6.9		10.8		13.7		8.2	
Median	5		11		14		9	

Source: prepared by the author on the basis of the 2003 National Household Survey (ENAHO).

[a] Population aged 25 or over (sample expansion).
[b] Small number of observations, 10% or less of the social group.

and despite the expansion of the education system, indigenous populations in Peru still suffer a significant degree of exclusion from the education system, both quantitatively and qualitatively.

H2. Hypothesis of "separability" and hierarchy in the relationship between education and income

The static sigma model predicts a positive relationship between mean years of education and mean incomes, but this relationship is separable and hierarchical by social group in ascending order Z, Y, A. This is precisely what the reduced form of the model claims, as shown in equation (3) and figure 1. Using the reduced-form equation to test the sigma model therefore has epistemological justification; there is no need to test it via the structural equations. In other words, if equation (3) were false, the assumptions contained in equations (1) and (2) could not be true.

Table 3 presents data on incomes by education level and social group. In this case, population refers to the economically active population (EAP) who are employed (the small unemployed group is not considered). The observed data suggest the existence of a direct relationship between mean income and education level both within social groups and between them. The question is whether the observed relationships are statistically separable for each social group and whether they display an A,Y, Z hierarchy as the hypothesis proposes.

Given the existence of quantitative educational exclusion, corroborating hypothesis *H1*, some cells in table 3 contain very small numbers. The test must then be carried out only between comparable levels of education. The members of group Z are concentrated in the initial education levels; only a small fraction, less than 10%, reach post-secondary levels. In group Y, there is a small illiterate population. In group A, no one was illiterate or had primary education only. The comparable or relevant levels of education and social groups are thus the following: at primary level, groups Y and Z; at secondary level, all three groups A, Y and Z; and at the technical and university levels, groups A and Y. The results of the statistical tests contained in the Appendix show that the observed differences are generally significant. In short, the empirical data do not refute hypothesis *H2*.

TABLE 3

Peru: mean income by education level and social group, 2003
(Thousands of people, soles per month and percentages)

Education level	Social group						Total		
	Z		Y		A				
	N[a] (percentage)	y[b]	N (percentage)	y	N (percentage)	y	N	(percentage)	y
None	9.6[c]	171	1.9[c]	297	–	–	604	7.1	180
Basic	41	330	13.1	479	1.0[c]	137	2 704	32	346
Secondary	32.7	592	44	710	22.7	798	2 998	35.4	637
Technical	9.6[c]	844	17.2	912	15.8	965	1 001	11.9	876
University	7.1[c]	1 529	23.8	2 005	60.5	2 721	1 148	13.6	1 943
Total	*100*	*535*	*100*	*1 015*	*100*	*1 981*	*8 464*	*100*	*717*
Total N	5 841		2 330		293		8 464		
Horizontal percentage	69		27.5		3.5		100		
Years of education									
Mean	7.6		11.4		14.2		9		
Median	8		11		15		11		

Source: prepared by the author on the basis of the 2003 National Household Survey (ENAHO).

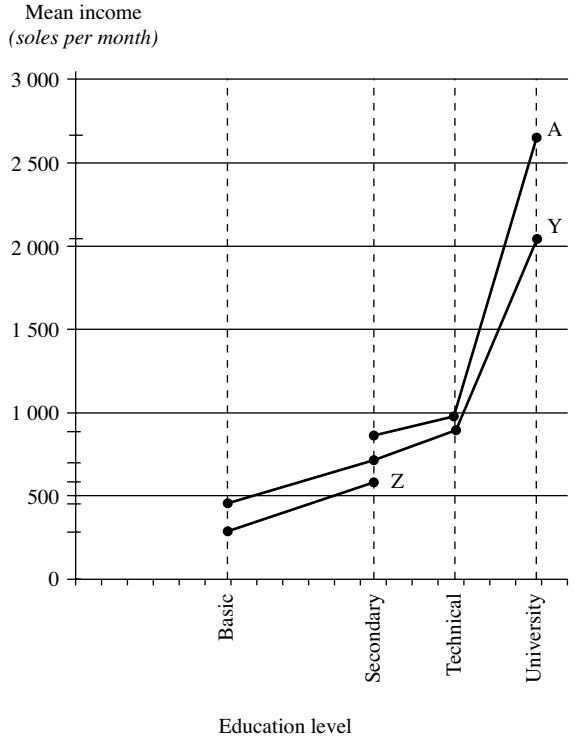
[a] Population aged 25 or over (sample expansion).
[b] Mean income in soles, Lima, October 2003.
[c] Small number of observations, 10% or less of the social group.
– No population exists in this category.

Figure 2 illustrates the nature of the empirical relationship found between income and education, showing the relevant levels for comparison. The gradients of the lines shown suggest a rising trend in the three social groups. Moreover, the slope of curve A rises faster than that of Y, and that of Y more rapidly than that of curve Z.

The standard regression method could be applied to these data to test the relationships between the slopes econometrically, using social groups A, Y and Z as dummy variables. However, the data reveal secondary education as the only relevant level for comparison. As is well known, an implicit assumption of the standard regression model is that dummy variables must cover the entire range of observations of the independent variables; if exclusion exists in the distribution of the regressors, as in the case shown in figure 2, the standard regression model is not applicable. In fact, standard econometric theory implicitly assumes the absence of exclusions in the relationships between variables.

FIGURE 2

Peru: Empirical relationship between mean income (y) and education level, by ethnic group



Source: prepared by the author on the basis of table 2 data.

H3. Hypothesis of "separability" and hierarchy in the relationships between education and wages

If wages are substituted for income on the vertical axis, figure 1 can also represent the relationships between education and wages. A number of assumptions will be introduced here. First, labour productivity depends upon the level of workers' human capital, not on their education level. Second, profit-maximizing firms operating in competitive labour markets will hire workers according to their labour productivity and will pay uniform wages for a given level of human capital.

The static labour-market equilibrium situation can be written as follows:

$$k_h = f(E, X), f_1 > 0, f_2 > 0 \quad (4)$$

$$w = h(K_h, X, p), h_1 > 0, h_2 > 0, h_3 > 0 \quad (5)$$

Hence,

$$w = h(f(E, X), X, p) = H(E, X, p), \quad H_1 > 0, H_2 > 0, H_3 > 0 \quad (6)$$

Equation (4) is the old equation (1) showing the transformation of education into human capital. Equation (5) is also an old friend, now showing the transformation of human capital into wages (w). Equation (6) is the reduced-form equation, now showing that wages ultimately depend upon education and social group, for a given value of p.

Tables 4A and 4B present data on education and pay in blue-collar and white-collar labour markets, respectively. The mean white-collar salary is twice the mean blue-collar wage. The difference in education goes in the same direction: mean years of schooling is 13.7 years for white-collar workers and 8.5 for blue-collar workers, with median values of 14 and 10 respectively. Education levels appear as a factor differentiating the two social groups. The proportion of Z workers engaged in the blue-collar labour market is 20%, but the figure for the white-collar market is only 14%. Within group Y, the equivalent proportions are 21% and 36%, while in group A they are 8% and 62%.

The relevant education levels for comparison among blue-collar workers (see table 4A) include basic and secondary, while the workers concerned are those from the Y and Z groups. In each case, the

TABLE 4A

Peru: mean wages (blue-collar workers), by education level and social group, 2003
(Thousands of people, soles per month and percentages)

Education level	Social group						Total	
	Z		Y		A			
	N[a] (%)	MW[b]	N (%)	MW[b]	N (%)	MW[b]	N (%)	MW[b]
None	6[c]	306	1.8[c]	349	–	–	4.7	310
Basic	37.5	489	21.6	518	–	–	32.4	490
Secondary	46.7	616	59.4	651	78.2	569	50.8	627
Technical	8.1[c]	695	13.8	712	21.8	743	10	703
University	1.7[c]	763	3.4[c]	796	–	–	2.1	778
Total	*100*		*100*		*100*	*606*	*100*	*579*
Total N	1 163		482		23		1 668	
Horizontal percentage	69.7		28.9		1.4		100	
Percentage of total EAP (from table 3)	19.9		20.7		7.8		19.7	
Years of education								
Mean	8		9.7		11.3		8.6	
Median	9		11		11		10	

Source: prepared by the author on the basis of the 2003 National Household Survey (ENAHO).

[a] Population of blue-collar workers aged 25 or over (sample expansion).
[b] Mean wage in soles, Lima, October 2003.
[c] Small number of observations, 10% or less of the social group.
EAP Economically active population.
– No population exists in this category.

TABLE 4B

Peru: mean salaries (white-collar workers), by education level and social group, 2003
(Thousands of people, soles per month and percentages)

Education level	Social group						Total	
	Z		Y		A			
	N[a] (%)	MS[b]	N (%)	MS[b]	N (%)	MS[b]	N (%)	MS[b]
None	0.1[c]	357	0.5[c]	431	–	–	0.3	415
Basic	4.6[c]	703	1.8[c]	641	–	–	2.8	685
Secondary	29.2	871	28.1	707	14.2	763	27.2	788
Technical	32.9	1 045	26.2	1 009	17.7	1 102	28.4	1 033
University	33.2	1 780	43.4	2 203	68.1	3 182	41.3	2 212
Total	*100*	*1 222*	*100*	*1 433*	*100*	*2 469*	*100*	*1 442*
Total N	807		833		181		1 821	
Horizontal percentage	44.3		45.7		10		100	
Percentage of total EAP (from table 3)	13.8		35.7		61.8		21.5	
Years of education								
Mean	13.5		13.7		14.9		13.7	
Median	14		14		16		14	

Source: prepared by the author on the basis of the 2003 National Household Survey (ENAHO).

[a] Population of white-collar workers aged 25 or over (sample expansion).
[b] Mean salary in soles, Lima, October 2003.
[c] Small number of observations, 10% or less of the social group.
EAP Economically active population.
– No population exists in this category.

observed average wage of Y workers is higher than that of Z workers. The statistical test shows that these differences are statistically significant at the secondary education level only (which represents the largest group in this labour market), as shown in the Appendix. The relevant education levels for comparison among white-collar workers (see table 4B) include the secondary, technical and university levels and the Z, Y and A social groups. The statistical test shows that the differences are statistically significant in five of the nine cases, and at the university level the difference is consistently significant, as shown in the Appendix.

Overall, the empirical evidence presented here shows that wages and salaries depend positively on workers' education level. As white-collar workers tend to have more education than blue-collar workers, the mean salary of the former is higher than the mean wage of the latter. Lastly, for a given and relevant level of education, pay in general depends on the social group hierarchy, given by the order A, Y, Z. The same relationship is observed in the case of white-collar salaries. In brief, the empirical data presented do not refute hypothesis *H3*.

H4. Hypothesis of the gap necessarily existing between wage and own-account income

How is labour discipline secured at the firm level in a sigma society? As stated in section II, sigma theory assumes that firms face conflict in their labour relations, so employers must find ways to extract effort from workers. Accordingly, they will seek to pay a market wage rate above the opportunity cost to the workers hired, so workers found shirking will be fired and will then suffer an economic cost, represented by this wage premium. The opportunity cost for hired workers will be equal to the income they can make in own-account (self-employed) activities. This differential is the labour discipline device used by capitalist enterprises to secure the highest level of labour productivity in overpopulated societies. For a given level of human capital, the sigma model thus predicts a gap between the price paid for labour services in the labour market (wages or salaries) and own-account incomes.

Suppose the equilibrium situation implies a 30% gap between wages and own-account income. As the average wage of social group Y is higher than the average wage of social group Z (shown in *H3* above), the own-account incomes of these social groups must follow the same ranking. The efficiency gap applies to each social group separately. This hypothesis differs from the standard Lewis model (Lewis, 1954), which ignores the ethnicity factor so that the efficiency gap is uniform for all workers.

Table 5 presents data on the mean incomes of wage earners, salary earners and own-account workers, separated by social group. In social group Z, aggregate own-account income is lower than salaries and wages.

TABLE 5

Peru: mean wage, mean salary and mean own-account income, by social group, 2003
(Soles per month)

Social group	Education level					Total
	None	Basic	Secondary	Technical	University	
Group Z						
Salary	357[a]	703[a]	871[a]	1 045[a]	1 780[a]	1 222
Wage	306[a]	483	618	695[a]	763[a]	557
Own-account income	136	260	451	568[a]	818[a]	327
Group Y						
Salary	431[a]	641[a]	707	1 009	2 203	1 433
Wage	349[a]	518[a]	651	712[a]	796[a]	630
Own-account income	221[a]	384[a]	550	696	1 180	621
Group A						
Salary	–	–	763	1 102	3 182	2 469
Wage	–	–	568[a]	743[a]	–	608
Own-account income	–	137[a]	1 032	508[a]	1 135	1 013

Source: prepared by the author on the basis of the 2003 National Household Survey (ENAHO).

[a] Small number of observations, 10% or less of the social group.
– No population exists in this category.

This relationship also holds true for the relevant education levels. Table 5 shows similar relationships for social group Y, but in group A the relationships seem less clear than in the other social groups. At the relevant education level, which is university, the relationship is fairly clear, however. The observed differences are all statistically significant, as shown in the Appendix, so the empirical data do not refute the model.

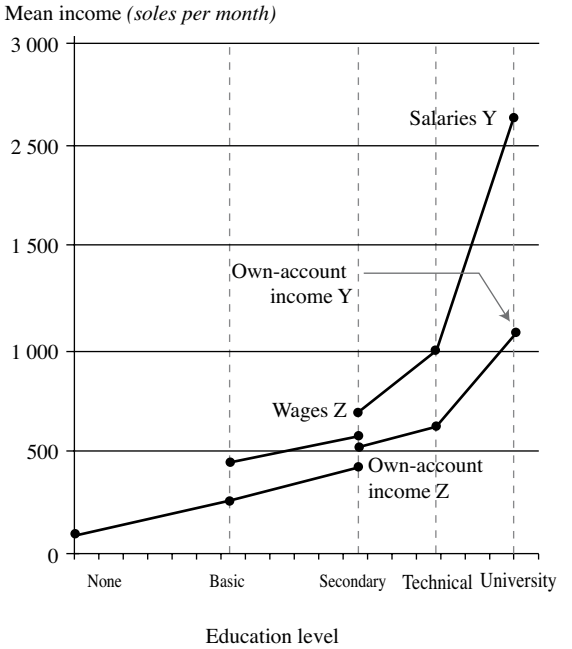
Figure 3 summarizes the results. In group Z, the average wage is higher than average own-account income at both the basic and secondary education levels, which are the relevant comparison categories. In group Y, the average salary is higher than average own-account income at the three relevant education levels: secondary, technical and university.

To summarize, Peru's labour markets seem to operate as predicted by the sigma model, so that the self-employed are mostly underemployed and hence form part of the excess labour supply: they would prefer employment in the labour market at current wage or salary rates.

Table 6 shows data on excess labour supply (unemployment and underemployment) by social group. In terms of differences between social groups, the results indicate very high rates of excess labour supply: 66% in social group Z, 45% in social group Y and 32% in social group A. The Appendix shows that the observed differences in the rate of excess supply by education level and social group, in the relevant categories, are statistically significant. As unemployment rates do not differ greatly, this difference mainly reflects rates of underemployment.

FIGURE 3

Peru: empirical relationship between wages or salaries and own-account income, by ethnic group



Source: prepared by the author on the basis of table 4 data.

H5. Hypothesis of high education mobility but low income mobility

The dynamic sigma model predicts that mean years of education tend to equalize between social groups over time, but mean incomes do not. This prediction can be tested statistically using age brackets as indicators of intergenerational differences.

Tables 7A and 7B present ratios of education and incomes for intergenerational groups of workers and social groups, respectively. Table 7A shows a large

TABLE 6

Peru: rate of excess labour supply (unemployment and underemployment), by education level and social group, 2003
(Percentages)

Education level	Social group			Total
	Z	Y	A	
None	88.1 *(11.7)*[b]	73.7[a] *(12.3)*	–	87.1 *(11.7)*
Basic	79.7 *(6.5)*	61.8 *(7.2)*	100.0[a] *0.0*	77.7 *(6.6)*
Secondary	57.6 *(5.4)*	49.8 *(8.2)*	39.3 *(11.4)*	54.4 *(6.5)*
Technical	35.0[a] *(5.6)*	31.1 *(7.3)*	25.7 *(12.4)*	32.9 *(6.6)*
University	29.3[a] *(7.0)*	32.7 *(8.7)*	28.8 *(2.7)*	31.0 *(7.2)*
Total	65.6 *(6.6)*	44.5 *(8.1)*	31.5 *(6.4)*	58.4 *(7.0)*

Source: prepared by the author on the basis of the 2003 National Household Survey (ENAHO).

[a] Small number of observations, 10% or less of the social group.
[b] Unemployment rate in parenthesis; underemployment is therefore the difference between the total rate of excess labour supply (shown in the table) and the unemployment rate (in parenthesis).
– No population exists in this category.

TABLE 7A

Peru: mean years of schooling by age bracket and social group, 2003

Social group	Age bracket (years)								Total	
	25-34		35-44		45-54		55-65			
	N[a]	E[b]	N	E	N	E	N	E	N	E
Z	1 551	11	1 636	8	1 260	7	850	5	5 297	8
Y	912	12	670	12	449	11	218	9	2 249	12
A	146	14	111	14	29	14	6	12	291	14
Total	*2 609*	*12*	*2 417*	*10*	*1 738*	*8*	*1 074*	*6*	*7 837*	*9*
Ratio Z/A	0.8		0.6		0.5		0.4		0.6	
Ratio Y/A	0.9		0.8		0.8		0.7		1.2	

Source: prepared by the author on the basis of the 2003 National Household Survey (ENAHO).

[a] Employed economically active population (thousands of people).
[b] Mean years of schooling.

TABLE 7B

Peru: mean income by age bracket and social group, 2003

Social group	Age bracket (years)								Total	
	25-34		35-44		45-54		55-65			
	N[a]	y[b]	N	y	N	y	N	y	N	y
Z	1 551	530	1 636	596	1 260	623	850	473	5 297	563
Y	912	771	670	987	449	1 195	218	1 626	2 249	1 003
A	146	1 456	111	2 455	29	2 982	6	1 648	291	1 991
Total	*2 609*	*666*	*2 417*	*790*	*1 738*	*810*	*1 074*	*713*	*7 837*	*743*
Ratio Z/A	0.4		0.2		0.2		0.3		0.3	
Ratio Y/A	0.5		0.4		0.4		1.0		0.5	

Source: prepared by the author on the basis of the 2003 National Household Survey (ENAHO).

[a] Employed economically active population (thousands of people).
[b] Mean income (soles per month).

increase in years of schooling between generations in social group Z: the "children" (in the social, not the biological sense) have twice as many years of schooling as the "parents", although the increase generated in other social groups is smaller. In social group A, education appears to come up against a ceiling at around 14 years. As a result, the educational gaps between group Z and the other groups have narrowed considerably.

The estimated changes in income gaps across generations are shown in table 7B. The mean income of social group Z in relation to social group A has not changed much: the relative income of the "parents" is 30% and that of the "children" is 40%, which is small compared to the changes in educational attainment shown in table 7A. The situation of social group Y relative to group A shows a similar pattern. The relative income of the "parents" and "children" is almost constant at around 40% and 50%, except for the eldest group, in which the figure is 100%. The latter result is certainly a paradox, possibly due to the small size of this group.

The conclusion regarding the convergence trends between social groups of different generations is that, while there is a tendency for differences in years of schooling to narrow, there is no similar trend towards lower income inequality. This empirical result is consistent with the hypothesis.

Regarding the degree of income inequality in Peru, the Gini coefficient estimated from the ENAHO 2003 dataset is 0.59, which is close to other estimates made in the past (0.60), based on national accounts. This figure must underestimate the true Gini coefficient because the incomes of Peru's economic élite are not included in the sample. It is well known that economic élites are absent or underrepresented in household sample surveys. Strictly speaking, the empirical social group A mostly consists of the middle class (executives, administrators and top professionals) and does not correspond to the capitalist class of the model. Theoretically, in the dual social structure of capitalists and workers, the middle class is usually included in the capitalist category (Wright, 1997). This is also the criterion followed in this study.

Peru still displays a relatively high degree of inequality and remains one of the most unequal countries in the world. As the sigma model predicts, education does not seem to be an income-equalizing system. Even if people have the same number of schooling years, average incomes will still depend on the social group to which they belong.

VII

Conclusions

The sigma model constructed in this paper predicts a set of empirically falsifiable hypotheses for the relationships between education and income. When compared against the Peruvian data, the statistical tests show that the hypotheses are not rejected. The sigma model also predicts the paradox observed in developing countries, in which a significant increase in the number of years of education has not been accompanied by less income inequality.

The conclusion that education is not an income-equalizing system is explained by two factors: the initial inequality in the distribution of economic and political assets between social groups (high degree of inequality) and society's factor endowments (overpopulation). Initial inequality and overpopulation are the ultimate factors explaining the paradox; in other words, they are the exogenous variables of the sigma model. Thus, as long as these exogenous variables remain unchanged, the education system will remain non-equalizing.

(Original: English)

APPENDIX

Summary of statistical tests

Hypothesis 1: Differences in mean years of schooling

Social groups	Variance equality test		Difference of means test				
	T-statistics	*p* value	Difference	Confidence interval		F	*p* value
Z - Y	354.56	0.00	-3.96	[-4.10	-3.81]	2 024.4	0
Z - A	150.21	0.00	-6.87	[-7.20	-6.55]	1 620.48	0
Y - A	53.21	0.00	-2.92	[-3.26	-2.58]	258.3	0

Social groups	Mann-Whitney U-test			
	Sum of ranks		Z-statistics	*p* value
	Z	Y		
Z – Y	490 300 000	218 000 000	-61.49	0.00
Z – A	420 400 000	10 058 291	-24.62	0.00
Y – A	38 984 586	2 479 585	-12.57	0.00

Source: prepared by the author.

Statistical testing relates to the differences in mean years of schooling between social groups. The first test uses F-statistics and the second the (non-parametric) Mann-Whitney U-test. The results show that the observed mean differences are statistically significant.

Hypothesis 2: Mean income differences

Level of education	Social groups	Variance equality test		Difference of means test				
		T-statistics	*p* value	Difference	Confidence interval		F-statistics	*p* value
Basic	Z - Y	4.24	0.04	-126.41	[-164.86	-87.97]	33.04	0.00
Secondary	Z - Y	0.63	0.43	-102.02	[-162.35	-41.68]	6.35	0.01
	Z - A	0.15	0.70	-217.56	[-513.60	78.47]	2.11	0.15
	Y - A	0.07	0.80	-115.54	[-416.41	185.32]	0.35	0.55
Technical	Y - A	0.96	0.33	-41.85	[-307.86	224.17]	0.15	0.69
University	Y - A	21.48	0.00	-938.12	[-1677.40	-198.83]	3.30	0.07

Level of education	Social groups	Mann-Whitney U-test			
		Sum of ranks		Z-statistics	*p* value
		Z	Y		
Basic	Z - Y	34 224 589	4 808 441	-9.83	0.00
Secondary	Z - Y	23 098 507	10 714 470	-7.61	0.00
	Z - A	17 017 958	289 828	-2.76	0.01
	Y - A	3 033 103	113 184	-1.20	0.23
Technical	Y - A	553 451	31 370	-0.72	0.47
University	Y - A	785 177	103 934	-0.66	0.00

Source: prepared by the author.

The test now relates to mean income differences between social groups. The parametric and non-parametric tests show that the observed differences are statistically significant at the basic and university levels of education. For the other levels of education, the parametric test does not pass the test of variance equality; hence, parametric tests are the relevant ones. These show that differences are not statistically significant at the technical level, whereas at the secondary level the differences are significant in all groups except Y-A. In four out of six cases, the hypothesis is not refuted by the facts.

Hypothesis 3: Differences in mean salaries and mean wages

White-collar workers: salaries

Level of education	Social groups	Variance equality test		Difference of means test				
		T-statistics	*p* value	Difference	Confidence interval		F-statistics	*p* value
Secondary	Z - Y	8.73	0.00	164.01	[93.99	234.03]	21.08	0.00
	Z - A	1.71	0.19	107.76	[-143.00	358.52]	0.71	0.40
	Y - A	0.39	0.53	-56.25	[-307.71	195.21]	0.19	0.66
Technical	Z - Y	4.75	0.03	35.85	[-67.13	138.83]	0.47	0.50
	Z - A	0.70	0.40	-56.94	[-378.31	264.43]	0.12	0.73
	Y - A	0.00	0.97	-92.79	[-421. 73	236.15]	0.31	0.58
University	Z - Y	21.64	0.00	-422.79	[-1 011.39	165.82]	1.98	0.16
	Z - A	60.13	0.00	-1 402.16	[-2 212.04	-592.28]	11.52	0.00
	Y - A	16.44	0.00	-979.37	[-1 822.26	-136.48]	5.19	0.02

Level of education	Social groups	Mann-Whitney U-test			
		Sum of ranks		Z-statistics	*p* value
		Z	Y		
Secondary	Z - Y	476 879	271 598	4.27	0.00
	Z - A	284 144	8 852	1.63	0.10
	Y - A	125 362	6 994	0.28	0.78
Technical	Z - Y	667 938	391 302	3.39	0.00
	Z - A	405 470	13 601	1.30	0.19
	Y - A	175 132	10 004	0.35	0.73
University	Z - Y	646 406	660 130	-1.85	0.06
	Z - A	363 848	56 139	-5.15	0.00
	Y - A	344 454	52 041	-4.10	0.00

Source: prepared by the author.

Given the relative size of the social groups (shown in the first hypothesis on exclusion), the relevant comparison is between groups Z-Y-A at the secondary, technical and university levels of education. The contrast for the university level can be made by parametric and non-parametric testing. Owing to the failure in the variance equality test, the other two can be done using non-parametric testing alone. The results show that the observed differences at the university level are consistently statistically significant; in general, the observed values are statistically significant in five of the nine cases.

Blue-collar workers: wages

Level of education	Social groups	Variance equality test		Difference of means test				
		T-statistics	*p* value	Difference	Confidence interval		F-statistics	*p* value
Basic	Z - Y	0.38	0.54	-35.08	[-115.23	45.07]	0.74	0.39
Secondary	Z - Y	4.01	0.05	-34.34	[-91.07	22.40]	1.41	0.24

Level of education	Social groups	Mann-Whitney U-test			
		Sum of ranks		Z-statistics	*p* value
		Z	Y		
Basic	Z - Y	1 139 292	212 899	-0.93	0.35
Secondary	Z - Y	1 653 321	774 385	-2.71	0.01

Source: prepared by the author.

The relevant comparison is between groups Z-Y at the basic and secondary levels of education. The parametric test cannot be used because the variance equality test fails. The non-parametric test shows that the observed differences are statistically significant for the secondary education level (the vast majority in the group), but not for the basic level.

Hypothesis 4: Mean differences between wage and salary incomes and own-account incomes

Social group Z

Level of education	Variance equality test		Difference of means test				
	T-statistics	*p* value	Difference	Confidence interval		F-statistics	*p* value
Basic	72.56	0.00	443.13	[287.90	598.36]	31.31	0.00
Secondary	24.64	0.00	420.17	[365.23	475.12]	224.64	0.00
Total	*495.77*	*0.00*	*894.33*	*[757.69*	*1,030. 97]*	*164.58*	*0.00*

Level of education	Mann-Whitney U-test			
	Sum of ranks		Z-statistics	*p* value
	Own-account	Employed		
Basic	16 417 007	603 688	-11.93	0.00
Secondary	4 943 900	2 010 686	-24.30	0.00

Social group Y

Level of education	Variance equality test		Difference of means test				
	T-statistics	*p* value	Difference	Confidence interval		F-statistics	*p* value
Secondary	0.42	0.52	157.72	[80.83	234.61]	16.17	0.00
Technical	0.21	0.65	313.21	[70.70	555.73]	6.41	0.01
University	20.77	0.00	1 022.85	[502.82	1 542.88]	14.86	0.00
Total	140.96	0.00	812.48	[586.29	1 038.68]	49.57	0.00

Level of education	Mann-Whitney U-test			
	Sum of ranks		Z-statistics	*p* value
	Own-account	Employed		
Secondary	772 113	492 733	-12.58	0.00
Technical	64 839	285 864	-13.97	0.00
University	84 954	479 825	-12.06	0.00

In the case of Z workers, the relevant comparisons include the one between the basic and secondary levels of education. Both the parametric and the non-parametric tests show that the observed differences are statistically significant. In the case of Y workers, the relevant comparisons include the secondary, technical and university levels of education. The non-parametric test shows that the observed differences are statistically significant in all cases. The parametric test is applicable to the university level only, where it also shows that the difference is significant.

Hypothesis 5: Mean differences in labour market exclusion rates

Level of education	Social groups	Chi-squared	*p* value
Basic	Z - Y	6 855	0.00
Secondary	Z - Y - A	4 855	0.00
Technical	Z - Y	1 034	0.00
University	Z - Y	819	0.00

Source: prepared by the author.

The chi-squared test is used to test differences in proportions of exclusion from the labour market between social groups. The results show that the observed differences are statistically significant for all levels of education and the relevant social groups at each education level.

Bibliography

Akerlof, G. and R. Kranton (2000), "Economics and identity", *The Quarterly Journal of Economics*, vol. 115, No. 3, Cambridge, Massachusetts, The MIT Press.

Becker, G. and N. Tomes (1979), "An equilibrium theory of the distribution of income and intergenerational mobility", *Journal of Political Economy*, vol. 87, No. 6, Chicago, University of Chicago Press.

Birdsall, N., A. de la Torre and R. Menezes (2008), *Fair Growth: Economic Policies for Latin America's Poor and Middle-Income Majority*, Washington, D.C., Brookings Institution Press.

Blom, A. and C. Vélez (2004), "The dynamics of the skill-premium in Brazil: growing demand and insufficient supply?", *Inequality and Economic Development in Brazil*, Washington, D.C., World Bank.

Bourguignon, F., F. Ferreira and M. Menendez (2007), "Inequality of opportunity in Brazil", *Review of Income and Wealth*, vol. 53, No. 4, Oxford, Blackwell Publishing.

Cueto, S., E. Jacoby and E. Pollit (1997), "Rendimiento escolar de niños y niñas en zonas rurales y urbanas del Perú", *Revista de psicología*, vol. 15, No. 1, Lima, Catholic University of Peru.

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

Durlauf, S. (1996), "A theory of persistent income inequality", *Journal of Economic Growth*, vol. 1, No. 1, New York, Springer.

Gardner, H. (1999), *Intelligence Reframed: Multiple Intelligences for the 21st Century*, New York, Basic Books.

Hall, G. and A. Patrinos (2006), *Indigenous Peoples, Poverty, and Human Development in Latin America*, London, Palgrave.

Hudson, R.A. (1996), *Sociolinguistics*, Cambridge, Cambridge University Press.

Lewis, A. (1954), "Economic development with unlimited supplies of labor", *The Manchester School of Economic and Social Studies*, vol. 22, Oxford, Blackwell Publishing.

Li, H., L. Squire and H. Zou (1998), "Explaining international and intertemporal variations in income inequality", *The Economic Journal*, vol. 108, No. 1, Oxford, Blackwell Publishing.

Marshall, T.H. and T. Bottomore (1992), *Citizenship and Social Class*, London, Pluto Press.

Mincer, J. (1974), *Schooling, Experience, and Earnings*, Cambridge, Massachusetts, National Bureau of Economic Research.

Ñopo, H., J. Saavedra and M. Torero (2004), "Ethnicity and earnings in urban Peru", *IZA Discussion Papers*, No. 980, Bonn, Institute for the Study of Labor (IZA).

Peru, Ministry of Education (2005), *Evaluación nacional 2004*, Lima.

Ratey, J. (2002), *A User's Guide to the Brain*, New York, Vintage Books.

Rivera, I. (1979), "Los determinantes de la calidad de la educación en el Perú. Un Estudio ECIEL", *Documento de trabajo*, No. 44, Lima, Catholic University of Peru.

Rousseau, J. (1984), *A Discourse on Inequality*, translated from French by Maurice Cranston, New York, Penguin Books. First published in 1755.

Searle, J. (1995), *The Construction of Social Reality*, New York, The Free Press.

World Bank (2001), *World Development Report. Attacking Poverty*, Washington, D.C.

Wright, E. (1997), *Class Counts: Comparative Studies in Class Analysis*, Cambridge, Cambridge University Press.

KEYWORDS

Financial crisis
Fiscal policy
Macroeconomics
Foreign exchange rates
Interest rates
Economic stabilization
Brazil

Fiscal policy in times of crisis: macroeconomic effects of the primary surplus

Manoel Carlos de Castro Pires, Fábio Goto and Bruno Rocha

The international financial crisis that hit the Brazilian economy in the third quarter of 2008 called forth several economic-policy responses. In the case of fiscal policy, expansionary measures rendered the existing fiscal target unviable, and it was formally reduced in April 2009. The purpose of this article is to analyse the macroeconomic effects of reducing the primary budget-surplus target and to evaluate its repercussions on the level of activity, time-structure of interest rates and the exchange rate.

Manoel Carlos de Castro Pires
Technical Officer of the Institute of Applied Economic Research (IPEA) and Coordinator General of Fiscal Policy at the Economic Policy Secretariat (SPE), Ministry of Finance
manoel.pires@ipea.gov.br

Fábio Goto
Coordinator General of Model Building at the Economic Policy Secretariat (SPE), Ministry of Finance
fabio.goto@fazenda.gov.br

Bruno Rocha
Finance and Control Analyst at the Economic Policy Secretariat (SPE), Ministry of Finance
bruno.r.silva@fazenda.gov.br

I
Introduction

The intensification of the international financial crisis that began in the third quarter of 2008 generated renewed interest in the countercyclical role of fiscal policy among macroeconomic theorists. Interest in that function —which was previously very clear and founded on the perceived need for active aggregate-demand-management (Keynesian) policies— faded in the mid-1970s for two main reasons. The first, recognized by Tobin (1982), is that the Keynesian analytical models of the time were static and, consequently, underestimated the repercussions of fiscal policy on the sustainability of the public debt. The second reason is that increases in public debt today would need to be matched by primary surpluses in the future. In that case, economic agents would anticipate the future tax increases, and fiscal expansion would have no macroeconomic effects, since it would be perfectly offset by a reduction in private aggregate demand. This argument is known as "Ricardian equivalence" (Barro, 1979).

For many years, concern about debt sustainability and theoretical doubts about the effectiveness of fiscal policy relegated the countercyclical fiscal tool to secondary status behind other measures, including monetary policy.[1]

Despite attempts to revive a number of aspects of fiscal policy, such as the notion that positive exogenous factors can increase the return on private capital — including public investment in infrastructure and education expenditure— the focus of the debate has changed, and fiscal policy is now viewed in a much more structural context than as a countercyclical tool.

Nonetheless, those policy recommendations were rejected by policymakers in the global economy when reacting pragmatically to the events caused by the financial crisis; and highly expansionary and unconventional fiscal-policy measures were adopted, such as giving financial assistance to firms and purchasing "toxic" assets from banks.

Although the policy measures adopted in Brazil to cope with the crisis were more conventional than elsewhere in the world, they have been subject to various criticisms. The first type of criticism is based on the general perception that the Brazilian economy faces a borrowing constraint, which limits the countercyclical potential of fiscal policy.[2] Consequently, reducing the primary surplus target would be constrained by other objectives, in which case the function of policymakers would be to calibrate the reduction to avoid compromising the sustainability of the public debt, even in the short term.[3] In view of that concern, it can be argued that the debate on the effect on activity levels of reducing the fiscal target has not yet run its course.

The second type of criticism is that Brazil was not in the limiting situation of setting a lower bound for the operation of monetary policy. Thus, the countercyclical function of fiscal policy would hinder a sharper reduction in interest rates (Parnes and Goldfajn, 2008). Moreover, it is well known that monetary-policy-transmission channels in Brazil are not perfect, so even a sharp cut in interest rates might not be sufficient to stimulate economic activity as quickly as desired.[4]

☐ The opinions and results contained in this article are the authors' exclusive responsibility and do not necessarily represent the view of the institution to which they are affiliated.

[1] Blinder (2006) discusses the evolution of that thinking and concludes that fiscal policy should be used only when monetary policy is exhausted. This occurs when the nominal interest rate falls to near-zero levels, thus reaching what the literature refers to the "lower bound on interest rates".

[2] For example, when it was announced that the target for the primary surplus would be lowered from 3.8% of gross domestic product (GDP) to 2.5% in April 2009, public-debt projections were published without estimating the repercussions of the fiscal stimulus on GDP, which was the main purpose of the measure.

[3] It is important to consider existing proposals for the measurement of fiscal indicators. Hemming and Ter-Minassian (2004) recognize that the primary-surplus concept can cause sacrifices that have repercussions for long term growth, including cutbacks in infrastructure investment. Nonetheless, they claim that many countries are not technically prepared to adhere to an alternative rule, such as the golden rule that seeks to balance the budget using the target for the current balance rather than the capital balance. One possibility is the proposal developed by Blanchard and Giavazzi (2004) to exclude investment from the concept of primary surplus. Here, Brazil has gained relevant experience with the procedure through the *Projeto Piloto de Investimentos-PPI*. For further information on the Brazilian case see Silva and Pires (2008).

[4] Nonetheless, the institutional characteristics of the domestic financial system make the lower bound on the interest rate different from zero. In reality, in Brazil it is around 8.5%, owing to the yield paid by savings banks, which impose a very firm limit on the action of monetary policy.

The third criticism is that countercyclical fiscal policy would worsen the Brazilian economy's external deficit by stimulating aggregate demand, causing an excessive exchange-rate devaluation and fuelling inflation (Bacha, 2008). The validity of this argument would depend on the response of exports and imports on the domestic and international market, the way in which global deflation is passed through to local price indices, and the combination of exchange-rate devaluation and lower commodity prices.

The purpose of this article is to discuss the macroeconomic effects of the reduction of the primary-surplus target in Brazil in 2009, focusing on the short-run repercussions. Apart from this introduction, the article is divided into four sections. The second section describes the main fiscal stimulus measures adopted by the federal government up to the first half of 2009, which led to a lowering of the primary-surplus target for that year, while the third describes the methodology of the study and database. The fourth section estimates the effects of the primary surplus on economic activity, the time structure of interest rates and the exchange rate; and the last section sets out the main conclusions.

II

Main fiscal stimulus measures

The speed with which the crisis took hold in Brazil as from the third quarter of 2008 provoked different reactions in the various economic policy instruments. Although initial uncertainty as to the future course of inflation constrained monetary-policy decision-making, the first steps to stimulate aggregate demand were taken through fiscal policy.[5]

In December 2008, the Ministry of Finance announced three tax-cutting measures, as follows:

i) Lowering of the rate of personal income tax (IRPF) in 2009 (estimated at 4.9 billion reais), to stimulate aggregate demand.
ii) Reduction, until March of the rate of industrialized products tax (IPI) for the purchase of motor vehicles (estimated at 1 billion reais), to reduce the sector's inventories, which had grown rapidly as a result of the slump in demand.
iii) Decrease in financial transactions tax (IOF) on consumer loans (estimated at 2.5 billion real), to revive private credit.

In late March 2009, the Ministry of Finance announced an extension of the lower IPI rate on automobiles, and extended it to cover motorcycles, including, in that case, the Social Security Funding Contribution (COFINS). To offset the concomitant loss of revenue, IPI was raised on tobacco. The net loss of revenue as a result of the measure was estimated at 700 million reais.

In April, two other tax cuts were announced: a reduction of IPI on home appliances (estimated at 170 million reais) and a lower rate on certain civil construction items (estimated at 90 million reais), both of which aim to revive aggregate demand. In June, the federal government extended all existing tax cuts and also lowered rates on certain categories of capital goods. The total loss of revenue thus caused was estimated at 12.5 billion real (see table 1).

In terms of expenditure-stimulus measures, when the depth of the crisis became clear, the federal government increased investments as a countercyclical measure, adopting the following initiatives for that purpose: removal of firms in the Petrobras group from the primary-surplus calculation, with the aim of reducing their investment constraints (estimated at 15 billion reais in 2009); an increase in the minimum wage to stimulate aggregate demand (estimated at 8.7 billion reais); and a housing package (estimated at 6 billion reais), which gave incentives to the civil construction sector. In addition, the social protection network

[5] The initial effects of the crisis caused liquidity problems in the interbank system and in foreign-exchange operations through financial derivatives issued by various companies. Although discussion of the interest rate was contaminated by inflationary risks, the central bank implemented nonconventional policies, such as using international reserves to finance firms in difficulties and lowering reserve requirements to alleviate liquidity problems in the interbank system.

TABLE 1

Fiscal stimulus through tax reductions

(Billions of reais)

Fiscal measures	Stimulus in 2009
Tax cuts in December 2008	
Personal income tax (IRPF)	4.90
Industrialized products tax (IPI) - automobiles	1.00
Financial transactions tax (IOF) - consumer credit	2.50
Total	*8.40*
Tax cuts in March 2009	
Extension of IPI - automobiles	1.00
Social Security Funding Contribution (COFINS) - motorcycles	0.15
IPI - civil construction	0.35
Restoration of IPI revenue - tobacco	-0.80
Total	*0.70*
Tax cuts in April 2009	
IPI - civil construction	0.09
IPI - white line	0.17
Total	*0.26*
Tax cuts in June 2009	
IPI –capital goods	0.41
Extension of IPI - automobiles	1.79
Extension of IPI - white line	0.20
Extension of COFINS - motorcycles	0.05
Extension of IPI - civil construction	0.69
Total	*3.15*

Source: Prepared by the authors.

was expanded by increasing access to unemployment insurance and the *Bolsa Família* programme The total value of those measures is estimated at 30,930 million reais (see table 2).[6]

As a result of that set of measures, in April the federal government announced that the primary surplus target for 2009 was being lowered from 3.8% to 2.5% of GDP, with the possibility of also making use of the fiscal slack generated by the PPI. Petrobras was excluded from the calculation of the target reduction, estimated at 0.5% of GDP. The federal government share shrank from 2.15% of GDP to 1.4% of GDP, while the share of the states and municipalities decreased from 0.95% to 0.90% of GDP.

TABLE 2

Fiscal stimulus through expenditure increases

(Billions of reais)

Fiscal measures	Stimulus in 2009
Increase in expenditure	
Petrobras[a] investments	15.00
Minimum wage	8.70
Housing package	6.00
Unemployment insurance	0.23
Bolsa Família program	1.00

Source: Prepared by the authors.

[a] Petróleo Brasileiro.

III

Methodology and description of the database

The model-building strategy used in this article follows what Hoover, Johansen and Juselius (2008) define as the probabilistic approach (Haavelmo, 1944). Nonetheless, contrary to what might be imagined, the probabilistic approach does not mean research without theoretical backing. The basic idea is that data can often be used to give orientation to the theory.

The structure chosen includes a highly stylized "core", model and auxiliary models to complement it. As Bardsen and others (2005) point out, building macroeconometric models in that format has the advantage of incorporating elements that are not yet sufficiently analysed by theory, and also provides flexibility to meet the demands of model users, particularly policy-makers. This methodology implies

[6] Although the rise in the minimum wage was a pre-crisis measure, it has been included in the analysis because of its clear expansionary effect. Although the same could be said of the wage increase paid to civil servants, this was not included since its expansionary effects would only be felt as from the second half of 2009, when the results already showed that the economy was recovering.

"encompassing models", proposed by Clements and Hendry (2008).[7]

A major concern in model building, particularly when using Brazilian data is the existence of structural breaks. The parameters obtained need to be constant and invariant with respect to certain types of intervention, including manipulations in the exogenous variables (Bardsen and others, 2005). Thus, robust parameters are sought through alternative forms of estimation, changes in the sample size and by using variable-parameter methods.

Applying that methodology, the analysis was divided into three parts. The first part (estimation of the core) evaluated the effect of the primary surplus on the level of activity through a model equivalent to that used by central banks, consisting of three equations: an IS curve, a Phillips curve, and a monetary-policy rule. Whereas the primary surplus is modelled directly in the IS curve (following Lambertini and Rovelli (2003)), the joint estimation of those equations is important for controlling the endogeneity of economic policy, which could bias the results.

The IS curve is described as follows:

$$y_t = c_y + \alpha_1 y_{t-1} + \alpha_2 (i_{t-1} - \pi_t) + \alpha_3 s_t + e_{y,t} \qquad (1)$$

where c_y is the intercept and $e_{y,t}$ is an error term with zero mean and constant variance. The IS curve, as specified in (1), shows how the primary surplus (s_t) and the real interest rate (i_{t-1}-p_t) affect the level of economic activity (y_t).

The Phillips curve is specified as:

$$\pi_t = \phi_1 \pi_{t-1} + (1 - \phi_1) E_t \pi_{t+1} + \phi_2 y_t + e_{\pi,t} \qquad (2)$$

where π_t is the inflation rate, $e_{\pi,t}$ is an error term with zero mean and constant variance.

Monetary policy reacts according to the following rules:

$$i_t = \lambda_1 i_{t-1} + \lambda_2 (E_t \pi_{t+1} - \bar{\pi}) + \lambda_3 y_t + e_{i,t} \qquad (3)$$

which indicates that the nominal interest rate responds to the deviation of inflation (π_t) from its target ($\bar{\pi}$) and the output gap, while allowing also for an interest-rate mitigation component.

A point to note about the proposed model is that it aims to evaluate the effect of fiscal policy in the short run, because only in this way can the primary surplus be treated as an exogenous variable. In medium-term evaluations, the primary surplus has to respect the government's budget constraint, as shown in the literature on public-debt sustainability (Bohn, 1997).

The database for estimating the core of the model consists of the output gap calculated using a linear and quadratic trend; the short-term nominal interest rate is the Special Settlement and Custody System (SELIC) rate; the primary surplus refers to the consolidated public sector as a percentage of GDP; and the inflation rate is measured by the national consumer price index in its broad definition (IPCA) accumulated over 12 months. Figure 1 shows the trend of those variables in the period running from the third quarter of 1999 to the fourth quarter of 2008.

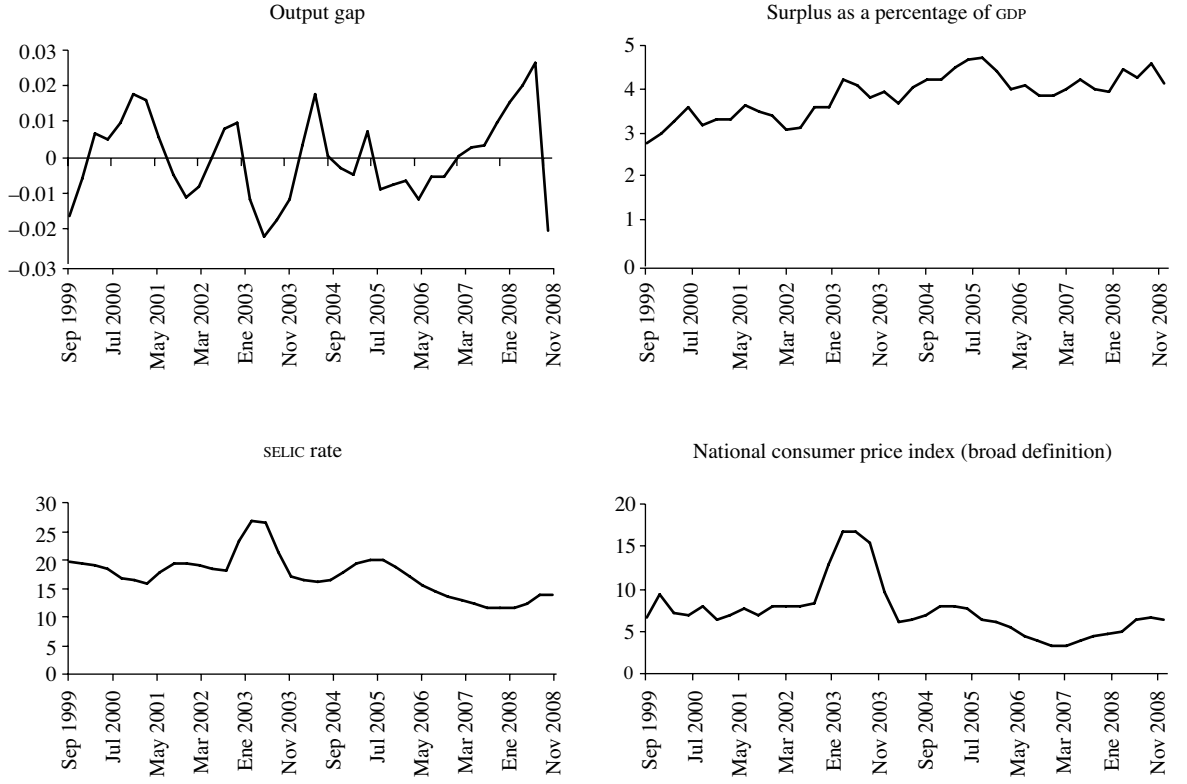
As figure 1 shows, the trend of the output gap displays mean reversion, which characterizes the series as stationary. This does not occur after 2005. In the case of the primary surplus as a proportion of GDP, the series displays a rising trend in the period under study, with a number of additional disturbances in 2003 and 2005. Given this growth trend, it can be said that the series displays mean reversion around the trend, so that variable would also be stationary. The SELIC interest rate is harder to characterize as a stationary series, because mean reversion is less clear. Nonetheless, this rate suffered a sharp disturbance in the second half of 2002, and when this is controlled for, reversion to mean is more evident. The inflation rate is analogous: when the observations of the second half of 2002 are considered as aberrations, it can be concluded that the series is also stationary.

To test those impressions regarding the characteristics of the variables, the Ng and Perron (2001) unit-root test was used, together with the Saikkonen and Lutkepohl (2002) test which allows for the existence of structural breaks in the dataset. Although the Ng and Perron test does not reject the null hypothesis of a unit root in the case of the primary surplus, the Saikkonen and Lutkepohl test shows that this primary surplus series can be considered stationary when the structural break in 2003 is controlled for (modelled with a dummy

[7] According to these authors, a good model should at least absorb or encompass already-existing models, to ensure that the research program moves in a positive direction.

FIGURE 1

Trend of the variables



Source: Prepared by the authors.

impulse variable). The two tests produced the same conclusions for the other variables.

The second part of the analysis evaluates the auxiliary models, the first of which estimates the effect of the primary surplus on the time-structure of interest rates. For this, a theoretical hypothesis for the behaviour of the time structure is needed, and traditional expectations hypothesis is adopted, although Brazilian data frequently reject the expectations hypothesis.[8] To adapt the model to existing empirical tests, a specification of the expectations theory is adopted, which assumes that the risk premium can vary through time (following Guillén and Tabak (2007)).

The time structure can thus be modelled through a state-space model of the following form:[9]

$$R_t^{swap,d} = i_t + r_t^d + \nu_t \quad (4)$$

(i_t = SELIC rate)

$$r_t^d = r_{t-1}^d + \mu_t \quad (5)$$

where R_t^{swap} is the interest rate swap at maturities of 30, 60, 90, 120, 180 and 360 days (d). The excess yield (given by $R_t^{swap} - i_t$) is defined as a risk premium (r_t) modelled as a time variable aggregated from a random variable $\nu t \sim (0, \omega^2)$. To estimate the variable risk premium over time, using the Kalman filter,

[8] For further details on the rejection of expectations theory in Brazil under the hypothesis that the risk premium is constant through time, see Tabak and Andrade (2003), and Lima and Issler (2003).

[9] For an analysis of space-state models see Commandeur and Koopman (2007).

a stochastic structure needs to be imposed. Our hypothesis is that it follows a random walk, in which $\mu_t \sim (0, \sigma^2)$ according to equation (5).

Next we aim to establish how the macroeconomic variables (inflation, output gap, and primary surplus) affect the risk premium through the system of equations:[10]

$$r_t^d = \rho^d r_{t-1}^d + \beta_0^d + \beta_\pi^d \pi_t + \beta_s^d s_t + \beta_y^d y_t + \varepsilon_t^d \qquad (6)$$

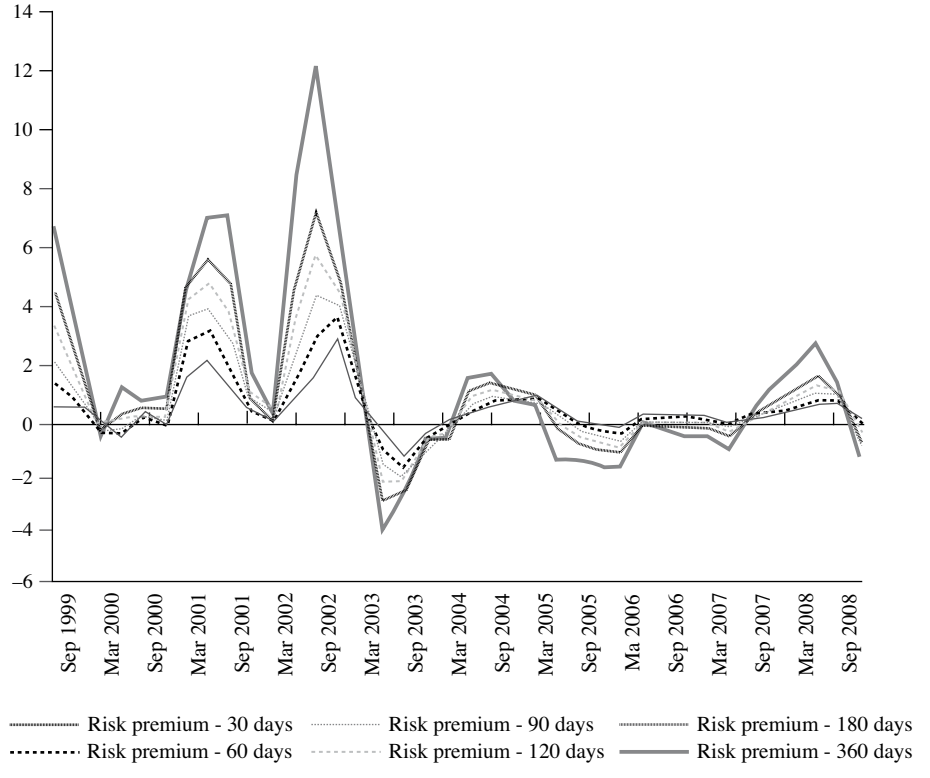
[10] By allowing the risk premium to be a random walk in equation (5), and then modelling the risk premium as a distributed lag (ADL) function in (6), one is assuming that the non-stationary nature originating in (5) occurs because of a possible problem of mis-specification. The time-series literature stresses that non-stationarity can arise as a result of a lack of linearity, structural breaks, or missing variables that are relevant in the data-generating process. This hypothesis can be rejected in the case where the parameter estimation ρ is less than 1, as is the case with the estimations obtained.

Figure 2 shows the trend of the risk premium estimated through the Kalman filter at various maturities, which displays high volatility in the period 1999-2003 before stabilizing. The second interesting feature is that the longer rates fluctuate by more than the shorter rates.

The second auxiliary model evaluates the repercussions of the primary surplus on the exchange rate. The most traditional exchange-rate model is purchasing power parity (PPP); and research into its validity generally consists of testing whether the real exchange rate can be modelled as a stationary variable (Rogoff, 1996). In that regard, the analysis closely follows the results of Juselius (2007), which tests PPP in Germany. These show that PPP is only upheld when the analysis includes the effect of the spread between the short- and long-term interest rate. The first finding of the study is that systematic deviations from PPP would be possible if the exchange-rate effects of capital movements are included in the analysis.

FIGURE 2

Trend of the risk premium



Source: Prepared by the authors.

Here the natural candidate for deviations from a PPP exchange rate would be the interest-rate spread.

Following that literature, an exchange-rate model is proposed that is similar to that used for the time-structure of interest rates. Exchange-rate movements are defined on the basis of fluctuations in domestic prices relative to the international price level. The resulting deviations are modelled as a function of the risk premium, as obtained in (5).[11] In addition to using the risk premium to model the real exchange rate, it should be noted that the equilibrium level of the exchange rate cannot be considered constant, since several factors can affect its value, including nominal shocks and productivity differentials.[12]

Figure 3 shows the trend of the real exchange rate in natural logarithm form, for the period running from the third quarter of 1999 to the fourth quarter of 2008. The first relevant feature of the series is the difficulty in observing reversion to mean. The second is that there were a number of significant positive shocks in 2001 and 2002. The Ng and Perron test was used to confirm the absence of mean reversion, because the null hypothesis of unit root is not rejected. When the real exchange rate is modelled with a structural break, the Saikonen and Lutkepohl test confirms the results obtained earlier.

Consequently, having rejected PPP, the proposed model incorporates the effect of the risk premium on the real exchange rate, and models the mean as a time variable process using the Kalman filter, as follows:

$$q_t = \psi_t + \omega r_t^d + \varepsilon_{q,t} \quad (7)$$

$$\psi_t = \psi_{t-1} + \varepsilon_{\psi,t} \quad (8)$$

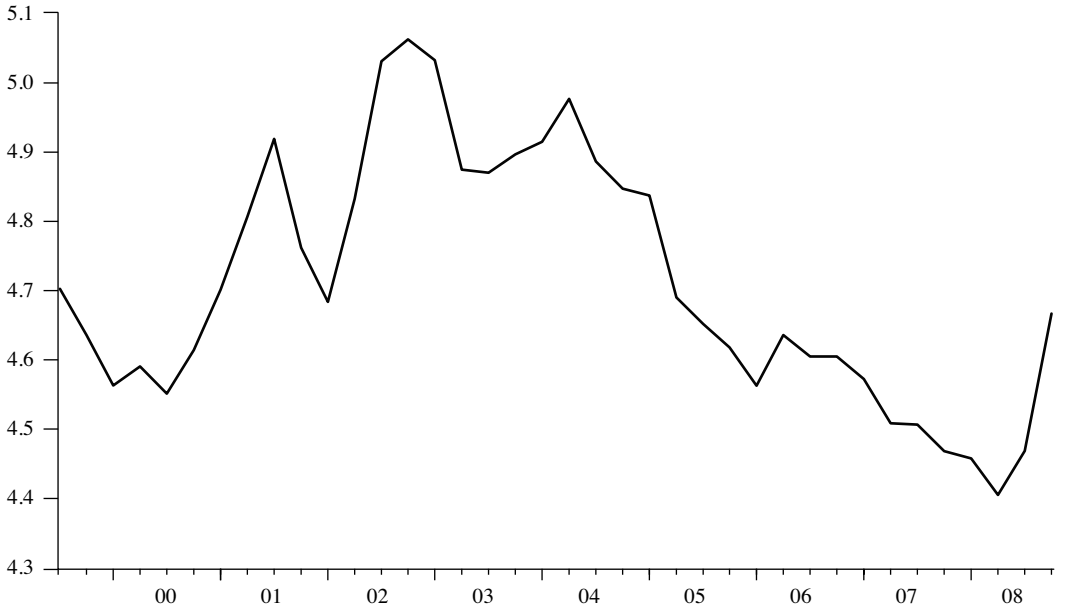
where q is the real exchange rate, ψ_t is the time-variable mean which follows the structure of a random walk, and $\varepsilon_{\psi,t}$ and $\varepsilon_{q,t}$ are error terms with zero mean and constant variance.

[11] With regard to the results obtained by (2007), it should be noted that that study used the risk premium rather than the interest-rate spread because of the possibility of multi-collinearity, since the two variables are functions of the short-term interest rate. It should also be noted that while for Germany it is reasonable to assume a risk premium close to zero, this is not true for Brazil; and this is a relevant factor in the composition of the yield spread between foreign and domestic assets.

[12] See Dornbusch (1976), Balassa (1964) and Samuelson (1964).

FIGURE 3

Trend of the logarithm of the real exchange rate

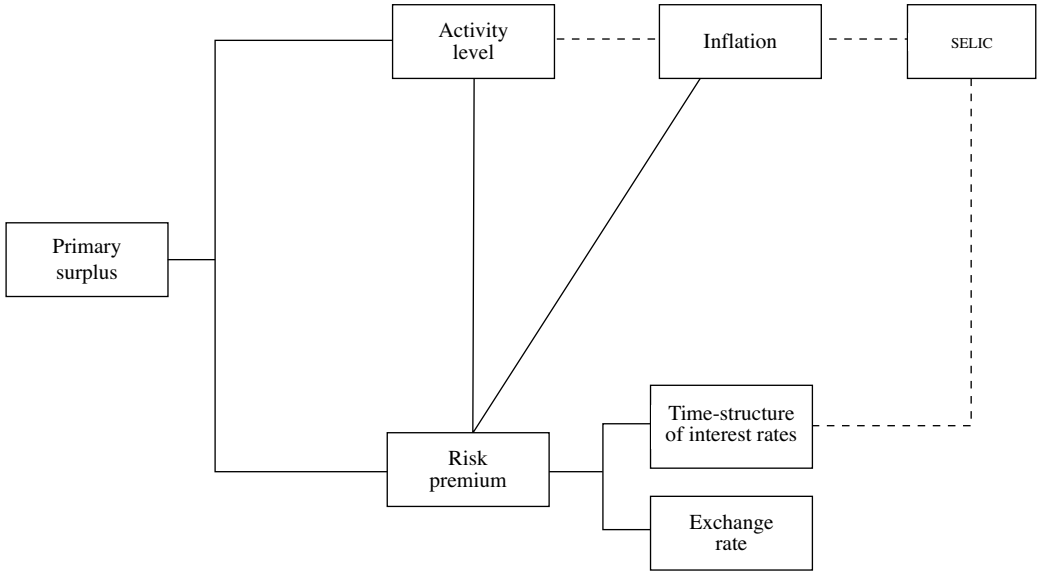


Source: Prepared by the authors.

In brief, the model consists of a core and two auxiliary models. The transmission mechanism to be evaluated thus shows how the primary surplus affects the levels of activity, time structure and exchange rate in the short run. Figure 4 illustrates the primary surplus macroeconomic transmission mechanism. In addition to the effects described, the indirect effects acting through economic activity should also be considered, including the rate of inflation, the short-term interest rate and the time structure.

FIGURE 4

Primary-surplus transmission mechanism



Source: Prepared by the authors.

SELIC: Special settlement and custody system.

IV
Estimations

The core of the model was estimated using ordinary least squares (OLS), seemingly unrelated regression (SUR) and the generalized method of moments (GMM). In the latter case, the list of instruments includes lags 1 to 3 in the model's own variables. The parameters are very similar in the three methods of estimation. For example, the effect of the primary surplus on the level of activity is estimated at between -0.377 and -0.430, although the effect is only significant according to GMM. In terms of statistical significance, the same is true for monetary policy, the effects of which are estimated at between -0.131 and -0.142.

The Phillips curve indicates a high degree of inflationary inertia, with an estimated coefficient of between 0.779 and 0.835, suggesting a highly regressive inflation. The coefficient of the output gap is significant in all three estimations (between 0.373 and 0.424).

The gradualism of monetary policy (measured by parameter λ_1) is quite high. In reality, that parameter approximates to a unit root, which may raise doubts as to the stationary nature of the process and, therefore, the validity of the inference. Although unit-root tests were applied to make it possible to undertake the

analysis, the fact is they tend to pose problems of low testing power. Nonetheless, the results obtained by Rothenberg and Stock (1997) show that conventional inference close to the unit root can be considered valid, so the results obtained are a good approximation.

The estimated effect of the primary surplus on the time structure suggest that this raises the interest rate at various maturities (see table 4). In both estimation periods, it is common to find that estimations using OLS and seemingly unrelated regression indicate significant effects, whereas the results in GMM estimations are not statistically significant at the shorter maturities (30 and 60 days). The test results also show that the effects vary by maturity. In general, the longer the maturity, the greater the estimated repercussion of fiscal policy. This can be clearly seen in figure 5, which shows the parameter measuring the effect of the primary surplus in the period running from the third quarter of 1999 to the fourth quarter of 2008, by maturity.

TABLE 3

Estimations of the small-scale model

Variables	Ordinary least squares		Seemingly unrelated regression		Generalized method of moments	
	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
c_{and}	0.026	1.42	0.027	1.60	0.028	4.01
α_1	0.471	2.97	0.482	3.22	0.493	10.69
α_2	-0.137	-1.51	-0.131	-1.55	-0.142	-3.269
α_3	-0.377	1.02	-0.430	-1.25	-0.400	-2.778
R^2	0.3107		0.3100		0.3096	
ϕ_1	0.779	9.75	0.801	10.88	0.835	29.60
ϕ_2	0.373	1.76	0.397	1.95	0.424	10.28
R^2	0.8221		0.8217		0.8196	
λ_1	0.965	79.42	0.965	87.31	0.966	205.7
λ_2	0.757	4.32	0.638	4.19	0.826	11.90
λ_3	0.439	2.72	0.477	3.13	0.405	10.06
R^2	0.9052		0.9033		0.9049	

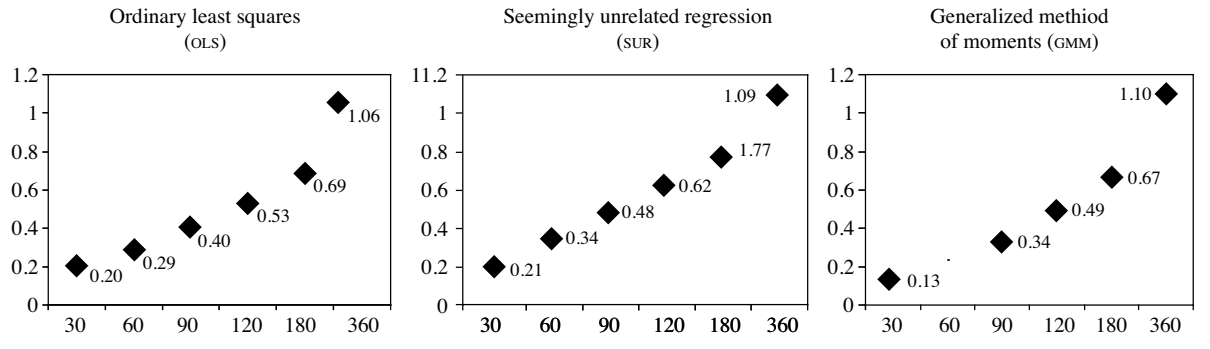
Source: Prepared by the authors.

Note: Total number: 100 observations. In the case of GMM, the J-statistic does not reject the null hypothesis of validity of the instruments tested. The instruments used were lags 1, 2 and 3 of the output gap and primary surplus, lags 1 and 2 of the broadly defined national consumer price index (IPCA) and lags 1, 2 and 3 of the interest rate.

R^2: Goodness of fit.

FIGURE 5

Effect of the primary surplus by maturity



Source: Prepared by the authors.

TABLE 4

Macroeconomic determinants of the risk premium

Variables	Ordinary least squares		Seemingly unrelated regression		Generalized method of moments	
	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
β_0^{30}	0.0097	2.82	0.0093	2.92	0.0062	1.63
ρ^{30}	0.8623	9.83	0.6860	14.09	0.8499	9.77
β_π^{30}	-0.0300	-2.11	-0.0180	-1.44	-0.0191	-1.87
β_s^{30}	0.1955	2.36	0.2060	2.68	0.1257	1.27
β_y^{30}	0.1260	3.49	0.1332	3.98	0.1979	3.73
R^2	0.80		0.78		0.77	
β_0^{60}	0.0142	2.84	0.0151	3.28	0.0104	1.79
ρ^{60}	0.8332	9.81	0.6726	15.26	0.8273	10.39
β_π^{60}	-0.0410	-1.96	-0.0236	-1.31	-0.0278	-1.89
β_s^{60}	0.2898	2.36	0.3396	3.03	0.2130	1.45
β_y^{60}	0.1923	3.65	0.2084	4.29	0.2797	3.91
R^2	0.82		0.80		0.81	
β_0^{90}	0.0196	2.98	0.0217	3.58	0.0165	2.13
ρ^{90}	0.8180	9.76	0.6771	16.08	0.8205	10.62
β_π^{90}	-0.0542	-1.99	-0.0340	-1.44	-0.0439	-2.09
β_s^{90}	0.4038	2.48	0.4833	3.28	0.3376	1.77
β_y^{90}	0.2379	3.43	0.2614	4.12	0.3154	3.64
R^2	0.83		0.82		0.83	
β_0^{120}	0.0258	3.03	0.0286	3.72	0.0242	2.40
ρ^{120}	0.7979	9.43	0.6990	16.90	0.8171	10.49
β_π^{120}	-0.0697	-2.02	-0.0519	-1.75	-0.0681	-2.30
β_s^{120}	0.5307	2.50	0.6237	3.33	0.4907	2.06
β_y^{120}	0.2689	3.05	0.2935	3.66	0.3123	3.00
R^2	0.84		0.83		0.85	
β_0^{180}	0.0332	3.09	0.0360	3.75	0.0333	2.59
ρ^{180}	0.7838	9.07	0.7154	16.93	0.8155	10.12
β_π^{180}	-0.0878	-2.05	-0.0727	-1.97	-0.0964	-2.35
β_s^{180}	0.6850	2.54	0.7745	3.30	0.6686	2.28
β_y^{180}	0.3005	2.72	0.3238	3.24	0.3035	2.30
R^2	0.84		0.83		0.85	
β_0^{360}	0.0517	3.12	0.0527	3.63	0.0561	2.90
ρ^{360}	0.7939	8.51	0.7789	16.23	0.8470	9.21
β_π^{360}	-0.1398	-2.10	-0.1343	-2.40	-0.1818	-2.37
β_s^{360}	1.0582	2.51	1.092	3.06	1.1003	2.60
β_y^{360}	0.3180	1.89	0.3268	2.17	0.1836	0.84
R^2	0.83		0.84		0.84	

Source: Prepared by the authors.

Note. Total number: 216 observations of the system. In the case of GMM, the J-statistic does not reject the null hypothesis of validity of the instruments tested. The instruments used were lags 1 and 2 of the explanatory variables.

R^2: Goodness of fit.

To evaluate the robustness of the results obtained, the model was reapplied to a subsample covering the period from the first quarter of 2003 to the fourth quarter of 2008, excluding the period of greatest volatility caused by the large number of supply shocks that occurred in 1999-2002 (see table 5). Figure 2 shows that the excluded time corresponds to the period when the various risk premia were most volatile.

The new results show that the primary surplus only affects the interest rate in the longer-term risk premia (particularly 360 days). In relation to the

TABLE 5

Macroeconomic determinants of the risk premium, first quarter of 2003 to fourth quarter of 2008.

Variables	Ordinary least squares		Seemingly unrelated regression		Generalized method of moments	
	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
β_0^{30}	0.0028	0.58	0.0029	0.74	0.0024	0.68
ρ^{30}	0.6326	7.50	0.6355	13.02	0.6257	11.29
β_π^{30}	-0.0339	-3.08	-0.0340	-4.64	-0.0326	-4.16
β_s^{30}	0.0314	0.28	0.0333	0.36	0.0231	0.28
β_y^{30}	0.0680	2.76	0.0680	3.10	0.0739	4.20
R^2	0.83		0.83		0.83	
β_0^{60}	0.0049	0.70	0.0045	0.80	0.0041	0.86
ρ^{60}	0.6277	7.24	0.6188	12.92	0.6187	10.11
β_π^{60}	-0.0421	-3.21	-0.0414	-3.92	-0.0400	-3.47
β_s^{60}	0.0737	0.45	0.0655	0.49	0.0588	0.53
β_y^{60}	0.1159	3.28	0.1162	3.71	0.1258	5.02
R^2	0.83		0.83		0.83	
β_0^{90}	0.0081	0.93	0.0074	1.06	0.0073	1.25
ρ^{90}	0.6187	7.05	0.6044	12.61	0.6089	9.16
β_π^{90}	-0.0488	-2.94	-0.0473	-3.56	-0.0461	-3.30
β_s^{90}	0.1467	0.73	0.1309	0.79	0.1292	0.99
β_y^{90}	0.1588	3.56	0.1598	4.05	0.1718	5.51
R^2	0.83		0.83		0.83	
β_0^{120}	0.0133	1.29	0.0121	1.44	0.0122	1.69
ρ^{120}	0.6083	6.96	0.5879	12.24	0.5971	8.15
β_π^{120}	-0.0558	-2.76	-0.0532	-3.29	-0.0524	-2.98
β_s^{120}	0.2652	1.09	0.2398	1.19	0.2431	1.53
β_y^{120}	0.2073	3.78	0.2100	4.36	0.2243	5.49
R^2	0.84		0.84		0.83	
β_0^{180}	0.0186	1.51	0.0171	1.69	0.0173	1.96
ρ^{180}	0.5945	6.80	0.5717	11.77	0.5824	7.53
β_π^{180}	-0.0636	-2.60	-0.0601	-3.06	-0.0593	-2.71
β_s^{180}	0.3867	1.34	0.3553	1.47	0.3673	1.90
β_y^{180}	0.2611	3.89	0.2655	4.53	0.2827	0.00
R^2	0.84		0.83		0.83	
β_0^{360}	0.0330	2.23	0.0329	2.57	0.0309	2.69
ρ^{360}	0.5552	6.93	0.5529	11.49	0.5380	8.14
β_π^{360}	-0.0695	-2.06	-0.0689	-2.58	-0.060	-1.94
β_s^{360}	0.7363	2.10	0.7338	2.39	0.6991	2.88
β_y^{360}	0.3729	4.11	0.3739	4.80	0.4132	4.61
R^2	0.84		0.84		0.84	

Source: Prepared by the authors.

Note. Total sample: 144 observations. In the case of GMM, the J-statistic does not reject the null hypothesis of validity of the instruments tested. The instruments used were lags 1 and 2 of the explanatory variables.
R^2: Goodness of fit.

other parameters, the results are equivalent to those obtained previously, with inflation having a negative effect and the output gap a positive one.

In conclusion it can be inferred that, apart from increasing the risk premium, the reduction in the primary surplus could increase the gradient of the time structure. This conclusion is very similar to the results obtained by Evans (1985, 1987a and 1987b) from a study of the effects of fiscal deficits on North American interest rates.

Table 6 shows the repercussion of the risk premium on the exchange rate, which seems to decrease with the length of maturity and is smaller in the three cases presented (360 days, 180 days or 120 days). The primary surplus thus affects the exchange rate through the pressure exerted on the risk premium. Nonetheless, unlike the results obtained for the time structure, where the primary surplus has greater effects as the maturity term lengthens, its repercussions on the exchange rate are offset by the fact that the risk premium at longer maturities has less influence of the behaviour of the exchange rate.

Figure 6 shows the trend of the exchange rate and the model's predictions with the 360-day risk premium.

TABLE 6

Effect of the risk premium on the exchange rate

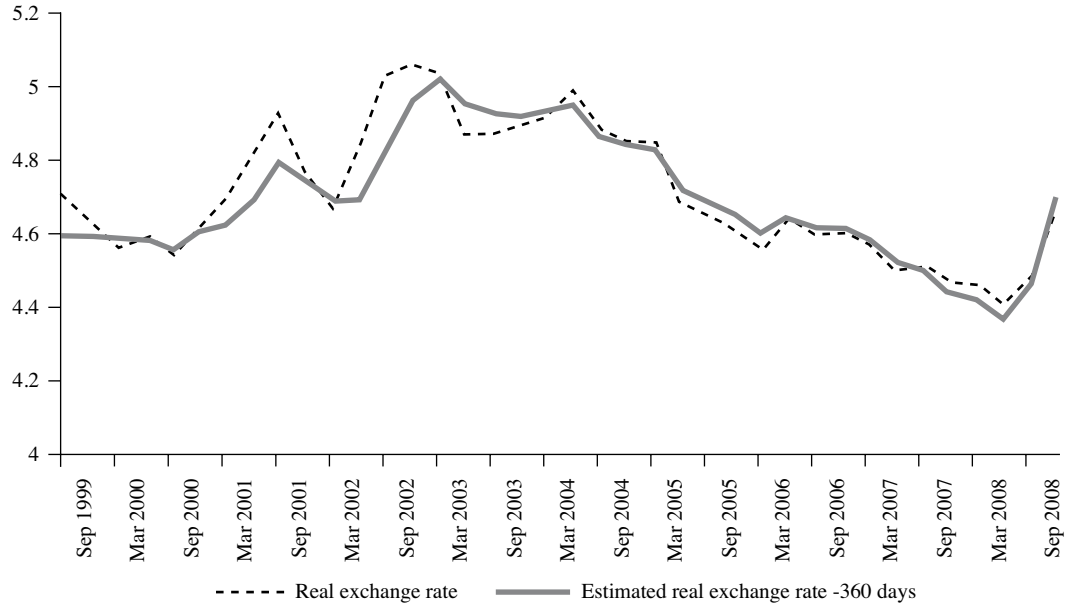
Model	Coefficient (t-statistic)	Coefficient (t-statistic)	Coefficient (t-statistic)	"Invariable" model
ψ_t	4.69 (743.9)	4.69 (766.7)	4.69 (751.2)	4.68 (81.74)
r_t^{360}	0.0172 (4.07)	-	-	0.0248 (1.99)
r_t^{180}	-	0.0253 (4.00)	-	-
r_t^{120}	-	-	0.0305 (3.85)	-
R^2	0.8451	0.8435	0.8399	0.1177

Source: Prepared by the authors.

R^2: Goodness of fit.

FIGURE 6

Comparison of prediction power of the 360-day model



Source: Prepared by the authors.

Although the model seems not to have a good fit at the start of the sample, its prediction power improves considerably through time. Much of an improvement is the result of modelling the equilibrium value as a time variable, although the statistical significance of the risk premium is not necessarily due to that. This borne out in the last column of table 6, which shows the performance of the model with a time-invariant intercept. Lastly, it should be noted that the effect on the exchange rate is qualitatively small, although statistically significant.

In general, the results suggest a markedly expansionary effect for fiscal policy, without major collateral repercussions on macroeconomic balances such as the external deficit or higher interest rates. In economic terms, that short-term strategy makes sense given the size of Brazilian domestic market, which absorbs the effect of the policies adopted almost entirely. In countries that are more dependent on international trade, it is possible that most of the fiscal stimulus will suck in imports and be less effective than estimated for the Brazilian case.

V
Conclusions

The speed with which the international financial crisis took hold in the Brazilian economy in the final quarter of 2008 had devastating effects on the real economy and required equally rapid economic policy responses, with major macroeconomic repercussions.

The purpose of this article was to describe the short-term macroeconomic effects of reducing the primary surplus target as an economic-policy response which, apart from maintaining the level of budgetary expenditure to sustain aggregate demand, also aimed to speed up public investment. The article therefore evaluated the effect of reducing the primary surplus on the level of activity, the time structure of interest rates and the exchange rate.

The results showed that reducing the primary surplus should increase the activity level significantly; and, with a three-equation model, commonly used in central banks, the reduction of the primary surplus should increase the level of economic activity, and thus serve as an effective policy for reviving the economy.

It was also seen that the reduction of the primary surplus could have a counterparts in an increase in the gradient of the time-structure of interest rates, particularly at the longest maturities, and in a devaluation of the exchange rate. Nonetheless, the exchange-rate effect can be considered small, as it would be an exaggeration to claim that countercyclical fiscal policy could worsen the external deficit.

In general, the results show that countercyclical fiscal policy should have satisfactory effects on the level of activity, without negative collateral repercussions. In short, although the activity level will need to be stimulated considerably, it is important to note that the effect on the time-structure of interest rates should only occur through the longer rates, and the effect on the external deficit —by exerting pressure on the exchange rate— should be small.

(Original: Portuguese)

APPENDIX

Ng and Perron (2001) unit-root test

Variable	Model	Significance level (5%)	Test statistic (MZa)
y	constant	-8.10	-26.74
π	constant	-8.10	-12.99
i	constant	-8.10	-17.71
s	constant and trend	-17.30	-10.46
q	constant and trend	-17.30	-10.71
d(q)	constant	-8.10	-36.95

Source: Prepared by the authors.

Saikkonen and Lutkepohl (2002) unit-root test

Variable	Model	Date of break	Significance level (5%)	Test statistic
y	constant	2004 (Q2)	-2.88	-3.77
π	constant	2003 (Q3)	-2.88	-2.91
i	constant	2002 (Q3)	-2.88	-2.97
s	constant and trend	2003 (Q1)	-3.03	-3.13
q	constant and trend	2003 (Q1)	-3.03	-2.08

Source: Prepared by the authors.

Bibliography

Andrade, J.P. and M.C.C. Pires (2009), "Uma análise da transmissão da política monetária e o canal da dívida pública com aplicação aos dados brasileiros", *Texto para discussão*, No. 1379, Rio de Janeiro, Institute of Applied Economic Research (IPEA).

Bacha, E. (2008), "O choque externo e a resposta possível", *Como reagir à crise? Políticas econômicas para o Brasil*, E. Bacha and I. Goldfajn (orgs.), Rio de Janeiro, Editora Imago.

Balassa, B. (1964), "The purchasing-power parity doctrine: a reappraisal", *Journal of Political Economy*, vol. 72, Chicago, University of Chicago Press.

Bardsen, G. and others (2005), *The Econometrics of Macroeconomic Modelling*, New York, Oxford University Press.

Barro, R. (1979), "On the determination of the public debt", *Journal of Political Economy*, vol. 87, No. 5, Chicago, University of Chicago Press.

Blanchard, O. and F. Giavazzi (2004), "Improving the SGP through a proper accounting of public investment", *CEPR Discussion Papers*, No. 4220, London, Centre for Economic Policy Research.

Blinder, A. (2006), "The case against the case against discretionary fiscal policy", *The Macroeconomics of Fiscal Policy*, R. Kopcke, G. Tootell and R. Triest, Cambridge, Massachusetts, The MIT Press.

Bohn, H. (1997), "The behavior of U.S. public debt and deficits", *The Quarterly Journal of Economics*, vol. 113, No. 3, Cambridge, Massachusetts, The MIT Press.

Clements, M. and D. Hendry (2008), *Forecasting Economic Time Series*, Cambridge, Cambridge University Press.

Commandeur, J.J.F. and S.J. Koopman (2007), *An Introduction to Space State Time Series Analysis*, New York, Oxford University Press.

Dornbusch, R. (1976), "Expectations and exchange rate dynamics", *Journal of Political Economy*, vol. 84, No. 6, Chicago, Chicago University Press.

Evans, P. (1987a), "Interest rates and expected future budget deficits in the United States", *Journal of Political Economy*, vol. 95, No. 1, Chicago, Chicago University Press.

_____ (1987b), "Do budget deficits raise nominal interest rates? Evidence from six countries", *Journal of Monetary Economics*, vol. 20, No. 2, Amsterdam, Elsevier.

_____ (1985), "Do large deficits produce high interest rates", *American Economic Review*, vol. 75, No. 1, Nashville, Tennessee, American Economic Association.

Guillén, O.T. and B. Tabak (2007), "Characterizing the Brazilian term structure of interest rates", *Anais do XXXV Encontro Nacional de Economia*, No. 108, Salvador, National Association of Centers for Post-graduation in Economics (ANPEC).

Haavelmo, T. (1944), "The probability approach in econometrics", *Econometrica*, vol. 12, Bethesda, Econometric Society.

Hemming, R. and T. Ter-Minassian (2004), "Making room for public investment", *Finance and Development*, Washington, D.C., International Monetary Fund, December.

Hoover, K.D., S. Johansen and K. Juselius (2008), "Allowing the data to speak freely: the macroeconomics of the cointegrated vector autoregression", *American Economic Review*, vol. 98, No. 2, Nashville, Tennessee, American Economic Association.

Juselius, K. (2007), "The PPP puzzle: what the data tell when allowed to speak", *The Handbook of Econometrics*, T. Mills and K. Patterson (eds.), London, Palgrave.

Lambertini, L. and R. Rovelli (2003), "Monetary and fiscal policy coordination and macroeconomics stabilization: a theoretical analysis", *Working Papers*, No. 464, Bologna, University of Bologna.

Lima, A. and J. Issler (2003), "A hipótese das expectativas na estrutura a termo da taxa de juros no Brasil: uma aplicação de modelos de valor presente", *Revista brasileira de economia*, vol. 57, No. 4, Rio de Janeiro, Getulio Vargas Foundation.

Ng, S. and P. Perron (2001), "Lag length selection and the construction of unit-root tests with good size and power", *Econometrica*, vol. 69, No. 6, Bethesda, Econometric Society.

Parnes, B. e I. Goldfajn (2008), "Como reagir à crise: política fiscal", *Como reagir à crise? Políticas econômicas para o Brasil*, E. Bacha and I. Goldfajn (orgs.), Rio de Janeiro, Editora Imago.

Rogoff, K. (1996), "The purchasing power parity puzzle", *Journal of Economic Literature*, vol. 34, No. 2, Nashville, Tennessee, American Economic Association.

Rothenberg, T. and J. Stock (1997), "Inference in a nearly integrated autoregressive model with nonnormal innovations", *Journal of Econometrics*, vol. 80, No. 2, Amsterdam, Elsevier.

Saikkonen, P. and H. Lutkepohl (2002), "Testing for a unit root in a time series with a level shift at unknown time", *Econometric Theory*, vol. 18, Cambridge, Cambridge University Press.

Samuelson, P. (1964), "Theoretical notes on trade problems", *Review of Economics and Statistics*, vol. 46, Cambridge, Massachusetts, The MIT Press.

Silva, A.M.A. and M.C.C. Pires (2008), "Dívida pública, poupança em conta corrente do governo e superávit primário: uma análise de sustentabilidade", *Brazilian Journal of Political Economy*, vol. 28, No. 4, São Paulo, Centro de Economía Política.

Tabak, B. and S. Andrade (2003), "Testing the expectations hypothesis in the Brazilian term structure of interest rates", *Revista brasileira de finanças*, vol. 1, No. 2, São Paulo, Sociedade Brasileira de Finanças.

Tobin, J. (1982), *Asset Accumulation and Economic Activity: Reflections on Contemporary Macroeconomic Theory*, Chicago, The University of Chicago Press.

KEYWORDS

Exports
Manufactures
Econometric models
Labour productivity
Economic policy
Mexico

The dynamics of Mexican manufacturing exports

Víctor M. Cuevas

This article evaluates several of the determinants of Mexican manufacturing exports, using two complementary econometric methods: a structural ARIMA model, which makes it possible to estimate elasticities; and a generalized VAR model, which provides a fully dynamic perspective by estimating impulse response functions. As some of the findings are robust to changes in the econometric methodology used, the article reaches the following conclusions. First, manufacturing exports are positively related to labour productivity and external demand; so the adverse effects of an international recession on Mexican exports could, to some extent, be offset by raising worker productivity. Second, real exchange-rate depreciation does not increase manufacturing exports, but actually reduces them, at least in the short run. These findings are consistent with the idea that a real depreciation not only affects demand, but also generates strong supply-side effects.

Víctor M. Cuevas
Research Professor, Department of Economics, Autonomous Metropolitan University (UAM) and member of the National Researchers System
victorcuevasahumada@yahoo.com.mx

I
Introduction

This article evaluates several of the variables that determine Mexico's manufacturing exports, using two complementary econometric methods: a structural Autoregressive Integrated Moving Average (ARIMA) model to estimate elasticities; and a Generalized Vector Autoregressive (GVAR) model, which makes it possible to estimate the dynamic responses of manufacturing exports to different types of shock.[1] Accordingly, both univariate and multivariate time series analyses are used to obtain two different perspectives on the factors driving manufacturing exports.

Several previous empirical studies have shown that exports are influenced not only by relative prices and external demand, but also by domestic demand and supply-side factors. In that context, this paper concludes that increased labour productivity and external demand expansions both have a significant impact on the growth of manufactured exports. Moreover, the evidence provided here suggests that a real exchange-rate depreciation could reduce the volume of exports in the short term, rather than increase it. A plausible explanation for this atypical result is that real currency depreciation generates two opposing effects, especially in developing countries: it makes their exports cheaper in terms of foreign currency; but it also raises the local-currency cost of imported intermediate inputs. The net effect on Mexico's international competitiveness appears to be negative, at least in the short term. Lastly, we present empirical results showing that strategic investment and production decisions are driven by the firm's desire to grow its exports.

An important economic-policy implication of this is that the adverse effects of an international recession on Mexican exports could, to some extent, be offset through higher labour productivity. Moreover, a comprehensive and coherent package designed to enhance labour productivity could prove more effective in stimulating manufacturing exports than a depreciation of the real exchange rate.

The rest of this paper is organized as follows. Section II briefly reviews the recent literature. Section III describes the model and the data set. Section IV conducts the integration and co-integration analyses. Sections V and VI present the estimates obtained from the univariate and multivariate time-series models, respectively; while Section VII summarizes the findings and examines the economic policy implications. Section VIII sets out the conclusions.

II
Literature review

The use of new econometric methods has revived interest in the short- and long-term determinants of exports. In general, export functions are usually specified under three basic approaches: (i) gravity models of trade, (ii) theoretical demand models, and (iii) theoretical models that combine demand- and supply-side variables. Although this is not an exhaustive classification, it provides the background needed to perform the empirical analysis and interpret the key findings.

The gravity model of trade has been relatively successful in modelling bilateral trade flows between countries since the early 1960s. Pioneering research in this field includes the work of Isard (1954), Tinbergen (1962) and Pöyhönen (1963). In the canonical specification of the gravity model, exports from one country to another are posited as an increasing function of economy size, measured by gross domestic product (GDP), and as a decreasing function of

☐ The author gratefully acknowledges financial support for this research from the National Science and Technology Council (CONACYT), and the comments and suggestions made by an anonymous referee.

[1] The GVAR method produces empirical evidence that is independent of the ordering of the equations, which is a major improvement over traditional recursive VAR models.

transport costs, measured by the distance between the two countries.

Over the years, the inclusion of other variables, such as population, common language and shared borders, has given rise to an "augmented gravity model". Two recent studies that use this version of the model are Boisso and Ferrantino (1997), and Bayoumi and Eichengreen (1997). Another line of research is represented by Bergstrand (1985) and Summary (1989), among others, who have improved the explanatory power of the gravity equation by adding variables such as prices and exchange rates, which are not directly related to geography or spatial factors.

In contrast, under the theoretical demand model, the volume of exports basically depends on the real exchange rate and the foreign income level, as measured by economic activity abroad or the physical volume of imports in the country's main trading partners. Two of the best known studies using this approach are Reinhart (1995), and Senhadji and Montenegro (1998).

Based on a sample of 12 developing countries (including Mexico), Reinhart (1995) shows that exports are more responsive to changes in external demand than to variations in relative prices. Senhadji and Montenegro (1998) enlarge the sample to encompass 53 countries, including both developing and industrialized economies, and they estimate long-term income- and price-elasticities of export demand at roughly 1.5 and -1, respectively.

Despite their popularity and widespread use, standard demand models have not been free from criticism. In particular, Riedel (1988) shows that neglecting supply-side factors leads to biased estimates of export-demand elasticities. As a result, some empirical models started to include variables relating to the supply of exports and the domestic demand for exportable goods, in addition to the traditional demand-side variables. This new strand of literature includes several papers focusing specifically on developing economies. In the case of Argentina, Catao and Falsetti (2002) show that manufacturing exports respond significantly to economic activity in Brazil, Uruguay and Paraguay (Argentina's trading partners in MERCOSUR), and also to the real exchange rate of the Argentine peso against the Brazilian real. These authors also provide evidence that an increase in net aggregate investment would stimulate manufacturing exports, while a rise in domestic consumption would have the opposite effect.

Berrettoni and Castresana (2007) analyse the impact of the real exchange rate, exchange-rate volatility, external demand and capacity utilization (among other explanatory variables) on Argentina's industrial manufacturing exports. They find that exchange-rate volatility is negatively related to exports, while external demand influences exports more than the real exchange rate.

For the Mexican economy, Cuevas (2008) shows that manufacturing exports are affected not only by changes in the real exchange rate and the level of output in the United States, but also by improvements in labour productivity. Similarly, Padilla and Juárez (2006) highlight the essential role of training in making the Mexican manufacturing sector more competitive. As that study measures competitiveness in the Mexican manufacturing sector through changes in total factor productivity,[2] a plausible hypothesis is that training enhances competitiveness (as measured by productivity), which, in turn, raises the level of manufacturing exports.

Lastly, a number of studies relate exports to other variables in developing economies. For example, Goldberg and Klein (1997) identify several positive effects of foreign direct investment (FDI) on exports in Latin American countries,[3] whereas Mbale and Golub (2002) show that a reduction in unit labour costs raises manufacturing exports in Senegal. Thus, according to these authors, labour productivity needs to outpace wages in order to stimulate export growth.

In brief, the econometric evidence suggests that exports depend on a variety of variables, including those specifically related to geography and spatial factors. Moreover, there is a basic consensus that exports are more responsive to external demand than to real exchange-rate adjustments. Lastly, the empirical evidence supports the claim that export volume depends not only on relative prices and foreign income, but also on the domestic demand for exportable goods and supply-side variables such as labour productivity, wages and fdi (Riedel, 1998).

[2] The competitiveness of a firm, industry or nation depends on numerous factors (such as wages, productivity, technological progress, infrastructure, human capital, and the like); and it can be defined and measured in a variety of ways. Given the difficulty of achieving general agreement on this issue, many authors have opted to define and even measure competitiveness on the basis of its results or consequences. For instance, according to Naby and Luthria (2002) competitiveness is the country's ability to maintain and expand its foreign-market share.

[3] There are two plausible explanations for this positive relation: first, the host nations are used as a platform to export a variety of goods to industrial countries; and second, FDI tends to promote broader-based trade in intermediate inputs between the parent company and subsidiary producers.

III
The model

The aim of this article is to assess the impact of various key variables on Mexican manufacturing exports, with a view to formulating policy recommendations. The choice of regressors has been determined by data availability, previous econometric work and economic theory. In principle the following expanded export equation will be estimated:[4]

$$X = f(\upsilon, W, Q, ED, FDI, R, CU, OP) \qquad (1)$$

where:
X = Volume of manufacturing exports.
υ = Labour productivity in manufacturing industry.
W = Average real hourly wages in manufacturing industry.
Q = Real effective exchange rate. This is a multilateral exchange rate, since it reflects changes in Mexico's international competitiveness, based on price indices in 111 of its trading partners.
ED = External demand for Mexican manufacturing exports. After performing various tests and estimations, it was decided use total United States manufacturing imports as a proxy for Mexico's external demand.[5]
FDI = Foreign direct investment in the manufacturing sector, measured in real terms.
R = Cost of domestic credit as measured by the weighted average of real interest rates on commercial paper.
CU = Percentage capacity utilization in manufacturing industry.
OP = Occupied personnel in manufacturing industry.

The variables X, υ, W, Q, ED, and OP and are measured as indices.[6] The next step was to gather quarterly data for each variable from January 1998 to December 2008,[7] during which period, in particular, the two models are stable and their residuals have well-behaved statistical properties. The data for all variables were seasonally adjusted using the *X12-ARIMA* procedure. Lastly, all series —except for the interest rate and capacity utilization— are expressed as natural logarithms.

In section III, equation (1) is estimated as a standard linear regression and then re-specified as a structural ARIMA model, once potentially redundant and/or omitted variables have been properly identified and residual tests performed. A structural ARIMA model can also be referred to as an ARIMAX (autoregressive integrated moving-average model with exogenous variables), since: (i) it differentiates the dependent variable (and the explanatory variables) by order of integration; (ii) it reflects a structural relation between the dependent and explanatory variables; and (iii) it includes autoregressive (AR) and moving-average (MA) terms to satisfactorily model the error process.

In section IV, a GVAR model is used to assess to what extent, if any, the empirical evidence is sensitive to changes in the econometric methodology applied. The use of both univariate and multivariate econometric techniques will allow two different perspectives of the response of manufacturing exports to changes in their basic determinants, such as labour productivity (υ), the real exchange rate (Q) and external demand (ED).

Installed capacity utilization (CU) and Occupied personnel (OP) are used as control variables.[8] Percentage installed capacity utilization helps to prevent a number of distortionary effects on the dependent variable, such as those generated when the gap between actual and potential output (in the manufacturing sector) narrows or widens. As a result, parameter estimates are more likely to discriminate between legitimate increases in manufacturing exports, reflecting higher labour productivity for

[4] This equation will later be subjected to a battery of specification tests, which will help identify potentially redundant and omitted variables.
[5] The external demand for Mexican manufactures can be adequately proxied by this variable, since 79.86% of Mexico's non-oil commodity exports went to the United States in 2008. Moreover, Mexican exports accounted for 10.28% of United States imports from the rest of the world in that year, making Mexico the United States' third most important supplier after China and Canada (Source: Banco de México, *Annual Report 2008*, pp. 47 and 48).
6 The remaining variables are measured as previously stated.
[7] See Appendix 1 for a detailed description of data sources and measurement units.
[8] Relevant control variables are usually included in an equation to reduce the risk of biased parameter estimates owing to omitted-variable problems.

instance, and spurious ones resulting from lower-than-expected domestic demand. Similarly, the employee index is intended to prevent the parameter estimates associated with labour productivity and wages from capturing the effects of possible mass layoffs, which could be triggered by trade liberalization policies, the acquisition of physical capital or the introduction of new technologies. See Jiménez et al (1998), Catao and Falsetti (2002) and Berretoni and Castresana (2007) for further details on this.

IV

Integration and co-integration analysis

Given that every unit root and stationarity test has arguments for and against, three different standard tests were used: the augmented Dickey-Fuller test (ADF, 1979), the Phillips-Perron test (PP, 1988), and the Kwiatkowski, Phillips, Schmidt, and Shin test (KPSS, 1992). In testing for unit roots (or for the presence of stationarity), an important issue is whether to include a constant and a linear trend in the test equation, or a constant only.[9] To resolve this matter satisfactorily the Hamilton procedure was used (Hamilton, 1994 p. 501), which consists of choosing the specification that provides the most realistic description of the data, under both the null and the alternative hypotheses. Each test equation was also subjected to a battery of *F*-type tests, which are based on the critical values that Dickey and Fuller (1981) and Dickey et al (1986) developed for that purpose.[10] The basic test results are reported in table 1.

As is well known, the ADF and PP tests contrast the null hypothesis of a unit root against the alternative hypothesis of stationarity, whereas the KPSS test compares the null hypothesis of stationarity against the alternative of non-stationarity. The rationale for including a stationarity test such as KPSS is that a failure to reject a unit root hypothesis is sometimes due to the lack of power of the ADF and PP tests.

It is not uncommon for unit root and stationarity tests to yield conflicting results, so the available empirical evidence needs to be globally assessed. Thus, in addition to formal tests, the consistency of the correlogram of a given time series with stationarity needs to be considered. In general terms, the picture that emerges is that real FDI and the real interest rate are stationary (or I(0)), whereas the other seven variables can reasonably be treated as integrated variables of order 1 (or I(1)) in levels, and stationary in terms of first differences.[11] This conclusion can also be validated to some degree by analysing the behaviour of the residuals in the univariate and multivariate models, which tends to improve when FDI and the rate of interest are expressed in terms of levels while the other variables are expressed as first differences.

Having determined the order of integration of each variable, the question of whether the I(1) variables share a long-run relation then needs to be considered. If the I(1) variables were in fact co-integrated, one could consider removing FDI and the interest rate from the system (as these variables are stationary), to estimate a seven-variable Vector Error-Correction (VEC) model. The Johansen (1995) co-integration tests will be used to check whether the non-stationary variables are co-integrated. Johansen uses two types of likelihood-ratio (LR) test statistics in testing for co-integration: the trace statistic, denoted by λ_{trace}, and the largest eigenvalue statistic, denoted by λ_{max}. Although both are *LR* statistics, they are not asymptotically distributed as a standard χ^2 distribution under the null hypothesis. Consequently, the non-standard critical values developed by MacKinnon, Haug and Michelis (1999) are used. The results of

[9] The third possibility is to omit both, but the KPSS test cannot be performed without the constant term.

[10] The null hypothesis of a unit root with no deterministic trend was tested against the alternative hypothesis of a stationary variable with a deterministic trend.

[11] Real FDI appears to be stationary for two reasons: (i) this particular variable does not display a clear growth trend in the 1998-2008 period, and (ii) its volatility does not seem to increase (or decrease) over time. By the same token, the tests performed consistently indicate that capacity utilization is a non-stationary (I(1)) variable. Although there are some atypical observations, neither their removal nor the use of moving-average filtering affects the outcome of the unit root and stationarity tests. Lastly, as shown below, the stability checks indicate that no structural change occurred during the sample period.

TABLE 1

Unit root and stationarity tests, 1998-2008

Variable	Specification of the test equation	ADF test statistic (Ho: unit root)	PP test statistic (Ho: unit root)	KPSS test statistic (Ho: stationarity)	Order of Integration
X_t	C and LT	-1.97	-1.97	0.11	I(1) or I(0)
ΔX_t	C	-4.66**	-4.63**	0.11	I(0)
υ_t	C and LT	-1.75	-1.92	0.13	I(1) or I(0)
$\Delta\upsilon_t$	C	-6.53**	-6.45**	0.28	I(0)
W_t	C and LT	1.24	0.54	0.23**	I(1)
ΔW_t	C	-5.97**	-6.30**	0.44	I(0)
Q_t	C	-1.90	-2.84	0.34	I(1) or I(0)
ΔQ_t	C	-4.89**	-4.97**	0.52*	I(1) or I(0)
ΔQ_t	None	-4.95**	-5.02**	N.D.	I(0)
ED_t	C and LT	-3.31	-2.13	0.11	I(1) or I(0)
ΔED_t	C	-4.35**	-4.99**	0.06	I(0)
FDI_t	C	-6.07**	-6.08**	0.22	I(0)
R_t	C	-3.43**	-4.65**	0.50*	I(1) or I(0)
ΔR_t	C	-7.40**	-7.96**	0.40	I(0)
ΔR_t	None	-7.58**	-7.88**	N.D.	I(0)
CU_t	C	-1.08	-1.19	0.57*	I(1)
ΔCU_t	C	-6.18**	-6.18**	0.16	I(0)
ΔCU_t	None	-6.25**	-6.28**	N.D.	I(0)
OP_t	C	-1.07	-0.52	0.61*	I(1)
ΔOP_t	C	-2.06	-2.04	0.33	I(1) or I(0)
ΔOP_t	None	-1.97*	-1.97*	N.D.	I(0)

Source: Author's estimations on the basis of quarterly data obtained from the databases of the National Institute of Statistics, Geography and Information of Mexico (INEGI), The Bank of Mexico, the United States Census Bureau and the consumer price index published by the United States Bureau of Labour Statistics, of the Department of Labor of the United States.
Notes.
ADF test: Augmented Dickey-Fuller test; pp: Phillips-Perron test; kpss: Kwiatkowski, Phillips, Schmidt and Shin test.
Ho: Null hypothesis.
C: Constant.
LT: Linear Trend.
— Asterisks * and ** denote rejection of the null hypothesis at the 5% and 1% significance levels, respectively.
— The symbols Δ and Δ^2 are the first- and second-difference operators, respectively.
— The ADF and PP test results are based on Mackinnon (1996) critical values and their associated one-sided *p*-values. In the ADF tests, the Schwarz Information Criterion is used to determine the lag length of each test equation. In the PP tests we control the bandwidth using the Newey-West bandwidth selection method and the Bartlett kernel.
— The KPSS test results are based on the critical values proposed by Kwiatkowski, Phillips, Schmidt and Shin (1992). To control the bandwidth, we use the Newey-West bandwidth selection method and the Bartlett kernel.

Johansen's co-integration trace and largest eigenvalue tests are reported in tables 2 and 3, respectively.

As shown in tables 2 and 3, these tests are done sequentially, starting at r=0 and ending at $r \leq k$-1, where r is the number of co-integrating equations, and k is the number of variables involved in the testing procedure. At the 5% significance level, the trace tests suggest that there are three co-integrating equations, whereas the largest eigenvalue tests suggest only two.[12] Despite these results, a VEC model is not feasible in this particular case for two reasons. First and foremost, after performing different normalizations for the co-integrating vectors, the conclusion is that they are not identifiable. In other words, given the sign and magnitude of the various parameter estimates, all of the candidate co-integrating equations are clearly inconsistent with economic theory, so they cannot be used for valid inference.[13] Secondly, under the Johansen methodology, manufacturing exports (X) turn out to be weakly exogenous with respect to the "adjustment parameters", which means that, in a VEC model, this particular variable can even be

[12] An important distinction between the trace and largest eigenvalue statistic is that the latter is based on a more restrictive alternative hypothesis, which is meant to increase the power of the test.

[13] This problem remained after eradicating potentially redundant variables through likelihood ratio tests.

TABLE 2

Johansen co-integration tests based on the trace statistic

Null hypothesis	Alternative hypothesis	λ_{trace}-statistic	5% critical value
$r = 0$	$r \geq 1$	171.9232*	125.6154
$r \leq 1$	$r \geq 2$	115.2237*	95.75366
$r \leq 2$	$r \geq 3$	70.36657*	69.81889
$r \leq 3$	$r \geq 4$	40.62333	47.85613
$r \leq 4$	$r \geq 5$	24.06045	29.79707
$r \leq 5$	$r \geq 6$	11.08441	15.49471
$r \leq 6$	$r \geq 7$	0.391198	3.841466

Source: Author's estimations on the basis of quarterly data obtained from the databases of the National Institute of Statistics, Geography and Information of Mexico (INEGI), The Bank of Mexico, the United States Census Bureau and the consumer price index published by the United States Bureau of Labor Statistics, of the Department of Labor of United States.
Notes:
— The letter *r* stands for the number of co-integrating equations.
— An asterisk * denotes rejection of the null hypothesis at the 5% significance level, given the critical values developed by MacKinnon, Haug and Michelis (1999).
— Trace tests indicate the existence of three co-integrating equations at the 5% significance level.
— The lag length of the VAR is 1. A constant in the co-integrating space and a linear trend in the data space were included.

TABLE 3

Johansen tests for co-integration based on the largest eigenvalue statistic

Null hypothesis	Alternative hypothesis	λ_{max}-statistic	5% critical value
$r = 0$	$r = 1$	56.69948*	46.23142
$r \leq 1$	$r = 2$	44.85710*	40.07757
$r \leq 2$	$r = 3$	29.74324	33.87687
$r \leq 3$	$r = 4$	16.56288	27.58434
$r \leq 4$	$r = 5$	12.97604	21.13162
$r \leq 5$	$r = 6$	10.69321	14.26460
$r \leq 6$	$r = 7$	0.391198	3.841466

Source: Author's estimations on the basis of quarterly data obtained from the databases of the National Institute of Statistics, Geography and Information of Mexico (INEGI), The Bank of Mexico, the United States Census Bureau and the consumer price index published by the United States Bureau of Labor Statistics, of the Department of Labor of the United States.
Notes:
— The letter *r* stands for the number of co-integrating equations.
— An asterisk * denotes rejection of the null hypothesis at the 5% significance level, given the critical values developed by MacKinnon, Haug and Michelis (1999).
— Largest eigenvalue tests indicate the existence of two co-integrating equations at the 5% significance level.
— The lag length of the VAR is 1. A constant in the co-integrating space and a linear trend in the data space were included.

removed from the system. Appendix 2 shows the weak exogeneity of this variable together with a brief technical explanation.[14]

Under these circumstances, a stationary GVAR model will be estimated in section VI. This means that the I(1) variables will be expressed in first differences, thus making every variable in the system I(0). A stationary GVAR model is statistically appropriate for short-term economic analysis.

[14] For further details, see Johansen (1995) and Patterson (2000, chapter 15).

V

Univariate time-series analysis

This section estimates a structural ARIMA model using least-squares algorithms, departing from an unadjusted regression equation, which in principle does not include autoregressive or moving average (ARMA) terms. The basic results are reported in table 4.

As noted above, except for FDI and the interest rate, all variables seem to be non-stationary, so FDI and the interest rate will enter the regression equation in levels, while the rest of the variables will be expressed in first differences. Moreover, except for the interest rate and percentage capacity utilization, all variables are expressed in natural logarithms, so their estimated coefficients should be interpreted as elasticities.

It is worth noting that including irrelevant variables reduces efficiency, while excluding relevant variables yields biased parameter estimates. But before dealing with potential specification errors, it needs to be stressed that both labour productivity (υ_t) and export demand (ED_t) have estimated coefficients that are positive and statistically significant at the 10% significance level. In contrast, the estimated coefficient of the real exchange rate is negative and statistically significant at the 5% level.

1. Adjusted regression equation

Several likelihood ratio tests for redundant and omitted variables were performed in an interactive process to make sure the final specification was appropriate. Although the primary strategy involved a general-to-specific search, several omitted-variables tests were performed to determine whether, and to what extent, any of the variables not originally considered contributed significantly to explaining the behaviour of manufacturing exports. The main candidate variables considered for possible inclusion were real-exchange rate volatility and gross fixed investment, but they failed to achieve statistical significance either individually or jointly. Similarly, to identify potentially redundant variables, the individual and joint statistical significance of the regressors already included in equation (1) were checked.[15]

[15] We also assessed alternative lag structures and interaction effects between the independent variables.

TABLE 4

Unadjusted equation for manufacturing exports

Dependent variable: ΔX_t				
Variable	Coefficient	Standard error	t-Statistic	Probability value
$\Delta\upsilon_t$	0.716876	0.384891	1.862545	0.0723
ΔW_t	-0.275609	0.243587	-1.131462	0.2668
ΔQ_t	-0.226815	0.094223	-2.407202	0.0224
ΔED_t	0.302333	0.151828	1.991284	0.0556
FDI_t	-0.006458	0.010792	-0.598433	0.5540
R_t	0.000538	0.001137	0.473169	0.6395
ΔCU_t	0.008229	0.005243	1.569348	0.1271
ΔOP_t	0.870729	0.593613	1.466829	0.1528
Intercept term	0.022091	0.038190	0.578452	0.5673
Adjusted R^2	0.501665	Standard error of regression		0.020093
DW statistic	2.249753	Probability value (F-statistic)		0.000172

Source: Author's estimations on the basis of quarterly data obtained from the databases of the National Institute of Statistics, Geography and Information of Mexico (INEGI), The Bank of Mexico, the United States Census Bureau and the consumer price index published by the United States Bureau of Labor Statistics, of the Department of Labor of the United States.

The general equation was thus gradually reduced to a range of specific equations, and the final model was chosen on the basis of diagnostic statistics and economic theory. Lastly, the procedure developed by Hannan and Rissanen (1982) was applied to identify a suitable ARIMA model for the regression residuals; and this eventually gave rise to a structural ARIMA (1,1,1) model with four explanatory variables, which seems to be well-grounded for three major reasons: (i) the regression residuals display a "normal" white-noise process; (ii) the regression equation is stable; and (iii) explanatory power as measured by the adjusted R^2 rises considerably. Such a model can be represented as follows:

$$\Delta X_t = b_0 + b_1 \Delta \upsilon_t + b_2 \Delta Q_t + b_3 \Delta ED_t + b_4 \Delta OP_t + u_t \quad (2)$$

where $u_t = \phi_1 u_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$ and ε_t is a normal white-noise process. Table 5 shows the basic estimation results.

The adjusted equation shows that the estimated coefficients for labour productivity and external demand are positive and statistically significant at the 5% and 1% levels, respectively. Assuming a variable's growth rate can be adequately proxied by its first difference, we could say that a one-percentage-point increase in the rate of labour productivity growth will raise the growth rate of manufacturing exports by 83 basis points (0.83 of a percentage point). Similarly, a one-percentage-point drop in the rate of growth of external demand will lower the manufacturing exports growth rate by 49 basis points. In view of these results, a coherent policy package designed to increase worker productivity might be effective in alleviating the negative effects of weaker external demand (resulting from an international recession) on Mexican manufacturing exports.

The tests also suggest that real exchange-rate depreciation may reduce the volume of exports in the short run. In other words, the estimated coefficient of the real exchange rate is negative and statistically significant at the 1% level. As noted above, this result is consistent with the notion that a real depreciation, especially in developing countries, produces two opposing effects: it makes exports cheaper in terms of foreign currency; but it also raises the local-currency cost of imported intermediate inputs. Consequently, international competitiveness might worsen if the latter effect dominates.

Lastly, the parameter associated with the number of employees occupied personnel is positive and significantly different from zero at the 1% level, thereby suggesting that manufacturing firms make a deliberate effort to sell their products abroad.

2. Diagnostic tests

A number of tests were done to ensure that the residuals from the adjusted regression equation were free from serial correlation, heteroscedasticity, autoregressive conditional heteroscedasticity (ARCH) and systematic departures from normality. Table 6 sets out the basic results for some of the tests performed.

TABLE 5

Adjusted equation for manufacturing exports

Dependent variable: ΔX_t				
Variable	Coefficient	Standard error	*t*-Statistic	Probability value
$\Delta \upsilon_t$	0.833277	0.314149	2.652487	0.0117
ΔQ_t	-0.193773	0.041048	-4.720681	0.0000
ΔED_t	0.486660	0.099567	4.887755	0.0000
ΔOP_t	1.277885	0.287783	4.440447	0.0001
Intercept term	0.000771	0.002877	0.267881	0.7903
AR(1)	0.428219	0.178590	2.397779	0.0217
MA(1)	-0.997398	0.125894	-7.922543	0.0000
Adjusted R^2	0.636163	Standard error of regression		0.018380
DW statistic	2.015594	Probability value (*F*-statistic)		0.000000

Source: Author's estimations on the basis of quarterly data obtained from the Institute of Statistics, Geography and Information of Mexico (INEGI), The Bank of Mexico, the United States Census Bureau and the consumer price index published by the United States Bureau of Labor Statistics, of the Department of Labor of the United States.

As can be seen in table 6, the behavior of the residuals from the adjusted regression model is broadly consistent with normal white noise. In other words, the Lagrange multiplier (LM) test for serial correlation indicates the absence of serial correlation up to lag order five,[16] whereas the White heteroscedasticity test shows that the errors are homoscedastic. Similarly, the ARCH test suggests that the magnitude of the residuals in the present is unrelated to their magnitude in the past; and, lastly, the outcome of the Jarque-Bera normality test rules out the presence of serious departures from normality.

To determine whether the adjusted manufacturing export equation remains unchanged throughout the period, we used Chow's test with multiple break-points, the results of which are shown in table 7.

Probability values corresponding to the three test statistics (the *F*-, log-likelihood ratio and Wald statistics) show that the null hypothesis of "no structural change" is also far from being rejected, which means that the parameters of the model are stable across the four sub-samples obtained from the break-points designated in the data set.

Similarly, the ARIMA process of the estimated regression equation is stationary and invertible. Stationarity stems from the fact that the inverse root of the autoregressive component (AR(1)) lies within the unit circle, whereas invertibility arises from the fact that the inverse root of the moving-average component (MA(1)) lies within the unit circle. See table 8 for details.

[16] The correlogram of the residuals is also consistent with the absence of autocorrelation up to lag order 20.

TABLE 6

Diagnostic tests for the residuals of the adjusted regression equation, 1998-2008

Type of test	Null hypothesis	Probability value
Serial correlation[a]	No serial correlation up to lag order five	0.2370
Heteroscedasticity[b]	Homoscedasticity	0.5880
ARCH[c]	No ARCH up to lag order five	0.2335
Normality[d]	Normality	0.6761

Source: Author's estimations on the basis of quarterly data obtained from the databases of the National Institute of Statistics, Geography and Information of Mexico (INEGI), The Bank of Mexico, the United States Census Bureau and the consumer price index published by the United States Bureau of Labor Statistics, of the Department of Labor of the United States.

Notes:
[a] Breusch-Godfrey test for serial correlation up to lag order five.
[b] White heteroscedasticity tests with two lags and no cross terms.
[c] Autoregressive Conditional Heteroscedasticity (ARCH) test with five lags.
[d] Jarque-Bera normality test.

TABLE 7

Chow's test with multiple break-points, 1998-2008
(Null hypothesis: no structural change)

F-statistic	0.367037	Prob. Value F(21,16)	0.9835
Log likelihood ratio statistic	17.30142	Prob. Value Chi-Squared (21)	0.6927
Wald statistic	20.44749	Prob. Value Chi-Squared (21)	0.4931

Source: Author's estimations on the basis of quarterly data obtained from the databases of the National Institute of Statistics, Geography and Information of Mexico (INEGI), The Bank of Mexico, the United States Census Bureau and the consumer price index published by the United States Bureau of Labor Statistics, of the Department of Labor of the United States.
Note: Break-points in the sample: 2000:03, 2003:02 and 2006:01.

TABLE 8

Inverse roots of the autoregressive and moving-average components of the estimated regression equation

Inverted autoregressive (AR(1)) root	0.428219
Inverted moving average (MA(1)) root	0.973980

Source: Author's estimations on the basis of quarterly data obtained from the databases of the National Institute of Statistics, Geography and Information of Mexico (INEGI), The Bank of Mexico, the United States Census Bureau and the consumer price index published by the United States Bureau of Labor Statistics, of the Department of Labor of the United States.
Note: All inverse roots lie within the unit circle, so the estimated ARMA process is stationary and invertible.

VI

Multivariate time-series analysis

Given the difficulty of identifying the co-integrating equations and the weak exogeneity of manufacturing exports, a generalized VAR (GVAR) model with stationary variables will be estimated. Such a model is specified as follows:

$$Y_t = \mathrm{B}_0 + \mathrm{B}_1 Y_{t-1} + \mathrm{B}_2 Y_{t-2} + ,...,+ \mathrm{B}_p Y_{t-p} + \varepsilon_t, \qquad (3)$$

where $Y_t = [\Delta X_t, \Delta \upsilon_t, \Delta W_t, \Delta Q_t, \Delta ED_t, FDI_t, R_t, \Delta CU_t, \Delta OP_t]'$ is a 9x1 vector of variables, B_0 is a 9x1 vector of intercept terms, and $\{B_i,\ i=1,\ 2,\ldots,\ \mathrm{p}\}$ are 9x9 coefficient matrices. Moreover, ε_t stands for a 9x1 vector of innovations that behaves according to the following assumptions: $E(\varepsilon_t)=0$ and $E(\varepsilon_t \varepsilon_t')=\Lambda$ for every t, where $\Lambda = \{\sigma_{ij},\ i, j = 1, 2, \ldots, 9\}$ is a non-diagonal positive definite matrix, and $E(\varepsilon_t \varepsilon_t')=0$ for every t and s, $t \neq s$, in the set 1, …, T. We can also view as being multivariate normally distributed, given the test results shown below.[17]

Since the covariance matrix of innovations (Λ) is non-diagonal, the elements of ε_t are "contemporaneously" correlated. Sims (1980) uses a Cholesky decomposition of matrix Λ to orthogonalize the VAR residuals; but the resulting impulse response functions and variance decompositions are sensitive to the ordering of the equations in the VAR model. The problem lies in the decomposition technique, which is recursive rather than structural. This produces an asymmetric structure, where a shock to a given variable will have a contemporaneous effect on that variable and those that follow it in the ordering. The variables that lead the hypothetical ordering, in contrast, will be affected only through the VAR lag structure.[18]

The generalized impulse-response function approach, developed by Pesaran and Shin (1998), does not have this drawback and can be used to construct an orthogonal set of innovations that is invariant to the ordering of the variables in the VAR model. For expositional convenience, a first-order VAR model is posited first:

$$Y_t = \mathrm{B}_0 + \mathrm{B}_1 Y_{t-1} + \varepsilon_t \qquad (4)$$

Next, we need to assume that all the inverse roots of the characteristic autoregressive polynomial (in other words, all the roots of $\left| I_9 - \sum_{i=1}^{P} \mathrm{B}_i L^i \right| = 0$, where L is the lag operator) lie within the unit circle, so the stability condition is satisfied. Equation (4) can then be rewritten as an infinite moving-average representation:

$$Y_t = \mu + \sum_{i=0}^{\infty} \mathrm{B}_1^i \varepsilon_{t-i} \qquad (5)$$

where $\mu = (I_9 - A_1)^{-1} \mathrm{B}_0$ stands for the mean of the process.

[17] Put briefly, it can be said that $\varepsilon_t \sim N_9(0, \Lambda)$ and is free from serial correlation. In principle, VAR residuals are uncorrelated across time but they are not uncorrelated across equations.

[18] Generally, a shock to Y_{it} will have a contemporaneous impact on Y_{jt} only if $j \neq i$.

According to Pesaran and Shin (1998), an impulse-response function can be interpreted as the difference between the expected value of a variable at time $t + n$, resulting from a shock that occurred at time t, and its expected value at time $t + n$ in the absence of such a shock. The expected value is derived from the economy's known history up to time t - 1. For instance, assuming that a magnitude- δ shock affects the jth equation of vector Y_t, then the vector of Generalized Impulse Response Functions (GIRF) is given by:

$$GIRF_Y(n,\delta,\Omega_{t-1}) = E(Y_{t+n}\left|\varepsilon_{jt}=\delta,\Omega_{t-1}\right.) - E(Y_{t+n}\left|\Omega_{t-1}\right.) \quad (6)$$

where the matrix Ω_{t-1} represents all available information on the economy's history up to time t-1. Combining (6) and (5) it can be inferred that $GIRF_Y(n,\delta,\Omega_{t-1}) = \mathrm{B}_1^n\delta$. Under the previously stated assumption that the vector of innovations (ε_t) is multivariate normally distributed, both Pesaran and Shin (1998) and Koop et al. (1996) show that:

$$E(\varepsilon_t\left|\varepsilon_{jt}=\delta\right.) = (\sigma_{1j},\sigma_{2j},...,\sigma_{9j})'\sigma_{jj}^{-1}\delta = \Lambda e_j\sigma_{jj}^{-1}\delta \quad (7)$$

where e_j is a hypothetical (9x1) vector of innovations with 1 in the jth row and zeros everywhere else. Consequently, the unscaled GIRF vector is given by:

$$\left(\frac{\mathrm{B}_1^n\Lambda e_j}{\sqrt{\sigma_{jj}}}\right)\left(\frac{\delta}{\sqrt{\sigma_{jj}}}\right) \quad (8)$$

Lastly, deriving the scaled GIRF vector, denoted $\psi_j^G(n)$, simply means setting $\delta = \sqrt{\sigma_{jj}}$. That is,

$$\psi_j^G(n) = \left(\frac{\mathrm{B}_1^n\Lambda e_j}{\sqrt{\sigma_{jj}}}\right) \quad (9)$$

Note that $\psi_j^G(n)$ measures the effect of a one-standard-deviation shock to the jth equation. Such a shock takes place at time t and affects the expected values of vector Y at time $t+n$, where n= 0, 1, 2, …,.

1. Empirical model

For the purpose of building an adequate empirical model a variety of specifications were tested, not only for the data set (the number and choice of variables in the system) but also for the lag structure of the model. The lag length of a VAR model is critical because the behaviour of the residuals and empirical results are sensitive to the order of the model (in other words the number of lags chosen). Moreover, there are complex trade-offs between the number of lags and the dimension of the VAR model. Thus, after using different lag-length selection criteria (and different parameterizations), the conclusion was that one lag for each variable in each equation allows for adequate dynamic adjustment and efficient estimation, since it is the smallest number of lags producing well-behaved residuals.

By the same token, the trade-off relationships between the data set and the model's lag structure led to the exclusion of wages, FDI and the interest rate. The rationale for excluding such variables is twofold: firstly, shocks to these three variables do not produce statistically significant impulse-response functions; and secondly, their inclusion generates serial correlation and departures from normality in the VAR residuals, which could not be solved through alternative lag structures and/or the use of dummy variables. In this perspective, the final specification is a six-variable VAR model: $Y_t = [\Delta X_t, \Delta \upsilon_t, \Delta Q_t, \Delta ED_t, \Delta CU_t, \Delta OP_t]'$. The VAR parameter estimates are shown in appendix 3.

2. Testing for model adequacy

As we shall see, a one-lag VAR model with the previous specification (or set of variables) eliminates serial correlation, heteroscedasticity and departures from normality in the VAR residuals. It will also be shown that this particular specification satisfies the stability condition. Table 9 shows the results of the multivariate serial correlation LM tests. The LM statistics and their corresponding p-values suggest the absence of serial correlation up to lag order five.[19]

Moreover, the multivariate version of the White heteroscedasticity test indicates that the null hypothesis of homoscedasticity cannot be rejected in any of the cases, at either the 5% or the 10% significance levels. See table 10 for further details.

[19] We also estimated the matrix of pair-wise cross-correlograms (with two-standard error bounds) for the VAR residuals, which are broadly consistent with the absence of autocorrelation.

TABLE 9

Multivariate serial correlation *LM* Tests, 1998-2008

(Null hypothesis: there is no serial correlation at lag order (p))

Lag order (*p*)	*LM*-Statistics	Prob.
1	34.55223	0.5375
2	46.14492	0.1198
3	34.80235	0.5255
4	30.63705	0.7213
5	38.66307	0.3503

Source: Author's estimations on the basis of quarterly data obtained from the databases of the National Institute of Statistics, Geography and Information of Mexico (INEGI), The Bank of Mexico, the United States Census Bureau and the consumer price index published by the United States Bureau of Labor Statistics, of the Department of Labor of the United States.
Note: Probabilities from the Chi-Squared distribution with 49 degrees of freedom.

LM: Lagrange multiplier.

TABLE 10

White heteroscedasticity tests for VAR residuals, 1998-2008

(Null hypothesis: homoscedasticity)

Joint test					
Chi-Squared statistic (χ^2)		Degrees of freedom		Prob.	
270.2510		252		0.2051	
Individual components					
Dependent	R-squared	F(12,31)	Prob.	Chi-sq(12)	Prob.
res1*res1	0.240216	0.816754	0.6323	10.56949	0.5661
res2*res2	0.255522	0.886659	0.5688	11.24296	0.5082
res3*res3	0.320257	1.217121	0.3155	14.09129	0.2949
res4*res4	0.367625	1.501792	0.1763	16.17548	0.1833
res5*res5	0.386347	1.626431	0.1351	16.99928	0.1496
res6*res6	0.346877	1.372021	0.2312	15.26259	0.2274
res2*res1	0.176493	0.553658	0.8613	7.765701	0.8032
res3*res1	0.255840	0.888144	0.5675	11.25697	0.5070
res3*res2	0.163653	0.505496	0.8951	7.200735	0.8441
res4*res1	0.349548	1.388266	0.2236	15.38013	0.2213
res4*res2	0.375232	1.551532	0.1587	16.51019	0.1690
res4*res3	0.233105	0.785227	0.6613	10.25661	0.5935
res5*res1	0.340077	1.331262	0.2512	14.96337	0.2434
res5*res2	0.336725	1.311480	0.2615	14.81589	0.2517
res5*res3	0.181716	0.573681	0.8462	7.995512	0.7855
res5*res4	0.386454	1.627160	0.1349	17.00396	0.1494

Source: Author's estimations on the basis of quarterly data obtained from the databases of the National Institute of Statistics, Geography and Information of Mexico (INEGI), The Bank of Mexico, the United States Census Bureau and the consumer price index published by the United States Bureau of Labor Statistics, of the Department of Labor of the United States.
Note: The test was performed with levels and squares only (no cross terms were included).

Res: residuals.

Table 11 shows the outcome of the multivariate normality tests. The Jarque-Bera statistics and related probability values indicate that, by and large, VAR residuals follow a multivariate normal distribution.

To show that the model satisfies the stability condition, the "inverse roots" of the characteristic autoregressive polynomial were calculated. As reported in table 12, all such roots have an absolute value (modulus) of less than 1 and lie within the unit circle, meaning that the "overall" model is stable and hence stationary. In conclusion, the estimated VAR model is adequate, since the residuals are well-behaved and the lag structure is stable.

TABLE 12

Stability condition test, 1998-2008

(Inverse roots of characteristic autoregressive polynomial)

Root	Modulus
0.788565	0.788565
0.364357	0.364357
-0.263775 - 0.217130i	0.341647
-0.263775 + 0.217130i	0.341647
0.153255	0.153255
0.000574	0.000574

Source: Author's estimations on the basis of quarterly data obtained from the databases of the National Institute of Statistics, Geography and Information of Mexico (INEGI), The Bank of Mexico, the United States Census Bureau and the consumer price index published by the United States Bureau of Labor Statistics, of the Department of Labor of the United States.
Note: All inverse roots have an absolute value (modulus) < 1, so the stability condition is fulfilled.

3. Generalized impulse-response functions

In this subsection a set of twelve-month impulse-response functions with 95% confidence intervals are presented, which depict the dynamic response of manufacturing exports to innovations in the different system variables. Each innovation (or shock) should be understood here as a one-standard-deviation increase in the variable in question, which is unexpected and transitory since it lasts for one period only. Moreover, the responses are generalized in that they do not depend on the VAR orderings, and the confidence bands are useful for establishing statistical significance. In other words, the impulse-response function is statistically significant at the 5% level only in the period for which the confidence interval excludes the value zero (see figure 1).

As shown in figure 1, a positive labour-productivity shock raises manufacturing exports upon impact, but the effect fades around the second month. In contrast, a real depreciation of the domestic currency

TABLE 11

Normality tests for VAR residuals, 1998-2008

(Null hypothesis: residuals follow a multivariate normal distribution)

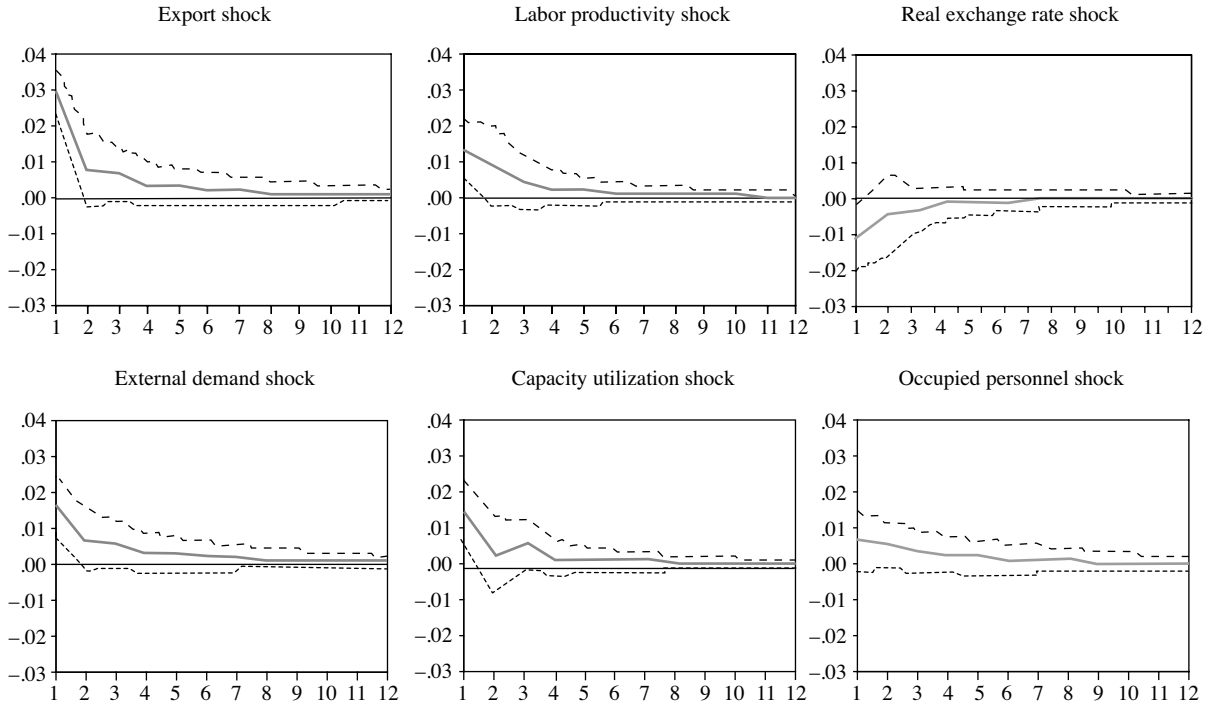
Joint test		
Jarque-Bera statistic	Degrees of freedom	Prob.
12.11528	12	0.4365

Individual components			
Component	Jarque-Bera Statistic	Degrees of freedom	Prob.
ΔX_t	0.145183	2	0.9300
$\Delta \upsilon_t$	2.580443	2	0.2752
ΔQ_t	2.407762	2	0.3000
ΔED_t	1.878920	2	0.3908
ΔCU_t	0.954520	2	0.6205
ΔOP_t	4.148453	2	0.1257

Source: Author's estimations on the basis of quarterly data obtained from the databases of the National Institute of Statistics, Geography and Information of Mexico (INEGI), The Bank of Mexico, the United States Census Bureau and the consumer price index published by the United States Bureau of Labor Statistics, of the Department of Labor of the United States.
H. Lütkepohl, *New Introduction to Multiple Time Series Analysis*, Nueva York, Springer-Verlag, 2006.
Note: The Cholesky Orthogonalization method was used (Lütkepohl, 2006: 174-181).

FIGURE 1

Dynamic response of manufacturing exports to shocks with 95% confidence intervals



Source: Author's estimations on the basis of quarterly data obtained from the databases of the National Institute of Statistics, Geography and Information of Mexico (INEGI), The Bank of Mexico, the United States Census Bureau and the consumer price index published by the United States Bureau of Labor Statistics, of the Department of Labor of the United States.

reduces the volume of exports, although this effect does not persist over time. As the reader may recall, these two findings are consistent with the univariate time-series results reported in table 5. As noted above, real exchange-rate depreciation affects exports through both demand- and supply-side channels. On the demand side it makes exports cheaper in terms of foreign currency, thereby strengthening international competitiveness; but, on the supply side, it raises the local-currency cost of imported capital goods and intermediate inputs, which inevitably undermines international competitiveness. Following this line of reasoning, the empirical evidence suggests that the net effect of a real depreciation on international competitiveness and exports is negative, at least in the short term.

External-demand shocks boost manufacturing exports, although the positive effect dissipates around the second month. Lastly, capacity-utilization shocks increase rather than reduce exports, which is consistent with the self-selection hypothesis in the sense that manufacturing firms make a conscious effort to sell in foreign markets.[20]

[20] In contrast, a rise in capacity utilization resulting from higher-than-expected domestic demand would be consistent with a declining trend in exports.

VII

Implications for economic policy

The estimation results consistently indicate that manufacturing exports respond positively not only to external demand but also to labour productivity. According to the elasticities obtained from the estimated structural ARIMA model, labour productivity has a larger impact on exports than external demand. Nonetheless, the impulse-response functions resulting from the GVAR model suggest that external demand is slightly more influential than worker productivity. In any event, labour productivity seems to be one of the key determinants of manufacturing exports; so a relevant economic-policy implication is that an external demand contraction (brought about by an international recession) could, to some extent, be offset by increasing worker productivity.

In this context, Mexico needs to develop a comprehensive, coherent and cost-effective policy package to enhance labour productivity in the manufacturing sector. This package should encompass temporary training programmes to develop certain skills at critical junctures (especially among the unemployed, to enable them to meet employers' demands in difficult times), along with continuous training programmes aimed at career advancement.

There are many empirical results showing that appropriate job-related training and quality formal schooling can yield significant productivity gains.[21] Generally speaking, education provides knowledge, skills and abilities that are useful in terms of raising wages, output per worker, and output per labour-hour. Notwithstanding the importance of education for coping with knowledge-based competition in the global marketplace, employment-related training courses are more likely to be successful as a short-term countercyclical measure in an international economic slowdown.[22] In other words, without diverting resources from longer-term human capital development through the formal education system, the Government could (in conjunction with other initiatives) intensify efforts through a wide range of training programmes to counterbalance lower-than-expected external demand for Mexican manufactures.

In this regard, it is essential to properly identify and categorize specific training needs, not only in the Mexican manufacturing sector as a whole but in every subsector and industry group. The underlying principle is that the content of training (which needs to target the correct knowledge, skills and abilities on a case-by-case basis) may prove to be as important as the amount of training provided to the workforce. Moreover, as employers' demands evolve along with technological change and innovation, an accurate follow-up system needs to be set up to gather precise information on new training requirements and future labour-market trends. Complementary research is needed to assess the effects of different types of training on workers' productivity: on-the-job/off-the-job, workshops, courses, training plans, and so forth. Black and Lynch (1996), for example, argue that off-the-job training may have a larger impact on productivity than on-the-job training, since workers who are trained outside the factory may be receiving more advanced skills. In short, a research endeavour of this sort could yield valuable insights into the right mix of policy instruments to cope with sudden slumps in external demand, thereby making labour-market policies more appropriate and effective.

More thorough knowledge of training-related productivity gains may also identify a framework for setting and attaining longer-term programmatic targets, in a joint effort involving all relevant stakeholders, including workers, employers, public- and private-sector training providers, sector bodies, and the like. As a final comment, even though short- and long-term labour-market policies should be designed and implemented in an integrated way, the econometric work and findings described in this paper basically relate to short-run policy-making.

Another critical issue concerns the adverse effects of real exchange-rate depreciation on manufacturing exports. As noted above, this evidence is consistent with the view that exchange-rate movements influence not only the demand side but also the supply side. In

[21] Authors such as Bartel (1992), Mincer (1994), Black and Lynch (1996), Dearden et al. (2005), Mungaray and Ramírez (2007), and Padilla and Juárez (2006) reach this conclusion on the basis of different productivity measures.

[22] According to Maglen (1995), training is more specific and more contemporaneously correlated with work-place performance, whereas education is more general and is usually imparted before (often long before) the individual in question joins the labour market.

such circumstances, the negative impact on exports of real currency-depreciation reflects the heavy reliance of manufacturing firms on foreign suppliers, mainly of intermediate inputs for which the local-currency cost rises as the Mexican peso weakens against the dollar. A well-known policy recommendation, which has had limited success thus far, is to improve productive capacity and efficiency in the import-substituting sector, with a view to reducing the import content of manufactured goods. Moreover, to raise the domestic content of manufactures, additional and more efficient production chains need to be formed and consolidated —between large manufacturing exporters and small and medium-sized local businesses.

VIII
Conclusions

This article has evaluated various determinants of Mexican manufacturing exports using two complementary econometric approaches: (i) the univariate time series approach, which makes it possible to estimate elasticities; and (ii) the generalized VAR approach, which provides a fully dynamic perspective by estimating impulse-response functions. Model adequacy was ensured in both cases through specification, residual and stability tests. The testing tools used varied according to the nature of each econometric modelling method.

As some of the findings are robust to changes in the econometric methodology used, two solid conclusions can be reached. Firstly, manufacturing exports are positively related to labour productivity and external demand. Secondly, real exchange-rate depreciation reduces rather than increases manufacturing exports, at least in the short term. This evidence is consistent with the notion that the Mexican economy is highly dependent on imported capital and intermediate goods. Consequently, real currency depreciation generates not only demand-side but also strong supply-side effects. On the demand side, it makes manufacturing exports cheaper in terms of foreign currency, and therefore strengthens international competitiveness. On the supply side, however, it raises the local-currency cost of imported intermediate inputs, thereby weakening international competitiveness. The empirical results show that the negative (supply-side) effects tend to dominate in the short run.

It is worth recalling that manufacturing exports display a positive relationship with occupied personnel under the univariate analysis, and with capacity utilization under the multivariate analysis. This was found to be consistent with the self-selection hypothesis in the sense that firms make a conscious effort to sell in foreign markets. Since a firm's desire to export determines many investment and production decisions, higher capacity utilization may lead to an increase (rather than a decrease) in manufacturing exports.

Two important economic policy implications can be drawn from this study. Firstly, real exchange-rate depreciation may worsen rather than improve export performance in the short term. Secondly, a comprehensive and coherent package to enhance labour productivity could significantly increase manufacturing exports. Moreover, the adverse effects of an international recession on Mexican exports could, to some degree, be offset by raising worker productivity. Thus, in the face of a recession in the United States it might be advantageous for Mexico to invest more in training and education.

APPENDIX 1

TABLE A1

Description of data sources and measurement units

Statistical Series	Source
Manufacturing exports quantity Index of (X).	National Institute of Statistics, Geography and Informatics of Mexico (INEGI).
Manufacturing industry labour productivity index (υ).	INEGI.
Real effective exchange rate index (Q)	Banco de México.
External demand for Mexican manufacturing exports (ED), measured by a quantity index of total U.S. manufacturing imports.	United States Census Bureau.
Real foreign direct investment (FDI).	Ministry of the Economy. The price index used as a deflator was the Consumer Price Index published by the United States Bureau of Labor Statistics.
Cost of domestic credit, as measured by the weighted average of real interest rates on commercial paper (R).	INEGI.
Percentage capacity utilization in manufacturing industry (CU).	INEGI.
Index of occupied personnel in manufacturing industry (OP).	INEGI.

Source: Prepared by the author.
Notes:
1. Time interval: January 1998-December 2008.
2. Frequency: Quarterly.

APPENDIX 2

A VEC model is a VAR model that is restricted to account for one or more co-integrating relations. Equation (3) can thus be re-parameterized to express the following VEC model:

$$\Delta Y_t = B_0 + \Pi Y_{t-1} + \Gamma_1 \Delta Y_{t-1} + \Gamma_2 \Delta Y_{t-2} + ,...,+ \Gamma_{(p-1)} \Delta Y_{t-(p-1)} + \varepsilon_t \quad (10)$$

where $\Pi = \sum_{i=1}^{p} B_i - I$ and $\Gamma_i = -\sum_{j=i+1}^{p} Bj$. The fourth implication of the Granger Representation Theorem (Engle and Granger, 1987) establishes that if a *k*-dimensional vector of I(1) variables involves one or more co-integrating relations, then a VEC model exists that can be correctly represented by the above equation. Formally, if the variables in vector Y_t are I(1) and the rank of the coefficient matrix Π (denoted *r*) is small (that is, if $r<k$), then it can be shown that $k x r$ matrices α and β (both with rank *r*) exist, such that (i) $\Pi = \alpha\beta'$ and (ii) $\beta' Y_{t-1}$ is a stationary system. Similarly, since each column of β is a co-integrating vector, it can be stated that $\beta' Y_{t-1}$ contains the *r* long-run equilibrium relations among the *k* variables. Such long-run relations in $\beta' Y_{t-1}$ are expressed in the form of error-correction terms (ECT). The matrix α, on the other hand, contains the adjustment or short-run coefficients of the VEC model, the values of which determine the speed at which equilibrium is restored following a disturbance. The purpose of capturing the short-run dynamics is also served by coefficient matrices $\Gamma_1, \Gamma_2, \ldots, \Gamma_{(p-1)}$.

In this context, if we assume that $\beta' Y_{t-1}$ contains two co-integrating relations (denoted $ECT1_{t-1}$ and $ECT2_{t-1}$, respectively), as the largest-eigenvalue tests suggest, then the resulting matrix of estimated adjustment coefficients is the following:

TABLE A2

Matrix of estimated adjustment coefficients
(The α matrix)

Equation	$ECT1_{t-1}$	$ECT2_{t-1}$
ΔX_t	0.071304 (0.23407) [0.30462]	-0.313378 (0.97010) [-0.01840]
$\Delta \upsilon_t$	0.306825*** (0.06186) [4.95999]	-1.275998*** (0.25639) [-4.99358]
ΔW_t	0.248226 (0.20551) [1.20785]	-1.01029 (0.85173) [-1.16447]
ΔQ_t	-0.503782 (0.35069) [-2.60101]**	2.097696 (1.45342) [1.44328]
ΔED_t	-0.236342 (0.19946) [-0.92653]	0.965415 (0.82666) [1.16785]
ΔCU_t	-1.010744 (4.77390) [-0.32095]	2.609343 (19.7853) [0.13188]
ΔOP_t	0.017312 (0.03268) [0.52974]	-0.068394 (0.13544) [-0.50498]

Source: Author's estimations on the basis of quarterly data obtained from the databases of the National Institute of Statistics, Geography and Information of Mexico (INEGI), The Bank of Mexico, the United States Census Bureau and the consumer price index published by the United States Bureau of Labor Statistics, of the Department of Labor of the United States.
Notes: Standard errors in ( ) and *t*-statistics in [ ].
Asterisks *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively.
The results shown in table A2 indicate that none of the adjustment parameters is statistically significant in the export equation (ΔX_t), which is why manufacturing exports were found to be weakly exogenous with respect to the α matrix.

APPENDIX 3

TABLE A3

Parameter estimates for the final VAR model specification

	ΔX_t	$\Delta \upsilon_t$	ΔQ_t	ΔED_t	ΔCU_t	ΔOP_t
ΔX_{t-1}	0.188165 (0.22829) [0.82422]	0.044354 (0.08525) [0.52026]	-0.203017 (0.37362) [-0.54338]	0.251443 (0.20776) [1.21028]	2.977106 (6.74198) [0.44158]	0.053435 (0.03536) [1.51128]
$\Delta \upsilon_{t-1}$	0.807990 (0.53519) [1.50973]	-0.019072 (0.19986) [-0.09543]	1.035099 (0.87588) [1.18178]	1.476432 (0.48704) [3.03144]	31.52161 (15.8052) [1.99439]	0.040187 (0.08289) [0.48484]
ΔQ_{t-1}	-0.044268 (0.12766) [-0.34678]	-0.000724 (0.04767) [-0.01519]	0.058313 (0.20892) [0.27912]	0.032161 (0.11617) [0.27684]	2.393695 (3.76994) [0.63494]	-0.027179 (0.01977) [-1.37470]
ΔED_{t-1}	0.114752 (0.22154) [0.51797]	0.015873 (0.08273) [0.19186]	0.366712 (0.36257) [1.01143]	0.137367 (0.20161) [0.68135]	2.527265 (6.54251) [0.38628]	0.018223 (0.03431) [0.53110]
ΔCU_{t-1}	-0.009063 (0.00671) [-1.35024]	0.001414 (0.00251) [0.56414]	-0.003511 (0.01099) [-0.31962]	-0.001987 (0.00611) [-0.32534]	-0.305305 (0.19823) [-1.54018]	-0.000396 (0.00104) [-0.38058]
ΔOP_{t-1}	1.209608 (0.80273) [1.50687]	-0.230256 (0.29977) [-0.76810]	-1.264143 (1.31373) [-0.96226]	0.329078 (0.73051) [0.45048]	1.342715 (23.7061) [0.05664]	0.640945 (0.12432) [5.15549]
C	0.005182 (0.00909) [0.57002]	0.005044 (0.00340) [1.48552]	-0.019301 (0.01488) [-1.29715]	0.005365 (0.00827) [0.64842]	-0.262657 (0.26850) [-0.97824]	-0.003210 (0.00141) [-2.27948]
R^2	0.254138	0.049825	0.092319	0.373294	0.157935	0.672209
Adjusted R^2	0.118527	-0.122934	-0.072714	0.259347	0.004833	0.612611

Source: Author's estimations on the basis of quarterly data obtained from the databases of the National Institute of Statistics, Geography and Information of Mexico (INEGI), The Bank of Mexico, the United States Census Bureau and the consumer price index published by the United States Bureau of Labor Statistics, of the Department of Labor of the United States.
Note: Standard errors in ( ) and t-statistics in [ ].

(Original: English)

Bibliography

Banco de México (2008), *Informe anual 2008 del Banco de México*, Mexico City.

Bartel, A. (1992), "Productivity gains from the implementation of employee training programs", *NBER Working Papers*, No. 3893, Cambridge, Massachusetts, National Bureau of Economic Research.

Bayoumi, T. and B. Eichengreen (1997), "Is regionalism simply a diversion? Evidence from the EC and EFTA", *Regionalism versus Multilateral Trade Arrangements*, T. Ito and A. Krueger (eds.), Chicago, The University of Chicago Press.

Bergstrand, J. (1985), "The gravity equation in international trade: some microeconomic foundations and empirical evidence", *Review of Economics and Statistics*, vol. 67, No. 3, Cambridge, Massachusetts, The MIT Press.

Berrettoni, D. and S. Castresana (2007), "Exportaciones y tipo de cambio real: el caso de las manufacturas industriales argentinas", *Revista del CEI: comercio exterior e integración*, No. 9, Buenos Aires, Centro de Economía Internacional.

Black, S. and L. Lynch (1996), "Human capital investment and productivity", *The American Economic Review*, vol. 86, No. 2, Nashville, Tennessee, American Economic Association.

Boisso, D. and M. Ferrantino (1997), "Economic distance, cultural distance, and openness in international trade: empirical puzzles", *Journal of Economic Integration*, vol. 12, Seoul, Sejong University.

Catão, L. and E. Falsetti (2002), "Determinants of Argentina's external trade", *Journal of Applied Economics*, vol. 5, No. 1, Buenos Aires, CEMA University.

Cuevas, V. (2008), "Efectos de la productividad laboral en las exportaciones manufactureras mexicanas", *Comercio exterior*, vol. 58, No. 6, Mexico City, Banco Nacional de Comercio Exterior.

Dearden, L., H. Reed and J.V. Reenen (2005), "The impact of training on productivity and wages: evidence from British panel data", *CEP Discussion Papers*, No. 674, London, Centre for Economic Performance.

Dickey, D., W. Bell and R. Miller (1986), "Unit roots in time series models: tests and implications", *The American Statistician*, vol. 40, Alexandria, Virginia, American Statistical Association.

Dickey, D. and W. Fuller (1981), "Likelihood ratio statistics for autoregressive time series with a unit root", *Econometrica*, vol. 49, No. 4, Cleveland, Ohio, The Econometric Society.

Engle, R. and C. Granger (1987), "Co-integration and error correction: representation, estimation, and testing", *Econometrica*, vol. 55, No. 2, Cleveland, Ohio, The Econometric Society.

Goldberg, L. and M. Klein (1997), "Foreign direct investment, trade and real exchange rate linkages in Southeast Asia and Latin America", *NBER Working Papers*, No. 6344, Cambridge, Massachusetts, National Bureau of Economic Research.

Hamilton, J. (1994), *Time Series Analysis*, Princeton, Princeton University Press.

Hannan, E. and J. Rissanen (1982), "Recursive estimation of mixed autoregressive-moving average order", *Biometrika*, vol. 69, No. 1, Oxford, Oxford University Press.

Isard, W. (1954), "Location theory and trade theory: short-run analysis", *Quarterly Journal of Economics*, vol. 68, No. 2, Cambridge, Massachusetts, The MIT Press.

Jiménez, F., G. Aguilar and J. Kapsoli (1998), "Competitividad en la industria manufacturera peruana" [online] http://www.pucp.edu.pe/departamento/economia/images/documentos/DDD148.pdf

Johansen, S. (1995), *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, Oxford, Oxford University Press.

Koop, G., M. Pesaran and S. Potter (1996), "Impulse response analysis in nonlinear multivariate models", *Journal of Econometrics*, vol. 74, No. 1, Amsterdam, Elsevier.

Kwiatkowski, D. and others (1992), "Testing the null hypothesis of stationary against the alternative of a unit root", *Journal of Econometrics*, vol. 54, No. 1-3, Los Angeles, California, John Wiley & Sons.

Lütkepohl, H. (2006), *New Introduction to Multiple Time Series Analysis*, New York, Springer-Verlag.

MacKinnon, J. (1996), "Numerical distribution functions for unit root and cointegration tests", *Journal of Applied Econometrics*, vol. 11, No. 6, Los Angeles, California, John Wiley & Sons.

MacKinnon, J., A. Haug and L. Michelis (1999), "Numerical distribution functions of likelihood ratio tests for cointegration", *Journal of Applied Econometrics*, vol. 14, No. 5, Los Angeles, California, John Wiley & Sons.

Maglen, L. (1995), "The role of education and training in the economy", *Working Paper*, No. 2, Clayton, Centre for the Economics of Education and Training, Monash University.

Mbaye, A. and S. Golub (2002), "Unit labor cost, international competitiveness, and exports: the case of Senegal", *Journal of African Economies*, vol. 11, No. 2, Oxford, Oxford University Press.

Mincer, J. (1994), "Investment in U.S. education and training", *NBER Working Papers*, No. 4844, Cambridge, Massachusetts, National Bureau of Economic Research.

Mortimore, M., S. Vergara and J. Katz (2001), "La competitividad internacional y el desarrollo nacional: implicancias para la política de inversión extranjera directa en América Latina", *Desarrollo productivo series*, No. 107 (LC/L.1586-P), Santiago, Chile, Economic Commission for Latin America and the Caribbean (ECLAC). United Nations publication, Sales No. S.01.II.G.130.

Mungaray, A. and M. Ramírez (2007), "Capital humano y productividad en microempresas", *Investigación económica*, vol. 66, No. 260, Mexico City, National Autonomous University of Mexico.

Nabi, I. and M. Luthria (2002), "Building competitiveness: a roadmap and policy guide", *Building Competitive Firms Incentives and Capabilities*, Washington, D.C., World Bank.

Padilla, R. and M. Juárez (2006), "Efectos de la capacitación en la competitividad de la industria manufacturera", *Estudios y perspectivas series*, No. 49, Mexico City, ECLAC Subregional Headquarters in Mexico.

Patterson, K. (2000), *An Introduction to Applied Econometrics: a Time Series Approach*, New York, Palgrave MacMillan.

Perron, P. and S. Ng (1996), "Useful modifications to unit root tests with dependent error and their local asymptotic properties", *Review of Economic Studies*, vol. 63, No. 3, London, Blackwell Publishing.

Pesaran, M. and Y. Shin (1998), "Generalized impulse response analysis in linear multivariate models", *Economic Letters*, vol. 58, No. 1, Amsterdam, Elsevier.

Phillips, P. and P. Perron (1988), "Testing for a unit root in time series regression", *Biometrika*, vol. 75, No. 2, Oxford, Oxford University Press.

Pöyhönen, P. (1963), "A tentative model for the volume of trade between countries", *Review of World Economics*, vol. 90, No. 1, Kiel, Schleswig-Holstein.

Reinhart, C. (1995), "Devaluation, relative prices, and international trade: evidence from developing countries", *IMF Staff Papers*, vol. 42, No. 2, Washington, D.C., International Monetary Fund.

Riedel, J. (1988), "The demand for LDC exports of manufactures: estimates from Hong Kong", *The Economic Journal*, vol. 98, No. 389, Royal Economic Society.

Senhadji, A. and C. Montenegro (1998), "Time-series analysis of export demand equations: a cross-country analysis", *IMF Working Papers*, No. 149, Washington, D.C., International Monetary Fund.

Sims, C. (1980), "Macroeconomics and reality", *Econometrica*, vol. 48, No. 1, Cleveland, Ohio, The Econometric Society.

Summary, R. (1989), "A political-economic model of U.S. bilateral trade", *The Review of Economics and Statistics*, vol. 71, No. 1, Cambridge, Massachusetts, The MIT Press.

Tinbergen, J. (1962), "Shaping the world economy-suggestions for an international economic policy", *The Twentieth Century Fund*, New York.

CEPAL Review referees 2009 and January-August 2010

During this period, the following academics and researchers acted as referees for the CEPAL Review

Dillon Alleyne
Hugo Altomonte
Roberto Álvarez
Rodrigo Aravena
Raúl Atria
Reynaldo Bajraj
Luis Felipe Beccaria
Ricardo Bielschowsky
Daniel Bitrán
Mauro Boianovsky
Guillermo Campero
Fernando Cantú
Rodrigo Cárcamo
Ricardo Carciofi
Pablo Carvallo
Simone Cecchini
Nelson Correa
Filipa Correia
Manlio Coviello
Armando Di Filippo
Martine Dirven
John Durston
Hubert Escaith
Ernesto Espíndola
Lorena Flores
Hernán Frigolett
Guillermo Fuentes
Juan Carlos Guajardo
Michael Hantke
René Hernández
Martín Hopenhayn
Luis Felipe Jiménez
Jorge Katz
Bernardo Kosacoff
Marcelo LaFleur
Patricio Leiva
Luis Felipe Lira
Rafael López-Monti
Aldo Madariaga
Carlos Maldonado
Xavier Mancero
Sandra Manuelito
Salvador Marconi
Manuel Marfán
Jorge Martínez
Carlos Massad
César Morales
Juan Carlos Moreno-Brid
Michael Mortimore
Nanno Mulder
Guillermo Mundt
Óscar Muñoz
Carlos Mussi
Lucas Navarro
Juan Francisco Pacheco
Wilson Peres
Esteban Pérez
Gabriel Pérez
Neil Pierre
Joe Ramos
Nieves Rico
Luis Felipe Riffo
Eugenio Rivera
Mônica Rodrigues
Marco V. Sánchez
Claudia Sanhueza
Marianne Schaper
Guilherme Schuetz
Jürgen Schuldt
Andrés Schuschny
Ana Sojo
Rogerio Studart
Guillermo Sunkel
Marcia Tavares
Víctor Tockman
Andras Uthoff
Cecilia Vera
Sebastián Vergara
Francisco Villarreal
Robert Vos
Jürgen Weller

CEPAL Review is grateful for their invaluable assistance.

Guidelines for contributors to the *CEPAL Review*

In order to facilitate the submission, consideration and publication of articles, the editorial board of the *CEPAL Review* has prepared the following information and suggestions to serve as a guide for future contributors.

The submission of an article implies an undertaking by the author not to submit it simultaneously to other publications. The copyright to all articles published in the *Review* shall be owned by the United Nations.

All articles will be submitted to external referees.

Papers should be submitted in the original language (English, French, Portuguese or Spanish). They will be translated into the appropriate language by the relevant ECLAC services.

Each article must be accompanied by a summary, no more than 150 words in length, giving a brief description of its subject matter and main conclusions.

Papers should be no longer than 10,000 words, including the summary, notes and bibliography. Shorter papers will also be considered.

Articles should be sent either by e-mail (to *revista@cepal.org*) or by regular mail, on cd or diskette, to: ***CEPAL Review***, Casilla 179-D, Santiago, Chile. They should not be submitted in pdf format.

Style guide:

Titles should not be excessively long.

Footnotes

- It is recommended that footnotes be kept to a minimum.
- It is recommended that footnotes not be used to cite bibliographical references; such references should preferably be incorporated into the text.
- Footnotes should be numbered consecutively using superscript Arabic numerals.

Tables and figures

- It is recommended that tables and figures be kept to a minimum, avoiding any redundancy with the text.
- Tables, figures and other elements should be inserted at the end of the text in the format in which they were designed; they should not be inserted as *"pictures"*. Figures in Excel should include the corresponding worksheets.
- The location of tables and figures in the body of the article should be indicated in the appropriate place as follows:

  Insert figure 1
  Insert table 1

- Tables and figures should include an explicit and complete reference to their sources.
- Tables should indicate the period covered at the end of the title, and should indicate the units in which the data are expressed in a subtitle (in italics and between brackets).
- The symbols referred to in the "Explanatory notes" which appear on the page preceding the table of contents should be taken into account in the preparation of tables and figures.
- Footnotes to tables and figures should be ordered consecutively using superscript lower-case letters.
- Figures should be prepared bearing in mind that they will be printed in black and white.

Acronyms and abbreviations

- Acronyms and abbreviations should not be used unless absolutely necessary, in which case the full name should be written out the first time it occurs in the article.

Bibliography

- Bibliographical references should be directly related to the content of the article and should not be excessively long.
- At the end of the article, under the title "Bibliography", all the necessary information should be included accurately and in alphabetical order by author: name of author(s), year of publication, full name of article (if any) and publication (including any subtitle), city of publication, publisher and, in the case of a periodical, month of publication.

The editorial board of the *Review* reserves the right to make any necessary 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 in Spanish and 30 in English, at the time of publication in each language.

NACIONES UNIDAS
UNITED NATIONS

CEPAL
ECLAC

Publicaciones de la CEPAL / *ECLAC publications*

Comisión Económica para América Latina y el Caribe / *Economic Commission for Latin America and the Caribbean*
Casilla 179-D, Santiago de Chile. E-mail: publications@cepal.org
Véalas en: www.cepal.org/publicaciones
Publications may be accessed at: www.eclac.org

Revista CEPAL / *CEPAL Review*

La Revista se inició en 1976 como parte del Programa de Publicaciones de la Comisión Económica para América Latina y el Caribe, con el propósito de contribuir al examen de los problemas del desarrollo socioeconómico de la región. Las opiniones expresadas en los artículos firmados, incluidas las colaboraciones de los funcionarios de la Secretaría, son las de los autores y, por lo tanto, no reflejan necesariamente los puntos de vista de la Organización.

La *Revista CEPAL* se publica en español e inglés tres veces por año.

Los precios de suscripción anual vigentes para 2010 son de US$ 30 para ambas versiones. El precio por ejemplar suelto es de US$ 15 para ambas versiones. Los precios de suscripción por dos años son de US$ 50 para ambas versiones.

CEPAL Review first appeared in 1976 as part of the Publications Programme of the Economic Commission for Latin America and the Caribbean, its aim being to make a contribution to the study of the economic and social development problems of the region. The views expressed in signed articles, including those by Secretariat staff members, are those of the authors and therefore do not necessarily reflect the point of view of the Organization.

CEPAL Review is published in Spanish and English versions three times a year.

Annual subscription costs for 2010 are US$ 30 for both versions. The price of single issues is US$ 15 in both cases. The cost of a two-year subscription is US$ 50 for both versions.

Informes periódicos institucionales / *Annual reports*

Todos disponibles para años anteriores / *Issues for previous years also available*

- *Balance preliminar de las economías de América Latina y el Caribe, 2009, 184 p.*
  ***Preliminary Overview of the Economies of Latin America and the Caribbean, 2009**, 166 p.*
- *Estudio económico de América Latina y el Caribe 2008-2009, 154 p.*
  ***Economic Survey of Latin America and the Caribbean 2008-2009**, 144 p.*
- *Panorama de la inserción internacional de América Latina y el Caribe, 2008-2009. Crisis y espacios de cooperación regional, 112 p.*
  ***Latin America and the Caribbean in the World Economy, 2008-2009. Crisis and opportunities for regional cooperation,** 108 p.*
- *Panorama social de América Latina, 2009*, 218 p.
  ***Social Panorama of Latin America, 2009**, 208 p.*
- *La inversión extranjera directa en América Latina y el Caribe*, 2009, 162 p.
  ***Foreign Direct Investment of Latin America and the Caribbean, 2009**, 158 p.*
- *Anuario estadístico de América Latina y el Caribe* / ***Statistical Yearbook for Latin America and the Caribbean*** (bilingüe/*bilingual*), 2009, 316 p.

Libros de la CEPAL

108 *Alianzas público-privadas. Para una nueva visión estratégica del desarrollo*, Robert Devlin y Graciela Moguillansky, 2010, 196 p.
105 *Regulation, Worker Protection and Active Labour-Market Policies in Latin America*, Jürgen Weller (ed.), 2009, 236 p.
104 *La República Dominicana en 2030: hacia una sociedad cohesionada*, Víctor Godínez y Jorge Máttar (coords.), 2009, 582 p.
103 *L'Amérique latine et les Caraïbes au seuil du troisième millénaire*, 2009, 138 p.
102 *Migración interna y desarrollo en América Latina entre 1980 y 2005*, Jorge Rodríguez y Gustavo Busso, 2009, 272 p.
101 *Claves de la innovación social en América Latina y el Caribe*, Adolfo Rodríguez Herrera y Hernán Alvarado Ugarte, 2009, 236 p.
100 *Envejecimiento, derechos humanos y políticas públicas*, Sandra Huenchuan (ed.), 2009, 232 p.
99 *Economía y territorio en América Latina y el Caribe. Desigualdades y políticas*, 2009, 212 p.

98 *La sociedad de la información en América Latina y el Caribe: desarrollo de las tecnologías y tecnologías para el desarrollo*, Wilson Peres y Martin Hilbert (eds.), 2009, 388 p.
97 *América Latina y el Caribe: migración internacional, derechos humanos y desarrollo*, Jorge Martínez Pizarro (ed.), 2008, 375 p.
96 *Familias y políticas públicas en América Latina: una historia de desencuentros*, Irma Arriagada (coord.), 2007, 424 p.
95 *Centroamérica y México: políticas de competencia a principios del siglo XXI*, Eugenio Rivera y Claudia Schatan (coords.), 2008, 304 p.
94 *América Latina y el Caribe: La propiedad intelectual después de los tratados de libre comercio*, Álvaro Díaz, 2008, 248 p.
93 *Tributación en América Latina. En busca de una nueva agenda de reformas*, Oscar Cetrángolo y Juan Carlos Gómez-Sabaini (comps.), 2007, 166 p.
92 *Fernando Fajnzylber. Una visión renovadora del desarrollo en América Latina*, Miguel Torres Olivos (comp.), 2006, 422 p.
91 *Cooperación financiera regional*, José Antonio Ocampo (comp.), 2006, 274 p.
90 *Financiamiento para el desarrollo. América Latina desde una perspectiva comparada*, Barbara Stallings con la colaboración de Rogério Studart, 2006, 396 p.
89 *Políticas municipales de microcrédito. Un instrumento para la dinamización de los sistemas productivos locales. Estudios de caso en América Latina*, Paola Foschiatto y Giovanni Stumpo (comps.), 2006, 244 p.
88 *Aglomeraciones en torno a los recursos naturales en América Latina y el Caribe: Políticas de articulación y articulación de políticas*, 2006, 266 p.
87 *Pobreza, desertificación y degradación de los recursos naturales*, César Morales y Soledad Parada (eds.), 2006, 274 p.

Copublicaciones recientes / *Recent co-publications*

Sesenta años de la CEPAL. Textos seleccionados del decenio 1998-2008, Ricardo Bielschowsky (comp.), CEPAL/Siglo Veintiuno, Argentina, 2010.
El nuevo escenario laboral latinoamericano. Regulación, protección y políticas activas en los mercados de trabajo, Jürgen Weller (ed.), CEPAL/Siglo Veintiuno, Argentina, 2010.
Internacionalización y expansión de las empresas eléctricas españolas en América Latina, Patricio Rozas, CEPAL/Lom, Chile, 2009.
Gobernanza corporativa y desarrollo de mercados de capitales en América Latina, Georgina Núñez, Andrés Oneto y Germano M. de Paula (coords.), CEPAL/Mayol, Colombia, 2009.
EnREDos. Regulación y estrategias corporativas frente a la convergencia tecnológica, Marcio Wohlers y Martha García-Murillo (eds.), CEPAL/Mayol, Colombia, 2009.
Desafíos y oportunidades de la industria del software en América Latina, Paulo Tigre y Felipe Silveira Marques (eds.), CEPAL/Mayol, Colombia, 2009.
¿Quo vadis*, tecnología de la información y de las comunicaciones?,* Martin Hilbert y Osvaldo Cairó (eds.), CEPAL/Mayol, Colombia, 2009.
***O Estruturalismo latino-americano,* Octavio Rodríguez, CEPAL/Civilização Brasileira, 2009.**
***L'avenir de la protection sociale en Amérique latine. Accessibilité, financement et solidarité, CEPALC/*Eska, France, 2009.**
Fortalecer los sistemas de pensiones latinoamericanos. Cuentas individuales por reparto, Robert Holzmann, Edward Palmer y Andras Uthoff (eds.), CEPAL/Mayol, Colombia, 2008.
Competition Policies in Emerging Economies. Lessons and Challenges from Central America and Mexico, Claudia Schatan and Eugenio Rivera Urrutia (eds.), ECLAC/Springer, USA, 2008.
Estratificación y movilidad social en América Latina. Transformaciones estructurales en un cuarto de siglo, Rolando Franco, Arturo León y Raúl Atria (coords.), CEPAL/Lom, Chile, 2007.
Economic growth with equity. Challenges for Latin America, Ricardo Ffrench-Davis and José Luis Machinea (eds.), ECLAC/Palgrave Macmillan, United Kingdom, 2007.
Mujer y empleo. La reforma de la salud y la salud de la reforma en Argentina, María Nieves Rico y Flavia Marco (coords.), CEPAL/Siglo XXI, Argentina, 2006.
El estructuralismo latinoamericano, Octavio Rodríguez, CEPAL/Siglo XXI, México, 2006.
Gobernabilidad corporativa, responsabilidad social y estrategias empresariales en América Latina, Germano M. de Paula, João Carlos Ferraz y Georgina Núñez (comps.), CEPAL/Mayol, Colombia, 2006.
Desempeño económico y política social en América Latina y el Caribe. Los retos de la equidad, el desarrollo y la ciudadanía, Ana Sojo y Andras Uthoff (comps.), CEPAL/Flacso-México/ Fontamara, México, 2006.

Coediciones recientes / *Recent co-editions*

***The United Nations Regional Commissions and the Climate Change Challenges*, ECLAC/ECA/ECE/ESCAP/ESCWA, 2009.**
Hacia un desarrollo inclusivo. El caso de Chile, Osvaldo Sunkel y Ricardo Infante (eds.), CEPAL/OIT/Fundación Chile 21, Chile, 2008.
Reformas para la cohesión social en América Latina. Panorama antes de la crisis, Alicia Bárcena y Narcís Serra (eds.), CEPAL/SEGIB/CIDOB, Chile, 2008.
El envejecimiento y las personas de edad. Indicadores sociodemográficos para América Latina y el Caribe, CEPAL/UNFPA, 2009.
Espacios iberoamericanos: la economía del conocimiento, CEPAL/SEGIB, Chile, 2008.
Hacia la revisión de los paradigmas del desarrollo en América Latina, Oscar Altimir, Enrique V. Iglesias, José Luis Machinea (eds.), CEPAL/SEGIB, Chile, 2008.
Por uma revisão dos paradigmas do desenvolvimento na América Latina, Oscar Altimir, Enrique V. Iglesias, José Luis Machinea (eds.), CEPAL/SEGIB, Chile, 2008.

Hacia un nuevo pacto social. Políticas económicas para un desarrollo integral en América Latina, José Luis Machinea y Narcís Serra (eds.) CEPAL/CIDOB, España, 2008.

Espacios iberoamericanos: comercio e inversión, CEPAL/SEGIB, Chile, 2007.

Espaços Ibero-Americanos: comércio e investimento, CEPAL/SEGIB, Chile, 2007.

Visiones del desarrollo en América Latina, José Luis Machinea y Narcís Serra (eds.), CEPAL/CIDOB, España, 2007.

Cohesión social: inclusión y sentido de pertenencia en América Latina y el Caribe, CEPAL/SEGIB, Chile, 2007.

Social Cohesion. Inclusion and a sense of belonging in Latin America and the Caribbean, ECLAC/SEGIB, Chile, 2007.

Cuadernos de la CEPAL

94 *El cuidado en acción. Entre el derecho y el trabajo*, Sonia Montaño Virreira y Coral Calderón Magaña (coords.), 2010, 236 p.

93 *Privilegiadas y discriminadas. Las trabajadoras del sector financiero*, Flavia Marco Navarro y María Nieves Rico Ibáñez (eds.), 2009, 300 p.

92 *Estadísticas para la equidad de género: magnitudes y tendencias en América Latina*, Vivian Milosavljevic, 2007, 186 pp.

91 *Elementos conceptuales para la prevención y reducción de daños originados por amenazas naturales*, Eduardo Chaparro y Matías Renard (eds.), 2005, 144 p.

Cuadernos estadísticos de la CEPAL

36 *Clasificaciones estadísticas internacionales incorporadas en el Banco de Datos de Comercio Exterior de América Latina y el Caribe de la CEPAL (Revisión 3)*. Solo disponible en CD, 2008.

35 *Resultados del Programa de Comparación Internacional para América del Sur.* Solo disponible en CD, 2007.

34 *Indicadores económicos del turismo*. Solo disponible en CD, 2006.

33 *América Latina y el Caribe. Balanza de pagos 1980-2005*. Solo disponible en CD, 2006.

32 *América Latina y el Caribe. Series regionales y oficiales de cuentas nacionales, 1950-2002*. Solo disponible en CD, 2005.

31 *Comercio exterior. Exportaciones e importaciones según destino y origen por principales zonas económicas. 1980, 1985, 1990, 1995-2002.* Solo disponible en CD, 2005.

Observatorio demográfico *ex Boletín demográfico* / *Demographic Observatory formerly Demographic Bulletin* (bilingüe/*bilingual*)

Edición bilingüe (español e inglés) que proporciona información estadística actualizada, referente a estimaciones y proyecciones de población de los países de América Latina y el Caribe. Incluye también indicadores demográficos de interés, tales como tasas de natalidad, mortalidad, esperanza de vida al nacer, distribución de la población, etc.

El Observatorio aparece dos veces al año, en los meses de enero y julio. Suscripción anual: US$ 20.00. Valor por cada ejemplar: US$ 15.00.

Bilingual publication (Spanish and English) proving up-to-date estimates and projections of the populations of the Latin American and Caribbean countries. Also includes various demographic indicators of interest such as fertility and mortality rates, life expectancy, measures of population distribution, etc.

The Observatory appears twice a year in January and July. Annual subscription: US$ 20.00. Per issue: US$ 15.00.

Notas de población

Revista especializada que publica artículos e informes acerca de las investigaciones más recientes sobre la dinámica demográfica en la región, en español, con resúmenes en español e inglés. También incluye información sobre actividades científicas y profesionales en el campo de población.

La revista se publica desde 1973 y aparece dos veces al año, en junio y diciembre.

Suscripción anual: US$ 20.00. Valor por cada ejemplar: US$ 12.00.

Specialized journal which publishes articles and reports on recent studies of demographic dynamics in the region, in Spanish with abstracts in Spanish and English. Also includes information on scientific and professional activities in the field of population.

Published since 1973, the journal appears twice a year in June and December.

Annual subscription: US$ 20.00. Per issue: US$ 12.00.

Series de la CEPAL

Comercio internacional / *Desarrollo productivo* / *Desarrollo territorial* / *Estudios estadísticos y prospectivos* / *Estudios y perspectivas* (Bogotá, Brasilia, Buenos Aires, México, Montevideo) / ***Studies and Perspectives*** (The Caribbean, Washington) / *Financiamiento del desarrollo* / *Gestión pública* / *Informes y estudios especiales* / *Macroeconomía del desarrollo* / *Manuales* / *Medio ambiente y desarrollo* / *Mujer y desarrollo* / *Población y desarrollo* / *Políticas sociales* / *Recursos naturales e infraestructura* / *Seminarios y conferencias*.

Véase el listado completo en: www.cepal.org/publicaciones / *A complete listing is available at*: www.cepal.org/publicaciones

Desarrollo Económico
Revista de Ciencias Sociales

ISSN 0046-001X

Vol. 49 enero-marzo de 2010 Nº 196

Desarrollo Económico es indizada, con inclusión de resúmenes, en las siguientes publicaciones: ***Current Contents*** (Social Science Citation Index, Institute for Scientific Information); ***Journal of Economic Literature*** (American Economic Association); ***Sociological Abstract*** (Cambridge Scientific Abstracts); ***International Bibliography of the Social Science*** (British Library of Political and Economic Science y UNESCO); ***Clase*** (UNAM, México); ***Hispanic American Periodicals Index*** (HAPI, Universidad de California, Los Angeles). También en otras ediciones de carácter periódico y en volúmenes especiales nacionales e internacionales, así como en índices en versión electrónica.

DESARROLLO ECONOMICO – Revista de Ciencias Sociales es una publicación trimestral editada por el Instituto de Desarrollo Económico y Social (IDES). Suscripción anual: R. Argentina, $ 120,00; Países limítrofes, U$S 130; Resto de América, U$S 150; Europa, U$S 170; Asia, Africa y Oceanía, U$S 180. Ejemplar simple: U$S 30 (recargos según destino y por envíos vía aérea). Más información en: desarrollo@ides.org.ar, o disponible en el WEB SITE: ***www.ides.org.ar.*** Pedidos, correspondencia, etcétera, a:

Instituto de Desarrollo Económico y Social
Aráoz 2838 ◆ C1425DGT Buenos Aires ◆ Argentina
Teléfono: 4804-4949 ◆ Fax: (54 11) 4804-5856
Correo electrónico: *desarrollo@ides.org.ar*

Sumario

Nº 119 Invierno 2010

ESTUDIOS PÚBLICOS

www.cepchile.cl

CENTRO DE ESTUDIOS PÚBLICOS
Monseñor Sótero Sanz 162. Santiago, Chile. Fono (56-2) 328-2400. Fax (56-2) 328-2440.

SUSCRIPCIONES: Anual $ 9.000 • Bianual $ 13.500 • Estudiantes $ 5.000

NUEVA SOCIEDAD
www.nuso.org

Director: Joachim Knoop
Jefe de redacción: José Natanson

¿El huevo o la gallina?
Economía y política en América Latina

229
SEPTIEMBRE-OCTUBRE 2010

COYUNTURA: **Francisco Leal Buitrago.** Colombia: elecciones atípicas y perspectivas inciertas. **Michael Dauderstädt.** Deudas de Estado y Estados con deuda. Crisis y crecimiento en Europa. APORTES: **Mercedes Isabel Botto.** El G-20 y la gobernanza global: ¿un cambio en la arquitectura o en los procedimientos? Una mirada desde América Latina. TEMA CENTRAL: **Alberto Acosta.** Maldiciones que amenazan la democracia. **Ludolfo Paramio.** Economía y política de las clases medias en América Latina. **Nelly Arenas.** La Venezuela de Hugo Chávez: rentismo, populismo y democracia. **Yann Basset.** Las relaciones entre las esferas política y económica. ¿Hacia el fin de los particularismos colombianos? **Federico Traversa.** ¿Qué hacer con los sectores medios? Coaliciones sociales, bienestar y socialdemocracia en la periferia capitalista. **Tomás Palau Viladesau.** La política y su trasfondo. El poder real en Paraguay.

PAGOS: Solicite precios de suscripción y datos para el pago a <info@nuso.org> o <distribucion@nuso.org>.

SECUENCIA

Revista de historia y ciencias sociales

Número 77 mayo - agosto 2010

ARTÍCULOS

Instituto Mora

Informes: Madrid 82, Col. del Carmen Coyoacán, CP 04100, México, D. F.Tel./Fax 5554 8946 ext. 3108
secuencia@mora.edu.mx
www.mora.edu.mx

Trayectorias

REVISTA DE CIENCIAS SOCIALES DE LA UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

Volumen 12 *Número 30* *enero-junio de 2010*

Suscripciones: en México (pesos M.N.): $ 160.00 (individual), $ 190.00 (institucional).
Números sueltos: $ 60.00.
América del Norte y El Caribe: USD $ 30.00.
Europa y Sudamérica: USD $ 40.00. Resto del mundo: USD $ 50.00.
Instituto de Investigaciones Sociales, UANL
Av. Lázaro Cárdenas Ote. y Paseo de la Reforma S/N, Campus Mederos, U.A.N.L.
C.P. 64930, Monterrey, N.L., México, C. P. 64440. Teléfono y fax: (52 81) 8329 4237.
Correo electrónico: trayectorias@uanl.mx. Internet: www.trayectorias.uanl.mx

EL TRIMESTRE ECONÓMICO

Vol. LXXVII (3) México, julio-septiembre de 2010 Núm. 307

EL TRIMESTRE ECONÓMICO aparece en los meses de enero, abril, julio y octubre. La suscripción en México cuesta $375.00. Número suelto $120.00.

Precios para otros países (dólares)

	Suscripciones
Centroamérica y el Caribe	80.00
Sudamérica y España	100.00
Canadá, Estados Unidos y resto del mundo	135.00

Fondo de Cultura Económica, Carretera Picacho Ajusco 227, Col. Bosques del Pedregal, 14738 México, D.F.
Suscripciones y anuncios: Tel. (52) (55) 5227-4671 Fax (52) (55) 5227-4640
Correo electrónico: suscripciones@fondodeculturaeconomica.com
Suscripciones y venta de ejemplares en: http://www.eltrimestreeconomico.com