

EL PROCESAMIENTO DE LOS CENSOS DE POBLACION EN
AMERICA LATINA Y EL CARIBE PARA LA DECADA DE 1990:
UN VISTAZO AL FUTURO



Art. Silva
(CELADE)
1985

SINTESIS

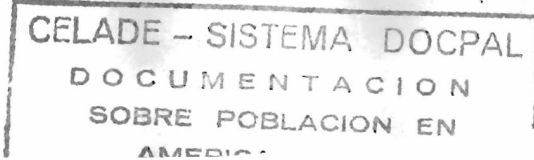
El objeto del documento es examinar los posibles métodos de procesamiento de datos a ser usados para los censos de población de la década de 1990, en los distintos países de América Latina y el Caribe. Para ello, se usa como variable principal la población estimada de cada país a la fecha, agregándose las experiencias recogidas durante el procesamiento de los censos de los años 80, y una previsión realista en los avances tecnológicos, de equipos y programas de computación, disponibles en la región.

Se dividen los países en cuatro grupos, y para cada uno de los ocho principales etapas del procesamiento censal, se sugieren los procedimientos y metodologías a ser usados.

Para las cinco primeras etapas (Planificación, Cartografía, Empadronamiento, Entrada de Datos y Control Primario), el microcomputador, con algunas variaciones, deberá atender a todos los países, exceptuándose la etapa de Entrada de Datos de los dos grupos más grandes (más de diez millones de habitantes), para los cuales se recomienda la utilización de las lectoras ópticas.

En las tres últimas etapas (Consistencia, Tabulación y Base de Datos), el microcomputador es indicado sólo para los países del grupo de menos de un millón de habitantes. Para los otros se propone la utilización de computadores grandes en forma centralizada.

Esta visión global debe ser aplicada tomándose en cuenta las variables particulares de cada país, y deberá variar sustancialmente de acuerdo al nivel de "acierto" en las previsiones acerca de las novedades computacionales.



EL PROCESAMIENTO DE LOS CENSOS DE POBLACION EN
AMERICA LATINA Y EL CARIBE PARA LA DECADEA DE 1990:
UN VISTAZO AL FUTURO

Ari Silva
(CELABE)

I. INTRODUCCION

Este documento tiene como objetivo principal suministrar una mirada previa a los métodos de procesamiento que deberán ser usados, en cada país, en la producción de los datos censales de la década de 1990. Se examinan las etapas más importantes en el flujo de procesamiento, y para cada una de ellas se proponen una o más alternativas, de acuerdo a sus características.

Un análisis estricto de las características que regirán el procesamiento de los datos censales en el decenio de los años 90 implicaría un franco ejercicio de adivinación. En efecto, el ritmo de cambio tecnológico en materia de computación es tan veloz que incluso en un lapso tan breve, como cinco años, pueden producirse modificaciones substanciales. Por lo tanto, lo que se pretende es efectuar una previsión, dentro de los márgenes de probabilidades conocidas, acerca de las condiciones en las que se desenvolverá el procesamiento de los datos de los censos de población, teniendo en cuenta la variación de los requerimientos involucrados según el distinto tamaño de la población de cada país.

En el capítulo II se comenta la tecnología computadorizada, presente y futura, desde la óptica del procesamiento censal. Como elementos de referencia se mencionan algunos aspectos ya históricos del procesamiento de los censos de los años 80, y se indican los nuevos equipos disponibles - o previsibles - en el mercado. El capítulo III tiene el mismo objetivo del capítulo anterior, pero se concentra en la discusión de los paquetes existentes o susceptibles de estar en operación en el futuro.

El capítulo IV está dedicado a establecer agrupaciones de los países de América Latina y el Caribe según el tamaño de sus poblaciones, y se discuten también algunas otras variables particulares de cada país, que pueden afectar esta tipificación. En el capítulo V se describe, para las agrupaciones definidas en el capítulo anterior, las técnicas que pueden ser usadas en cada etapa del procesamiento censal.

Finalmente, se hace un resumen de los puntos más importantes, con comentarios generales, de la "futurología" aplicada.

II. EVOLUCION DE LA TECNOLOGIA EN COMPUTACION

Antes de considerar las propuestas propiamente tales, es necesario presentar algunos aspectos que, ciertamente, tendrán gran influencia en el tipo de procesamiento censal que deberá utilizarse.

En los censos de América Latina de la década de 1980, los equipos más utilizados fueron los de gran tamaño y los minicomputadores. Hubo alguna utilización de los microcomputadores, sobre todo en la etapa de digitación, y se puede mencionar también la lectora óptica, usada en varios países. 1/

Para los años 90, se espera que la 'estrella' sea el microcomputador; si no se le usa para todo el procesamiento, al menos se lo empleará en gran parte del mismo. Su capacidad de memoria central ha venido aumentando rápidamente: ya hay microcomputadores en el mercado con 4 Mb (Megabytes) de memoria real, siendo difícil prever un límite para el valor que eventualmente se alcanzará. El espacio disponible en los discos duros 'convencionales' también ha experimentado un aumento considerable: los más comunes están alrededor de los 20 Mb; existen, además, los de 70 Mb, y, tal como la memoria central, no se tiene una idea clara de sus límites.

La nueva tecnología 'láser' de grabación ya promete discos removibles de 400 Mb, lo que torna prácticamente ilimitado el espacio disponible para almacenamiento. Para tener una idea de lo que significaría esta capacidad traducida a archivos censales, se puede mencionar el caso de Chile, donde está siendo usado para probar un sistema de base de datos en elaboración en el CELADE, llamado REBATAN. El censo de Chile de 1982 tiene cerca de 3 millones 750 mil viviendas, con más o menos 12 millones de habitantes, lo que ocuparía alrededor de 1200 Mb. Con el método de compresión de variables usado en REBATAN, este espacio se reduce a 320 Mb. 2/

1/ Para mayores detalles, véase Silva, A., Procesamiento de Datos Censales: Lecciones de la Experiencia Latinoamericana. Santiago, CELADE, serie A, número 170, junio de 1985.

2/ CELADE. Resumen del Proyecto REBATAN, presentado al Centro Internacional de Investigación y Desarrollo de Canadá, 1985.

Las unidades de cinta magnética ya están apareciendo incorporadas a los microcomputadores, a un costo aún relativamente alto, pero sólo es una cuestión de tiempo para que se tengan cintas a precios accesibles. Esta adaptación es necesaria, principalmente para aumentar la capacidad de almacenamiento y para facilitar la conexión y transferencia de información con los equipos de gran tamaño.

La comunicación de los microcomputadores con los computadores grandes ya es un hecho consumado, ya sea en conexión directa o a través de la telecomunicación. Los microcomputadores también se comunican fácilmente entre ellos, haciendo prever una tendencia al establecimiento de redes locales y al procesamiento distribuido especialmente.

Otro tipo de equipo que jugará un papel importante en los futuros censos es la lectora óptica. Sus características negativas, tales como la necesidad del uso de un papel especial, con una impresión técnicamente perfecta, están siendo eliminadas a pasos muy rápidos. Hoy día ya existen, por ejemplo, tabletas de lectura óptica acopladas a microcomputadores, uniéndose así las dos tecnologías.

Y, por supuesto, los precios y costos operativos de todo lo señalado han seguido bajando inexorablemente; a medida que se domina la técnica, aumenta la competencia y la producción masiva de dichos equipos.

III. SISTEMAS PARA EL PROCESAMIENTO CENSAL

Los denominados "paquetes" de computación son programas genéricos prefabricados, que se ejecutan a través de comandos o parámetros escritos por el usuario. El objetivo principal de los paquetes es ahorrar mano de obra especializada y producir resultados en corto plazo y con un nivel mínimo de errores.

Para los países de América Latina, es muy importante la existencia de estos paquetes. La falta de mano de obra especializada, la gran "migración" de personal de los institutos de estadística hacia la empresa privada, los sueldos generalmente bajos, la frecuencia de procesamiento censal muy reducida (promedio de 10 años), son algunas de las razones que demandan la casi obligatoriedad del procesamiento censal a través de paquetes.

En el procesamiento de los censos de los años 80 se usaron, con mucha frecuencia y éxito, algunos programas que estaban disponibles, tales como el CONCOR para la crítica y corrección, el SISTHARG para el cálculo de frecuencias, el CENTS para la tabulación y, en menor escala, el SPSS o el SAS para cruces básicos y análisis de variables. 3/

Para la etapa de entrada de datos ya existen también algunos programas que, dentro de sus limitaciones (y precios), podían ser usados en los censos.

Sin embargo, estos esfuerzos no serán suficientes para hacer frente a la labor involucrada en los próximos censos, no sólo porque no tocan todas las etapas del procesamiento, sino también porque los requisitos y características deseables para los programas deberán tener mayor amplitud, disminuyendo aún más la tarea de programación. Este es un tema muy importante, que merece ser discutido más profundamente. 4/

La existencia o no de estos paquetes para las diversas etapas censales, principalmente versiones para ser instaladas en los microcomputadores, será la variable fundamental que operará como limitante o potencial de expansión para la utilización de los micros en los censos de los 90.

De acuerdo a lo señalado anteriormente, se espera un gran aumento de las potencialidades de los microcomputadores, pero en el caso de América Latina, es necesario que este desarrollo - para que sea totalmente provechoso - esté acompañado por la existencia de paquetes para procesamiento estadístico.

IV. TIPIFICACION SEGUN EL TAMAÑO DE LA POBLACION

Aunque la experiencia censal en América Latina pueda ser transmitida de uno a otro país, hay problemas específicos difíciles de generalizar. Radicalizando el raciocinio, no se puede tomar los conceptos de procesamiento utilizados en México, por ejemplo, y aplicarlos en Santa Lucía, ya que la diferencia básica está en el tamaño de sus

3/ Silva, A., Procesamiento de Datos op. cit., capítulo IV, pgs. 31-32.

4/ Silva, A., La Necesidad del Desarrollo de Sistemas para los Censos de América Latina, Santiago, inédito.

respectivas poblaciones.

Evidentemente, éste no es el único factor que se debe tomar en cuenta, pero desde el punto de vista computacional es, sin duda, el más importante. Existen otros elementos diferenciales, como, por ejemplo, el grado de avance tecnológica de los respectivos institutos de estadística, las diferentes posibilidades de acceso a equipos más poderosos, la situación económica local, la importancia que los respectivos gobiernos asignen a la información estadística, o más bien, la cantidad de fondos que será reservada para el proyecto censal, etc. Todas estas consideraciones tienen un mayor o menor grado de impredecibilidad, que puede variar significativamente desde ahora hasta la época del censo; además, esos factores serán afectados, directamente, por una eventual variación de la política de los organismos internacionales que tradicionalmente prestan su aporte financiero y tecnológico a los proyectos censales.

Por consiguiente, de todos los aspectos susceptibles de considerar en estas proyecciones acerca del procesamiento de los datos censales en la década del 90, el tamaño de la población constituye la variable menos aleatoria. Usando las cifras proyectadas para 1990 de los distintos países de América Latina, es posible identificar cuatro grupos de países que se describen sucintamente a continuación. 5/

En el capítulo siguiente, estos grupos serán determinantes para el establecimiento de los procesos que se usarán en las diferentes etapas de los censos.

1. Países muy grandes (población de más de 20 millones)

En este grupo están Argentina, Brasil, Colombia, México, Perú y Venezuela. En rigor, ésta es una agrupación algo disímil, pues si se consideran los extremos (Brasil con 159 millones y Venezuela con 21 millones, proyectados), se advierte un rango muy amplio. Quizás sería preferible aislar a Brasil, o crear un grupo separado sólo para Brasil y México (89 millones). Sin embargo, para efectos de especificación de los métodos de trabajo en las etapas de computación que se mencionan, estos dos países se comportan de la misma manera que los demás del grupo. Además, como ésta es una agrupación puramente teórica, sujeta a adaptaciones de acuerdo a la futura realidad de cada país, se estima prudente considerar al grupo tal como aquí se le ha identificado.

5/ CELADE. Boletín Demográfico, Santiago, año XVIII, número 35, enero de 1985.

2. Países de gran población (entre 10 y 20 millones)

Es el caso de Chile, Cuba y Ecuador. Se podría también integrar a Guatemala (9,6 millones) en este grupo, sin mayores problemas de homogeneidad.

3. Países de tamaño medio (población entre 1 y 10 millones)

Forman parte de este grupo, entre otros, Bolivia, Honduras, Paraguay y Uruguay. Dentro de este grupo, la ayuda de los organismos internacionales juega un rol importante, lo que puede diferenciar el tipo de procesamiento entre un país y otro. De la misma manera, la dependencia del desarrollo de paquetes es más acentuada que en los grupos anteriores.

4. Países de poca población (menos de 1 millón)

Básicamente, este grupo está conformado por la mayoría de los países del Caribe de habla inglesa, con la excepción de Jamaica (perteneciente al grupo 3). Los casos de Guyana y Trinidad y Tobago (1 millón y 1,3 millones, respectivamente) son discutibles, pudiendo incorporárselos al grupo 3 ó 4, de acuerdo al avance de la tecnología de microcomputadores. Como la elección de los límites entre los grupos no es rígida, se incluyen ambos países en el grupo 4.

Tal como se ha sostenido reiteradamente, estos agrupamientos son puramente teóricos y tienen un carácter provisional; otros factores, como los mencionados anteriormente, pueden reubicar a un país dentro de los diferentes grupos, lo que también puede suceder con los cambios tecnológicos, que por su vez introducirían cambios en estos conceptos.

V. METODOS DE PROCESAMIENTO

No se trata aquí de mencionar todas las etapas del procesamiento censal, ni tampoco de describirlas o definir las en profundidad. Más que nada, la labor está centrada en los elementos o etapas que tienen una gran correlación con el volumen de datos que deben procesarse. Estas etapas son:

- Planificación
- Cartografía
- Empadronamiento
- Entrada de datos
- Control primario
- Consistencia y limpieza de datos
- Tabulación
- Base de datos

V.1. Planificación.

Nunca está demás insistir en la necesidad de una buena planificación en un proyecto tan complejo como los censos de población. Esta etapa, en el pasado muchas veces olvidada, o acortada, puede verse notablemente fácil con la utilización de los computadores.

Independientemente de la tipificación descrita en el capítulo anterior, todos los países deberán usar los microcomputadores en la planificación para el diseño de cronogramas, elaboración y seguimiento del presupuesto, preparación de toda la documentación, etc.

V.2. Cartografía.

La introducción de los procedimientos computacionales en el trabajo cartográfico es una necesidad. Sin la ayuda del computador va a ser muy difícil garantizar una cobertura plena y sistemática para todo el país en los próximos censos. La creación de una base cartográfica computerizada permitirá la producción automática de mapas con mucho mayor precisión y rapidez, en cualquier nivel político-administrativo, y permitirá también la conjugación de los archivos censales con los datos cartográficos,

generando resultados en forma gráfica.

Igual que en la planificación, los microcomputadores deben ser usados en esta fase, en todos los países, de manera centralizada por las oficinas de estadística.

9.3. Demás etapas.

La agrupación de los países según el tamaño de su población no tuvo ninguna influencia en el establecimiento de los microcomputadores como instrumentos de trabajo para las etapas anteriores. A partir del Empadronamiento se hacen notar las diferencias, y por eso, se presentan estas etapas dentro de cada grupo.

1. Países muy grandes (población de más de 20 millones)

1.1 Empadronamiento.

Ya en la década de 1980, cuatro de los cinco países de este grupo (Argentina, Brasil, Colombia y Perú), levantaron sus censos, en parte, por muestra, variando el porcentaje de muestreo y el método de selección según cada país.

En todos los casos, el procedimiento de recolección de datos fue el mismo: se usaron dos tipos de cuestionario, uno llamado "Básico", para toda la población, que contenía pocas preguntas (parentesco, sexo y edad, por ejemplo), y otro llamado "Ampliado" o "Extendido", que contenía, además de las preguntas del cuestionario básico, las otras consultas tradicionales de un censo de población. El cuestionario "Ampliado" fue aplicado a una fracción (la muestra) de las viviendas, y el cuestionario Básico se empleó en las viviendas restantes.

La razón fundamental por la cual los países de este grupo aplicaron el muestreo fue, sencillamente, el volumen total de información a procesarse, lo que significa, en última instancia, tiempo y costos. Aparentemente, este procedimiento dió resultados satisfactorios y no habría razón para no repetirlo en el futuro.

El hecho de que se use o no un procedimiento de muestreo en el censo afecta al procesamiento de datos en varios puntos; por ejemplo:

a) disminución del volumen total de procesamiento, en una proporción que es dependiente de los factores de muestreo;

b) existencia de dos tipos de cuestionarios, lo que significa dos subsistemas de producción; y

c) necesidad de una etapa adicional de ponderación, para el cálculo de los factores de expansión, cuya complejidad depende del método de selección de la muestra.

De todas maneras, la decisión debe tomarse en términos de la calidad de la información final versus la rapidez y el costo de su procesamiento.

1.2 Entrada de datos.

La utilización de la técnica muestral en el relevamiento es importante, porque puede variar fundamentalmente la cantidad de datos que debe ingresarse. Por ejemplo, el uso de una muestra del 20 por ciento significaría una reducción de alrededor de 2/3 en la entrada de datos, dependiendo del número de preguntas en cada cuestionario. En el caso de Venezuela, por ejemplo, se podría decir que el volumen censal correspondería a alrededor de 7 millones de registros, lo que, a primera vista, podría significar un cambio de grupo de procesamiento, pero no hay que olvidarse de los otros efectos colaterales de la muestra.

Independientemente del uso del muestreo en la recolección, se puede prever dos decisiones básicas para el procedimiento de captura de la información: lectora óptica versus digitación; y centralización versus descentralización. Estos dos criterios son combinables entre sí, generando cuatro hipótesis distintas para la entrada de datos.

1.2.1 Lectora óptica versus digitación.

La definición del método que se empleará para ingresar los datos se restringe a una comparación de las ventajas y desventajas de cada uno. Sin profundizar demasiado en el tema, se podría caracterizar a la lectora óptica actual como un equipo caro, rápido, que exige un frecuente ajuste y mantenimiento, necesita un papel y una impresión de buena calidad para el cuestionario, y que exige un mayor cuidado en el manejo de las boletas en todas las etapas previas a la lectura.

Por otro lado, la digitación es un proceso relativamente más barato, más lento (dependiendo del número de operadores), con un mayor porcentaje de errores, y como involucra una mayor cantidad de personal, necesita más

entrenamiento, espacio físico para el trabajo, administración, etc. 6/

Como regla general, se puede deducir entonces que la lectora óptica es adecuada cuando se tiene un gran volumen de información, siempre y cuando sea posible controlar los problemas de su utilización. La digitación sería el método más indicado en los otros casos.

Los avances tecnológicos hacen prever, para este grupo de países, una clara opción a favor de las lectoras ópticas. La utilización del ingreso por medio de la digitación dependería de criterios secundarios, tales como el aprovechamiento de la mano de obra disponible en el país, etc.

Por incongruente que pueda parecer, los dos mayores países de este grupo (Brasil y México) no usaron lectoras ópticas en sus censos de los años 80.

1.2.2 Centralización versus descentralización.

En el caso de las lectoras ópticas, el criterio más razonable sería el de la centralización, para concentrar los recursos de entrada, usarlos como respaldo entre ellos, etc. En los casos de Brasil y México (los dos países más grandes), dependiendo de los recursos disponibles, se podría pensar excepcionalmente en la descentralización de las lectoras ópticas. Para los demás, esta descentralización dependería de la cantidad de equipos que se usarán en el censo y, lógicamente, de sus costos.

En el caso del ingreso de datos por digitación, la descentralización de actividades parece ser lo más adecuado, siempre que se mantenga el número de centros de operación dentro de cantidades manejables - y sólo en las grandes ciudades - para evitar o disminuir los factores negativos del proceso (administración y coordinación de los centros, disponibilidad de mano de obra calificada, uniformidad de procedimientos y costos, etc.)

1.3 Control primario (cantidades).

La ejecución del subsistema de Control Primario, también llamado de Control de Cantidades, es necesaria para verificar, entre otros asuntos, el volumen de material que se procesará (cuestionarios y personas) y garantizar las

6/ Para mayores detalles, véase Silva, A., Procesamiento de Datos para CENSA, capítulo III, pgs. 23-24.

variables de Ubicación Geográfica en la boleta censal.

Como el objetivo de esta etapa es controlar el ingreso de todas las boletas y personas, garantizando que no haya duplicación, este proceso debe realizarse lo más cerca posible del local del ingreso de datos, con el fin de facilitar las correcciones. Así, la decisión de descentralizar o no la entrada de datos será la que, a la postre, localizará también el Control Primario.

Con relación al tipo de equipo que se utilizará, esto dependerá también de la posibilidad de conexión entre una lectora óptica y un microcomputador, o de que las propias lectoras ópticas aumenten sus capacidades de procesamiento o funciones. Como ambas posibilidades son perfectamente previsibles, esta etapa puede ser ejecutada con microcomputadores.

1.4 Consistencia y limpieza.

En la década de 1980, en la casi totalidad de los países (de todos los grupos), esta etapa fue realizada con procedimientos automáticos, especificándose las reglas de consistencia y de corrección, con buenos resultados. Para los años 90, estos procedimientos deben ser repetidos. Como se trata de modalidades de operación puramente automáticas, ellas deben ser centralizadas, de modo que se pueda usar la mayor velocidad de los computadores de mayor tamaño, y que se consiga maximizar el refinamiento y la calidad en los procesos de asignación y análisis de la información recolectada.

En los casos en que la entrada de datos es descentralizada, es necesario definir el método de transmisión de los datos al computador central. El avance en los sistemas de telecomunicación, la utilización de satélites, etc, aparentemente llevarían la decisión por este camino. Sin embargo, el procesamiento censal no exige una transmisión de datos en línea y "al minuto", ni mucho menos. Es perfectamente razonable una demora de 5 a 10 días, a veces más, entre el ingreso de la información (incluyendo el Control Primario), y el inicio del procesamiento centralizado. Por lo tanto, es recomendable que esta transferencia de datos sea hecha a través del transporte físico de algún medio magnético (cintas, cartuchos, etc.). Esta es una práctica más barata y más segura, que no exige los sofisticados controles necesarios para garantizar la transmisión segura de los datos por medio de la telecomunicación, a no ser que se produzcan cambios substanciales en las telecomunicaciones, acompañados de una disminución de sus costos, y que, lógicamente, esté disponible en los países de América Latina.

Finalmente, faltaría decidir entre la utilización de los computadores de gran tamaño o los microcomputadores. Aunque se tome en cuenta todo el avance de la tecnología, actual y del futuro próximo, es difícil imaginar estos procedimientos en microcomputadores, sobre todo en países en vías de desarrollo. Todo lo anterior lleva a pensar que lo más apropiado sea la utilización de los grandes computadores.

1.5 Tabulación.

Una vez decidida la centralización del proceso anterior, es razonable que esta etapa también lo sea. El gran volumen de información (magnética), los procesos automatizados de procesamiento y el tamaño de los programas de tabulación, son las razones principales para preferir la centralización.

Adicionalmente se puede mencionar la sofisticación de estos programas, con la posibilidad de conexión directa con equipos de fotocomposición y producción de microfichas. En rigor, es más que una posibilidad; se trata, más bien, de una obligatoriedad, toda vez que ella permite aumentar la calidad de los volúmenes publicados y facilita tanto el almacenamiento como las consultas de otros resultados no publicados.

1.6 Base de datos.

La creación de una base de datos censales también sería centralizada, decisión que se basa en los mismos motivos de la centralización del proceso anterior. Eso no significa, sin embargo, que no se pueda tener copias de una parte (o del todo) en otros computadores. Por otro lado, la comunicación entre computadores facilita al investigador el acceso a la base de datos casi en cualquier región del país.

2 Países de gran población (entre 10 y 20 millones)

2.1 Empadronamiento.

Como se mencionó anteriormente, el empadronamiento no corresponde a una decisión que se adopte como parte del procesamiento de datos.

2.2 Entrada de datos.

Igual que para el grupo 1 de países, hay que decidir entre la lectora óptica y/o la descentralización.

2.2.1 Lectora óptica.

En este caso, no parece tan obvia la ventaja de la lectora óptica sobre la digitación. Su utilización dependerá básicamente de los costos de esta tecnología en 1990. Como ejemplo, en este grupo está el caso de Chile, que usó la lectora óptica en su censo de 1982 y cuyos resultados, desde el punto de vista de la entrada de datos, fueron buenos. Los problemas encontrados fueron consecuencia de que las etapas posteriores al ingreso de datos no tuvieron la misma velocidad. 7/

2.2.2 Descentralización.

Aquí tampoco la decisión es clara en favor de la descentralización. Su adopción dependerá de la disponibilidad de los recursos existentes en las ciudades principales. Como precedente, en este grupo está el caso del Ecuador, que descentralizó su entrada de datos (por digitación) en dos ciudades, Quito y Guayaquil, con algunos problemas iniciales de ajuste de procedimientos que luego se solucionaron), pero cuyos resultados finales fueron satisfactorios.

2.3 Control primario.

Valen las mismas observaciones hechas en el punto 1.3.

2.4 Consistencia, tabulación y base de datos.

Estos procesos deben ser centralizados y ejecutados en equipos de gran tamaño, por las mismas razones mencionadas en los puntos 1.4, 1.5 y 1.6.

7/ Tacla, O., Breves comentarios, según etapas censales, con especial énfasis en ingreso de datos, Buenos Aires, mayo de 1985, presentado en el Taller de Análisis y Evaluación de los Censos del 80.

3 Países de tamaño medio (población entre 1 y 10 millones)

3.1 Estandarización.

El beneficio obtenido al reducirse los volúmenes de datos para procesar no justifica los problemas y dificultades encontrados en el manejo de dos tipos de cuestionarios, como ocurre cuando se hace uso del muestreo, excepto en el caso en que el número de preguntas "extras" sea muy grande. Aun así, quizás sería más práctico realizar una encuesta paralela al censo que tuviera sólo las preguntas específicas adicionales.

3.2 Entrada de datos.

No tiene sentido pensar en una descentralización de estas actividades, sino más bien en concentrar los recursos para su mejor aprovechamiento.

La utilización de una lectora óptica no sería recomendable, excepto en el caso, ya comentado anteriormente, en que los costos operativos y las exigencias técnicas del equipo desciendan hasta alcanzar magnitudes que los países de este grupo puedan solventar.

3.3 Control primario.

Deberá ser ejecutado centralizadamente, en un microcomputador. La elección de esta tecnología para el grupo considera el hecho de que, primero, la entrada de datos es por medio de microcomputadores y, segundo, que los volúmenes son perfectamente manejables si son trabajados por lotes de procesamiento.

3.4 Consistencia y tabulación.

Estos dos procesos se realizarán, probablemente, en un equipo grande. La posibilidad de utilización de microcomputadores dependerá de un futuro aumento en su capacidad de memoria secundaria (discos duros y/o cintas magnéticas), y de la existencia de paquetes de programación razonablemente eficientes.

3.5 Base de datos.

La creación de la base de datos se haría en el mismo computador usado para las etapas anteriores. Para su máximo aprovechamiento, se puede prever la utilización de microcomputadores, apoyados en los paquetes específicos para

tales trabajos. 8/

4 Países de poca población (menos de 1 millón)

Para estos países, se prevé la utilización de los microcomputadores en todas las etapas, con un procesamiento centralizado.

Con el objeto de facilitar el uso de microcomputadores, es imprescindible la existencia de un conjunto de paquetes destinados al procesamiento censal (véase el capítulo III), sobre todo porque se trata de países pequeños que no pueden asignar recursos a proyectos de esta naturaleza.

La utilización de una lectora óptica dependerá de consideraciones relativas a costos y a la capacidad de acoplamiento a un microcomputador (o una microlectora óptica).

5 Cuadro resumen.

Los cuadros 1 y 2, que se presentan a continuación, proporcionan la composición de los grupos de países y una visión resumida de las opciones más probables para cada etapa del procesamiento censal, según el grupo de países.

CUADRO 1
COMPOSICION DE LOS GRUPOS DE PAISES
(y población estimada, en miles, para la década de 1990)

PAISES MUY GRANDES (más de 20 millones)

Argentina	30.277
Brazil	153.271
Colombia	31.820
México	89.012
Perú	23.355
Venezuela	21.284

PAISES GRANDES (entre 10 y 20 millones)

Chile	13.061
Cuba	10.540
Ecuador	10.949
Guatemala	9.676

PAISES MEDIANOS (entre 1 y 10 millones)

Bolivia	7.314
Costa Rica	2.766
El Salvador	6.484
Haiti	7.509
Honduras	5.105
Jamaica	2.535
Nicaragua	3.778
Panamá	2.346
Paraguay	4.231
Puerto Rico	4.747
República Dominicana	6.971
Uruguay	3.128

PAISES PEQUEÑOS (menos de 1 millón)

Antillas Neerlandesas	313		
Bahamas	295	Guyana Francesa	73
Barbados	292	Guyana	1.069
Belice	205	Martinica	337
Dominica	95	Santa Lucía	136
Grenada	125	Surinam	527
Guadalupe	339	Trinidad y Tobago	1.337

CUADRO 2

OPCIONES PROBABLES PARA EL PROCESAMIENTO DE LOS CENSOS DE 1990, POR ETAPAS, SEGUN TIPO DE PAIS

ETAPAS	TIPOLOGIA DE PAISES SEGUN TAMAÑO			
	MOY GRANDES -- en millones (más de 20)	GRANDES (10 - 20)	MEDIANOS - (1 - 10)	FLOJEROS (menos de 1)
1-Planif.	Microcomp.	Microcomp.	Microcomp.	Microcomp.
2-Cartogr.	Microcomp.	Microcomp.	Microcomp.	Microcomp.
3-Empedr.	Muestra	Universo	Universo	Universo
4-Entrada de datos	Lectores descentr.	Lectores descentr. ----- Digitación descentr.	Digitación central.	Digitación microcomp. central.
5-Control primario	Microcomp. descentr.	Microcomp. c/ Entrada	Microcomp. central.	Microcomp. central.
6-Consist.	Computador grande central.	Computador grande central.	Computador grande central.	Microcomp. central.
7-Tabulac.	Computador grande central.	Computador grande central.	Computador grande central.	Microcomp. central.
8-Base de datos	Computador grande central.	Computador grande central.	Generación Computador grande ----- Explotación Microcomp.	Microcomp. central.

VI. CONCLUSION

La tipificación de los países según el tamaño de su población es una simplificación muy gruesa, y por lo tanto peligrosa. Sin embargo, el volumen de material es uno de los más importantes factores en la elección del método de procesamiento.

Adicionalmente, se deben "ajustar" las decisiones adoptadas en base a los volúmenes de información, tomando en cuenta otros elementos específicos de cada país, como, por ejemplo, el grado de sofisticación tecnológica alcanzada por sus respectivas oficinas de estadística.

Finalmente, deberán estudiarse los dos grandes imponderables del futuro, que son las nuevas tecnologías en computación (principalmente los microcomputadores) y la disponibilidad de paquetes estadísticos. La existencia de estos programas, dirigidos al procesamiento censal, deberá ser el punto neurálgico que dirigirá las decisiones sobre la utilización masiva de los microcomputadores en los años 90.