



Comisión Sectorial de Población

# Estimación de Población en áreas menores con métodos que utilizan variables sintomáticas

Informe final  
Diciembre 2011

Mariana Cabrera







## **Estimación de Población en áreas menores con métodos que utilizan variables sintomáticas**

Este trabajo se realizó en el marco del Programa Conjunto "Unidos en la Acción" - Sistema Naciones Unidas en Uruguay "Construyendo capacidades para el desarrollo", Proyecto H "Apoyo a los procesos de modernización de la gestión pública en el marco de la democratización del Estado y la promoción de los derechos de los ciudadanos", coordinado por la Organización Internacional para las Migraciones (OIM) y con el apoyo del Fondo de Población de Naciones Unidas (UNFPA).

El análisis y las recomendaciones contenidos en esta publicación no reflejan necesariamente las opiniones de la Comisión Sectorial de Población ni la de los Ministerios y organismos que integran la misma.





La Comisión Sectorial de Población (CSP) fue creada por la Oficina de Planeamiento y Presupuesto a través de la Resolución N° 180/2010 del 3 de agosto de 2010, la cual establece tanto sus objetivos como la integración institucional.

Para su funcionamiento operativo, la CSP cuenta con un Plenario, con representantes del Poder Ejecutivo, Parlamento, Academia y organismos de cooperación internacional, y un Comité Ejecutivo, el cual está integrado por representantes de los siguientes organismos:

Oficina de Planeamiento y Presupuesto (OPP)  
Ministerio de Desarrollo Social (MIDES)  
Ministerio de Economía y Finanzas (MEF)  
Ministerio de Educación y Cultura (MEC)  
Ministerio de Salud Pública (MSP)  
Ministerio de Trabajo y Seguridad Social (MTSS)  
Ministerio de Turismo y Deporte (MINTUR)  
Ministerio de Vivienda, Ordenamiento Territorial y Medio Ambiente (MVOTMA)  
Ministerio de Relaciones Exteriores (MRREE) – en calidad de invitado

La CSP tiene los siguientes cometidos:

- a. Proponer políticas que promuevan la equidad intergeneracional en las condiciones de vida y acceso a oportunidades.
- b. Proponer políticas que aborden el tema de la natalidad así como la compatibilización de la vida productiva con la reproductiva.
- c. Proponer la articulación de las políticas de vinculación con la población uruguaya residente en el exterior y retorno de nacionales con el Plan Nacional de Desarrollo.
- d. Analizar la conveniencia de una política activa de promoción de la inmigración y de sus características.
- e. Proponer políticas que permitan una distribución territorial más adecuada de la población y de los procesos migratorios internos, incluyendo la migración temporal transfronteriza, en armonía con los sistemas ambientales.
- f. Proponer una política de producción de datos, información y análisis sociodemográficos y poblacionales, cuyos productos sean insumos para el diseño, implementación y monitoreo de políticas públicas.
- g. Asesorar en materia de política demográfica poblacional toda vez que le sea requerido.





# Indice

Introducción .....	10
1. Estimaciones de población en áreas menores .....	11
1.1 ¿Para qué estimar la población en áreas menores? .....	11
1.2 Estimaciones y proyecciones de población.....	12
1.3 Dificultades para estimar y proyectar el volumen de población en áreas menores .....	15
1.4 Uso de variables sintomáticas .....	17
1.5 Encuestas de hogares y estimaciones de población en áreas menores.....	19
2. Métodos para la estimación de población en áreas menores .....	20
2.1 Panorama general.....	21
2.2 Métodos que utilizan variables sintomáticas como auxiliares .....	21
2.3 Métodos que utilizan variables sintomáticas para modelar el cambio y el volumen de la población .....	23
2.4 Herramientas para evaluar los resultados de los métodos.....	29
2.5 Aplicaciones a nivel mundial y latinoamericano .....	32
3. Uruguay: fuentes disponibles y delimitaciones de áreas .....	38
3.1 Fuentes disponibles en Uruguay .....	38
3.2 Delimitación de áreas.....	41
3.3 Hacia la aplicación de los métodos .....	42
4. Exploración de modelos para estimar poblaciones departamentales.....	44
4.1 Métodos aplicados, variables sintomáticas utilizadas y herramientas de evaluación de modelos .....	45
4.2 Resultados obtenidos.....	49
4.3 Estimaciones y proyecciones al 2011.....	53
5. Comentarios finales .....	55
5.1 Sobre los resultados.....	55
5.2 Pasos a seguir a mediano y largo plazo.....	57
Bibliografía.....	60
Anexos.....	62

## ***Introducción***

El presente informe sintetiza el trabajo realizado durante la consultoría para la exploración de modelos para estimar población en áreas pequeñas utilizando variables sintomáticas y aplicado a áreas del Uruguay. Incluye una primer parte en la que se exponen diferentes métodos y experiencias internacionales así como las fuentes disponibles en Uruguay y diferentes tipos de delimitaciones de áreas menores para las que podría interesar estimar volúmenes poblacionales. Estos dos grandes apartados habían sido presentados ya en un informe de avance de la consultoría. Se agrega finalmente la exploración realizada en base a un conjunto de variables sintomáticas y métodos para la estimación de población a nivel departamental. Se contrastan las estimaciones obtenidas con los datos censales de 2004, con las proyecciones al 2011 y con los resultados –aún preliminares- del Censo 2011.

La primera parte contiene una breve aproximación conceptual de distintos métodos propuestos para estimar volúmenes poblacionales en áreas geográficas menores en períodos inter y post-censales. Se presenta en primer lugar la justificación del desarrollo de este tipo de herramientas como insumo pertinente y útil para el diseño de políticas y la gestión pública, así como para la actividad privada. Se definen los principales conceptos involucrados y se delimita en relación a otras herramientas, en particular a las proyecciones demográficas, mostrando sus utilidades particulares pero también la potencialidad de su combinación para analizar la dinámica poblacional en áreas geográficas menores. Existe una larga tradición de desarrollo de diferentes métodos para estimar la población en áreas menores. Actualmente existe un conjunto de métodos que son los utilizados más frecuentemente, cada uno con especificidades teóricas y prácticas. En el apartado sobre métodos se presentan algunas de estas herramientas con los principales supuestos que utilizan así como las ventajas y desventajas que han sido detectadas por quienes las utilizan. Se repasan también algunas experiencias a nivel mundial y regional. En América Latina el desarrollo está en una etapa más experimental pero hay diversos estudios a nivel de exploración y prueba de los distintos métodos. De las experiencias extrarregionales se presentan algunos casos donde el uso de estos métodos es sistemático y forma parte de la información provista por las oficinas de estadísticas.

En el capítulo 3 el informe presenta un avance sobre las posibilidades de utilización de estos métodos en Uruguay. Se listan en primer lugar las fuentes de información

que han sido consideradas, con una primera evaluación sobre las posibilidades de su uso para aplicar algunos de los métodos propuestos. En segundo lugar se presentan diferentes niveles de divisiones territoriales para las cuales podría ser pertinente obtener indicadores sobre volumen poblacional y a su vez de acuerdo a las posibilidades que brinda la información existente, que es diferente según las variables utilizadas.

El capítulo 4 contiene el trabajo realizado entorno a la exploración de métodos para la estimación de poblaciones departamentales. Se exponen los puntos principales del proceso realizado: variables seleccionadas y acondicionamiento de las mismas, métodos aplicados y resultados obtenidos al comparar con valores poblacionales provenientes de otras fuentes (censos, proyecciones). En particular, se analizan distintos métodos basados en regresión lineal múltiple dado que son los que la experiencia internacional muestra como de mayor precisión para estimar volúmenes poblacionales en áreas menores. Sin embargo en este caso, como se verá, no generan modelos adecuados.

Como cierre, se presentan algunos comentarios sobre los resultados obtenidos y un plan de trabajo con metas a mediano y largo plazo. El objetivo final de este proceso de trabajo es obtener una herramienta para obtener estimaciones de volúmenes de población en distintos niveles de divisiones territoriales del país, herramienta que quede suficientemente validada, aunque como muestra la experiencia internacional, es posible y necesaria su revisión y mejora permanente.

## ***1. Estimaciones de población en áreas menores***

### ***1.1 ¿Para qué estimar la población en áreas menores?***

La caracterización de la población en términos demográficos y de sus condiciones de vida es un insumo fundamental para el diseño, implementación y evaluación de políticas públicas. Esta afirmación es especialmente válida cuando se considera la gestión en su dimensión territorial, sea de organismos de jurisdicción nacional como y particularmente de gobiernos locales. En este sentido, un insumo mínimo imprescindible es el número de personas que residen en las distintas áreas

geográficas del país (departamentos, municipios, localidades y zonas rurales, u otras divisiones del territorio).

Acciones como la localización de servicios públicos, planes de apoyo alimentario o de capacitación, la distribución de recursos presupuestales entre distintas divisiones político administrativas, la promoción de desarrollo de inversión privada o la recaudación fiscal que puede asociarse a determinadas áreas son ejemplos puntuales que muestran el uso intensivo de la información sobre el volumen de población, acompañada de otros elementos de caracterización de la misma. Esta información es también demandada desde el ámbito académico y desde el sector privado como insumo para la investigación y la planificación.

Por otra parte, el volumen de población en distintas áreas geográficas puede utilizarse como información auxiliar para ajustar o controlar los marcos muestrales de encuestas que permitan conocer otras características de la población residente, como la Encuesta Continua de Hogares del Instituto Nacional de Estadística. En este sentido, sirven también como insumo (al proveer un estimado del volumen poblacional) para otras herramientas estadísticas que buscan caracterizar la población y los hogares en áreas no cubiertas (o con representatividad insuficiente) por las muestras de encuestas.

La aplicación de estos instrumentos permite alertar sobre posibles dinámicas poblacionales locales que requieren mirar esas áreas con más detenimiento, aun en el caso que se utilicen modelos simples y/o que existan dificultades con la calidad de los datos.

## **1.2 *Estimaciones y proyecciones de población***

Los censos de población son instancias fundamentales para obtener un mapa de la población, los hogares y las viviendas del país en un momento dado. En particular, es la herramienta más adecuada para obtener cifras sobre la distribución y el volumen de personas que residen en cada área geográfica del país.

Sin embargo es una herramienta costosa por lo que su periodicidad es habitualmente no menor a 10 años, o más, como ha sucedido en Uruguay.

Los conteos de población (como el de 2004), aunque habilitan una actualización del volumen de población por áreas geográficas, tienen un alcance limitado en cuanto al resto de información que proveen por lo cual tampoco parece razonable una mayor periodicidad dada la relación costos-resultados.

Por tanto, el volumen de población para las distintas áreas geográficas se conoce con un grado importante de precisión para los años censales. Pero la distancia temporal entre censos implica que el mapa censal deje de representar en forma realista el volumen de la población según su distribución espacial, con lo cual se desactualizan estos insumos que son fundamentales para gestionar las políticas públicas en su dimensión territorial.

En países como Uruguay, en los que no se cuenta con otro tipo de registros de población que pretendan asociar en forma universal a las personas con su lugar de residencia (como existen en otros países), para los períodos intercensales y postcensales es necesario estimar el volumen de la población y su distribución en el territorio utilizando otro tipo de registros.

En la literatura demográfica se distinguen distintos tipos de estimaciones de volumen de población según las referencias temporales y las bases que utilizan para los cálculos (Shryock y Siegel 1973)<sup>1</sup>:

- a. **Estimaciones intercensales:** refieren a un período entre dos censos y consideran para los cálculos los resultados de esos censos
- b. **Estimaciones postcensales:** refieren a un período pasado o presente que es posterior a un censo y consideran para los cálculos los resultados de ese censo y también puede considerar resultados de censos anteriores.
- c. **Proyecciones:** refieren a un período posterior al último censo, habitualmente fechas futuras, también consideran para sus cálculos los resultados de ese censo pero no considera otras informaciones contemporáneas a los períodos a estimar.

Otros autores distinguen las estimaciones intercensales y postcensales y las proyecciones en base al período al cual hacen referencia (pasado y presente para las estimaciones; futuro para las proyecciones) y/o al tipo de supuestos que se realizan

---

<sup>1</sup> Además de las estimaciones enumeradas que son las más habituales, estos autores mencionan las estimaciones precensales (que refieren a períodos anteriores a un censo y no toman en cuenta censos anteriores sino el o los censos siguientes) y las estimaciones no-censales (que no se basan en resultados de ningún censo de población).

para calcularlas. En el caso de las estimaciones son supuestos sobre las posibilidad de reconstruir el volumen población con el apoyo de informaciones contemporáneas al período de la estimación y en el caso de las proyecciones utilizando supuestos sobre el comportamiento postcensal del crecimiento poblacional o de cada componente involucrado en esa dinámica -nacimientos, fallecimientos y migración.

En este contexto entonces, una **estimación** del volumen de población presente o pasada es una aproximación al número de población que reside en un área geográfica específica calculado utilizando una técnica analítica no censal ni muestral. Esta técnica utiliza como población base uno o más censos de población y datos obtenidos de registros administrativos contemporáneos al período a estimar como estadísticas vitales y datos de migración (que permiten una estimación directa) pero también puede utilizar otras variables que reflejen cambios en la población, habitualmente llamadas variables sintomáticas (estimación indirecta). Su utilidad radica en brindar aproximaciones sensibles a cambios demográficos no previstos, cambios de tendencias, etc. y por tanto pueden ser especialmente adecuados a niveles subnacionales y áreas con menores volúmenes de población.

Un **indicador o variable sintomática** es una variable que cambia en el tiempo en concordancia con cambios en el volumen de la población en el mismo período.

Las **proyecciones** de volumen de población en cambio tienen como objetivo obtener una aproximación al número de población utilizando como base también un censo de población y haciendo supuestos sobre el comportamiento para períodos posteriores al censo de los componentes de la dinámica (natalidad, mortalidad y migración). No incorpora en los cálculos datos contemporáneos al período para el cual se obtienen los valores. Tiene un papel fundamental como insumo para escenarios prospectivos.

Existen a su vez otras formas de clasificar o identificar estimaciones y proyecciones. Por ejemplo, por el área geográfica sobre la cual se calculan las cifras (países o niveles subnacionales de distinto orden); por el tipo de población para la que se realizan los cálculos (población total, determinados grupos de edad o sexo, población activa, etc.). (Shryock y Siegel 1973)

Una distinción importante que se deriva de las características de los censos utilizados como base es si se trata de población con residencia habitual (de jure) o

población presente físicamente al momento del censo (de facto). (Shryock y Siegel 1973)

### **1.3 Dificultades para estimar y proyectar el volumen de población en áreas menores**

A partir del uso de la ecuación compensadora<sup>2</sup>, la demografía provee distintas herramientas que permiten realizar estimaciones a nivel del total del país (considerando nacimientos, defunciones y una estimación del saldo migratorio internacional). A su vez, haciendo supuestos sobre el comportamiento futuro de nacimientos, defunciones y migración internacional se construyen proyecciones de población.

Pero las estimaciones del volumen y la distribución de población a nivel subnacional requieren tener información actual o hacer supuestos específicos sobre los valores actuales (al momento que se quiere estimar) de cada componente (nacimientos, defunciones y migración) para cada área geográfica para la que se quiere obtener valores poblacionales. En el caso de las proyecciones implica además supuestos sobre el comportamiento futuro de estos componentes en cada área.

Las cifras de nacimientos y defunciones pueden ser obtenidas a partir de los registros públicos, en los cuales se identifica habitualmente el lugar de residencia de la madre del recién nacido o del fallecido. Las dificultades en este caso son fundamentalmente: la omisión o imprecisión de este dato en los registros (o el utilizar codificaciones del lugar de residencia no estandarizadas), el subregistro de los eventos, la demora en el registro de los mismos y –a otro nivel- la demora en el procesamiento de la información por parte de los organismos públicos responsables y por tanto también el acceso oportuno a las cifras. En el caso de Uruguay el subregistro es casi inexistente y la demora de la inscripción pública de los eventos no es tampoco problemática. Sin embargo, persisten aún dificultades asociadas a los otros factores planteados.

Sin duda, el componente más complejo para modelar en cualquier estimación o proyección es la migración, pero en caso de aplicar estos modelos a divisiones

---

<sup>2</sup> La ecuación compensadora expresa que la población en un período dado es igual a la población en un momento anterior mas el saldo entre nacimientos y fallecimientos entre los dos periodos (diferencia por dinámica vegetativa) mas el saldo migratorio (diferencia entre inmigrantes y emigrantes internacionales en el período)

subnacionales es aún más complejo ya que es necesario considerar la migración interna entre distintas áreas del país (agregada a la internacional que impacta en forma diferencial en el territorio) para lo cual no hay ninguna fuente registral sistemática.

Uruguay cuenta actualmente con proyecciones de población de las áreas urbanas y rurales y departamentales, utilizando supuestos sobre cada uno de los componentes. Estas proyecciones son elaboradas por el Instituto Nacional de Estadística. La última revisión fue realizada en 2005.

Sin embargo, es mucho más difícil hacer supuestos fuertes sobre el comportamiento futuro de nacimientos, defunciones y sobre todo migración a un nivel más desagregado (por ejemplo, localidades, municipios).

En el caso de las estimaciones postcensales, es posible obtener los registros de nacimientos y defunciones por residencia (aunque como se mencionó es necesario evaluar la calidad de esta información a un nivel tan desagregado) pero no hay información sobre los movimientos migratorios a nivel local. Esta ausencia de registros impide considerar los cambios que se generan por factores coyunturales en la distribución espacial de la población (migración interna e internacional y los cambios en las tasas de natalidad y mortalidad que pueden verse afectadas si se modifica la composición de la población por migración interna, entre otros). Inversiones o cierres de emprendimientos que generan/cierran puestos de trabajo directos e indirectos en una localidad pequeña, otros cambios en la matriz productiva de la zona –como el turismo, la forestación, el cultivo de soja-, la implantación de un plan de viviendas o la promoción de un área para fines residenciales, etc., son ejemplos de factores que pueden afectar en forma drástica la composición de la población localmente. Además, en una perspectiva dinámica, estos cambios poblacionales modifican a su vez el entorno económico, social y ambiental local. Es un proceso de retroalimentación específico de ese territorio.

Estas características de la dinámica de la población en áreas menores han llevado al desarrollo de métodos alternativos o complementarios a los mencionados anteriormente tanto para realizar estimaciones como proyecciones.

En nuestro país existe una experiencia de elaboración de proyecciones por localidades realizada por J.J.Calvo como insumo de un estudio en convenio de UTE con la Universidad de la República, que abarca el período 2000-2030. Para esta



proyección se tomó como base las proyecciones nacionales realizadas por el método de los componentes y se aplicó el método de Duchesne<sup>3</sup> para proyectar población por departamentos; finalmente se aplicó el método del parque habitacional para proyectar secciones y localidades. Este método considera las viviendas particulares existentes al último censo y la información sobre permisos de construcción de viviendas particulares.

Es decir, la elaboración de estas proyecciones requirió utilizar información adicional sobre viviendas y permisos de construcción y además el supuesto sobre que no hay diferencias en el movimiento migratorio entre las distintas localidades de un departamento.

#### **1.4 Uso de variables sintomáticas**

La idea básica es aprovechar el hecho que los cambios en el volumen de la población modifican el comportamiento de otros aspectos a nivel local (por ejemplo: el número de niños en las escuelas, las conexiones y consumo de electricidad, el uso de servicios de salud, etc.). El paso fundamental estriba en buscar algún modelo matemático que permita cuantificar los cambios o el volumen de población utilizando esa información auxiliar.

Como se verá en las experiencias relevadas y en la situación en Uruguay las dificultades para modelizar en base a variables sintomáticas refieren a dos grandes temáticas: la delimitación de las áreas geográficas y las características de las variables sintomáticas.

En relación a las áreas geográficas, más allá de las posibilidades de encontrar información de las variables sintomáticas para los desagregados deseados, existe un problema asociado a los cambios en los límites de las mismas. En particular, esto afecta tanto a las áreas delimitadas en los censos (y cambios de un censo a otro) como en la identificación de las áreas para las fuentes que relevan las variables sintomáticas (y sus cambios en el tiempo). Tanto las diferentes definiciones de límites según fuente como los cambios de las mismas en el tiempo inciden (y pueden

---

<sup>3</sup> El método de Duchesne es llamado también el método de la "relación entre cohortes". Consiste fundamentalmente en proyectar coeficientes que miden, para cada grupo de cohortes de cada área pequeña, el diferencial de crecimiento en relación a un área mayor de referencia para la cual se dispone de una proyección previa, en este caso cada departamento.

inhabilitar incluso) la modelización. Requiere un análisis cuidadoso de yuxtaposición de límites y búsqueda de agregados de áreas (hasta obtener “mapas” compatibles entre fuentes y censos) o supuestos sobre el comportamiento de las áreas que no logran yuxtaponerse.

Por su parte, las variables sintomáticas son básicamente registros administrativos. Un elemento central a considerar es la disponibilidad tanto en cuanto a su oportunidad (es decir, que se pueda acceder en forma relativamente rápida en relación al momento en que se genera el registro) como en cuanto a los desagregados geográficos con los que se presenta la información. En relación a este último punto, es necesario saber si es un problema de presentación solamente (con lo cual podría llegar a accederse a una mayor desagregación) o si el relevamiento o la naturaleza del registro impide obtener datos más desagregados geográficamente (que pueden no ser pertinentes para la variable). Cuando se utilizan series es fundamental evaluar la consistencia de las mismas y realizar ajustes si es necesario (por ejemplo, en caso que la evaluación muestre determinados cambios en las variables que pueden atribuirse a factores no asociados con cambios poblacionales). Otra dificultad refiere a la calidad de la información: puede haber subregistros o registros tardíos, puede haber déficits en el llenado del registro (en particular los campos que identifican áreas de residencia, etc.); por otra parte, las herramientas utilizadas para registrar pueden cambiar (por mejoras, por cambios de objetivos, etc.) y por tanto hacer variar el dato por factores no asociables a cambios poblacionales. Lo mismo sucede en relación a la cobertura (cambios en la población atendida por un programa, apertura de otros centros educativos u otras oficinas locales, etc.).

Finalmente es necesario evaluar las magnitudes de los valores de las variables involucradas en la modelización ya que al tratarse de áreas pequeñas es posible que algunas variables tengan oscilaciones muy fuertes (en términos relativos) por el bajo número de casos que se registran (por ejemplo, nacimientos o defunciones). En estos casos puede ser necesario utilizar por ejemplo, promedios móviles entre años contiguos, etc.

Los métodos que se presentan en este informe forman parte de la familia de herramientas que utilizan estas informaciones auxiliares (llamadas habitualmente variables sintomáticas) para estimar la población de áreas menores. Cada modelo utiliza un conjunto de datos y supuestos diferentes pero tienen en común el supuesto

que el volumen de población en cada área menor guarda una relación cuantificable con las variables sintomáticas.

Son métodos de estimación y no de proyección de población. Es decir, no realizan supuestos sobre el comportamiento futuro de cada componente de la dinámica (nacimientos, fallecimientos y migraciones) sino que utilizan información simultánea al momento de la estimación para modelar la misma y en algunos modelos se incorpora información pasada, bajo el supuesto que determinadas relaciones que se constatan durante el período anterior se mantendrán estables hasta el momento de la estimación.

Otra diferencia con las técnicas de proyección de población habituales es que no utilizan ni generan cuantificaciones para conocer la estructura por edad y sexo de la población. Sin embargo, existen algunas propuestas de estimaciones con variables sintomáticas que tienen por objetivo estimar el volumen de poblaciones por grandes grupos de edad.

### ***1.5 Encuestas de hogares y estimaciones de población en áreas menores***

Las encuestas de hogares (que relevan información en una muestra de la población) no son un instrumento adecuado para obtener estimaciones de totales poblacionales, dado que la muestra se toma en base a un listado inicial de viviendas que corresponde al último censo o conteo y, en todo caso, actualizaciones parciales sobre volúmenes y distribuciones de viviendas y/o personas. Es decir, suponen para su diseño muestral un volumen fijo de viviendas o personas (una “foto” tomada en un momento anterior a la encuesta).

Sí pueden dar indicios de cambios en la composición de la población (por edad y sexo, por ejemplo) y de los hogares (tamaño y tipos de hogar), elementos que son utilizados en algunos métodos de estimación o proyección.

Por otra parte, habitualmente estas encuestas pueden no considerar en su muestra a todas las áreas o localidades (como lo hacía la Encuesta de Hogares en Uruguay hasta 2005, sólo representativa de localidades de 5000 habitantes o más) o realizar una primer selección de sólo algunas áreas o localidades (como es el diseño actual de la Encuesta). A su vez, incluso cuando las localidades quedan comprendidas en la

muestra, el diseño puede no permitir desagregar territorialmente los resultados hasta ese nivel, por no estar prevista la representatividad estadística de los mismos.

Existe un desarrollo importante de un conjunto de herramientas estadísticas para obtener estimaciones sobre características relevadas en las encuestas para áreas pequeñas que intentan superar los problemas de representatividad y/o no selección de estas áreas debido a los diseños muestrales. Estas estimaciones tienen un interés particular por las mismas razones que se han esgrimido en relación a estimar valores poblacionales en áreas menores: proveer de información específica para cada área en una diversidad de aspectos que son relevados por las encuestas y que son pertinentes tanto para uso gubernamental como académico y privado. Estos métodos son “primos” de las estimaciones de volumen poblacional en áreas menores. Utilizan también información auxiliar como variables sintomáticas y en algunos casos realizan supuestos similares a los planteados para los métodos presentados en este informe. Sin embargo habitualmente para la aplicación de estas herramientas el volumen de población en cada área es un valor dado. Más aún, varias de ellas requieren como insumo tener cifras del volumen de población sobre cada área, por lo cual deben apelar a valores censales o estimados. Justamente aquí radica otra de las razones por las cuales interesa tener métodos que permitan estimar volúmenes poblacionales en áreas menores.

## ***2. Métodos para la estimación de población en áreas menores***

Existe una larga tradición de desarrollo de diferentes métodos para estimar la población en áreas menores, que parten principalmente del interés por la gestión de políticas públicas a nivel territorial pero también por la necesidad de este tipo de insumos para la actividad económica privada. En algunos países se comienzan a hacer estimaciones a fines del siglo XIX. Investigadores provenientes de distintas disciplinas (demografía, estadística, economía, entre otras) fueron aportando desde inicios del siglo XX mejoras y ampliaciones al abanico de métodos propuestos inicialmente. Este proceso ha estado apoyado tanto en un mayor conocimiento de la dinámica demográfica y su vínculo con otras dimensiones que afectan el

comportamiento poblacional, la aparición de nuevas herramientas estadísticas (y refinamiento de las existentes), el aumento en la facilidad y velocidad de cálculos aportadas por el desarrollo informático. A su vez, otro avance fundamental ha sido la consolidación y ampliación enorme de los sistemas de información de registros administrativos de diversa índole y la aparición de bases de datos informatizadas de los mismos.

## **2.1 *Panorama general***

Smith y Morrison (2005) mencionan un conjunto de métodos que se utilizan asociados predominantemente a estimaciones intercensales y a estimaciones postcensales. Las estimaciones intercensales pueden estar basadas en interpolaciones matemáticas entre los dos censos o pueden ser calculadas a partir de series de datos de variables sintomáticas que reflejan los cambios intercensales.

Las estimaciones postcensales pueden basarse en extrapolaciones de las tendencias históricas (hasta el último censo) o pueden vincularse a datos que reflejan los cambios poblacionales hasta el momento de la estimación. Los autores plantean que los métodos para estimaciones postcensales son los que habitualmente tienen más relevancia para los usuarios por reflejar la información más reciente disponible.

Las estimaciones postcensales de la población total no basadas en extrapolaciones utilizan métodos que incorporan variables sintomáticas del cambio poblacional. Una primera diferencia entre estos métodos refiere al papel que cumplen estas variables en el modelo: si se utilizan como auxiliares para la estimación de algunos componentes del modelo para los cuales no es posible obtener información o si son en sí mismos los parámetros con los cuales se estima el cambio del volumen poblacional o la distribución de la población entre las áreas geográficas en un determinado momento.

## **2.2 *Métodos que utilizan variables sintomáticas como auxiliares***

Dos de los métodos más utilizados que apelan en ocasiones a variables sintomáticas como auxiliares para estimar componentes para los cuales no se cuenta con información son el método de unidades de vivienda y el método de componentes.

El **método de unidades de vivienda** es el más utilizado predominantemente para estimar poblaciones en áreas pequeñas en EUA y algunos países europeos. En este método la población de un área es calculada como el número de unidades de viviendas particulares ocupadas (asimiladas a hogares) multiplicadas por el promedio de personas por hogar, a lo cual se suma el número de personas viviendo en hogares colectivos (prisiones, casas de salud, etc) o sin vivienda, de acuerdo con la siguiente fórmula:

$$P_t = (HP_t \times PPH_t) + PHCSV_t$$

P: población total  
HP: número de hogares particulares  
PPH: promedio de personas por hogar  
PHCSV: personas en hogares colectivos y sin vivienda  
t: período de estimación

El número de viviendas particulares ocupadas puede estimarse con datos de permisos de construcción, usuarios de energía eléctrica, impuestos vinculados a la propiedad y otros registros que reflejan cambios en el stock de viviendas particulares ocupadas. El promedio de personas por hogar se estima con datos censales o de encuestas de hogares o usando variables sintomáticas que estén asociadas a cambios en el tamaño de los hogares. La población residual (en hogares colectivos y personas sin hogar) se puede estimar con registros administrativos públicos o aportados por los responsables de los hogares colectivos.

El **método de componentes** está basado en la ecuación compensadora de población en la cual la población en el año t se expresa como la población contada en el censo más reciente ( $P_0$ ), más el número de nacimientos (N) e inmigrantes (I) menos el número de fallecimientos (D) y emigrantes (E) desde ese censo:

$$P_t = P_0 + N - D + I - E$$

Dos variaciones del método de componentes desarrollados por el US Census Bureau son **componentes II** y **retornos de impuestos** (tax returns), que se utilizan para estimar la población menor a 65 años (la estimación de población de 65 y más se estima utilizando los cambios en datos de cuidados médicos); ambos usan estadísticas vitales para medir el crecimiento natural de la población (N-D) y datos de

migración internacional del Servicio de inmigración y naturalización de EUA. Difieren en los datos y técnicas usadas para estimar el saldo de migración interna. El método de componentes II estima el saldo de migración interna a partir de cambios en la matriculación escolar; estas estimaciones se convierten en tasas de migración para todas las personas menores de 65 años usando datos migratorios del censo más reciente. El método de retorno de impuestos (llamado también método de registros administrativos) basa las estimaciones de la migración interna en los cambios de dirección de las personas que se registran para recibir la devolución de impuestos federales sobre los ingresos.

### ***2.3 Métodos que utilizan variables sintomáticas para modelar el cambio y el volumen de la población***

Los métodos mencionados anteriormente pueden requerir el uso de variables sintomáticas como auxiliares para estimar algún componente del modelo. Otros métodos sin embargo, utilizan como supuesto básico que el cambio del volumen de población entre dos momentos dados implica cambios en otras variables para las cuales es posible obtener información en períodos inter o postcensales y por tanto conocer cómo cambian estas variables permite estimar el volumen de población. En otros casos el supuesto se asocia no al cambio en el tiempo sino a una proporcionalidad en la distribución de la población entre áreas y la distribución de otras variables para las cuales se cuenta con información.

Existe un conjunto amplio de modelos construidos utilizando estos supuestos o variaciones de los mismos. En este apartado se presentan algunos de estos métodos, caracterizándolos de acuerdo a los supuestos específicos que utilizan, la información que requieren para las estimaciones, las fórmulas de cálculo y las ventajas y desventajas que tienen en términos tanto matemáticos como de factibilidad de uso.



<b>Métodos</b>	<b>I. Distribución por prorrateo</b> Distribución de la población por área es idéntica a la distribución de la variable sintomática	<b>II. Distribución proporcional</b> Población varía en igual proporción que la variable sintomática	<b>III. Razón censal</b> Población de las AP tienen una razón de cambio proporcional a la del área mayor entre c y t
<b>Supuestos</b>	S <sub>(t,c)</sub> : variable sintomática en año t para AP, S <sub>(t,t)</sub> : variable sintomática en año t para AM, P <sub>(t,c)</sub> : estimación de población en año t para AM	S <sub>(t,c)</sub> : variable sintomática en año c para AP, S <sub>(t,t)</sub> : variable sintomática en año t para AP, P <sub>(t,c)</sub> : población en año c para AP, P <sub>(t,t)</sub> : población en año t para AM, P <sub>(t,c)</sub> : estimación de población en año t para AM	S <sub>(t,c)</sub> : variable sintomática en año c para AP, P <sub>(t,c)</sub> : población en año c para AP, S <sub>(t,t)</sub> : variable sintomática en año t para AM, P <sub>(t,c)</sub> : estimación de población en año c para AM, S <sub>(t,t)</sub> : variable sintomática en año t para AM, P <sub>(t,t)</sub> : estimación de población en año t para AM, S <sub>(t,t)</sub> : variable sintomática en año t para AP,
<b>Algoritmo</b>	$P_{(t,t)} = \frac{S_{(t,t)}}{S_{(c,t)}} * P_{(c,t)}$	$P_{(t,t)} = P_{(t,c)} \frac{S_{(t,t)}}{S_{(t,c)}} * F_{(t)}$ ajuste para que suma de las estimaciones de población de AP sea igual a la estimación de población de AM en t	$P_{(t,t)} = \frac{S_{(t,t)}}{r_{(t,t)}} \text{ con } r_{(t,t)} = \phi * r_{(t,c)}$ siendo $\phi$ la razón de cambio de S entre c y t: $\hat{\phi} = \frac{r_{(t,t)}}{r_{(c,c)}} = \frac{S_{(t,t)}/P_{(t,t)}}{S_{(c,c)}/P_{(c,c)}} \text{ y } r_{(t,c)} = \frac{S_{(t,c)}}{P_{(t,c)}}$
<b>Fortalezas</b>	- Sencillez de cálculo - Necesita una sola variable sintomática - Sólo necesita información del momento t: * no necesidad de compatibilizar áreas en el tiempo * no afectado por cambios de definición, recolección o cobertura de la variable sintomática en el tiempo	- Sencillez de cálculo - Necesita una sola variable sintomática - Supuesto menos débil que método anterior: no requiere igual distribución sino variación proporcional a la de la población entre c y t	- Sencillez de cálculo - Necesita una sola variable sintomática
<b>Debilidades</b>	- Supuesto débil (distribución proporcional de la variable sintomática igual a la distribución proporcional de la población) - Afectada artificialmente si la variable sintomática tiene diferencias en cobertura o calidad de relevamiento de un área a otra	- Mayor cantidad de información - Compatibilización geográfica si hay cambios en la definición de áreas entre c y t (tanto en población sintomática como en la variable) - Afectada artificialmente por cambios en la cobertura o calidad de relevamiento de la variable sintomática en el tiempo - Afectada artificialmente por cambios en la cobertura o calidad de relevamiento de la variable sintomática en el tiempo	- Mayor cantidad de información - Compatibilización geográfica si hay cambios en la definición de áreas entre c y t (tanto en población como en la variable sintomática) - Afectada artificialmente por cambios en la cobertura o calidad de relevamiento de la variable sintomática en el tiempo - La dirección de las series sintomáticas podría estar afectada por factores distintos a los que justifican el cambio en la población (razón en una AP puede variar en una proporción e incluso en una dirección diferente a la del AM). No se cumple supuesto

AP: área pequeña (a estimar); AM: área mayor (universo o conjunto de todas las AP). Tiempo c: año censal (c0 y c1, en caso de utilizar dos censos); t: año de la estimación



<b>Métodos</b>	<b>IV. Diferencia de tasas</b>	<b>V. Correlación de razón</b>	<b>VI. Correlación de diferencia</b>
<b>Supuestos</b>	Variación del método de razón censal. Iguales supuestos. Se diferencia en la forma de estimación de la tasa de ocurrencia de S	Evolución de la población está correlacionada con la variación de un conjunto de variables sintomáticas. Existe una relación lineal entre las razones de cambio de las proporciones poblacionales y las de las variables sintomáticas entre dos periodos.	Evolución de la población está correlacionada con la variación de un conjunto de variables sintomáticas Similar a V, pero calcula las variaciones de proporción por medio de diferencias
<b>Información necesaria</b>	S <sub>(i,c)</sub> : variable sintomática en año c para AP, P <sub>(i,c)</sub> : población en año c para AP, S <sub>(c,c)</sub> : variable sintomática en año c para AM, P <sub>(c,c)</sub> : estimación de población en año c para AM, S <sub>(i,t)</sub> : variable sintomática en año t para AM, P <sub>(i,t)</sub> : estimación de población en año t para AM, S <sub>(i,t)</sub> : variable sintomática en año t para AP, P <sub>(i,t)</sub> : estimación de población en año t para AP,	S <sub>j(i,c0)</sub> : variable sintomática j en año c0 para AP, S <sub>j(i,c1)</sub> : variable sintomática j en año c1 para AP, S <sub>j(c,c0)</sub> : variable sintomática j en año c0 para AM, S <sub>j(c,c1)</sub> : variable sintomática j en año c1 para AM, P <sub>j(i,c0)</sub> : población en año c0 para AP, P <sub>j(i,c1)</sub> : población en año c1 para AP, P <sub>j(c,c0)</sub> : población en año c0 para AM, P <sub>j(c,c1)</sub> : población en año c1 para AM,	S <sub>j(i,c0)</sub> : variable sintomática j en año c0 para AP, S <sub>j(i,c1)</sub> : variable sintomática j en año c1 para AP, S <sub>j(c,c0)</sub> : variable sintomática j en año c0 para AM, S <sub>j(c,c1)</sub> : variable sintomática j en año c1 para AM, P <sub>j(i,c0)</sub> : población en año c0 para AP, P <sub>j(i,c1)</sub> : población en año c1 para AP, P <sub>j(c,c0)</sub> : población en año c0 para AM, P <sub>j(c,c1)</sub> : población en año c1 para AM
<b>Algoritmo</b>	$P_{(i,t)} = \frac{S_{(i,t)}}{\hat{r}_{(i,t)}}$ con $\hat{r}_{(i,t)} = \hat{r}_{(i,c)} + \hat{r}_{(c,t)} - r_{(c,c)}$ $\hat{r}_{(i,c)} = \frac{S_{(i,c)}}{P_{(i,c)}}$ , $\hat{r}_{(c,t)} = \frac{S_{(c,t)}}{P_{(c,t)}}$ y $\hat{r}_{(c,t)} = \frac{S_{(c,t)}}{P_{(c,t)}}$	$Y_{(i,c1)} = a_0 + a_1 X_{1(i,c1)} + \dots + a_n X_{n(i,c1)} + e_i$ con $Y_{(i,c0)} = \frac{P_{(i,c0)}}{P_{(i,c1)}}$ y $X_{j(i,c1)} = \frac{S_{j(i,c0)}/S_{j(c,c0)}}{S_{j(i,c1)}/S_{j(c,c1)}}$ Estimación en momento t: $P_{(i,t)} = Y_{(i,t)} * \frac{P_{(i,c1)}}{P_{(c,c1)}} * P_{(c,t)}$ con $Y_{(i,t)}$ : razón de proporción de la población del área i estimada para el periodo t por el modelo de regresión	$W_{(i,c1)} = a_0 + a_1 Z_{1(i,c1)} + \dots + a_n Z_{n(i,c1)} + e_i$ con $W_{(i,c1)} = \frac{P_{(i,c1)}}{P_{(c,c1)}} - \frac{P_{(i,c0)}}{P_{(c,c0)}}$ y $Z_{j(i,c1)} = \frac{S_{j(i,c1)}}{S_{j(c,c1)}} - \frac{S_{j(i,c0)}}{S_{j(c,c0)}}$ Estimación en momento t: $P_{(i,t)} = \left[ W_{(i,t)} + \frac{P_{(i,c1)}}{P_{(c,c1)}} \right] * P_{(c,t)}$ con $W_{(i,t)}$ : diferencia de proporciones de la población del área i estimada para el periodo t por el modelo de regresión
<b>Fortalezas</b>	Similares a III	<ul style="list-style-type: none"> <li>- Utiliza una o más variables sintomáticas con el supuesto que su evolución explica la variación de la población</li> <li>- Estimaciones no afectadas por errores en calidad de información si los datos se mantienen en el tiempo o cambian en calidad en forma similar en todas las áreas</li> <li>- Puede usarse para estimar áreas no involucradas en la estimación del modelo</li> <li>- Mayor cantidad de información</li> <li>- Compatibilización geográfica si hay cambios en la definición de áreas entre c0 y c1 (tanto en población como en la variable sintomática)</li> <li>- Compatibilización de información entre c0 y c1 y entre c1 y t</li> </ul>	<ul style="list-style-type: none"> <li>- Similares a V</li> <li>- Logra mayor intercorrelación entre W y Z<sub>j</sub> y estimaciones estarían menos afectadas por cambios temporales en las variables</li> </ul>
<b>Debilidades</b>	Similares a III	<ul style="list-style-type: none"> <li>- Mayor cantidad de información</li> <li>- Compatibilización geográfica si hay cambios en la definición de áreas entre c0 y c1 (tanto en población como en la variable sintomática)</li> <li>- Compatibilización de información entre c0 y c1 y entre c1 y t</li> </ul>	<ul style="list-style-type: none"> <li>- Similares a V</li> </ul>

AP: área pequeña (a estimar); AM: área mayor (universo o conjunto de todas las AP). Tiempo c: año censal (c0 y c1, en caso de utilizar dos censos); t: año de la estimación

<b>Métodos</b>	<b>VII. Tasa de correlación</b>	<b>VIII. Método de tasas vitales</b>	<b>IX. Método compuesto</b>
<b>Supuestos</b> Evolución de la población está correlacionada con la variación de un conjunto de variables sintomáticas. Similar a V, pero utiliza una aproximación exponencial de cambio en vez de lineal.	Similar a V, pero utiliza una aproximación exponencial de cambio en vez de lineal.	Hay una relación inversa entre las variaciones anuales de las tasas de natalidad y de mortalidad de las AP y las desviaciones proporcionales de esas tasas con respecto a las de las AM se mantiene constante en el tiempo.	Estimaciones por grupos de edad utilizando variables sintomáticas específicas que pueden asociarse a la dinámica de cada grupo de edad.
<b>Información necesaria</b>	$Y^*_{(i,c1)} = a_0 + a_1 X^*_{(i,c1)} + \dots + a_n X^*_{n(i,c1)} + e_i$ $Y_{(i,c1)} = \frac{P_{(i,c0)}}{P_{(i,c1)}} / P_{(c0)}, \quad X_{(i,c1)} = \frac{S_{j(i,c0)}}{S_{j(i,c1)}} / S_{j(c0)}$ $Y^*_{(i,c1)} = \frac{\ln(Y_{(i,c1)})}{k} \quad \text{y} \quad X^*_{(i,c1)} = \frac{\ln(X_{(i,c1)})}{k}$ <p> <math>S_{j(i,c0)}</math>: variable sintomática j en año c0 para AP,  <math>S_{j(i,c1)}</math>: variable sintomática j en año c1 para AP,  <math>S_{j(c0)}</math>: variable sintomática j en año c0 para AM,  <math>S_{j(c1)}</math>: variable sintomática j en año c1 para AM,  <math>P_{(i,c0)}</math>: población en año c0 para AP,  <math>P_{(i,c1)}</math>: población en año c1 para AP,  <math>P_{(c0)}</math>: población en año c0 para AM,  <math>P_{(c1)}</math>: población en año c1 para AM         </p>	$\hat{P}_{(i,t)} = a * \left( \frac{B_{(i,t)}}{\hat{b}_{(i,t)}} \right) + (a-1) * \left( \frac{D_{(i,t)}}{\hat{d}_{(i,t)}} \right)$ <p>Con:</p> $\hat{b}_{(i,t)} = \frac{B_{(i,t)} / P_{(i,c)} * \left( \frac{B_{(i,t)}}{\hat{P}_{(i,t)}} \right)}{B_{(c,t)} / P_{(c,c)}} \quad \text{y} \quad \hat{d}_{(i,t)} = \frac{D_{(i,t)} / P_{(i,c)} * \left( \frac{D_{(i,t)}}{\hat{P}_{(i,t)}} \right)}{D_{(c,t)} / P_{(c,c)}}$ <p>           a: coeficiente de ponderación. Puede utilizarse 0.5 (igual peso para nacimientos y defunciones);            propuesta: inverso del coeficiente de variación de nacimientos y de defunciones.         </p>	Como ilustración: para el caso de 5 a 14, sean: $M_{(i,c)}$ : matrícula escolar en año c para AP, $M_{(i,t)}$ : matrícula escolar en año t para AP, $P_{(i,c)5a14}$ : población de 5 a 14 años en año c para AP, $P_{(c)5a14}$ : población estimada de 5 a 14 años en año t para AM
<b>Algoritmo</b>	$\hat{P}_{(i,t)5a14} = \frac{P_{(i,c)} * \frac{M_{(i,t)}}{M_{(i,c)}}}{\sum_t P_{(i,c)5a14} * \frac{M_{(i,t)}}{M_{(i,c)}}} * \hat{P}_{(c,t)5a14}$	Como ilustración: para el caso de 5 a 14, sean: $M_{(i,c)}$ : matrícula escolar en año c para AP, $M_{(i,t)}$ : matrícula escolar en año t para AP, $P_{(i,c)5a14}$ : población de 5 a 14 años en año c para AP, $P_{(c)5a14}$ : población estimada de 5 a 14 años en año t para AM	Como ilustración: para el caso de 5 a 14, sean: $M_{(i,c)}$ : matrícula escolar en año c para AP, $M_{(i,t)}$ : matrícula escolar en año t para AP, $P_{(i,c)5a14}$ : población de 5 a 14 años en año c para AP, $P_{(c)5a14}$ : población estimada de 5 a 14 años en año t para AM
<b>Fortalezas</b>	- Similares a V. Supera crítica a Método V: coeficientes se calculan tomando como referencia tiempo entre 2 censos (c0 y c1), pero estimaciones habitualmente están a un intervalo de tiempo entre c1 y t menor que el intercensal. Puede afectar estabilidad de coeficientes y estimaciones. Mantiene tendencia similar a la estimación del método V (es transformación de ella). Transformación reduce varianzas y permite mayor exactitud. Se reduce efecto de la autocorrelación espacial y es esperable estimaciones más precisas.	- Utiliza solo estadísticas vitales sobre las que es posible evaluar en forma relativamente sencilla su calidad, nivel de omisión y de registro tardío.	- Obtener estimaciones por grupos de edad - Extensible a estimaciones desagregadas para otro tipo de grupos, si se obtienen variables sintomáticas sobre las que se pueda suponer relaciones específicas con los grupos.
<b>Debilidades</b>	- Similares a V	- No considera cambios en las estructuras de edades que puedan afectar las relaciones en el periodo - Indeterminación sobre ponderación a utilizar - Depende de calidad de las estadísticas vitales	Requiere identificar variables sintomáticas disponibles sobre las que se pueda suponer relaciones específicas con los grupos

AP: área pequeña (a estimar); AM: área mayor (universo o conjunto de todas las AP). Tiempo c: año censal (c0 y c1), en caso de utilizar dos censos); t: año de la estimación

Los modelos presentados no agotan el universo de métodos que se han desarrollado para la estimación de población en áreas pequeñas, pero son los utilizados con mayor frecuencia en los estudios que se han relevado para este informe.<sup>4</sup>

En particular, los modelos de regresión –en sus distintas versiones- han mostrado una muy buena capacidad para estimar volúmenes poblacionales. Los primeros modelos elaborados fueron presentados por Snow en 1911 en la *Journal of Royal Statistical Society*<sup>5</sup> del Reino Unido. Desde ese período han sido revisados y mejorados en forma sustantiva por distintos investigadores y cuentan con un desarrollo importante de fundamentos de estadística inferencial. Sin embargo, como se menciona en la tabla, las dificultades para su utilización derivan principalmente de los requerimientos de datos para distintos años y compatibilización de las áreas para esos períodos.<sup>6</sup>

Por otra parte, existen diversas variaciones de estos modelos que buscan mejorar la precisión de las estimaciones o para hacerlos aplicables a otros tipos de estimaciones (como subpoblaciones).

**Uso de estratos para mejorar la precisión de las estimaciones.** Dado que las áreas menores que componen el área mayor pueden tener características muy diferentes que estén asociadas a las dinámicas poblacionales, algunos estudios proponen formar estratos que contengan áreas más homogéneas entre sí. Si la estratificación es adecuada, se obtienen estimaciones más precisas para las áreas de los distintos estratos. Algunas variables utilizadas para estratificar son el tamaño poblacional, la densidad de población en el último censo o la tasa de crecimiento intercensal (se forman 3 o 4 estratos), el carácter urbano o rural (2 estratos), etc.

**Modelos con variables retardadas.** Algunos autores plantean que uno de los problemas de estos métodos es la oportunidad con la cual se accede a los datos correspondientes a las variables sintomáticas. Por lo general aún en casos en que el flujo de consolidación y difusión de los registros es rápido, no es posible obtener las

---

<sup>4</sup> Una descripción y valoración de estos modelos puede consultarse en Bay(1998), Texeira Jardim (2001), Alvarez (2001), Chavez Esquivel (2003),

<sup>5</sup> “The application of the method of multiple correlation to the estimation of post-censal populations”, citado en Swanson y Tayman (s/f)

<sup>6</sup> Un mayor detalle de los fundamentos estadísticos de estos modelos puede consultarse en Swanson y Tayman (s/f), Australia Bureau of Statistics (2000). Una comparación de estos modelos con los de unidades de vivienda y componentes puede consultarse en Smith y Mandell (1984) y en Smith y Morrison (2005).

cifras con menos de un semestre o un año de retraso. Por ello proponen una modificación de los métodos, utilizando valores de las variables sintomáticas de un período determinado para estimar población a una fecha posterior (por ejemplo, un año después o más). Esta propuesta tiene igualmente la dificultad de suponer que en el período para el cual se hace la estimación se mantiene el comportamiento (sea tasa de cambio, razones, distribuciones proporcionales) similar al del período en que se toman las variables sintomáticas en el modelo. Si bien puede ser una alternativa (sobre todo cuando el acceso a la información es lento) pierde sensibilidad a cambios bruscos en el volumen de población.

**Combinación de estimaciones y proyecciones y estimación indirecta de saldos migratorios.** Otras propuestas apuntan a utilizar proyecciones de población a nivel de áreas menores junto con las estimaciones. Para ello proponen elaborar varias proyecciones de población utilizando supuestos sobre tasas de natalidad y mortalidad construidas para las áreas menores (en base al comportamiento previo de nacimientos y fallecimientos) y dos o más hipótesis sobre comportamiento migratorio (que como se ha dicho es el componente que presenta mayor incertidumbre para la elaboración de proyecciones y que puede tener cambios bruscos e imprevistos a nivel local). A su vez se generan estimaciones para esas áreas utilizando variables sintomáticas. La idea central es aprovechar las ventajas de cada forma de modelizar el volumen de población: en las proyecciones, la estructura por edad y sexo y en las estimaciones, la mayor sensibilidad de las mismas a los cambios reales que se producen en las áreas a considerar. Por ello se trabaja con las diferencias entre el volumen de población entre las proyecciones y las estimaciones en cada localidad para el mismo período, se analiza qué proyecciones (es decir, con cuáles hipótesis de migración) hay menores y mayores diferencias y se focaliza en aquellos casos en que aparecen diferencias importantes para ver si hay efectivamente cambios en esos lugares que no hayan sido “previstos” por las proyecciones. (González y Celton 2008)

**Estimaciones de subpoblaciones en áreas pequeñas.** Finalmente, otros investigadores han propuesto métodos para estimar la población por grandes grupos de edad. La idea central de estos métodos es que existen variables sintomáticas cuya dinámica sigue en forma aproximada la de grupos de población específicos (por ejemplo: para los niños en edad escolar, la matrícula en educación

primaria; para adultos mayores, registros de los servicios de salud o de la seguridad social, etc.). (Texeira Jardim 2001)

## **2.4 Herramientas para evaluar los resultados de los métodos**

La evaluación que se realiza de los métodos aplicados compara los resultados obtenidos (población estimada para cada AP) en relación a la información censal. Esta evaluación permite identificar los modelos y las variables sintomáticas que generan estimaciones más precisas. Algunos de los estudios analizados comparan con los resultados de las proyecciones de población, con lo cual es posible comparar la precisión y potencialidades de ambas herramientas.

Los indicadores utilizados para evaluar los modelos se basan en las diferencias entre los valores estimados y los obtenidos en los censos de población. Para cada estimación (valor obtenido en una AP) es posible construir los siguientes indicadores:

➤ **Error relativo o diferencia porcentual:**

$$\text{Error relativo o diferencia porcentual} = \frac{(\text{Población estimada} - \text{Población censada}) * 100}{\text{Población censada}}$$

Este es el indicador de mayor uso y es de interpretación sencilla. Sin embargo es complejo para comparar estimaciones de AP con diferencias muy amplias en tamaño de población, ya que generalmente las diferencias porcentuales son mayores para áreas con poca poblaciones (un pequeña variación en números genera un cambio porcentual importante, con lo cual es necesario tener precauciones al interpretar).

➤ **Diferencia numérica o absoluta:**

$$\text{Diferencia numérica} = \text{Población estimada} - \text{Población censada}$$

Permite ver la diferencia en términos de volumen de población en cada AP, pero no es un buen indicador para comparar entre AP, excepto que se trabaje por estratos de tamaño de AP.

➤ **Diferencia en logaritmos:**

$$\text{Diferencia en logaritmos} = \frac{\log(\text{Población estimada} - \text{Población censada}) * 100}{\log(\text{Población censada})}$$

Permite “suavizar” la mayor sensibilidad de los cambios en las AP de poca población frente a lo que sucede en las de mayor tamaño. Sin embargo, es de difícil interpretación.

Estos indicadores se calculan para cada una de las AP. Sin embargo es necesario contar con criterios para **evaluar los modelos** en base al conjunto de resultados obtenidos. Básicamente se utilizan dos **criterios** diferentes:

- Elegir el modelo que genere el menor promedio de errores relativos
- Elegir el modelo que minimiza el número de “malas” estimaciones (diferencias mayores a un determinado valor del indicador)

En ambas estrategias es posible utilizar cualquiera de los indicadores de diferencias presentadas (porcentual, numérica, logarítmica).

Algunos estudios presentan un conjunto de estadísticos del indicador que permiten utilizar una u otra estrategia y generar una evaluación más completa (media, mediana y otros n-tilas, desvío estándar, distribución de frecuencias).

Además de la evaluación en base a los indicadores que reflejan los errores cometidos en las estimaciones en relación a los datos censales, es necesario considerar **otros elementos para la elección de los modelos**:

- estabilidad de los coeficientes de regresión: cuando es posible construir modelos para dos períodos intercensales, es posible comparar en qué medida los coeficientes se mantienen en valores similares (el supuesto para usar



regresión es justamente que es previsible que las relaciones expresadas en el modelo se mantengan estables en años postcensales)

- en los métodos multivariados es posible comparar modelos que incluyen/excluyen cada variable sintomática en términos de los errores relativos que se producen. Este análisis permite identificar variables que pueden estar influyendo o comportándose de una manera que no es la supuesta o que son pertinentes para algunos estratos de AP y no otros.
- Es fundamental tener en cuenta qué expectativa hay a futuro en relación a cada variable sintomática: si seguirá relevándose sistemáticamente, su periodicidad, posibles cambios en su definición o cobertura, etc. ya que estos elementos pueden inhabilitar el uso de alguno de los métodos elegidos.

En los países en que se implementa en forma sistemática las estimaciones, la evaluación de modelos se realiza al momento de tener los datos censales. Una vez elegido el o los modelos, se mantienen los mismos como herramientas para estimaciones hasta el siguiente censo, excepto que existan evidencias por ejemplo sobre cambios en la calidad de una variable que pueda distorsionar las estimaciones, o que se deje de contar con una variable.

La atención y evaluación sobre los indicadores utilizados debe ser permanente para mantener la vigencia y calidad de los modelos, aún cuando ya se han validado y se ha sistematizado su uso. Este aspecto es uno de los puntos neurálgicos del uso de estos métodos y si puede considerarse un aspecto problemático de los mismos, es también una de sus fortalezas. Como plantea Texeira Jardim *“al contrario de lo que ocurre con los métodos empleados para proyecciones demográficas, el dinamismo del proceso de estimación de la población hace que se vaya perfeccionando cada vez que se añade nueva información relevante al acervo de datos disponibles. En este sentido, la mejora de las estimaciones está íntimamente ligada a la calidad de la información utilizada”*. (Texeira Jardim 2001: 29)

Más allá de esta evaluación global, es importante focalizar en aquellas AP en las cuales las discrepancias con los datos censales son mayores y buscar elementos que ayuden a entender estas diferencias (fallas de los supuestos, errores en las variables sintomáticas- de datos, cobertura, relevamiento, etc.) o comportamientos

de la dinámica poblacional muy específicos en esas AP que no siguen los patrones obtenidos en la modelización y que se adecuan a la gran mayoría de las áreas.

En forma complementaria, cuando se compara con proyecciones, las discrepancias con los valores proyectados que no son atribuibles a problemas de la calidad de la información de las variables sintomáticas, permiten identificar procesos específicos que “escapan” a los supuestos utilizados para elaborar las proyecciones.

Estos comportamientos diferenciados deberían ser analizados en forma particular, lo cual implica conocer y entender las dinámicas locales, para lo cual es imprescindible el aporte de quienes tienen un conocimiento sobre los procesos que se están dando localmente.

## **2.5 Aplicaciones a nivel mundial y latinoamericano**

Este apartado tiene como objetivo presentar experiencias de otros países y tomar nota de algunos elementos comunes que aparecen en la evaluación de la aplicación de estas herramientas. En América Latina el desarrollo está en una etapa más experimental pero hay diversos estudios a nivel de exploración y prueba de los distintos métodos. De las experiencias extrarregionales se presentan algunos casos donde el uso de estos métodos es sistemático y forma parte de la información provista por las oficinas de estadísticas.

Se presentan algunas experiencias de aplicación de métodos de estimación de población con variables sintomáticas en distintos países de **América Latina**. En la región, este tipo de trabajos tuvo un impulso importante de parte de un programa conjunto del BID y el CELADE. En el año 2000 este programa propulsó un “Taller sobre estimaciones e indicadores demográficos y económicos para áreas menores”, en Lima, Perú, donde se expusieron varias de las experiencias que se mencionan en este apartado.

En **Argentina** un equipo coordinado por Gustavo Alvarez del INDEC (Instituto Nacional de Estadísticas y Censos), con apoyo del programa BID/CELADE realizó la estimación de la población por departamento para el año 1996 con el fin de comparar sus resultados con proyecciones a ese mismo año. Los departamentos son un nivel administrativo intraprovincial; el país está dividido en 509



departamentos. Las variables sintomáticas utilizadas fueron: matrícula en nivel primario; nacimientos; defunciones; eventos vitales (nacimientos+defunciones). Se utilizaron datos censales de 1980 y 1991. Los métodos aplicados fueron: distribución por prorrateo; distribución proporcional; método de tasas vitales de Bogue; correlación de razones; correlación de diferencias; correlación de tasas. En los últimos tres, se propusieron también modelos estratificados (tres niveles definidos en terciles; variables de estratificación: crecimiento poblacional 1980-1991, densidad poblacional 1991, urbanización 1991). Una descripción de este estudio es accesible a través de la publicación de Alvarez, G (2001).

Posteriormente, Leandro González y Dora Celton, investigadores de la Universidad de Córdoba, desarrollan una propuesta metodológica para ajustar proyecciones de población menores con variables sintomáticas. Dado que a nivel de áreas menores el componente de migración es muy difícil de modelar se elaboran 4 proyecciones utilizando diferentes supuestos sobre el comportamiento migratorio. Por su parte se estima la población con las variables sintomáticas y se comparan los resultados con las distintas. Bajo el supuesto que fecundidad y mortalidad están bien modelados, la proyección más cercana a las estimaciones es la que tiene implícita la tasa de migración más ajustada a la dinámica demográfica real. El estudio se realiza para las localidades del "Gran Córdoba" (ciudad de Córdoba y 12 localidades que conforman un área metropolitana). Utiliza las variables: nacimientos, defunciones, hechos vitales (nacimientos+defunciones), matrícula en primaria y número de electores. Se realizan estimaciones con los métodos: distribución por prorrateo, distribución proporcional, tasas vitales, razón censal, diferencia de tasas, método compuesto, correlación de razón, correlación de tasas y correlación de diferencias.

En **Brasil** hay una acumulación importante de experiencias de estimaciones de población para distintas divisiones administrativas al interior de los Estados brasileiros. Una reseña de estos trabajos es presentada por Texeira Jardim (2001), quien recoge a su vez un inventario de este tipo de estudios realizado por Laura Wong (1987). Dentro de las experiencias relevadas se encuentra una estimación de la población del Distrito Federal realizada por Irvin y Fialho en 1972, quienes utilizan el método compuesto, el método de componentes II y el método de tasas vitales y como variables sintomáticas, defunciones, nacimientos, matrimonios y matrícula escolar (Irvin y Fialho, 1977). Según el inventario desarrollado por Wong, solamente en Paraná y el Mato Grosso se aplicaron metodologías basadas en el uso de

variables socioeconómicas como indicadores de crecimiento demográfico (Wong, 1987). Texeira reseña otros estudios donde se realizaron estimaciones de población de áreas menores de Brasil a partir de variables sintomáticas (Sawyer, 1969; Sawyer y Carvalho, 1981; Santa Catarina, 1989; IBGE, 1990 y Bega et al., 1991).

El estudio de Texeira busca evaluar la eficacia de distintas variables sintomáticas para estimar la población municipal del estado de Río Grande del Sur a lo largo de la década de 1990. Utiliza como variables sintomáticas: número de electores registrados, matrícula inicial en la educación básica, nacimientos registrados, defunciones registradas. Aplica los métodos de distribución por prorratio, tasas vitales, una variación del método compuesto de Bogue y Duncan y el método de correlación de razón. Los resultados son comparados con los datos censales de 1991.

Para **Chile**, Guiomar Bay del CELADE, realiza un estudio (Bay 1998) para estimar la población de las comunas en 1992 y comparar los resultados con los proporcionados por el Censo de ese año. Uno de los problemas encontrados es el importante número de comunas que modificaron sus límites entre 1970 y 1982, años censales que se utilizan tanto para obtener valores poblacionales como datos en las variables sintomáticas. Enumera a su vez un conjunto importante de dificultades asociadas a ausencias de registros o cambio en la forma de relevar la información, o falta de actualización en algunas fuentes. La selección final de variables incluye: nacimientos, defunciones, matrícula en educación básica, inscripción de automóviles, acontecimientos vitales (nacimientos+defunciones). Utiliza como métodos: la distribución por prorratio, distribución proporcional, correlación de razón y correlación de diferencia. A su vez, se prueba también los métodos utilizando dos tipos de estratificación de las áreas: tres estratos según tamaño poblacional en el censo anterior y dos estratos para incorporar la condición urbana o rural de las áreas. Las estimaciones obtenidas para 1992 se comparan con los resultados censales pero también con las proyecciones de población por comuna para ese año, concluyéndose que algunos de los métodos logran una precisión mayor que las obtenidas mediante proyecciones.

En el caso de **Costa Rica**, los estudios coordinados por Chávez Esquivel (2001, 2003) han dado una continuidad al trabajo sobre estimación de población con variables sintomáticas para ese país. Un primer estudio (Chávez Esquivel 2001) compara estimaciones en cantones (división administrativa intraprovincial) con

proyecciones de población que utilizan crecimiento vegetativo, para el año 1984. En este caso, debido a las dificultades para la recopilación de datos (dado que era necesario contar con datos comparables con el Censo de 1963), se aplican los métodos a un área piloto compuesta por 22 cantones de dos provincias para realizar una selección de variables, quedándose con las más correlacionadas con la variación poblacional entre los Censos 1973 y 1984). Se seleccionan siete fuentes de datos: matrícula escolar en primer y segundo ciclos, nacimientos, defunciones, permisos de construcción de viviendas, registros de asegurados de la Caja Costarricense del Seguro Social, registros de atención en los centros de salud facilitados por el Ministerio de Salud y número de electores inscritos en el padrón electoral, proporcionado por el Tribunal Supremo de Elecciones. Se prueban: método de razón censal, método de diferencia de tasas, método de correlación de razón, método de tasa de correlación (aproximación por modelo exponencial). Posteriormente se aplican los métodos y variables seleccionados para realizar estimaciones de todos los cantones del país.

Un estudio posterior del investigador (Chávez Esquivel 2003) permite validar el uso de los métodos y variables sintomáticas con los resultados del Censo del año 2000. En este caso se vuelven a tomar datos de 1963 y 1973 y se validan con los resultados del Censo de 1984. Posteriormente se estiman valores poblacionales para el año 2000 a partir de los modelos generados y se compara con los resultados censales para ese año. Las variables utilizadas son los electores, nacimientos y matrícula escolar. Se utilizan: método de razón censal, método de diferencia de tasas, método de razón de correlación, método de tasa de correlación. Concluye que los resultados del estudio para el año 2000 son consistentes con los obtenidos en el estudio para 1984 a pesar del tiempo transcurrido entre los Censos. Para el año 2000, sin embargo, algunas variables reducen su correlación con el volumen poblacional (específicamente, el número de electores).

En relación a **experiencias extrarregionales**, se presentan dos casos para los cuales cobra interés el proceso de consolidación en el uso de este tipo de métodos y su institucionalización en oficinas estadísticas.

**Australia**<sup>7</sup>. En Australia las estimaciones de población para períodos intercensales tienen un desarrollo de varias décadas. Hasta 1996 estas actividades se realizaban en forma descentralizada en las oficinas estadísticas de sus estados y territorios, en la mayoría de los casos utilizando métodos de regresión. Desde 1997 se centralizan parcialmente estos procesos en una unidad operativa (Small Area Population Unit) de la Australian Bureau of Statistics, en Adelaida. Esta centralización permitió no sólo una mayor eficiencia en el uso de recursos humanos y materiales sino una estandarización de procedimientos y una mayor coherencia de la información utilizada y sobre todo de los resultados. Actualmente las estimaciones se realizan utilizando métodos de regresión. Las variables utilizadas se seleccionan para cada área administrativa según la disponibilidad y estabilidad de los registros en los diez años anteriores al último censo (que son realizados cada cinco años). Estas estimaciones son controladas por parte de las oficinas regionales, donde se utilizan otras fuentes de datos disponibles a nivel local y métodos complementarios para contrastar los resultados, además de incorporar el conocimiento sobre la dinámica demográfica específica de cada territorio.

**British Columbia, Canadá**<sup>8</sup>. Las distintas oficinas estadísticas canadienses han implementado métodos de estimación de áreas menores desde hace varias décadas. Como ilustración se presenta el caso de British Columbia. Desde 1978 la oficina estadística de British Columbia (actualmente BC Stats) comenzó a trabajar en un programa para estimar cifras de población a nivel de áreas sub-provinciales. Se testearon un conjunto de modelos y en 1979 se optó por un método de regresión (el de diferencias de correlaciones) como método primario y por el método de distribución proporcional como complementario para determinadas áreas. Los modelos se testearon con los resultados censales de 1976. Posteriormente han sido revisados y testeados con varios censos (realizados cada 5 años), lográndose un refinamiento de los modelos iniciales y la incorporación de nuevas variables sintomáticas para el cálculo. Las estimaciones son elaboradas anualmente como parte de las estadísticas presentadas por la BC Stats. Las áreas para las cuales se realizan las estimaciones son los municipios de British Columbia y se utilizan como variables sintomáticas el número de conexiones eléctricas residenciales, número de beneficiarios de la seguridad social de 65 años y más.

---

<sup>7</sup> Por mayores detalles, ver Howe, Andrew (2000).

<sup>8</sup> Por mayores detalles, ver BC Stats (1998).

En todas estas experiencias la evaluación que realizan los autores sobre la aplicación de los métodos pone énfasis en que la clave está en la calidad y cantidad de información a utilizar como variables sintomáticas. Esto es, en términos de la forma de relevamiento y la cobertura (así como los cambios de estos aspectos durante los períodos considerados). Y a su vez, la posibilidad de obtener información para las mismas áreas (delimitación equiparable entre variables y de éstas con las áreas censales y a su vez la equiparación en el tiempo).

Uno de los indicadores que se utiliza con más frecuencia y que se evalúa como una buena variable sintomática es la matrícula de enseñanza primaria. Presenta como ventaja que es un indicador que se obtiene de registros administrativos que habitualmente tienen series históricas largas, que es de uso para la gestión y por tanto está bastante sistematizado y estandarizado y que puede desagregarse hasta el nivel de cada centro de enseñanza, lo cual remite a territorios muy acotados. Por otra parte, es sensible a los cambios en la composición de población de edades activas (porque en muchos casos se trata de núcleos familiares con niños en edad escolar) además de los cambios en la población infantil. Y a su vez, el área de influencia de los centros educativos está delimitada territorialmente en forma bastante clara, por lo cual el comportamiento de la población escolar es consistente con el de la población en edad escolar del área de influencia.

Las estadísticas vitales (nacimientos y defunciones o la suma de ambas: “hechos vitales”) son también variables relevantes en los modelos seleccionados. Al igual que la matrícula de primaria, hay un esfuerzo creciente por parte de las administraciones de mejorar estos registros para la gestión. Son sensibles a cambios en poblaciones de determinados grupos de edad.

En las distintas experiencias aparece un abanico amplio de variables sintomáticas utilizadas. La pertinencia de usar unas u otras se relaciona no sólo a las particularidades de cada país, sino la región para la cual se hacen las estimaciones y el tipo de áreas pequeñas en las cuales se desagrega el área mayor.

Otra evidencia que surge de estas experiencias es que los problemas de calidad/cobertura de los indicadores construidos a partir de registros administrativos están presentes en todos los países y como ya se mencionó constituyen una de las debilidades más fuertes para trabajar con métodos de estimación apoyándose en información auxiliar.

Esto también tiene repercusiones en los métodos que es posible utilizar en cada caso ya que, como se vio, algunos requieren un volumen mayor de información auxiliar que otros.

Los métodos que la mayoría de los documentos consideran más robustos son los que estiman los volúmenes poblacionales en base a modelos de regresión (en sus distintas versiones). Estos métodos son justamente los que tienen mayores requisitos en cuanto a la información auxiliar necesaria (en dos momentos del tiempo coincidentes con los censos y a su vez, con delimitaciones de áreas similares para los datos censales y las variables auxiliares y en los dos momentos del tiempo). Por esta razón, la mayor debilidad para su aplicación de acuerdo a las experiencias ha sido la disponibilidad de información confiable para construir las variables sintomáticas.

Un tema importante es que los datos no requieran demasiados procesamientos específicos para quienes relevan la información. Esto implica un esfuerzo adicional y además puede insumir tiempos adicionales, en contra de la oportunidad de las estimaciones.

Otra fuente de complejidad al realizar las estimaciones es la necesidad de utilizar distintos métodos para afinar las estimaciones. Por un lado, se puede usar los resultados de un método solo o promedios (combinaciones de resultados de distintos métodos). Por otra parte, los métodos pueden ser más o menos apropiados de acuerdo al nivel de desagregación y también a los tamaños poblacionales.

### ***3. Uruguay: fuentes disponibles y delimitaciones de áreas***

#### ***3.1 Fuentes disponibles en Uruguay***

Como ya se mencionó existen diversas dificultades asociadas al uso de registros administrativos como variables sintomáticas. Sin embargo tienen la ventaja fundamental de constituir flujos continuos y sistematizados de información, relevados por diversas fuentes para sus propios objetivos y por tanto los costos de obtención son sensiblemente menores que la implementación de relevamientos propios. En Uruguay existen buenas bases de datos registrales en algunas fuentes y en los últimos años se ha desplegado un esfuerzo importante por mejorar la calidad de los

datos en distintos ámbitos sobre todo de la administración pública. Esto permite suponer que a futuro será posible obtener un conjunto mayor de variables con cobertura adecuada y confiabilidad que las disponibles actualmente.

### 3.1.1 Variables relevadas

El relevamiento inicial realizado sobre posibles variables sintomáticas a utilizar en este estudio se sintetiza en la siguiente matriz:

Datos	Fuente	Cobertura geográfica	Cobertura histórica	Desagregado factible	Confiabilidad y otros problemas	Acceso oportunidad	Acceso: facilidad y mecanismos
Nacimientos	MSP	Nacional	1996-2009	Localidades Secciones judiciales	Buena cobertura y registro	Variable: al menos 3 semestres de diferencia para la publicación	Información pública a través de página web (aunque no está publicada la última información disponible)
Defunciones	MSP	Nacional	1996-2009	Secciones judiciales	Buena cobertura y registro	Variable: al menos 3 semestres de diferencia para la publicación	Información pública a través de página web (aunque no está publicada la última información disponible)
Consumo eléctrico y conexiones residenciales	UTE	Nacional	1995-2010	Oficinas comerciales; localidades (a consultar factibilidad)	Dificultades por cambios de cobertura, alcances muy heterogéneos de las oficinas comerciales, variaciones de consumo por razones de servicio, etc.	Información actualizada	Información pública, vía solicitud en UTE; para localidades, es necesario un procesamiento específico
Matrícula primaria	ANEP-OBS	Nacional	1991-2010	Localidades y zonas censales(para escuelas rurales)	Cambios de cobertura (aperturas y cierres de establecimientos ); matrícula incluye repeticiones	Información al cierre del año anterior	Información pública; Monitor
Consultas Centros externos de atención ASSE	ASSE	Nacional	1985-2009	Localidades de mayor tamaño	Cambios de cobertura (aperturas y cierres de establecimientos ); cambios asociados a ciclos económicos y actualmente a FONASA (a evaluar)	Variable (a evaluar)	Información histórica impresa; información reciente por solicitud



### **3.1.2 Otras variables consideradas:**

Dentro de las fuentes que se evaluaron se han descartado por presentar mayores problemas:

*Consumo de agua de OSE:* Dado que la gestión y administración del agua potable quedó en manos de OSE en forma exclusiva a partir de 2004 y a partir de esa fecha se abrió un proceso de traspaso de usuarios de empresas privadas a OSE, es una variable con una dinámica propia no asociable a cambios poblacionales. Sin embargo, podría convertirse en una variable sintomática a incorporar en futuras revisiones de los modelos.

*Padrones electorales:* un alto porcentaje de votantes no modifica su credencial cívica al cambiar de residencia, por lo cual es una variable que probablemente se asocia en forma muy débil con los cambios poblacionales.

*Beneficiarios de asignaciones familiares:* Durante la última década se modificó reiteradamente la cobertura de beneficiarios por cambios en normativas de seguridad social. Por tanto es una variable que ha tenido una dinámica propia no asociable a cambios poblacionales.

*Beneficiarios de prestaciones para pasivos (jubilaciones, pensiones):* La descentralización de los pagos ha modificado el registro de los usuarios y de acuerdo a informantes del BPS no se cuenta con un registro actualizado que permita conocer el lugar de residencia de los beneficiarios. De cualquier manera, se seguirá explorando en esta fuente.

*Matrícula en enseñanza secundaria y UTU.* Es una variable que podría ser utilizada en las estimaciones, pero existen dificultades tanto en relación a una mayor dispersión del área de influencia de las instituciones y un mayor peso de la enseñanza privada, así como en términos operativos dado que la propia información está dispersa al menos en dos entes autónomos (Consejo de Enseñanza Secundaria y UTU). Por otra parte, los estudios analizados muestran



que la matrícula en enseñanza primaria es un indicador con mayor capacidad predictiva que el de nivel secundario.

### **3.2 Delimitación de áreas**

Las necesidades de información a nivel territorial pueden llevar a considerar diversas delimitaciones de áreas geográficas, sea a partir de criterios político-administrativos o surgidos a partir de la consideración de otros aspectos (por ejemplo: regionalizaciones a partir de criterios de afinidades económicas, sociales, etc.).

Por otra parte, como ya se ha mencionado, las divisiones que pueden realizarse para la aplicación de métodos de estimación de población, están restringidas en forma muy importante según la disponibilidad de información, es decir, la posibilidad de desagregar territorialmente los datos de las variables sintomáticas.

A ello se agrega las dificultades que pueden surgir por cambios a lo largo del tiempo en los límites político-administrativos así como en la cartografía que el INE utiliza de un censo de población a otro (delimitaciones de localidades y secciones y segmentos censales).

En forma similar, los cambios en la cobertura territorial o en las delimitaciones utilizadas en el relevamiento de la información es también una restricción que debe considerarse con atención para que las áreas consideradas en distintos momentos del tiempo y en relación a distintas variables sean las mismas (o muy aproximadas para no generar errores importantes al asimilarlas).

Es necesario analizar para cada variable y a lo largo del tiempo las posibilidades de desagregación considerando las siguientes delimitaciones de áreas pequeñas:

- Departamentos/Capitales departamentales/Resto de localidades (o más de un agrupamiento de localidades) y áreas rurales de cada departamento
- Departamentos/Municipios (o agrupamiento de municipios al interior de cada departamento)
- Departamentos/Secciones judiciales (o agrupaciones de secciones judiciales al interior de cada departamento): este conjunto de áreas pequeñas permite

utilizar en forma bastante precisa la información proveniente de las estadísticas vitales (nacimientos y defunciones). Sin embargo, los límites difieren de los utilizados por la cartografía censal del INE. Existe al menos un antecedente de evaluación de estas diferencias (trabajo inédito del especialista en Estadística Omar Prat (ASSE-MSP)) que se intentará recuperar y actualizar para analizar las posibilidades de uso.

Es también necesario evaluar como divisiones de interés algunas propuestas de regionalización existentes.

### **3.3 *Hacia la aplicación de los métodos***

Para aplicar en Uruguay, la presente propuesta consideró sólo algunos de los métodos propuestos como forma de iniciar una primer aproximación tanto a distintos modelos como variables sintomáticas pasibles de uso y diferentes opciones de división del territorio nacional.

El primer paso fue el análisis y acondicionamiento de los datos para su uso, proceso en el que es necesario seguir trabajando. Se mencionan algunos resultados de este análisis.

#### **Nacimientos:**

Existe un porcentaje bajo sin información de localidad de residencia de la madre (promedio: 4 %). Otra opción es utilizar sección judicial, que reduce el porcentaje sin información a cifras cercanas o inferiores al 1%. Dificultades: la sección judicial no se asimila a localidad; tampoco tienen límites asimilables a la cartografía censal.

Anualmente pueden no ocurrir nacimientos en algunas localidades pequeñas o tener un número muy pequeño de nacimientos; dificulta la implementación de algunos de los métodos (se estudiará la posibilidad de utilización de promedios móviles).

#### **Fallecimientos:**

No se cuenta con localidad de residencia del fallecido sino con sección judicial. El porcentaje sin información de sección judicial de residencia es muy bajo, sobre

todo si se excluye Montevideo (promedio: 1 %). La información no está incluida en las bases obtenidas para la década del 90.

### **Matrícula en primaria:**

La identificación de localidad que tiene ANEP para las escuelas no coincide con la utilizada por el INE. El Área de Investigación de ANEP realizó una equiparación de localidades pero no ha sido revisada. Por tanto se está realizando un chequeo de ubicación de las escuelas según localidades (los resultados serán devueltos al Área de Investigación para su validación).

Otro aspecto a considerar es la aparición/cierre de establecimientos en el período, que modifica las coberturas de cada centro.

La información obtenida es sobre matriculación y por tanto contiene casos de repetición en cada grado escolar. Fue necesario tomar decisiones entorno a como “depurar” repeticiones: promedio departamental pero ajustado a la tasa de cambio a nivel nacional. Es posible obtener la información depurada pero requiere una solicitud específica; se consideró no necesario en esta etapa.

### **Conexiones residenciales y consumo de energía eléctrica:**

La información obtenida es sobre consumo y contadores residenciales por oficinas comerciales. Las oficinas tienen coberturas muy heterogéneas (parciales de localidades, más de una localidad, incluso cubren en algún caso áreas de más de 1 departamento). En una segunda etapa será necesario ver con UTE la posibilidad de desagregados mayores, pero en esta etapa se optará por utilizar la información de esta forma que tiene como principal ventaja la oportunidad (es la información con más actualización encontrada en las fuentes).

Otro aspecto a analizar es la evolución que ha tenido la cobertura de UTE en el período, sobre todo en algunas áreas rurales. Se considerarán algunas posibles alternativas (como ajustar tasa de cambio según tasa nacional o regional, etc.)

### **Consultas en centros de atención de ASSE:**

Se obtuvo información histórica de consultas en los centros de atención primaria de ASSE en formato impreso, por lo cual, fue necesario digitar la información.

Es necesario analizar la cobertura de estos centros que están distribuidos en las capitales departamentales y sólo en algunos casos en otras localidades. A su vez, es necesario evaluar la aparición/cierre de servicios en el período.

### ***Combinación de fuentes y variables***

El nudo crucial sobre el cual es necesario trabajar y tomar decisiones a mediano plazo es la combinación de variables y fuentes tanto para la aplicación de los métodos multivariados como cuando se utilizan en forma individual las variables y luego se promedian resultados.

En particular aparecen dificultades para la apertura de las áreas rurales. En nacimientos y defunciones la opción es utilizar sección judicial; en el caso de consumo eléctrico las oficinas comerciales atienden áreas rurales agrupadas. En la matrícula de primaria: es necesario evaluar agrupamientos de zonas en áreas rurales ya que las escuelas cubren hogares en más de una zona. Puede ensayarse un nivel de secciones censales (yuxtaponer secciones judiciales y áreas atendidas por UTE para encontrar agrupamientos óptimos).

En general las variables que se han obtenido pueden ser utilizadas a nivel de áreas administrativas mayores (departamentos). Más aún, a este nivel existen otras variables que podrían utilizarse y que se analizarán para considerar la consulta en las fuentes sobre posibles desagregaciones a nivel intradepartamental.

## ***4. Exploración de modelos para estimar poblaciones departamentales***

Con el fin de hacer un primer testeo sobre métodos y sobre todo fuentes que se han obtenido, se aplicaron diferentes combinaciones de los mismos para la estimación de población a nivel de las divisiones político-administrativas mayores (departamentos).

Esto permitió una primer exploración y evaluación de las diferentes herramientas y fuentes para el caso uruguayo, sin dejar de considerar que la robustez y oportunidad de las estimaciones a nivel departamental son muy diferentes a las de niveles territoriales más desagregados. Este es el alcance del trabajo cuyos resultados se exponen en este capítulo.

Aplicar los modelos para estimar datos a nivel departamental como contenido para esta etapa del proyecto es una meta acotada pero factible dada la ausencia de trabajos previos para Uruguay, las dificultades para armonizar las áreas para las que se cuenta con información según el indicador e incluso en algunos casos, las dificultades para obtener los datos a niveles más desagregados.

Sin embargo, la estimación de población departamental tiene un interés propio. El INE presenta proyecciones de población por departamento y puede aportar en el futuro el contar también con estimaciones a partir de variables sintomáticas que permitan visualizar cambios de tendencia, etc. no captados por las proyecciones.

Por otra parte, tiene dos ventajas: contar con una mayor cantidad de variables sintomáticas que pueden utilizarse en los modelos. Contar con las proyecciones por departamentos que elaboró el INE en 2005 para contrastar las estimaciones en 2011 y –a último momento- contar también con los volúmenes poblacionales departamentales presentados por el INE en su página de acuerdo al Censo 2011 (aunque en varios departamentos el dato es aún parcial).

En una próxima etapa en la que se avance en la estimación a niveles más desagregados se podrá contrastar resultados tanto con datos censales (incluyendo 2011) como con las proyecciones por localidad realizadas por J.J. Calvo (hasta 2010).

En este estudio se utilizaron variables que pueden llegar a desagregarse en áreas menores a departamentos. Sin embargo se podrían testear otras variables que no han sido consideradas en este caso pero que están disponibles si el nivel de desagregación es el departamento.

#### ***4.1 Métodos aplicados, variables sintomáticas utilizadas y herramientas de evaluación de modelos***

Los **métodos** que se evaluaron fueron los expuestos con los números I a VIII (ver páginas 18-20), es decir: método de prorratio (I), método de razón censal (II), método de diferencia de tasas (III), método de distribución proporcional (IV), método de correlación de razones (V), método de correlación de diferencias (VI), método de tasa de correlación (VII) y método de tasas vitales (VIII).

Se evaluaron también los métodos utilizando la información de variables sintomáticas del año anterior (es decir, con un año de rezago), procedimiento que ha sido aplicado en algunas experiencias tal como se mencionó en el capítulo de presentación de los métodos. Estos últimos modelos son seguramente los de mayor utilidad, dado que el interés de obtener estimaciones en forma oportuna implica tener que utilizar variables sintomáticas obtenidas para algún período anterior al año para el que se hacen las estimaciones. Sobre todo es importante ver la “estabilidad” en cuanto a la capacidad de estimación de las variables que funcionan mejor en cada método (si es buena para estimar en el mismo período que también lo sea para la situación con rezago).

El método de prorrateo (I) requiere información base de un solo período, el de la estimación. La información necesaria es la población para el área mayor y valores de la variable sintomática para las áreas menores y el área mayor.

Los métodos II a IV y VIII requieren información base para dos períodos: un período 0 para el cual requiere conocer la población y la variable sintomática para cada área menor y para el área mayor y para el período de la estimación precisa el volumen de población del área mayor y los valores de la variable sintomática para las áreas menores.

Finalmente, los métodos V a VII requieren información de población por áreas menores para dos momentos en el tiempo por lo cual sólo pueden generarse modelos en base a dos censos que provean esos datos.

La información obtenida para la mayoría de las variables sintomáticas es de 1990 en adelante, por lo cual se descartó en esta etapa la utilización de datos censales de 1985. Esto restringe al Censo de 1996 y al Censo Fase I de 2004 la posibilidad de acceder a datos poblacionales por departamento (áreas menores). En cuanto al Censo 2011 recién se obtienen (en algunos casos parcialmente) los volúmenes de población departamental a comienzos de diciembre.

Dado que el interés está centrado en comparar las precisiones de las estimaciones obtenidas según las variables y los métodos, se resolvió centrar las **estimaciones en el año 2004** para los métodos I a IV y VIII y comparar con los resultados censales de ese año.

En el caso de los métodos V a VII, la exploración se restringió a evaluar los modelos obtenidos utilizando como base los datos censales de 1996 y 2004.

Tanto para el Censo de 1996 como el 2004, la información utilizada sobre población refiere al lugar de residencia habitual (no a la pernoctación la noche anterior, dato que puede obtenerse en el caso del Censo 1996). Esto permite acompañar también el trabajo futuro con el Censo de 2011 que se levantó con el criterio de residencia habitual.

De cualquier manera, dado que a la finalización de la consultoría se cuenta datos del Censo 2011- aunque aún no definitivos, se agregó un proceso más de **estimaciones al 2011** para hacer algunas comparaciones con los datos censales y también en este caso con las proyecciones al 2011.

Las **variables sintomáticas** utilizadas fueron: ***nacimientos***, ***defunciones***, ***hechos vitales*** (suma de nacimientos y defunciones), ***matrícula en educación primaria ajustada***, ***consultas externas de ASSE***, ***conexiones residenciales de UTE***, ***consumo de clientes residenciales de UTE*** y ***afiliados activos al BPS***.

**Nacimientos:** se trabajó con bases de nacimientos provistas por el Banco de Datos de la Facultad de Ciencias Sociales, con fuente en el Ministerio de Salud Pública. Estas bases se presentan en forma anual (según fecha de registro del nacimiento). Para obtener los nacimientos efectivamente sucedidos en un año se identificaron los nacimientos según fecha en la base del año de interés y en la del año subsiguiente. La variable que identifica el área geográfica es “departamento de residencia de la madre”.

**Defunciones:** también se trabajó con bases de defunciones provistas por el Banco de Datos de la Facultad de Ciencias Sociales, con fuente en el Ministerio de Salud Pública. Se utiliza como variable de identificación del área geográfica el “departamento de residencia del fallecido”.

**Matrícula educación primaria ajustada (pública y privada):** se obtuvieron los datos del Monitor de Educación Primaria y del Observatorio de ANEP para 2004 y 2010 y del Anuario Estadístico del INE (1995, 1996, 2003). La matrícula se ajustó extrayendo un número de casos equivalente a la tasa de repetición departamental para el año de interés. Para las matrículas obtenidas en los Anuarios no se contó con tasas de repetición de esos años

por lo cual se utilizó la tasa de repetición del año 2004, suponiendo estabilidad hacia atrás en esas tasas.

**Consultas externas de los centros de atención de ASSE (Administración de Servicios de Salud del Estado):** se utilizaron datos extraídos de Informes Estadísticos de ASSE (1995 a 2004) y del 2005 en adelante se obtuvieron los datos por solicitud al Área estadística de ASSE. Se contabilizaron datos departamentales como suma de las consultas externas de los Centros departamentales y Centros departamentales auxiliares.

**Conexiones residenciales UTE:** se utilizaron datos obtenidos vía solicitud al Área Comercial de UTE. Estos datos están agrupados por Oficinas Comerciales que si bien pertenecen a una localidad, atienden áreas que incluso en varios casos trascienden la frontera departamental. En estos casos fue necesario prorratear las conexiones entre departamentos utilizando criterios de distribución según se tratara de zonas urbanas o rurales, etc. Sin embargo el criterio dista de ser el más adecuado, sobre todo para los departamentos del área metropolitana (San José y Canelones) con varias oficinas comerciales que involucran a Montevideo. En una siguiente etapa se buscará mejorar la información si la misma puede obtenerse con un mayor desglose en la fuente.

**Consumo por conexiones residenciales UTE:** se obtuvieron de la misma fuente que las conexiones residenciales y presentan los mismos problemas. Adicionalmente, se encontró un problema con los datos de 1995: para algunas oficinas comerciales aparecen consumos muy bajos en relación a los demás años. Por ello cuando se requirió utilizar datos de 1995 en los modelos, para esta variable se usaron los valores de 1996.

**Afiliados al BPS (activos):** se obtuvo del sitio del INE. Esta variable permite trabajar en las estimaciones a nivel departamental; sin embargo, se requerirá una consulta en el BPS para saber si es posible un desglose a un nivel intradepartamental.

La evaluación sobre la eficacia para estimar los volúmenes poblacionales por cada método y variable(s) se realiza en base a las características de la distribución



de los errores (distancia entre la estimación y el parámetro poblacional en cada departamento –básicamente censal aunque en algún caso se utilizó como parámetro la proyección INE. Estos errores pueden trabajarse de distinta manera. En este caso la evaluación se realiza en base al *error relativo de valores absolutos*: esto es, tomando los errores sin considerar su signo (si es una sobre o subestimación del parámetro) y calculando qué porcentaje es ese error del volumen de población censal. Con fines descriptivos se presenta también el error en número de personas (*diferencia absoluta*), es decir, la diferencia en número de personas entre la estimación y el censo para cada departamento.

Para ambas medidas del error se calcularon los estadísticos: promedio, mediana, desvío, coeficiente de variación, porcentaje de casos con error mayor al 10% y porcentaje de casos con error menor al 5%. Los resultados se presentan en los cuadros del Anexo.

El promedio y la mediana permiten evaluar la precisión de las estimaciones: cuanto más cercano a cero, mejores estimaciones. Es de notar que en el caso de la diferencia absoluta, el promedio incluye valores positivos y negativos (sobre y subestimación). En cambio en el error relativo, el promedio “suma” valores positivos ya que se toma el error en valores absolutos, es decir, prescindiendo de los signos. Por tanto permite evaluar en qué medida los errores se alejan de los valores censales, sea por sub o sobreestimación.

Los demás estadísticos apuntan sobre todo a evaluar la precisión del método para generar estimaciones que en cada área estén cercanos a cero (más allá del promedio o la mediana de la distribución). Es decir, medir la variabilidad de la distribución de errores.

## **4.2 Resultados obtenidos**

Como ya se mencionó la aplicación de los métodos giró entorno a estimar el volumen de población para el año 2004. Se aplicaron los métodos I a IV y VIII (métodos que utilizan una sola variable sintomática) usando las 8 variables. Una vez evaluados los métodos, se volvieron a aplicar pero utilizando las variables con un año de rezago (1995 y 2003) para aquellos métodos que mostraron tener buen rendimiento con al menos una de las variables sintomáticas.

En el caso de los métodos que utilizan más de una variable sintomática (métodos V a VII) se analizaron los modelos de regresión lineal múltiple para explorar qué variables y modelos podían resultar adecuados de acuerdo a la significación de los coeficientes obtenidos y al comportamiento en relación a los supuestos que requiere el uso de estas técnicas estadísticas. El objetivo era comparar estimaciones obtenidas con esos modelos y los valores proyectados (ya que no se contaba con otro parámetro para comparar, aunque al finalizar el trabajo se pudo utilizar también los primeros resultados del Censo 2011 -parciales).

#### 4.2.1 Métodos que utilizan una sola variable sintomática

Se evaluó dentro de cada método cuáles variables generaron las estimaciones con menores errores y qué variables presentaron mejores estimaciones en los distintos métodos. Se testearon modelos y variables sintomáticas del mismo año que el dato poblacional pero también con rezago de un año.

**Cuadro 1. “Promedio” de los estadísticos para las 8 aplicaciones de cada método (una por variable sintomática) y resultados de las mejores estimaciones de cada uno**

Método	Promedio de los promedio	Promedio de los porcentajes por encima de 10%	Promedio de los porcentajes por debajo de 5%	Mejores estimaciones
Prorratio	15.3	54	24	Hechos vitales: promedio 3.4%; 0 deptos encima de 10%; 12 deptos menos de 5%
Prorratio con rezago	15.3	52	27	Hechos vitales: promedio 3.3%; 0 deptos encima de 10%; 15 deptos menos de 5%
Razón censal	6.2	17	58	Conexiones UTE: promedio 2.7%; 0 deptos encima de 10%; 14 deptos menos de 5%
Razón censal con rezago	10.1	32.4	37.5	Matrícula en primaria: promedio 3.5%; 0 deptos encima de 10%; 12 deptos menos de 5%
Diferencia de tasas	6.4	20.8	58	Conexiones UTE: promedio 3.1%; 0 deptos encima de 10%; 15 deptos menos de 5%
				Matrícula en primaria: promedio 3.1%; 1 depto encima de 10%; 14 deptos menos de 5%
Diferencia de tasas con rezago	7.7	22.2	47.9	Matrícula en primaria: promedio 3.4%; 0 deptos encima de 10%; 12 deptos menos de 5%
				Conexiones UTE: promedio 3.4%; 0 deptos encima de 10%; 14 deptos menos de 5%
Distribución proporcional	7.9	29	42	Conexiones UTE: promedio: 5.6%;

				2 deptos encima de 10%; 9 deptos menos de 5%
				Matrícula en primaria: promedio: 5.7%; 3 deptos encima de 10%; 9 deptos menos de 5%
Distribución proporcional con rezago	16.8	41	39	Conexiones UTE: promedio 5.0%; 3 deptos encima de 10%; 11 deptos menos de 5%
Tasas vitales	7.9	28	39	
Tasas vitales con rezago	7.9	39	56	

Los métodos que dieron mejores resultados fueron los de razón censal y diferencia de tasas, tanto utilizando las variables sintomáticas del año censal como usando el año anterior. (Ver Cuadro 1)

Las variables que generaron mejores resultados en al menos uno de los métodos fueron *hechos vitales*, *matrícula en primaria* y *conexiones de UTE*. (Ver Cuadro 1)

El análisis de los mejores resultados por método y variable muestra que en el caso del método de distribución por prorrato, la variable que permitió las mejores estimaciones fue *hechos vitales* con un promedio del porcentaje de error (en valores absolutos) de 3.4%, ninguna estimación con más de 10% de error y 12 de los 18 departamentos (67%) con un error menor al 5%. (Ver Cuadro 1 y Cuadro A.1 en Anexo)

Al utilizar las variables con rezago de un año en el método de prorrato, también la variable que genera errores menores es *hechos vitales*, con un promedio de 3.3%, ningún caso con más de 10% de error y sólo 3 departamentos con errores mayores a 5%. (Ver Cuadro 1 y Cuadro A.1 en Anexo)

En el método de razón censal la variable *conexiones de UTE* es la que tiene menor promedio de porcentajes de error absoluto (2.7%), ninguna estimación con más del 10% y 14 departamentos con errores menores al 5%; la estimación con mayor error relativo es de 5.8%. (Ver Cuadro 1 y Cuadro A.2 en Anexo)

En segundo lugar, la *matrícula de primaria* tiene un promedio de error relativo de 3.2%. También 14 departamentos se estiman con errores menores al 5% y solo en un caso la estimación es mayor al 10% (13%).

Al realizar las estimaciones con rezago de un año por el método de razón censal, la *matrícula de primaria* logra un promedio de porcentajes de error de 3.5%, ningún

departamento tiene un error relativo mayor al 10% y 12 departamentos se estiman con un error menor al 5%. (Ver Cuadro 1 y Cuadro A.2 en Anexo)

En el método de las diferencias de tasas, las variables *conexiones de UTE* y *matrícula de primaria* tienen promedios de porcentajes de error relativo de 3.1%. En el caso de las *conexiones de UTE*, ningún departamento tiene errores mayores al 10% y 15 departamentos logran estimaciones con errores menores al 5%. En la *matrícula en primaria*, 2 departamentos tienen errores mayores al 10% (pero menores a 11%) y 14 departamentos se estiman con errores menores al 5%. (Ver Cuadro 1 y Cuadro A.3 en Anexo)

Al utilizar las variables con rezago, estas dos variables empatan nuevamente en el promedio de error relativo: 3.4%, ninguna de las dos tiene estimaciones por encima del 10% y los departamentos con errores menores al 5% son 14 (*conexiones UTE*) y 13 (*matrícula de primaria*). (Ver Cuadro 1 y Cuadro A.3 en Anexo)

El método de distribución porcentual no generó buenas estimaciones con ninguna de las variables utilizadas, así como tampoco al utilizar las variables del año anterior al de la estimación. (Ver Cuadro A.4 en Anexo)

En el caso del método de tasas vitales, no se obtuvieron buenas estimaciones, aunque mejora al utilizar los datos con un año de rezago. (Ver Cuadro 1)

#### **4.2.2 Métodos que combinan variables en modelos de regresión**

Estos métodos, que son los de mayor uso para la estimación de población en áreas menores no pudieron aplicarse en este caso debido a las dificultades encontradas en la construcción de los mismos.

Para utilizar estos modelos es necesario construir un conjunto de variables “independientes” (en base a cada variable sintomática en dos años censales) y la dependiente (que se construye en base a los datos censales de esos dos años). En el caso de las variables sintomáticas se elaboraron usando los años censales y también los años anteriores –modelos con rezago de un año).

Para cada modelo se elaboraron las matrices de correlaciones de las variables independientes y la dependiente. En los modelos (sin y con rezago) de correlación de diferencias la mayoría de las variables independientes aparecieron con una alta correlación lineal tanto con la variable dependiente como entre ellas. Este hecho

hacía prever que para estos modelos la ecuación de regresión iba a tener problemas de alta multicolinealidad entre regresores. A su vez, el análisis de las gráficas de dispersión mostraron la influencia de los departamentos de mayor volumen poblacional y crecimiento (Canelones y Maldonado), lo cual constituyó un obstáculo adicional por la influencia importante de estos dos departamentos.

La elaboración de modelos de regresión para los distintos métodos sin y con rezago mostró la imposibilidad para encontrar un modelo óptimo multivariado.

En todos los casos fue posible construir un modelo que utilizaba solamente la variable matrícula en primaria pero, en varios de los mismos, presentaba problemas con los supuestos sobre residuos (normalidad e independencia).

Dadas estas dificultades para obtener modelos multivariados óptimos no se realizaron posteriormente estimaciones y cálculo de errores. Posiblemente este tipo de métodos pueda ser de mayor utilidad con un número sensiblemente más grande de áreas involucradas en la estimación de la ecuación de regresión. En estos casos, sin embargo, será probablemente necesario estratificar según el tamaño de población –al menos para utilizar el método de correlación de diferencias.

### ***4.3 Estimaciones y proyecciones al 2011***

Se aplicaron los métodos y variables con mejor resultado en las estimaciones la población en 2004, las variables sintomáticas con un año de rezago (2010) y el total poblacional obtenido en el Censo 2011 hasta la fecha. En los métodos que requieren valores poblacionales anteriores (total y por departamento), se utilizaron los datos censales 2004 y las variables sintomáticas al 2003.

Cuando se aprovecha el valor poblacional obtenido en el Censo 2011, la variable matrícula en primaria logra buenas estimaciones tanto con el método de razón censal como con el de diferencia de tasas (promedio: 2.5% y 2.4% de error relativo.). En ambos casos, se estima la población de 15 departamentos con un error menor al 5% y en ningún caso hay errores mayores al 10%. (Ver Cuadro A.5 en Anexo)

Sin embargo, cuando se estima la población departamental usando como total poblacional el brindado por la proyección del INE al 2011, el error relativo promedio se eleva a 6.0 para el método de razón censal y 6.1% para el de diferencia de tasas.

A su vez si bien es bajo el porcentaje de departamentos con errores mayores al 10% (ninguno en el caso de razón censal y 1 en el caso de diferencia de tasas), es bajo el número de estimaciones con errores menores al 5% (6 en el caso de razón censal, 7 al aplicar diferencia de tasas).

En cambio, mejora la estimación realizada con el método de prorrateo y las conexiones residenciales de UTE: el error promedio en este caso es 3%, solo un departamento se estima con un error mayor al 10% y 15 departamentos se estiman con errores menores al 5%.

Finalmente se compararon los resultados obtenidos por estimación con variables sintomáticas (usando como valor poblacional del área mayor el censo y el de la proyección al 2011) y los valores generados por la proyección del INE para el año 2011. Se utilizó la *matrícula en primaria* y *hechos vitales* por ser las variables sintomáticas que generaron mejores estimaciones en los distintos métodos aplicados para 2004. Se evaluaron para cada departamento cuál de los 3 métodos se acercaba más a los valores poblacionales que se han obtenido (al 17 de diciembre) para cada departamento.

Cuando se utiliza como parámetro para la población total del país el valor de la proyección, la estimación por razón censal (con rezago) de *matrícula en primaria* estima mejor que la proyección del INE en 12 departamentos; y con el método de diferencia de tasas, en 11.

Cuando se utiliza como total de población el obtenido en el censo, las estimaciones de razón censal con *matrícula en primaria* logran menores errores que la proyección, excepto en un departamento. En cambio a pesar de utilizar la cifra censal total de población, el método de diferencia de tasas estima mejor que los valores de proyección sólo en 11 departamentos.

Con la variable *hechos vitales* la estimación tanto por razón censal como por diferencia de tasas da resultados más cercanos a los datos censales en 10 departamentos cuando se utiliza como total poblacional el de la proyección. Cuando el total usado es el censal, 14 departamentos son estimados en forma más precisa que por medio de las proyecciones.

## 5. Comentarios finales

### 5.1 Sobre los resultados

En primer lugar es necesario destacar la dependencia que tiene cualquiera de estos métodos para estimar poblaciones de áreas menores en períodos postcensales de un valor proyectado de población para el área mayor (en este caso, el volumen de población del país). Esto es visible, por ejemplo, al comparar los resultados para 2011 usando como población la proyectada y la obtenida por censo.

En relación a los métodos analizados, el *método de razón censal* y el de *diferencias de tasas* fueron los que lograron estimar con mayor precisión y menor dispersión las poblaciones departamentales para el año 2004 en al menos alguna de las variables utilizadas. El *método de prorrateo* permitió buenas estimaciones utilizando la variable hechos vitales, aspecto a destacar por la simplicidad y poca información que requiere su aplicación.

Los *hechos vitales* mostraron una mayor capacidad para estimar población que los *nacimientos* y las *defunciones* por separado en los distintos métodos aplicados. Esta conclusión ratifica lo encontrado en otros estudios a nivel latinoamericano.

Con la variable *afiliados activos al BPS* no se obtuvieron buenas estimaciones. Probablemente está afectada por los procesos de informalización-formalización asociados tanto a coyunturas económicas como a intervenciones gubernamentales (crecimiento del número de cotizantes en los últimos años) y que son independientes de los cambios poblacionales.

Por su parte, la variable *consultas de ASSE* tampoco generó buenas estimaciones. Es posible que existan diversos factores asociados tanto a las formas de registrar consultas a través de los años y entre centros, así como funcionamientos diferentes, áreas de influencia que se modifican a través de los años en un mismo centro, etc. En este caso, será importante analizar a futuro la posibilidad de obtener otro tipo de información vinculada a los servicios de salud (en particular, el número de afiliados a través de FONASA a algún servicio de salud público o privado).



En cuanto a la variable *consumo residencial de UTE* será necesario también un análisis más profundo ya que puede haber factores asociados al funcionamiento de cada centro y sus áreas de influencia que expliquen parte de su variación independientemente de los cambios poblacionales y –como ya se ha mencionado– es fundamental obtener la información a un nivel más desagregado que el de oficinas comerciales.

Llama la atención en cambio que las *conexiones residenciales de UTE* generen buenas estimaciones (a pesar de haberse utilizado un criterio muy grueso para distribuir en casos de oficinas que atienden más de un departamento). Tal como se ha visto en otros estudios esta variable puede ser fundamental a la hora de obtener estimaciones de población en áreas menores, por lo cual se profundizará especialmente en conocer cómo es el movimiento de este indicador, los factores que influyen en su cambio y sobre todo la obtención de la información a un nivel más desagregado.

Las dificultades encontradas con los métodos que utilizan modelos de regresión lineal multivariada obstaculizaron el avance en la búsqueda de modelos óptimos para estimar poblaciones departamentales. De cualquier manera, estos métodos serán explorados en forma intensiva en los próximos meses, cuando se trabaje con conjuntos de áreas menores con mayor número de integrantes.

Se reitera la utilidad de utilizar las estimaciones y proyecciones en forma combinada para obtener valores de población –en este caso departamentales para períodos postcensales. El uso de ambos procedimientos permite identificar dinámicas demográficas que no serían visibles estadísticamente si no se complementaran ambas miradas: la tendencia prevista en base a supuestos de comportamiento de nacimientos, defunciones y migración y la estimación utilizando datos de coyuntura (variables sintomáticas). La detección de diferencias importantes entre ambos resultados es una alerta para focalizar la atención en procesos locales. Queda por estudiar –a la luz de los resultados censales 2011– qué decisiones se pueden tomar cuando aparecen estas discrepancias (usar la estimación, la proyección, un promedio de ambas, etc.)



Finalmente, dado que los datos censales de 2011 recién se han podido emplear en los últimos momentos del trabajo realizado, será necesario ampliar el trabajo con los mismos y evaluar con mayor detenimiento los resultados obtenidos con los distintos métodos y variables empleadas, como se menciona en el siguiente apartado.

## **5.2 Pasos a seguir a mediano y largo plazo**

Se deberá seguir profundizando a partir de los resultados de este estudio con el objetivo de encontrar los modelos más adecuados de acuerdo a las posibilidades de desagregación y cobertura de las distintas variables sintomáticas y también a las divisiones del territorio que se consideren más pertinentes.

A **mediano plazo** es necesario trabajar en diversos aspectos:

1. Para avanzar en la aplicación a niveles más desagregados es necesario seguir trabajando con los datos ya disponibles pero también dialogar con responsables de distintas fuentes para saber si es posible obtener datos desagregados a un nivel más adecuado (por ejemplo, UTE: desagregar datos de oficinas comerciales según residencia), si es posible obtener datos más precisos (por ejemplo, ANEP: matrícula de primaria sin repeticiones), si es posible dilucidar referencias confusas o mejorarlas a futuro (por ejemplo: MSP: registros de nacimientos y defunciones sin datos claros de residencia), etc.
2. En simultáneo, avanzar en cuál o cuáles son las delimitaciones de áreas que pueden tener mayor interés y para las cuales a su vez es posible adecuar los datos disponibles. Este proceso involucra directamente a los usuarios potenciales de este tipo de información. Implica también tomar decisiones sobre hasta dónde se “fuerza” a los datos y cuando no hay otra opción que tomar en forma agrupada algunas áreas con el fin de poder trabajar un determinada delimitación general del territorio (por ejemplo, para trabajar con localidades puede ser necesario considerar grupos de localidades en pro de mantener la integridad de la base de datos utilizada).

3. Si estos dos procesos requieren un tiempo considerable de trabajo, también lo tiene el evaluar las aplicaciones concretas que se realicen: cada variable o variables utilizadas y cada método y combinaciones de métodos, así como distintos tipos de estratificaciones y demás decisiones operativas que abren un abanico muy amplio de alternativas y resultados. Cada una de estas alternativas deben evaluarse utilizando los diferentes criterios planteados en el apartado de métodos.

4. Otro paso es combinar estimaciones con proyecciones a corto plazo (esto requiere el desarrollo de proyecciones a nivel más desagregado que el departamental) dado que entre ambas herramientas es posible obtener una visión más completa sobre en qué medida determinadas áreas pueden estar imbuidas en dinámicas poblacionales específicas.

5. Como pasos a seguir concretamente en una etapa inmediata se elegirán algunos departamentos para aplicar los métodos a grupos de localidades y áreas rurales de los mismos. La selección de departamentos y agrupamientos estarán basados en al menos dos criterios: a partir de la evaluación sobre las delimitaciones geográficas (estabilidad en el tiempo, posibilidad de desagregación de al menos algunas variables) y la evaluación de la aplicación a nivel departamental (departamentos cuyas estimaciones se consideren con menores errores).

Estas son metas a **mediano plazo**, idealmente al **primer trimestre del 2012**. Es decir, queda por delante un proceso de trabajo largo y complejo. Sin embargo, dado que el objetivo final es obtener herramientas lo más adecuadas posibles para la estimación de población en áreas menores en años postcensales y dado que este año se va a contar con información censal, es posible pensar que se tiene un tiempo suficiente para “madurar” todo este proceso, lograr resultados adecuados y a su vez precisar el tipo de recomendaciones a realizar sobre mejoras en los registros, etc.

Finalmente, a **largo plazo** el interés práctico está centrado en la institucionalización de un espacio de trabajo que permita obtener estimaciones a partir de **2012-2013** en forma periódica y sistemática.

Para ello son necesarios algunos procesos que exceden largamente la etapa inicial de evaluación, selección y ajustes de modelos:

- Dado que la elaboración y publicación de estas estimaciones debería tener como actor fundamental al INE, es necesario además de una apropiación y apoyo por parte del Instituto de esta actividad, generar una institucionalización que permita el sostenimiento del proceso de trabajo en forma continua (conexión permanente con las fuentes de datos para generar procedimientos sistematizados de obtención de registros en forma adecuada y oportuna) y a su vez un intercambio con los usuarios que retroalimente el trabajo en términos de demandas y sugerencias y consolide la actividad.
- Asociado a lo anterior, es necesario un trabajo con las fuentes de datos que tenga como objetivo apoyar la mejora en la calidad de los registros (cuando lo requiera) y la elaboración oportuna de resultados; a su vez requiere un seguimiento permanente de los cambios en la cobertura, formas de registro, etc.
- La consolidación de este espacio de trabajo en el INE permitiría avanzar también en otras estimaciones de poblaciones de interés (por grupos de edad y sexo, por ejemplo).
- Como se mencionó al inicio de este informe, contar con estimaciones sobre tamaños poblacionales en áreas menores permitiría a su vez utilizar estos datos para la aplicación de herramientas para estimar otros indicadores en áreas menores (provenientes de Encuestas de Hogares en particular). El Instituto de Estadística de la Facultad de Ciencias Económicas está trabajando actualmente en estos desarrollos.

## **Bibliografía**

- Álvarez, Gustavo (2001). *Estimación de población en áreas menores mediante variables sintomáticas: una aplicación para los departamentos de la República Argentina (1991 y 1996)*. Serie Población y Desarrollo N° 13. Santiago de Chile: CELADE
- Bay, Guiomar (1998). “El uso de variables sintomáticas en la estimación de la población de áreas menores”, en *Revista Notas de Población N° 67/68*. Santiago de Chile: CELADE (p. 181-208)
- Bay, Guiomar (2001). “Estimaciones indirectas de indicadores demográficos para áreas menores. Situación en América Latina”. en *Revista Notas de Población N° 71*. Santiago de Chile: CELADE (p. 9-20)
- BC Stats (1998). *Generalized Estimation System (GES). Small Area Population Estimation Methodology*. Canada: Ministry of Finance and Corporate Relations, Government of British Columbia
- Bega, M.T., M.V.D. Soares y R.N. Azzolini (1991), "Populaciio paranaense nos anos 80: evolugio e distribuiciio espacial", Análise de conjuntura, vol. 13, No 3, Curitiba, Instituto Paranaense de Desarrollo Económico y Social (IPARDES).
- Chávez Esquivel, Edwin (2001). “Variables sintomáticas en las estimaciones poblacionales a nivel cantonal en Costa Rica”, en *Revista Notas de Población N° 71*. Santiago de Chile: CELADE (p. 51-72)
- Chávez Esquivel, Edwin (2003). “Validación de los indicadores sintomáticos en las estimaciones poblacionales: censo del 2000” en *Revista Población y Salud en Mesoamérica (revista electrónica)*, Vol.1, N° 1, Artículo 1. San José de Costarrica: Centro Centroamericano de Población.
- Chi, Guangqing (2009). “Can Knowledge improve population forecasts at subcounty levels?” en *Demography*, Vol. 46, N° 2 (mayo). Baltimore: Population Association of America (p. 405-427)
- Duque, Ignacio (2009). “El desafío de la medida de la población en cada momento: las Estimaciones de la Población Actual” en *Revista Índice N°34* (mayo). Madrid: INE-UAM. (p. 10-13)
- González, Jesús; Enrique González (2007). *Estimación en áreas pequeñas de la encuesta de población activa en Canarias*. Ponencia en taller Estimación en áreas pequeñas. Tenerife: ISTAC
- González, Leandro; Dora Celton (2008). *Aplicación combinada de variables sintomáticas y proyecciones en poblaciones menores*, ponencia presentada en el XVI Encuentro nacional de estudios poblacionales, ABEP, Caxambú, MG, Brasil (29/9 al 3/10). Córdoba: CONICET-Universidad de Córdoba
- González, Leandro (2007). “Ajuste de proyecciones de poblaciones menores con variables sintomáticas. El caso del Gran Córdoba (Argentina) 2001-2016”, en *Revista Notas de Población N° 91*. Santiago de Chile: CELADE (p. 105-127)
- Howe, Andrew (2000). *Methods and procedures for estimating small area populations in Australia*. Demography Working Paper 2000/3. Adelaide: Australian Bureau of Statistics, Small Area Population Unit

- Irvin, Richard; Eduardo Fialho (1977). "Estimativa provisória da população do Distrito Federal: 15 de novembro de 1972", en Revista *Brasileira de Estatística*, No 35, No 146
- Kish, Leslie; Vijay Verma (1986). "Complete Censuses and Samples" en *Journal of Official Statistics*, Vol.2, N° 4. Suecia: Statistics Sweden (p. 381 – 395)
- Mancho, J (2002). *Técnicas de estimación en áreas pequeñas*. España: EUSTAT.
- Sawyer, Donald R. (1969), *Methods of estimating current population size and distribution, Penetration Roads and Settlement on the Belem-Brasilia Highway*, Harvard: Department of Social Relations, Harvard College.
- Sawyer, Diana Oya y José Alberto Magno de Carvalho (1981), *Projeções da população brasileira: 1980-2000*. Belo Horizonte: Centro de Desarrollo y Planificación Regional (CEDEPLAR)/ Universidad Federal de Minas Gerais (UFMG)/ Ministerio de Transporte.
- Shryock, H. S. y J.S. Siegel (1973). *The Methods and Materials of Demography*. Washington, D.C.: United States Bureau of the Census
- Smith, Stanley y Marylou Mandell (1984). "A Comparison of Population Estimation Methods: Housing Unit Versus Component II, Ratio Correlation, and Administrative Records" en *Journal of the American Statistical Association*, Vol. 79, N° 386, *Applications Section (junio)*. American Statistical Association
- Smith, Stanley y P. A. Morrison (2005). "Small-area and business demography" en Dudley Poston y Michael Micklin (editores.), *Handbook of Population*. New York: Springer Publishers
- Swanson, David; Jeff Tayman (2011). *On the ratio-correlation regression method*. California: University of California Riverside
- Texeira Jardim, María L. (2001): "Uso de variables sintomáticas para estimar la distribución espacial de población. Aplicación a los municipios de Río Grande do Sul, Brasil", en *Revista Notas de Población N° 71*. Santiago de Chile: CELADE (p. 21-50)
- Wong, Laura (1987). "Situação das projeções de população a nível de unidades federativas e algumas grandes regiões", en *Futuro da população brasileira: projeções, previsões e técnicas*, .R. Wong, R. Hakkert y R. A. Lima (comps.). Sao Paulo: ABEP

# Anexos

**Cuadro A 1 Método de prorrateo. Resultados de la distribución de errores en las estimaciones para 2004**

VARIABLE	DIFERENCIA ABSOLUTA (número de personas)				ERROR RELATIVO (en valores absolutos) (%)								
	promedio	mediana	desvío	Coef. de variación	máximo	mínimo	promedio	mediana	Coef. de variación	máximo	mínimo	Porcentaje por encima de 10%	Porcentaje por debajo de 5%
<b>Prorrateo</b>													
AFILIADOS BPS	5775	62532	111952	-218733	20.7	13.7	97.6	79.9	1.3	67	17		
CONSUMO UTE	-14529	59920	184358	-52349	30.5	30.3	65.2	90.5	6.0	89	0		
CONEXIONES UTE	-11052	41649	133355	-43947	21.3	22.6	59.0	53.9	1.5	83	17		
CONSULTAS ASSE	11870	34978	35363	-117254	23.3	25.6	42.7	34.3	0.5	89	6		
MATRICULA PRIMARIA	1265	12010	15875	-38294	7.2	6.6	55.9	14.6	0.5	28	22		
HECHOS VITALES	-77	5460	8007	-15818	3.4	3.0	69.6	7.1	0.1	0	67		
DEFUNCIONES	-733	9206	13653	-15287	9.0	7.3	68.9	20.7	0.2	39	28		
NACIMIENTOS	610	8976	13633	-19705	7.0	7.4	59.4	11.7	0.6	39	39		
<b>Prorrateo con rezago 1 año</b>													
AFILIADOS BPS	6390	60161	97019	-213612	20.7	14.5	89.3	69.2	1.9	67	6		
CONSUMO UTE	-14580	61287	200972	-52396	30.0	29.7	60.8	80.2	4.0	94	6		
CONEXIONES UTE	-11602	42167	135942	-42182	21.4	22.3	58.7	54.0	1.6	83	17		
CONSULTAS ASSE	9132	36188	49660	-117977	23.6	22.1	60.9	58.7	2.1	83	11		
MATRICULA PRIMARIA	1784	11520	13177	-38260	6.7	6.2	54.7	14.7	0.7	22	33		
HECHOS VITALES	-281	4460	6122	-10470	3.3	3.1	68.6	7.5	0.2	0	83		
DEFUNCIONES	-1860	12364	30642	-24684	10.3	8.8	70.8	25.7	0.5	50	22		
NACIMIENTOS	847	6784	10707	-11736	6.5	6.0	59.6	14.3	1.1	17	39		

**Cuadro A 2 Método de razón censal. Resultados de la distribución de errores en las estimaciones para 2004**

VARIABLE	DIFERENCIA ABSOLUTA (número de personas)				ERROR RELATIVO (en valores absolutos) (%)								
	promedio	mediana	desvío	Coef. de variación	máximo	mínimo	promedio	mediana	Coef. de variación	máximo	mínimo	Porcentaje por encima de 10%	Porcentaje por debajo de 5%
<b>Razón censal</b>													
AFILIADOS BPS	-2359	1294	30351	-1287	28032	-118343	8.3	6.4	87.1	24.4	0.3	28	44
CONSUMO UTE	-1147	1052	6751	-589	9615	-15314	5.9	4.9	68.5	13.7	0.7	17	50
CONEXIONES UTE	-397	147	3821	-964	5228	-9541	2.7	1.9	63.3	5.8	0.2	0	78
CONSULTAS ASSE	2292	-598	17702	772	44925	-22952	14.5	8.9	105.9	56.1	0.2	50	28
MATRICULA PRIMARIA	173	11	3457	1996	5816	-5101	3.2	2.8	86.5	10.9	0.3	6	78
HECHOS VITALES	619	-1270	11458	1852	43352	-10064	4.7	3.7	75.0	12.9	0.0	11	67
DEFUNCIONES	563	-281	6492	1153	19977	-7136	4.7	4.1	65.7	14.1	0.7	6	61
NACIMIENTOS	1003	-763	14401	1436	54454	-12652	5.7	4.4	83.3	16.2	0.4	22	56
<b>Razón censal con rezago 1 año</b>													
AFILIADOS BPS	-1696	1405	27332	-1612	24925	-105791	8.3	6.0	86.5	23.8	0.7	22	39
CONSUMO UTE	-966	1238	7934	-821	15562	-15188	6.4	5.5	51.7	12.3	1.1	17	44
CONEXIONES UTE	-340	-8891	33474	-9847	104973	-52593	20.6	18.6	64.8	55.5	1.6	78	11
CONSULTAS ASSE	2583	-2048	21195	821	40805	-30442	20.9	17.3	68.8	46.9	0.1	78	17
MATRICULA PRIMARIA	352	-221	4586	1302	11074	-6647	3.5	3.4	72.5	7.9	0.1	0	67
HECHOS VITALES	596	-1325	10444	1752	36912	-10764	5.9	6.2	64.0	15.5	0.5	11	33
DEFUNCIONES	781	-1499	8498	1089	16489	-13777	7.0	6.5	71.1	21.1	1.6	22	44
NACIMIENTOS	1308	284	14559	1113	48416	-18043	7.9	5.3	83.3	20.9	0.5	33	44

**Cuadro A 3 Método de diferencia de tasas. Resultados de la distribución de errores en las estimaciones para 2004**

VARIABLE	DIFERENCIA ABSOLUTA (número de personas)				ERROR RELATIVO (en valores absolutos) (%)								
	promedio	mediana	desvío	Coef. de variación	máximo	mínimo	promedio	mediana	Coef. de variación	máximo	mínimo	Porcentaje por encima de 10%	Porcentaje por debajo de 5%
<b>Diferencia de tasas</b>													
AFILIADOS BPS	-1524	1035	23584	-1548	22854	-89973	7.4	4.5	89.9	21.8	0.9	28	56
CONSUMO UTE	-2618	-2559	9711	-371	19919	-20009	6.9	4.8	77.6	16.8	0.8	28	56
CONEXIONES UTE	-1024	-992	3735	-365	4812	-11238	3.1	2.5	80.6	9.1	0.1	0	83
CONSULTAS ASSE	800	1538	17160	2145	28976	-34024	13.2	7.8	99.0	49.6	0.1	44	33
MATRICULA PRIMARIA	152	65	3382	2221	5640	-5255	3.1	2.6	87.2	10.9	0.2	6	78
HECHOS VITALES	787	-1605	13362	1698	50987	-10987	5.3	4.3	72.8	13.7	0.1	22	67
DEFUNCIONES	517	-329	6255	1210	19178	-7136	4.6	4.0	66.8	13.7	0.1	6	67
NACIMIENTOS	1629	-1723	18860	1157	71517	-14742	7.6	6.8	69.9	18.9	1.5	33	39
<b>Diferencia de tasas con rezago 1 año</b>													
AFILIADOS BPS	134	130	18460	13743	41506	-59581	7.3	5.0	129.1	39.6	0.1	17	50
CONSUMO UTE	-2345	-2053	10977	-468	31033	-19520	6.3	4.2	88.2	16.2	0.0	22	56
CONEXIONES UTE	-1078	-420	3835	-356	4771	-9312	3.4	3.3	68.5	7.6	0.1	0	78
CONSULTAS ASSE	885	-1442	21708	2452	42059	-36832	18.7	13.9	78.5	49.7	2.6	67	22
MATRICULA PRIMARIA	313	-178	4250	1357	9394	-6499	3.4	3.2	72.5	7.7	0.0	0	67
HECHOS VITALES	722	-1792	11718	1624	41869	-12047	6.4	6.6	63.7	16.4	0.2	11	33
DEFUNCIONES	716	-1476	8277	1156	15626	-13345	6.9	6.3	68.2	20.0	1.7	22	39
NACIMIENTOS	1905	45	17716	930	59011	-21628	9.4	6.2	82.5	26.5	0.5	39	39



**Cuadro A 4 Métodos de distribución proporcional y de tasas vitales. Resultados de la distribución de errores en las estimaciones para 2004**

VARIABLE	DIFERENCIA ABSOLUTA (número de personas)				ERROR RELATIVO (en valores absolutos) (%)								
	promedio	mediana	desvío	Coef. de variación	máximo	mínimo	promedio	mediana	Coef. de variación	máximo	mínimo	Porcentaje por encima de 10%	Porcentaje por debajo de 5%
<b>Distribución proporcional</b>													
AFILIADOS BPS	2393	18509	35958	-55546	10.8	7.1	83.8	31.8	0.2	44	22		
CONSUMO UTE	2238	14730	19437	-50275	8.1	8.9	64.1	16.3	0.1	44	39		
CONEXIONES UTE	1503	10038	11712	-33668	5.6	4.7	77.9	15.3	0.5	11	50		
CONSULTAS ASSE	527	15875	39869	-25526	12.2	8.4	107.6	47.1	0.7	39	33		
MATRICULA PRIMARIA	1069	12415	17935	-42100	5.7	5.2	84.1	15.2	0.2	17	50		
HECHOS VITALES	2071	17120	16518	-65205	6.6	6.8	79.1	19.1	0.3	28	44		
DEFUNCIONES	594	8126	17543	-16154	6.1	5.3	67.0	15.5	0.3	17	50		
NACIMIENTOS	2499	20430	19495	-78085	8.0	6.9	77.7	22.5	1.4	28	44		
<b>Distribución proporcional con rezago 1 año</b>													
AFILIADOS BPS	1880	17074	23926	-53204	10.7	9.1	82.0	34.2	1.0	44	28		
CONEXIONES UTE	368	9684	13540	-33604	5.0	3.2	82.8	14.4	0.7	17	61		
CONSULTAS ASSE	-2702	20511	33334	-31266	18.1	19.8	77.5	48.5	0.2	61	22		
MATRICULA PRIMARIA	851	13637	19225	-47569	5.7	5.2	86.0	17.0	0.2	17	50		
HECHOS VITALES	790	13243	13593	-47060	5.8	3.9	90.5	16.4	0.2	22	61		
DEFUNCIONES	-1083	15320	55988	-14723	6.4	5.6	67.1	13.5	0.3	28	44		
NACIMIENTOS	1793	21486	24871	-77852	8.8	5.2	94.2	29.4	0.4	39	50		
<b>Tasas vitales</b>													
	2008	795	15155	755	51886	-17360	7.9	4.7	82.1	21.6	0.2	39	56
	214	-1572	9234	4320	22499	-12840	7.9	8.4	69.4	18.9	0.4	28	39

**Cuadro A 5 Resultados de la distribución de errores en las estimaciones para 2011 (Censo y proyecciones). Métodos y variables con mejor desempeño en las estimaciones 2004**

VARIABLE	DIFERENCIA ABSOLUTA (número de personas)				ERROR RELATIVO (en valores absolutos) (%)				Porcentaje por encima de 10%	Porcentaje por debajo de 5%					
	promedio	median a	desvío de	coeficiente de variación	máximo	mínimo	promedio	median a			desvío de	coeficiente de variación	máximo	mínimo	
<b>Contraste: CENSO 2011</b>															
<b>Prorrateo (rezago 1 año)</b>															
HECHOS VITALES		733	8208		12073	-26414		4.0	2.9	4.3	106.9	18.5	0.1	6	72
<b>Razón censal (rezago 1 año)</b>															
CONEXIONES UTE	281	3500	12567	4468	10163	-47706		5.5	5.0	2.8	50.1	11.4	1.0	6	50
MATRICULA PRIMARIA	179	871	2037	1140	4130	-2822		2.5	2.1	1.7	65.3	6.0	0.1	0	83
HECHOS VITALES	151	1395	5717	3787	9867	-16927		4.4	3.5	3.5	78.9	15.1	0.6	6	83
<b>Diferencia de tasas (rezago 1 año)</b>															
CONEXIONES UTE	-460	1679	9798	-2131	7881	-37376		3.8	3.4	2.8	73.9	10.4	0.5	6	72
MATRICULA PRIMARIA	224	762	2089	932	4217	-2959		2.4	2.1	1.8	74.3	6.4	0.3	0	83
HECHOS VITALES	173	1617	5644	3266	9716	-16245		4.5	3.7	3.4	75.7	14.9	0.7	6	83
<b>Contraste: PROYECCIONES 2011</b>															
<b>Prorrateo (rezago 1 año)</b>															
HECHOS VITALES	-5520	-4466	11047	-200	7035	-45467		5.6	5.6	3.3	58.6	11.1	0.5	11	50
<b>Razón censal (rezago 1 año)</b>															
CONEXIONES UTE	-5239	-1342	15606	-298	2051	-66759		3.0	2.1	3.0	99.9	12.6	0.1	6	83
MATRICULA PRIMARIA	-5342	-5382	5175	-97	7223	-18668		6.0	6.2	2.7	44.2	10.0	1.6	0	33
HECHOS VITALES	-5369	-3256	9431	-176	6482	-35980		6.3	6.8	3.6	56.2	12.6	0.0	17	33
<b>Diferencia de tasas (rezago 1 año)</b>															
CONEXIONES UTE	-5980	-2889	13245	-221	3460	-56429		4.4	4.6	3.1	70.4	10.6	0.2	6	67
MATRICULA PRIMARIA	-5296	-5546	4885	-92	7289	-16257		6.1	6.7	2.7	44.0	10.3	1.8	6	39
HECHOS VITALES	-5348	-3208	9398	-176	7402	-35298		6.5	6.6	3.6	55.7	12.9	0.3	17	33



PRESIDENCIA  
REPÚBLICA ORIENTAL DEL URUGUAY  
Oficina de Planeamiento y Presupuesto



Naciones Unidas Uruguay  
Unidos en la Acción



Fondo de Población  
de las Naciones Unidas

