

## GUÍA METODOLÓGICA

# SEDLAC

## BASE DE DATOS SOCIOECONÓMICOS PARA AMÉRICA LATINA Y EL CARIBE (CEDLAS y Banco Mundial) \*

---

### ÍNDICE

0. INTRODUCCIÓN.....	2
1. ENCUESTAS DE HOGARES .....	4
2. INGRESOS .....	14
3. POBREZA .....	27
4. DESIGUALDAD .....	34
5. DEMOGRAFÍA.....	39
6. EDUCACIÓN.....	41
7. EMPLEO .....	46
8. VIVIENDA.....	52
9. INFRAESTRUCTURA .....	54
<i>Referencias.....</i>	<i>56</i>

---

\* Citar esta guía como CEDLAS (2025). "Guía metodológica. SEDLAC. Base de datos socioeconómicos para América Latina y el Caribe (CEDLAS y Banco Mundial)."

CEDLAS es el Centro de Estudios Distributivos, Laborales y Sociales de la Universidad Nacional de La Plata.  
E-mail: [cedlas@depeco.econo.unlp.edu.ar](mailto:cedlas@depeco.econo.unlp.edu.ar)

## 0. INTRODUCCIÓN

SEDLAC es una base de datos de estadísticas socioeconómicas, construida a partir de microdatos provenientes de encuestas de hogares de América Latina y el Caribe (ALC), desarrollada por el CEDLAS (Centro de Estudios Distributivos, Laborales y Sociales de la Universidad Nacional de La Plata) y el Banco Mundial. Esta guía describe el contenido de cada tabla en la base de datos y explica brevemente las principales decisiones metodológicas tomadas para construir las estadísticas.

SEDLAC está dividida en 9 secciones: encuestas de hogares, ingresos, pobreza, desigualdad, características demográficas, educación, empleo, vivienda e infraestructura. Cada sección contiene al menos un archivo Excel que incluye varias hojas. Cada una de estas hojas presenta una tabla con estadísticas con datos específico para cada uno de los 24 países de ALC (o para todos aquellos en que existen datos). Todas las estadísticas son obtenidas a partir de microdatos de encuestas de hogares a través de rutinas documentadas en *Stata do files*.

Una base de datos a gran escala, como lo es SEDLAC, implica un minucioso y periódico trabajo de revisión y control de la construcción de estadísticas y variables, identificando errores y buscando formas más eficientes de procesar los datos y presentar las estadísticas. En este sentido, es de esperar que SEDLAC presente cambios en el tiempo. Las sugerencias de usuarios e investigadores son un aporte importante para mejorar la base de datos. SEDLAC intenta ser un espacio de interacción entre investigadores sobre temas sociales en ALC.

### Comparabilidad de las estadísticas

Las encuestas de hogares no son uniformes entre países de ALC. En especial difieren significativamente en cobertura geográfica y cuestionarios. También se diferencian a través del tiempo dentro de un mismo país. Muchos países de ALC han mejorado sus encuestas de hogares en la década pasada, con cambios en la cobertura y los cuestionarios de las mismas. Debido a ello, la cuestión de la comparabilidad es de gran importancia. En particular, ¿cuán comparables (entre países y en el tiempo) son las estadísticas presentadas en SEDLAC?

Hacemos todos los esfuerzos posibles para lograr que las estadísticas sean comparables entre países y en el tiempo. Con ese fin, usamos definiciones similares de variables en cada país/año y aplicamos métodos consistentes de procesamiento de datos. Sin embargo, no es posible asegurar la perfecta comparabilidad, en la medida que la cobertura y los cuestionarios de las encuestas de hogares difieren entre países y, frecuentemente, en el tiempo para un mismo país. Por lo tanto, surge un *trade-off* entre cobertura y precisión. Si queremos ser suficientemente abarcativos en el análisis,

pagamos el costo de perder precisión y aparecen problemas de comparabilidad. A veces dichos problemas son demasiados severos y es conveniente restringir el análisis. Esta guía y otros documentos, publicados en nuestra página web, proveen a los usuarios la información relevante para decidir sobre dicho *trade-off*. La decisión final sobre si realizar o no una comparación la podrá tomar cada usuario de acuerdo a sus preferencias y necesidades específicas.

Aún a pesar de tener varios problemas, las encuestas de hogares son la mejor fuente de información para obtener estadísticas socioeconómicas nacionales. Pensamos que deberían evitarse las dos posiciones extremas hacia las encuestas de hogares: descartarlas completamente o usarlas sin ninguna precaución. Con todas sus limitaciones las encuestas de hogares proveen información valiosa, siendo la fuente más confiable, entre las disponibles, para generar estadísticas representativas de la población. Sin embargo, es importante ser consciente de sus desventajas. A pesar de que los gobiernos de ALC y los organismos internacionales han dado importantes pasos en las últimas décadas, falta recorrer un largo camino para tener un conjunto de encuestas de hogares más confiables, valiosas y homogéneas.

Esta guía está organizada en 9 capítulos. En cada capítulo describimos las tablas disponibles en cada sección de la base de datos, proveemos la información básica para leer las tablas y discutimos las decisiones metodológicas tomadas para construirlas.

## Citas

La información tomada de la base de datos debe citarse como “Fuente: SEDLAC (CEDLAS y Banco Mundial)” o “Fuente: Base de Datos Socioeconómicos para América Latina y el Caribe (CEDLAS y Banco Mundial)”. Recomendamos hacer referencia a la fecha en la cual la base de datos fue consultada debido a que las estadísticas pueden cambiar.

## 1. ENCUESTAS DE HOGARES

El archivo [AÑO\\_ActX\\_surveys\\_LAC.xls](#) posee información sobre las principales características de las encuestas de hogares incluidas:

- País
- Nombre de la encuesta
- Acrónimo
- Año en que fue realizada
- Fecha del trabajo de campo
- Cobertura geográfica
- Número de hogares encuestados
- Número de individuos encuestados
- Número de hogares representados (expandidos)
- Número de individuos representados (expandidos)
- Contenidos de la encuesta

En particular, el archivo reporta si una encuesta específica tiene información básica sobre:

- Demografía
- Vivienda
- Educación
- Empleo
- Ingreso
- Gasto/Consumo
- Salud
- Programas sociales
- Antropometría
- Migración

También se indica si la encuesta tiene otros módulos especiales y si se realizaron encuestas paralelas (*Ejemplo*: para recabar precios locales).

## **Aspectos Metodológicos**

### **Países**

En este momento, SEDLAC incluye información para los siguientes países: Argentina, Bolivia, Brasil, Colombia, Costa Rica, Chile, Ecuador, El Salvador, Guatemala, Honduras, México, Nicaragua, Panamá, Paraguay, Perú, República Dominicana, Uruguay y Venezuela. La muestra cubre todos los países de América Latina continental. En cada periodo la muestra de países considerada representa más del 97% de la población total de ALC. La base de datos abarca principalmente las décadas de 1990s, 2000s, y 2010s y 2020s, no obstante también presenta información de décadas previas en algunos países.

La mayoría de las encuestas incluidas en la muestra son representativas a nivel nacional. La principal excepción es Argentina, donde las encuestas cubren sólo población urbana, la cual representa más del 85% de la población de este país. Además, trabajamos con algunas encuestas que cubren sólo áreas urbanas en Bolivia, Colombia y Paraguay a comienzos de los 90s, y Uruguay hasta el año 2005.

A continuación se describen brevemente las principales características de las fuentes de información de cada país.

### **Argentina**

Todas las estimaciones de Argentina son obtenidas a partir de la Encuesta Permanente de Hogares (EPH). Esta encuesta es llevada a cabo por el Instituto Nacional de Estadística y Censos (INDEC) desde comienzos de los 70s en el Gran Buenos Aires, y desde los 80s en las ciudades más habitadas del país (aquellas con más de 100.000 habitantes). Antes de 2003 era implementada en dos rondas: mayo y octubre. Durante 2003 el INDEC implementó un importante cambio metodológico, realizando modificaciones en los cuestionarios y en el período de captación de la información. La nueva encuesta (conocida como Encuesta Permanente de Hogares Continua o EPHC) es ahora realizada a lo largo de todo el año. En 2003 el INDEC también comenzó a realizar imputaciones a las observaciones sin respuesta de ingreso. Para evaluar el impacto de estos cambios metodológicos, en varias tablas presentamos 2 conjuntos de estadísticas para 2003: uno computado a partir de la EPH de Mayo, y otro computado sobre la EPHC del segundo semestre de 2003 con ponderadores que consideran el problema de no-respuesta. La EPHC cubre las 31 mayores áreas urbanas del país, en las cuales habita alrededor del 70% de la población urbana de Argentina. Dado que el porcentaje de población urbana en el país es del 87%, la muestra de la EPH representa cerca del 60% de la población argentina total. Desde 2003 en adelante SEDLAC incluye información de 2 encuestas para cada año, una para cada semestre. La última encuesta armonizada en el proyecto corresponde al segundo semestre de 2024.

### ***Bolivia***

La encuesta de hogares que se lleva a cabo en la actualidad en Bolivia es la Encuesta de Hogares (EH). La misma tiene representatividad nacional y es realizada desde 1999. Bolivia tuvo anteriormente otras dos encuestas de hogares: la Encuesta Integrada de Hogares (EIH) hasta 1995, que cubría sólo ciudades capitales de regiones y la ciudad de El Alto, y la Encuesta Nacional de Empleo (ENE), realizada en tres ocasiones entre 1996 y 1997. La última encuesta disponible en el proyecto es la EH correspondiente al año 2023.

### ***Brasil***

Hasta 2015, la principal encuesta de hogares de Brasil fue la Pesquisa Nacional por Amostra de Domicílios (PNAD), la que se efectuaba una vez por año, en septiembre, por el Instituto Brasileiro de Geografia y Estadísticas (IBGE). La PNAD se llevó a cabo por primera vez en el estado de Río de Janeiro en 1967. A partir del año 2004, con la incorporación de las áreas rurales de algunos estados del norte del país (Rondonia, Acre, Amazonas, Roraima, Pará y Amapá), la PNAD alcanzó la cobertura total del territorio de Brasil. La encuesta no se realizó en 1970, 1980, 1991, 2000 y 2010 (años censales) y 1994. Desde 2012, la PNAD comenzó a realizarse durante todo el año, en forma continua (entre 2012 y 2015 se llevaron a cabo ambas encuestas, la PNAD puntual en septiembre y la PNAD Continua durante todo el año). Precisamente, la PNAD Continua 2024 es la última encuesta de Brasil armonizada en el proyecto SEDLAC.

### ***Chile***

La Encuesta de Caracterización Socioeconómica Nacional (CASEN) es una encuesta de hogares representativa a nivel nacional y regional llevada a cabo por el Ministerio de Planeamiento (MIDEPLAN) a través del Departamento de Economía de la Universidad de Chile, que es responsable de por la recolección de datos, la digitalización y el chequeo de consistencia de la base de datos. Hasta 2013, y antes de poner las bases de datos a disponibilidad del público general, CEPAL llevó a cabo los ajustes por no respuesta, ingresos faltantes y el subreporte o sobreporte de diferentes categorías de ingresos. Desde entonces, el equipo del Observatorio Social del Ministerio de Desarrollo Social y Familia se encarga de dichos ajustes.

La encuesta se implementó regularmente cada dos años desde 1985 durante Noviembre y en algunos casos, hasta mediados de Diciembre. A partir del año 2000 la frecuencia de realización disminuyó, realizándose la encuesta cada tres años. A partir del año 2009 se retomó la frecuencia bienal. La última CASEN disponible es la realizada durante el año 2022, en los próximos meses se incorporará la de 2024..

## **Colombia**

La Encuesta Nacional de Hogares-Fuerza de Trabajo (ENH) fue, inicialmente, la principal encuesta de hogares en Colombia. Esta era llevada a cabo por el Departamento Administrativo Nacional de Estadística (DANE) hasta el 2000. Inicialmente, la encuesta sólo cubría las cuatro principales áreas urbanas del país. Progresivamente se fueron incorporando nuevas áreas urbanas y en 1992 en la onda de septiembre se incluyeron las áreas rurales. En 2000 la ENH dejó su lugar a la Encuesta Continua de Hogares (ECH). La ECH tiene diferencias significativas con la ENH en términos de frecuencia, cobertura y cuestionarios. A partir del año 2006, la principal encuesta de hogares de Colombia es la Gran Encuesta Integrada de Hogares (GEIH), la que también se realiza en forma continua, a lo largo de todo el año. En la actualidad SEDLAC reporta información para los años 2001-2005 (ECH) y 2008-2024 (GEIH). Si bien se dispone de la información de la ENH la misma no es comparable con las series mencionadas, por lo que no se incluye en SEDLAC.

## **Costa Rica**

Hasta el año 2009, la encuesta de hogares más importante de Costa Rica era la Encuesta de Hogares de Propósitos Múltiples (EHPM), la que se comenzó a realizar en el año 1987 por el Instituto Nacional de Estadísticas y Censos (INEC).<sup>1</sup> La EHPM se levantaba una vez al año, en julio, y era nacionalmente representativa, cubriendo áreas urbanas y rurales del país. La EHPM, que entrevistaba aproximadamente a 40,000 individuos, no tenía un cuestionario muy rico. Aunque esta encuesta se mantuvo prácticamente inalterada en el tiempo, los cambios en los ponderadores que se produjeron después del Censo de Población 2000 generan algunos problemas de comparabilidad entre encuestas anteriores y posteriores a dicho año. A partir del año 2010 la EHPM ha sido reemplazada por una nueva encuesta: la Encuesta Nacional de Hogares (ENAH). Esta no es directamente comparable con las encuestas anteriores, particularmente en lo referido a la captación del ingreso, ya que cuenta con un cuestionario mucho más detallado (diferenciando entre ingresos brutos y netos, ingresos monetarios y en especie) del que se aplicaba en la EHPM.

En el proyecto SEDLAC se ha procesado y se presenta la serie completa de encuestas de hogares de Costa Rica para el período 1989-2024 (serie EHPM para el período 1989-2009 y serie ENAH para el período 2010-2024). La serie de estadísticas de SEDLAC para Costa Rica presenta dos cortes: uno en el año 2001, dado que recién a partir de ese año se

---

<sup>1</sup> Las encuestas de hogares en Costa Rica fueron implementadas en 1966. La Encuesta Centroamericana de Hogares (ECH) cubrió el período 1966-1971 y fue llevada a cabo en marzo, julio y noviembre. Desde 1976 a 1986 el INEC realizó la Encuesta Nacional de Hogares, Empleo y Desempleo (ENHED), en las mismas tres ondas: marzo, julio y noviembre.

cuenta con la información sobre vivienda de la EHPM, lo que hace posible estimar la renta imputada de la vivienda propia e incluir la misma en las estimaciones de ingresos del proyecto SEDLAC; y otro en 2010, cuando la serie cambia de la EHPM a la ENAHO.

### ***República Dominicana***

Hasta 2016, la principal encuesta de hogares de República Dominicana fue la Encuesta Nacional de Fuerza de Trabajo (ENFT). La ENFT se realizó desde 1991 por el Banco Central de la República Dominicana. En dicho año y en 1992 la encuesta se llevó a cabo en cuatro rondas: enero, abril, julio y octubre. Desde 1993 la ENFT se llevó a cabo en dos ondas: abril y octubre. La ENFT era nacionalmente representativa con un tamaño muestral de más de 20,000 habitantes. Algunos cambios de importancia significativa en la encuesta, realizados a partir del año 2000, provocan serios problemas de comparabilidad con las encuestas previas. A partir de 2016, la ENFT se transformó en la ECNFT: la Encuesta Continua Nacional de Fuerza de Trabajo, la que se realiza durante todo el año. La ECNFT 2024 es la última encuesta de República Dominicana que se ha incorporado en la base de datos del proyecto.

### ***Ecuador***

La Encuesta de Condiciones de Vida (ECV) es la versión ecuatoriana de las Encuestas de Mediciones de Niveles de Vida. Hasta ahora, han sido llevadas a cabo seis encuestas nacionales: 1994, 1995, 1998, 1999, 2006 y 2013-2014 (las cinco primeras están incorporadas en el proyecto SEDLAC). Si bien comenzó a realizarse en la década de los 90s, a partir del año 2003 la Encuesta de Empleo, Desempleo y Subempleo (ENEMDU) se ha convertido en la principal encuesta de hogares a nivel nacional en Ecuador. En la actualidad la ENEMDU se realiza en forma continua a lo largo del año, pero en SEDLAC se armonizan los datos correspondientes a diciembre de cada año.

En la actualidad SEDLAC presenta la serie 2003-2024 de la ENEMDU, además de las cinco primeras ECV (1994-1995-1998-1999-2006) y la Encuesta Periódica de Empleo y Desempleo (EPED) entre 1995 y 2000. Debido a cuestiones metodológicas, la información proveniente de las distintas encuestas no es estrictamente comparable.

### ***El Salvador***

La Encuesta de Hogares de Propósitos Múltiples (EHPM) es la encuesta de hogares de El Salvador. Antes de que el país se incorporara al programa MECOVI, las encuestas de hogares eran reducidas en términos de cobertura y cuestionarios. La EHPM se realiza bajo el programa MECOVI desde 1997. Las mayores contribuciones de este programa han sido la expansión y la profundización de la cobertura temática del cuestionario. En las



estadísticas de pobreza y desigualdad de SEDLAC únicamente se incluyen las EHPM del período 2000-2023 (con la excepción de la EHPM 2003), debido a que para las encuestas anteriores no se dispone de uno de los módulos originales donde se releva el ingreso de trabajadores agropecuarios por cuenta propia, lo que dificulta la comparabilidad. Las encuestas correspondientes a dichos años se encuentran en proceso de revisión.

### ***Guatemala***

El programa MECOVI ha promovido las encuestas de hogares en Guatemala. La principal encuesta de hogares es la ENCOVI (Encuesta Nacional de Condiciones de Vida) llevada a cabo en los años 2000, 2006, 2011, 2014 y 2023. Otra encuesta disponible es la ENEI (Encuesta Nacional de Empleo e Ingresos), la que fue realizada durante el período 2002-2004, para luego ser retomada y realizada anualmente a partir del año 2010. La base de datos SEDLAC incluye estadísticas para Guatemala que han sido estimadas a partir de la información de las ENCOVIs 2000, 2006, 2011, 2014 y 2023 y de las ENEIs 2002, 2003 y 2004.

### ***Honduras***

La principal encuesta de hogares en Honduras es la Encuesta Permanente de Hogares de Propósitos Múltiples (EPHPM). Durante 1990 y 1991 la EPHPM fue realizada una vez al año. Desde 1992 se realizan dos rondas (abril-mayo y setiembre-octubre), con la excepción de 2000. Los ingresos no monetarios y no laborales no fueron incluidos en la EPHPM hasta 1996. Hasta el año 1999 la encuesta estuvo a cargo de la Dirección General de Estadísticas y Censos, pero desde 2000 es realizada por el Instituto Nacional de Estadística. La última EPHPM incorporada en la base de datos SEDLAC es la correspondiente al año 2024, pero debe tenerse en cuenta que no se armonizaron las encuestas correspondientes al período 2020-2022, debido a problemas de comparabilidad o falta de datos.

### ***México***

La principal encuesta de hogares en México es la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH). Esta encuesta contiene información sobre características socioeconómicas, ingresos e información detallada sobre diferentes categorías de gastos. Es nacionalmente representativa, cubriendo áreas urbanas y rurales. La ENIGH es realizada por el Instituto Nacional de Estadística y Geografía (INEGI) y se dispone de información de la misma para 1984, 1989, 1992, 1994, 1996, 1998, 2000, 2002, 2004, 2005, 2006, 2008, 2010, 2012 y 2014. Desde 2016 la ENIGH se denomina ENIGH Nueva

Serie (ENIGH-NS) y se incluyen en SEDLAC las encuestas correspondientes a 2016, 2018, 2020, 2022 y 2024 de esta nueva serie.

### ***Nicaragua***

La encuesta de hogares de Nicaragua es la Encuesta Nacional de Hogares sobre Medición del Nivel de Vida (EMNV), una encuesta con cobertura nacional que se realizó en seis ocasiones: 1993, 1998, 2001, 2005, 2009 y 2014.<sup>2</sup> Ha sido desarrollada por el Instituto Nacional de Estadística y Censos de Nicaragua (INEC) en el marco del programa MECOVI. Desde 2014 no se han incorporado nuevas encuestas de Nicaragua al proyecto SEDLAC.

### ***Panamá***

La principal encuesta de hogares de Panamá es la Encuesta de Hogares (EH), efectuada en agosto de cada año. Es una encuesta de hogares representativa a nivel nacional. Panamá realizó encuestas del tipo LSMS en 1997, 2003 y 2008. Nuestro análisis se restringe sólo a la información provista por la EH. Aunque ésta es una buena encuesta con más de 50,000 observaciones, la información para la mayoría de los años se limita a variables demográficas, educativas, de empleo, de migración e ingreso. A partir del año 2008 se cuenta con información más completa, en particular en lo referido a acceso a programas sociales y propiedad de la vivienda. La serie de la EH que se presenta en SEDLAC presenta un corte en 2008, dado que en dicho año se comenzó a incluir en los ingresos la información de la renta imputada de la vivienda propia, práctica que no se podía realizar en las encuestas anteriores por falta de información sobre la propiedad de la vivienda. Desde 1997 la EH está disponible en todos los años, excepto 2022. La EH del año 2024 es la última encuesta de Panamá armonizada en SEDLAC.

### ***Paraguay***

El estudio de variables sociales a nivel nacional en Paraguay sólo puede realizarse desde 1995, a partir de los microdatos de la Encuesta de Hogares-Mano de Obra (EH-MO, 1995), la Encuesta Integrada de Hogares (EIH, 1997-1998 y 2000-2001) y la Encuesta Permanente de Hogares (EPH, 1999 y anualmente desde 2002). Una perspectiva histórica de la evolución de los indicadores sociales en el país puede computarse solamente para el Área Metropolitana de Asunción desde 1983 con datos de la Encuestas de Hogares-Mano de Obra. La EPH armonizada en el proyecto SEDLAC incluía únicamente información del último trimestre del año hasta 2021, pero a partir de 2022 la EPH Continua (EPHC) comenzó a recoger información de ingresos a lo largo de todo el año, por lo que la

---

<sup>2</sup> Una encuesta específica para analizar las consecuencias del Huracán Mitch se introdujo en 1999.

armonización de la EPH ahora toma como referencia el año completo. La última EPHC disponible en SEDLAC es la correspondiente a 2024.

### ***Perú***

En el pasado, la principal encuesta de hogares de Perú era la ENNIV, perteneciente al grupo de las LSMS, fue llevada a cabo en 1985, 1991, 1994, 1997 y 2000. Por su parte, la ENAHO se realiza en cuatro ondas a partir de 1997 y es la principal encuesta de hogares del país en la actualidad. La cuarta onda de la encuesta es nacionalmente representativa y es la utilizada para estimar las tasas de pobreza oficial. SEDLAC presenta estadísticas computadas con la ENAHO para el período 1997-2024. Después de 2000 se empezó a utilizar un nuevo marco muestral y aumentó el número de observaciones de la encuesta, lo que genera algunos problemas de comparabilidad con encuestas previas. Para enfatizar esos problemas de comparabilidad, en la mayoría de las tablas se presentan separadamente las estadísticas para los períodos 1997-2000, 2001-2003 y 2003-2024.

### ***Uruguay***

La Encuesta Continua de Hogares (ECH) es la principal encuesta de hogares en Uruguay. La ECH es realizada por el Instituto Nacional de Estadística (INE). Se implementó por primera vez en 1968 en Montevideo y por un largo periodo de tiempo cubrió sólo las áreas urbanas. De todas maneras, la muestra de la ECH representaba cerca del 80% del total de la población del país. A partir de 2006 la encuesta empezó a tener cobertura geográfica nacional. En 2021 ocurrió el último cambio metodológico importante en la encuesta, ya que se introdujeron distintas modificaciones cuyo objetivo fue capturar la información con mayor precisión. El principal cambio fue que a partir de ese momento la ECH pasó a ser un panel rotativo con periodicidad mensual, donde cada hogar participa durante 6 meses consecutivos en la encuesta. La ECH 2024 es la última que se armonizó en la base de datos SEDLAC.

### ***Venezuela***

La Encuesta de Hogares por Muestreo (EHM) es la encuesta de hogares de Venezuela. La EHM tiene cobertura nacional de la población no indígena residente en el país. Se lleva a cabo desde 1967, dos veces al año. El período del trabajo de campo es de seis meses. El principal objetivo de esta encuesta es proveer información general sobre las características socioeconómicas y sociodemográficas de la fuerza laboral. La última EHM procesada para este proyecto es la del año 2006.

### Cambios en la cobertura geográfica

Los cambios en la cobertura geográfica de una encuesta de hogares determinada puede afectar severamente la comparabilidad de los resultados dentro de un país dado a lo largo del tiempo. Debido a que las regiones de un país difieren en sus situaciones económicas y sociales, agregar una nueva región en la encuesta generalmente afecta de manera significativa las estadísticas nacionales.

En la mayoría de los países donde hubo cambios en la cobertura geográfica de la encuesta en la última década, usamos toda la información disponible en la encuesta más reciente (la cual incluye áreas no cubiertas en las encuestas previas) y proveemos información para valorar el impacto del cambio en la cobertura. Por ejemplo, en Bolivia la encuesta de hogares era solamente urbana en 1993 y desde 1997 su representatividad es nacional. Entonces, presentamos dos conjuntos de estadísticas para Bolivia 1997: uno para la muestra completa y otro más restringido, incluyendo sólo aquellas áreas urbanas también encuestadas en 1993. En Argentina, por ejemplo, el INDEC ha incorporado ciudades adicionales dentro de la muestra de la EPH durante los 90s. De ese modo, computamos estadísticas para la muestra más grande disponible en las últimas rondas de la encuesta, y también estadísticas para una muestra más pequeña de ciudades que están presentes en la encuesta desde los 80s. Para evitar la multiplicación de información en las tablas, reportamos estadísticas para ambas muestras sólo para un año.

### Encuestas nacionales y urbanas

Las encuestas de hogares en Argentina cubren sólo áreas urbanas. Sin embargo, Argentina es uno de los países más urbanos del mundo, con más del 85% de la población viviendo en ciudades. Cuando reportamos estadísticas *nacionales* incluimos estadísticas de Argentina, aunque dichas estadísticas provengan de encuestas de hogares urbanas.

### Cambios en las encuestas

La mayoría de los países en ALC han experimentado mejoras significativas en sus encuestas de hogares en las últimas décadas. En particular, los mayores cambios se realizaron en muchos países desde mediados de los 90s luego del programa MECOVI. Aunque esos cambios son bienvenidos, ellos suponen problemas significativos para propósitos de comparabilidad dentro de países a través del tiempo. Debido a que este proyecto va a continuar en el futuro, tratamos de explotar toda la información de las nuevas encuestas, aunque esta decisión implique dificultades cuando comparamos con encuestas previas en el país. Por ejemplo, la encuesta de hogares de Honduras sólo incluye preguntas sobre ingresos monetarios e ingresos no laborales a partir de 1996.

Podríamos ignorar el ingreso no laboral para comparar las viejas encuestas con las nuevas, o incluir el ingreso no laboral para tener una medida más precisa de los ingresos, pagando el costo de perder comparabilidad con las encuestas previas. Optamos por la segunda opción, aunque en algunos casos (e.g. indicadores de pobreza y desigualdad en Honduras) incluimos estimaciones usando las dos alternativas.

### **Ponderadores**

Todas las oficinas nacionales de estadística (ONEs) proveen ponderadores individuales y familiares en las bases de datos. Usamos esos ponderadores oficiales para computar todas las estadísticas. La totalidad de las estadísticas presentadas en SEDLAC son calculadas con los correspondientes pesos poblacionales.

## 2. INGRESOS

Esta sección comprende dos archivos: *AÑO\_ActX\_construction\_incomes\_LAC.xls* contiene información sobre los distintos conceptos incluidos en la construcción de las variables de ingresos, mientras que *AÑO\_ActX\_incomes\_LAC.xls* reporta el ingreso per cápita por deciles, áreas y regiones.

El archivo *AÑO\_ActX\_construction\_incomes\_LAC.xls* incluye tres tablas que proveen información sobre los ítems incluidos en (o excluidos de) las variables de ingreso familiar e individual construidas en el proyecto SEDLAC.

### *total*

- Información general sobre la construcción del ingreso individual y familiar.

### *labor*

- Información sobre la construcción del ingreso laboral.

### *non\_labor*

- Información sobre la construcción del ingreso no laboral.

## Códigos

Cada celda en las tablas de este archivo admite varias respuestas posibles:

- (1) *Yes* = El ítem está en la encuesta, tiene valor monetario y es incluido en la definición de ingreso.

Puede suceder que haya encuestas donde no existe una pregunta específica separada para reportar un ítem, pero estamos seguros que la persona encuestada incluyó este en la respuesta de una pregunta más general. Suponga que la encuesta pregunta: ¿recibió usted rentas, pagos de intereses, beneficios o pensiones durante el último mes? En la tabla *non\_labor* colocamos un *YES* en la celda de Pensiones (y también en la de Beneficios y Dividendos, Rentas e Intereses), aunque no podamos separar el monto exacto correspondiente a cada concepto de ingreso.

- (2) *W.S.Q.* (without specific question) = No hay una pregunta específica sobre el ítem, pero suponemos que es incluido dentro de una variable más general.

Ejemplo: La EPH de Argentina tiene una pregunta general sobre ingreso laboral. Esta variable incluye, en principio, ítems como viáticos y comisiones, pero la encuesta no pregunta explícitamente por ellos. Creemos que ellos están incluidos en el ingreso laboral

reportado por el individuo encuestado, pero no estamos completamente seguros. Por lo tanto, en el archivo colocamos *W.S.Q.*

Note la diferencia con la situación previa. Si la EPH preguntara, ¿cuál es su ingreso laboral incluyendo salarios, viáticos y comisiones?, escribiríamos en la celda de viáticos y comisiones *YES*, en lugar de *W.S.Q.*

(3) *Yes, but ignored* = hay una pregunta específica en la encuesta con un valor monetario, pero la ignoramos en nuestra definición de ingreso.

Suponga que la encuesta recoge información sobre el ingreso extraordinario que produjo la venta de algún activo. Dado que ignoramos esta fuente de ingreso en nuestra definición de ingreso familiar, escribimos *Yes, but ignored*.

(4) *Only binary* = existe una pregunta sobre el ítem, pero el cuestionario permite sólo respuestas binarias (sí o no) o tal vez categóricas. No hay un valor monetario para el ítem, de manera que no es incluido en las variables de ingreso.

La EPH de Argentina pregunta si un trabajador ha recibido pagos no monetarios, pero las posibles respuestas son sólo sí o no. Luego, escribimos *Only binary* en las celdas correspondientes en las hojas *total* y *labor*.

(5) *No* = El ítem no es investigado en la encuesta, ni se incluye en una variable más general. Nuestra variable de ingreso familiar no incluye este concepto.

(6) *No, but estimated* = La encuesta o pregunta sobre el ítem, ni es incluido en una variable más general. Sin embargo, estimamos el ítem y lo incluimos en nuestra definición de ingreso familiar.

Ejemplo: La EPH de Argentina no incluye preguntas sobre la renta implícita de la vivienda propia. Sin embargo, por propósitos de comparación con otros países incluimos una estimación aproximada de este concepto en nuestra definición de ingreso familiar (ver abajo).

El archivo *incomes\_LAC.xls* reporta ingreso familiar per capital medio en moneda local corriente. Este contiene dos tablas:

#### *deciles\_pci*

- Ingreso per cápita familiar promedio, en moneda local corriente, por deciles de la distribución del ingreso per cápita familiar. Estadísticas nacionales, urbanas y rurales.

### **Cuestiones Metodológicas**

#### **El hogar**

Durante una onda de una encuesta de hogares un agente de la ONE visita una vivienda. Luego, el agente identifica cada hogar en la vivienda, en el caso de que más de una familia viva en la misma. Cada hogar está compuesto por un jefe y un número variable de miembros. Las ONEs de distintos países difieren en el tratamiento de los miembros “externos” de un hogar: pensionistas, empleados domésticos y sus familias. En este estudio asumimos que esas personas no comparten el mismo presupuesto que los miembros primarios del hogar y, por lo tanto, no son considerados como parte del hogar. Por supuesto, en la práctica puede haber casos donde, por ejemplo, los empleados domésticos son realmente parte del hogar. Sin embargo, como la mayoría de las encuestas de hogares no tienen información del comportamiento al interior del hogar, decidimos excluir los miembros externos del hogar. Un análisis exploratorio para algunos países sugiere que la mayor parte de los resultados no son afectados por esta decisión.

#### **¿Por qué ingreso y no consumo?**

Es aceptado en la literatura económica que el consumo familiar es una mejor medida del bienestar que el ingreso familiar.<sup>3</sup> Tres principales razones justifican esta preferencia: (i) si se pueden pedir y prestar libremente recursos, el consumo corriente es más cercano al ingreso (o consumo) permanente que el ingreso corriente, (ii) el sub-reporte diferencial por estratos es, usualmente, un problema más severo para ingreso que para consumo, y (iii) los ingresos se reportan frecuentemente antes de impuestos, mientras que el consumo es un concepto después de impuestos.

A pesar de esta dominancia, casi todos los estudios distributivos en ALC usan el ingreso como indicador de bienestar. Una simple razón justifica esta práctica: pocos países en la región realizan rutinariamente encuesta de hogares nacionales con cuestionarios de consumo o gasto, mientras que todos los países incluyen preguntas sobre ingreso

---

<sup>3</sup> Ver por ejemplo Deaton y Zaidi (2003).



individual y familiar. La mayoría de los países tienen encuestas de gastos, usadas principalmente para la construcción de los ponderadores de los Índices de Precios al Consumidor. Sin embargo, dichas encuestas son comúnmente realizadas a intervalos prolongados de tiempo (e.g. 10 años), de manera que no son apropiadas para monitorear pobreza, desigualdad y otros indicadores sociales.

En esta etapa del proyecto seguimos la práctica común en la literatura especializada en cuestiones distributivas en ALC y usamos el ingreso como indicador de bienestar. Esperamos incluir gastos en etapas posteriores del proyecto, en la medida que los países empiecen a recoger información de los gastos de los hogares con mayor frecuencia.

### Construcción de variables de ingreso

El concepto principal de ingresos que se utiliza en las estadísticas del proyecto SEDLAC es el **ingreso per cápita familiar**. En cada país esta variable es calculada usando una metodología donde diferentes agregados de ingresos se construyen en un proceso que involucra varias etapas sucesivas.

Inicialmente se definen los ingresos de cada individuo entrevistado en la encuesta. El **ingreso individual** se construye adicionando tanto **ingresos de fuentes laborales** como **ingresos de fuentes no laborales**. Siempre que sea posible, dentro de los ingresos laborales se distinguen los **ingreso por trabajo asalariado**, los **ingreso por actividades por cuenta-propia** y los **ingresos de patrones y empleadores**, siendo estas las tres categorías que conforman el ingreso laboral. Alternativamente, el ingreso laboral también se construye como la suma de los ingresos laborales en la ocupación principal y el ingreso laboral que proviene de ocupaciones secundarias.

En relación al **ingreso individual no laboral**, el mismo se divide en tres grupos: (i) **ingreso por jubilaciones y pensiones**; (ii) **ingresos de capital**; e (iii) **ingresos por transferencias públicas y privadas**. Los países difieren en la cantidad de preguntas dedicadas a capturar los ingresos de capital, intereses, beneficios, rentas y dividendos. Por propósitos de comparabilidad preferimos reunir todas esas preguntas en una sola categoría. El mismo criterio se aplica a transferencias, aunque también construimos una variable que, si es posible, identifica aquellas transferencias hechas por el gobierno y otra variable que captura transferencias claramente asociadas a programas de alivio de la pobreza.

Una vez que se construyen los **ingresos laborales** y los **ingresos no laborales** a nivel individual, los mismos se adicionan para obtener el **ingreso individual total**. Sumando los ingresos individuales de todos los individuos del hogar principal se obtiene el **ingreso total familiar**, concepto al que posteriormente se adiciona la **renta implícita de la vivienda propia** en los casos de los hogares que son dueños de la vivienda en que habitan (ver más abajo por una explicación más detallada). Al **ingreso per cápita familiar** se llega luego de

dividir el **ingreso total familiar que incluye la renta implícita** por el **número de miembros del hogar principal**. Posteriormente, se computan otros conceptos de ingreso familiar per cápita, al aplicar diversos ajustes relacionados a factores demográficos (véase abajo para una discusión sobre escalas de equivalencia).

Un par de ajustes adicionales se realizan sobre el ingreso construido en el proyecto SEDLAC. Si los ingresos corresponden a individuos y hogares entrevistados en diferentes momentos del tiempo, los mismos se deflactan para llevar todos los ingresos a valores de un mismo período. Adicionalmente, y para tener en cuenta las diferencias de precios que existen entre áreas urbanas y rurales, los ingresos familiares de los hogares rurales se incrementan en un 15%.

### Período de reporte

La mayoría de las encuestas de hogares reportan los ingresos obtenidos durante el mes previo a la encuesta. Algunas encuestas también incluyen información sobre los ingresos ganados en los últimos seis meses (por ejemplo, la ENIGH en México). En esos casos, y por cuestiones de comparabilidad, optamos por incluir solamente el ingreso del último mes.

Los ingresos son transformados en ingresos mensuales si los valores no son reportados en base mensual. Sin embargo, surgen algunas ambigüedades ya que la información sobre la frecuencia de los ingresos es a veces escasa. Suponga por ejemplo que un individuo reporta ingresos por \$100 recibidos en base semanal. Si no se dispone de otra información, dos situaciones son compatibles con este reporte: (i) el individuo gana \$100 semanales, de modo que su ingreso mensual es de \$400, o (ii) el individuo trabajó sólo esa semana en el mes y su ingreso mensual es de \$100. Cuando no existe otra información para conocer cuál es la alternativa correcta asumimos que es la primera opción.

El uso del *ingreso corriente* como indicador de bienestar merece algunas observaciones importantes. Si el ingreso es volátil y la gente puede ahorrar o tener acceso a los mercados de créditos, el bienestar individual puede no estar demasiado asociado al ingreso corriente. Cuando usamos el ingreso corriente en lugar del ingreso permanente, la desigualdad es generalmente sobreestimada y las estimaciones de pobreza también están sesgadas. En ese caso, el signo del sesgo depende de la localización de la línea de pobreza en la distribución del ingreso. Suponga una economía con dos individuos: A gana 100 en  $t_1$  y 0 en  $t_2$ , mientras B gana 0 en  $t_1$  y 100 en  $t_2$ . Suponiendo que no existe tasa de descuento, el ingreso “permanente” es 50 para ambos individuos: existe completa igualdad. Sin embargo, si la encuesta solo captura ingreso corriente, las estadísticas de desigualdad mostrarán una sociedad muy desigual en  $t_1$  y  $t_2$ . Si la línea de pobreza es 40 y la gente tiene algún tipo de acceso a los mercados de crédito (puede usar sus ahorros o tomar préstamos de parientes o vecinos durante los meses malos), la pobreza real en esta

economía es 0. Sin embargo, las estadísticas basadas en ingreso corriente reportarán una tasa de incidencia de la pobreza de 50%, es decir, habrá una sobreestimación de la pobreza. En cambio, si en la misma economía la línea de pobreza fuera de 60, la pobreza real sería 100% pero las estadísticas mostrarían una tasa de incidencia de sólo 50%, en este caso existe una subestimación de la pobreza.

Formas de aliviar este problema son usar consumo corriente y computar ingresos promedios a partir de paneles. Explotaremos estas dos alternativas en las próximas etapas de este proyecto. Por ahora, la mayoría de los países de ALC no tienen encuestas de consumo ni paneles, o en los casos que los hay, las encuestas de consumo son realizadas muy esporádicamente y los paneles son muy cortos y en muchos casos los microdatos no están disponibles al público.

### Ingreso no corriente

Debido a que estamos interesados en capturar ingreso corriente, los ítems de ingreso no corriente no son incluidos en nuestra definición de ingreso. Muchas veces ítems extraordinarios son explícitamente incluidos en las encuestas. Por ejemplo, en la EPH de Argentina se pregunta por “bonificaciones *no habituales*” y “pagos de indemnizaciones”. Nuestra definición de ingreso excluye estos conceptos, diferenciándose entonces de la definición oficial, que los incluye.

El mismo criterio lleva a la exclusión del ingreso por venta de bienes y activos como ser autos, casas o animales. Tampoco son incluidos los ingresos provenientes de regalos, seguros de vida, herencias y juegos y apuestas.

### Renta implícita de la vivienda propia

El concepto de ingreso se refiere al flujo de recursos obtenidos como remuneración por el uso de todos los activos que el individuo o el hogar posee. De acuerdo a esa definición, el ingreso debería incluir no sólo los retornos por el uso del capital y el trabajo, sino también cualquier otra renta producida por la posesión de bienes durables, tales como autos y casas. Las familias que viven en sus propias viviendas reciben implícitamente un flujo de ingreso equivalente al valor de mercado del servicio que el uso de esa propiedad representa para ellos. Esta remuneración debería computarse como parte del ingreso familiar, aun cuando nunca se formalice en el mercado.

En algunas encuestas se pide a los propietarios que estimen la renta que ellos tendrían que pagar si tuvieran que alquilar la casa que ocupan. Las respuestas a esta pregunta son utilizadas para imputar la renta de la vivienda propia, aunque existen problemas de confiabilidad en las respuestas, en particular en áreas donde no existen mercados de vivienda bien desarrollados. En aquellas encuestas donde esta información no está

disponible, o es claramente de poca confiabilidad, incrementamos el ingreso familiar de las familias propietarias en un 10%, un valor que es consistente con las estimaciones de renta implícita en la región.

### No respuesta e ingresos *missings*

No todos los individuos seleccionados para responder la encuesta de hogares contestan todas las preguntas. La no respuesta suele ser más usual en las preguntas de ingresos. La existencia de ingresos no declarados puede sesgar las estadísticas si la no respuesta está correlacionada con el ingreso. Aun cuando esto ocurra, las tendencias pueden no verse afectadas si la gente que no responde las preguntas de ingreso comparte características similares.

La forma típica de mitigar este problema es imputar ingresos a quienes no responden con técnicas de *matching* o aplicando los coeficientes de una ecuación de Mincer. Esta ecuación puede ser estimada por OLS, Heckman ML, regresiones por cuantiles u otro método y su especificación depende de la información disponible en la encuesta. La práctica de imputar ingresos tiene dos clases de problemas. En primer lugar, el  $R^2$  de las ecuaciones de Mincer son típicamente bajos: muchos determinantes de los ingresos son inobservables (i.e., talento individual) o no mensurables (i.e., calidad de la educación), y entonces imputar puede implicar errores importantes. Este problema es particularmente grave para los ingresos de capital, que es la fuente con mayores tasas de no respuesta. Las técnicas de *matching* tienen la misma desventaja. El segundo problema se relaciona con la confiabilidad del procedimiento de imputación. Como mencionáramos arriba, varias decisiones deben tomarse para implementar el ajuste. El investigador debe optar por un procedimiento de estimación, elegir la variable dependiente y las independientes, seleccionar un método para imputar el término de error, etc. Aun cuando todos los pasos sean claramente documentados, mucha gente puede tener sospechas sobre la forma en que los datos son tratados y sobre la elección de una estrategia de imputación en particular. Trabajar con los datos originales tiene la ventaja de ofrecer más transparencia.

En esta versión de SEDLAC computamos estadísticas con las bases de datos oficiales, como se hace en la mayoría de los estudios académicos y oficiales.<sup>4</sup>

Supongamos que el ingreso de la fuente  $s$  para el individuo  $i$  no es reportado (es *missing*). ¿Deberíamos considerar como *missing* el ingreso total de dicho individuo? Si optamos por esa alternativa, ¿deberíamos considerar entonces como *missing* el ingreso total de la familia a la que pertenece ese individuo? En este trabajo tomamos las decisiones que se comentan a continuación, las cuales son necesariamente arbitrarias:

---

<sup>4</sup> En algunos países las bases de datos incluyen algunas imputaciones básicas por no respuesta (e.g. Chile y Paraguay).

Si *s* no es la principal fuente de ingresos para *i*, computamos el ingreso total individual ignorando la fuente *s*.<sup>5</sup> Si por el contrario *s* es la principal fuente, consideramos el ingreso total como *missing*. Esta alternativa tiene la ventaja de no excluir de la base de datos a individuos quienes no responden preguntas sobre fuentes de ingreso de importancia secundaria. El costo que pagamos es el de subestimar los ingresos de esos individuos.

Con respecto al ingreso familiar, consideramos a este como *missing* si el ingreso del jefe del hogar es *missing* o, en el caso que el jefe no sea ocupado, si el ingreso del individuo ocupado con mayor nivel educativo es *missing*. En cualquier otro caso, computamos el ingreso familiar asignando ingreso cero a miembros del hogar, distintos del jefe o del ocupado no jefe con mayor nivel educativo, con ingreso *missing*.

### Ingresos cero

En muchas encuestas una fracción no despreciable de la población trabajadora reporta ingresos cero. Esta respuesta puede ser consecuencia de diferentes situaciones: (i) el individuo no recibió ningún ingreso durante el período cubierto por la encuesta, (ii) sólo recibió ingreso no monetario, el cual no es registrado en la encuesta, o (iii) reportó de forma errónea su ingreso.

El ingreso mensual familiar es usado como una medida de bienestar. Como hemos argumentado arriba, uno de los principales problemas de ello surge de la alta volatilidad del ingreso mensual comparado con el bienestar. El caso del ingreso cero es quizás el más ilustrativo de ello. Aunque hay gente que puede tener ingreso familiar cero en un mes en particular, este no puede ser una buena medida de su bienestar, si es que la familia puede obtener ingresos (monetarios o no monetarios) de algunas fuentes (caridad, transferencias, ahorros, etc.). Por esa razón, la existencia de ingresos cero es un caso particularmente importante, ya sea de reportar en forma errónea o de falla en la medida de bienestar.

En esta etapa computamos las tablas aceptando los ingresos cero en las estadísticas de pobreza e ignorándolos en los indicadores de desigualdad, como se hace mayormente en los trabajos académicos.<sup>6</sup> El tratamiento diferencial de los ingresos cero para pobreza y desigualdad surge del supuesto que el ingreso familiar cero proviene mayormente de hogares con todos sus miembros desocupados, y/o de mal reporte de personas con bajos ingresos quienes olvidan o a las que no se les pregunta sobre ciertas fuentes de ingreso (e.g. caridad, pagos en especie). Bajo este supuesto aquellos que responden ingreso cero deberían ser considerados como pobres. Sin embargo, algunas medidas de desigualdad

---

<sup>5</sup> La mayoría de las encuestas de hogares identifican la principal fuente de ingreso individual, aun cuando el individuo no responda las preguntas de ingreso. En aquellos casos en que también carecemos de dicha información suponemos que el ingreso laboral es la principal fuente de ingresos para los adultos.

<sup>6</sup> Ver, por ejemplo, la discusión en Székely y Hilgert (1999).

colapsan cuando consideramos ingresos cero. Los indicadores de desigualdad son invariantes a la escala, de modo que funcionan con diferencias proporcionales de ingreso. Aceptar ingresos cero implica dividir por cero, lo cual genera problemas computacionales. Dado este hecho, y la probable poca confianza en los ingresos familiares cero, las familias que reportan ingresos cero son habitualmente ignoradas cuando se computan indicadores de desigualdad.

Suponga que el ingreso de la fuente  $s$  es reportado como cero por el individuo  $i$ . Si  $s$  no es identificada como la principal fuente de ingreso para  $i$ , computamos el ingreso individual total ignorando la fuente  $s$ , aun cuando exista información que nos lleve a creer que el individuo debería haber tenido algún ingreso de la fuente  $s$  (e.g. horas de trabajo positivas). Si en realidad  $s$  es la principal fuente, consideramos el ingreso total como *missing*. Hay una sola excepción: si el individuo se identifica a sí mismo como “trabajador con ingreso cero”, mantenemos su ingreso como cero. Con respecto al ingreso familiar, consideramos a este como *missing* si el ingreso total que reporta el jefe del hogar es cero o si el jefe no es ocupado si el ingreso total que reporta el individuo ocupado de mayor nivel educativo es cero.

La discusión anterior se aplica al caso de ingresos negativos, los cuales pueden ser consecuencia de beneficios negativos en una actividad económica. Aunque poco frecuentes, algunos granjeros en las encuestas de hogares reportan ingresos negativos. Hasta el momento tratamos ingresos negativos como ingresos *missings*.

### Ingresos incoherentes

Algunas respuestas de ingresos son claramente no confiables. Ello puede ser consecuencia de errores de medición o de declaraciones de ingresos deliberadamente erróneas. Algunas ONEs identifican las respuestas incoherentes dentro de la base de datos basándose en sus conocimientos. Cuando ello ocurre, aceptamos esas decisiones. Como en el caso de los ingresos no declarados, en esta etapa del proyecto también aceptamos las imputaciones de las ONEs para reemplazar respuestas incoherentes.

En algunos casos especiales también excluimos observaciones que parecen claramente no confiables, aun cuando la ONE del país no lo haga. Por ejemplo, en Paraguay 1995 el mayor ingreso reportado por un individuo es juzgado de ser implausible por expertos que han trabajado con los datos de Paraguay (ver Székely y Hilgert, 1999). En esos casos marcamos las observaciones como incoherentes y las ignoramos en el análisis. Esas observaciones son claramente *outliers*, y su número suele ser muy pequeño.

## Sub-reporte

El sub-reporte puede ser consecuencia de una decisión deliberada del encuestado quien prefiere reportar mal sus ingresos, o de la ausencia de preguntas que capturen algunas fuentes de ingreso, o de las dificultades de recordar o estimar el ingreso de ciertas fuentes (ingresos de actividades informales, pagos en especie, ingreso de capital, bienes producidos en el hogar, etc.). Este problema probablemente implique un sesgo negativo en la medición de los estándares de vida de la gente pobre, quienes dependen mayormente de una combinación de actividades informales y/o de producción para el propio consumo, y de la gente rica quienes derivan una gran proporción de sus ingresos de fuentes no laborales y son, además, más propensos al sub-reporte. La conducta desigual de sub-reporte entre encuestados y los esfuerzos diferenciales en el diseño de la encuesta pueden distorsionar las comparaciones entre países. Si estas conductas y esfuerzos cambian a través del tiempo también pueden distorsionarse las tendencias.

Los investigadores aplican tres clases de estrategias para mitigar estos problemas. La primera es restringir el análisis a variables más homogéneas, menos afectadas por el problema de errores en el reporte. Típicamente, la gente mira la distribución del ingreso laboral, o aún más restringida, la distribución de salarios monetarios del trabajo asalariado en áreas urbanas. Por supuesto, el costo de esto es ignorar una parte muchas veces importante de la distribución total de los ingresos.

La segunda estrategia es aplicar algún procedimiento de ajuste. En este caso, el ingreso proveniente de una fuente dada en la encuesta de hogares se ajusta para coincidir con el valor correspondiente en Cuentas Nacionales. Este ajuste usualmente lleva a aumentar los ingresos de capital relativamente más que los provenientes de otras fuentes. Esto se basa en el dudoso supuesto de que los datos provenientes de Cuentas Nacionales están libres de error (Deaton, 2003).

Finalmente, una tercera estrategia es estimar los ingresos sub-reportados a partir de otras piezas de información en la encuesta. Pueden correrse regresiones de Mincer para estimar salarios de trabajadores que evidentemente sub-reportan sus salarios, pero reportan confiables características individuales.

Aquellos investigadores que han hecho diferentes tipos de ajustes han encontrado que, en general, la mayoría de los resultados para las tendencias de desigualdad son robustos a estos ajustes. Gasparini y Sosa Escudero (2001), por ejemplo, encuentran que la tendencia creciente en la desigualdad de los ingresos en Argentina es robusta a los tres tipos de ajustes mencionados arriba. Los resultados entre países son un tanto menos robustos a estos cambios metodológicos. Székely e Hilgert (1999) encuentran que algunos rankings de desigualdad entre países de ALC varían al realizarse un amplio rango de ajustes. Los resultados para pobreza difieren entre países, en algunos casos los ajustes por sub-reporte disminuyen significativamente las estimaciones de pobreza.

Como en el caso de no respuesta, hasta ahora hemos computado estadísticas con los datos originales como en la mayoría de los estudios académicos y oficiales.

Es importante notar que computamos algunas estadísticas para un amplio rango de variables, algunas de las cuales tienen presumiblemente menos problemas de sub-reporte (e.g. ingreso para trabajadores asalariados urbanos formales). Los usuarios pueden restringir las comparaciones a esas variables si están particularmente preocupados por problemas de sub-reporte.

### Ausencia de gente “muy rica”

Una observación común entre usuarios de encuestas de hogares es que estas normalmente no incluyen individuos “muy ricos”: millonarios, grandes terratenientes, poderosos empresarios y grandes capitalistas usualmente no aparecen en las encuestas. Los mayores ingresos en las encuestas de ALC corresponden mayormente a profesionales en áreas urbanas. Este hecho puede ser una consecuencia natural del muestreo aleatorio (hay muy pocos millonarios, por lo que es improbable que ellos sean elegidos para responder la encuesta en un proceso de selección muestral aleatorio), de no respuesta o de alto sub-reporte. El hecho es que los ricos en las encuestas son “profesionales altamente educados obteniendo ingresos laborales, más que capitalistas viviendo de beneficios” (Székely e Hilgert, 1999). La omisión de este grupo no afecta las estimaciones de pobreza pero seguramente implica una subestimación de la desigualdad de una magnitud difícil de precisar. En otras regiones algunos estudios han usado información impositiva para estimar el ingreso de individuos ricos (Piketty y Saez, 2003). Por lo menos en esta primera etapa no planeamos realizar ningún ajuste para aliviar este problema.

### Ingreso real e ingreso nominal

En cualquier análisis distributivo debería utilizarse el ingreso real en lugar del nominal. Sin embargo, si los precios enfrentados por todos los hogares fueran los mismos, esta distinción se vuelve irrelevante. Pero los precios suelen diferir según el lugar de residencia: si dos hogares situados en diferentes regiones tienen el mismo ingreso nominal pero enfrentan diferentes precios, tendrán diferentes niveles de vida. A pesar de que muchos autores han señalado la importancia de considerar las variaciones espaciales de precios en los estudios distributivos (e.g. Deaton (1997), Ravallion y Chen (1997)), rara vez se realizan ajustes de precios en países que no recogen, de manera rutinaria, información sobre precios regionales como parte de la encuesta de hogares. Desafortunadamente, la mayoría de los países de ALC entran en esta categoría. Todos los países tienen algún estudio de precios regionales, lo cual no resuelve completamente este problema, ya que la dispersión de precios puede ser aún alta dentro de una misma región,



especialmente entre áreas urbanas y rurales. Más importante para nuestros fines comparativos es el hecho de que estos estudios difieren sustancialmente entre países en resultados y metodologías utilizadas.

En esta base de datos los ingresos rurales son incrementados un 15% para capturar diferencias entre precios urbanos y rurales. Ese valor es un promedio de algunos de los estudios detallados de precios regionales disponibles en la región. Aunque ciertamente arbitraria, creemos que esta alternativa es mejor que (i) ignorar completamente el problema de precios regionales, o (ii) usar para cada país la información de precios disponible, a pesar de las enormes diferencias en metodologías, alcance y resultados. En etapas futuras del proyecto se realizarán ejercicios de robustez que permitan reevaluar esta decisión metodológica.

Otro problema surge en aquellos países donde la encuesta es llevada a cabo a lo largo de varios meses. Si hay inflación, los ingresos nominales reportados en diferentes meses deberían ser deflactados para hacer que ellos sean comparables. En todos los países donde esto sucede, usamos el índice de precios al consumidor oficial para ajustar los ingresos nominales.

### Escalas de equivalencia

Comúnmente los individuos viven en hogares y comparten gastos. Este hecho implica que el bienestar de un individuo depende de los recursos disponibles en el hogar así como del tamaño, la estructura y la regla usada para compartir recursos dentro del hogar. Probablemente, el indicador más común de bienestar individual es el ingreso per cápita familiar: ingreso total familiar dividido por el número de miembros del hogar. Aunque utilizado ampliamente, esta variable ignora tres hechos relevantes: (i) la existencia de economías de escala en el consumo dentro del hogar, lo que permite, por ejemplo, que una pareja viva con un presupuesto menor al doble del presupuesto que necesitaría una persona viviendo sola, (ii) las diferencias en necesidades entre individuos, básicamente en función de su edad y género (estas diferencias están detrás de los ajustes por *adulto equivalente*), y (iii) la asignación desigual de recursos dentro del hogar.<sup>7</sup> Mientras los puntos (i) y (ii) son considerados en esta etapa del proyecto, dejamos el punto (iii) para una exploración potencial en etapas futuras del proyecto, dada la escasez de datos sobre este tema.

Definimos el ingreso familiar *equivalente* de un individuo como el ingreso total familiar dividido por  $(A + \alpha_1 K_1 + \alpha_2 K_2)^\theta$ , donde  $A$  es el número de adultos,  $K_1$  el número de niños menores a 5 años, y  $K_2$  el número de niños entre 6 y 14 años. Los parámetros  $\alpha$  permiten la ponderación diferencial entre niños y adultos, mientras que  $\theta$  determina el

---

<sup>7</sup> Ver Buhmann, Rainwater, Schmaus y Smeeding (1988) y Deaton (1997) por discusiones en estos puntos.

grado de economías de escala internas al hogar. Deaton y Zaidi (2002) sugieren valores intermedios de los  $\alpha$ 's ( $\alpha_1 = 0.50$  y  $\alpha_2 = 0.75$ ), y un mayor valor de  $\theta$  (0.90) para países similares a los de nuestra región. Tomamos esta sugerencia como caso de referencia, pero también experimentamos con vectores de parámetros alternativos. Aunque probablemente sería más correcto asignar diferentes parámetros a países de ALC en diferentes etapas de desarrollo, preferimos usar la misma escala en todos los países en cada simulación, por razones de transparencia en las comparaciones.

### 3. POBREZA

La primera parte de esta sección reproduce medidas de pobreza a partir de fuentes oficiales de cada país, mientras que la segunda parte reporta nuestras propias estimaciones de pobreza.

#### 3.1. Pobreza usando la metodología CEDLAS

Esta sección incluye estimaciones de pobreza calculadas con la metodología CEDLAS; *i.e.*, usando como indicador de bienestar el ingreso construido siguiendo la metodología descrita en la sección 2. El archivo [AÑO\\_poverty\\_LAC.xls](#) muestra nuestras propias estimaciones de pobreza a nivel país y sus regiones, usando las líneas internacionales de USD diarios.

#### *Estimaciones por país*

El archivo [AÑO\\_poverty\\_LAC.xls](#) tiene varias hojas:

##### *poverty USD 3.00*

- Tasa de pobreza, brecha de pobreza y FGT (2) usando la línea de pobreza de 3.00 dólares diarios. Estadísticas nacionales, urbanas y rurales.

##### *intervals USD 3.00*

- Estimación puntual, error estándar, coeficiente de variación e intervalos de confianza del 95% para la tasa de pobreza usando la línea de pobreza de 3.00 dólares diarios. Estadísticas nacionales, urbanas y rurales. Computados con la técnica de bootstrap con 100 reposiciones.

##### *poverty USD 4.20*

- Tasa de pobreza, brecha de pobreza y FGT (2) usando la línea de pobreza de 4.20 dólares diarios. Estadísticas nacionales, urbanas y rurales.

##### *intervals USD 4.20*

- Estimación puntual, error estándar, coeficiente de variación e intervalos de confianza del 95% para la tasa de pobreza usando la línea de pobreza de 4.20 dólares diarios. Estadísticas nacionales, urbanas y rurales. Computados con la técnica de bootstrap con 100 reposiciones.

##### *poverty USD 8.30*

- Tasa de pobreza, brecha de pobreza y FGT (2) usando la línea de pobreza de 8.30 dólares diarios. Estadísticas nacionales, urbanas y rurales.

### ***intervals USD 8.30***

- Estimación puntual, error estándar, coeficiente de variación e intervalos de confianza del 95% para la tasa de pobreza usando la línea de pobreza de 8.30 dólares diarios. Estadísticas nacionales, urbanas y rurales. Computados con la técnica de bootstrap con 100 reposiciones.

### ***poverty median***

- Medidas de pobreza relativa. Tasa de pobreza, brecha de pobreza y FGT (2), usando la línea de pobreza del 50% de la mediana. Estadísticas nacionales, urbanas y rurales.

### ***intervals median***

- Estimación puntual, error estándar, coeficiente de variación e intervalos de confianza del 95% para la tasa de pobreza usando la línea de pobreza del 50% de la mediana. Estadísticas nacionales, urbanas y rurales. Computados con la técnica de bootstrap con 100 reposiciones.

### ***regions\_3.00***

- Tasa de pobreza. Línea de pobreza de 3.00 dólares diarios. Por región.

### ***regions\_4.20***

- Tasa de pobreza. Línea de pobreza de 4.20 dólares diarios. Por región.

### ***regions\_8.30***

- Tasa de pobreza. Línea de pobreza de 8.30 dólares diarios. Por región.

### ***NBI\_1***

- Tasa de pobreza computada usando un indicador de necesidades básicas insatisfechas (NBI) (ver abajo). Áreas urbanas. Porcentaje de gente que cumple cada uno de los criterios usados para definir la pobreza como NBI. Proporción de todos los criterios NBI cumplidos por los pobres.

### ***NBI\_2***

- Tasa de pobreza computada usando un indicador de NBI por quintiles de ingreso equivalente. Proporción de todos los criterios NBI cumplidos por quintiles. Áreas urbanas.

### NBI\_3

- Tasa de pobreza computada usando una combinación del indicador de NBI e ingreso (ver abajo). Áreas urbanas.

### Cuestiones metodológicas

La mayoría de las estadísticas de pobreza mostradas en SEDLAC son de *pobreza de ingresos*, definida como la imposibilidad de alcanzar un cierto nivel de ingreso mínimo, conocido como línea de pobreza. Aún dentro de este marco limitado, la medición de la pobreza enfrenta dos problemas principales: la identificación y la agregación de los pobres.

### Identificación

El problema de identificación es ciertamente el más complicado de resolver. No existen claros argumentos normativos u objetivos para fijar una línea debajo de la cual todos son pobres y sobre la cual todos son no pobres (Deaton, 1997). A pesar de esta ambigüedad conceptual central, la reducción de la pobreza es un objetivo político deliberado para los gobiernos alrededor del mundo. La comunidad internacional se ha puesto de acuerdo con esta meta y ello se refleja en los Objetivos de Desarrollo del Milenio (reducir la pobreza a la mitad para 2015), los Poverty Reduction Strategy Papers, y los World Bank Poverty Assessments.

Ya que existe una arbitrariedad fundamental en la definición de pobreza, diferentes autores y agencias usan diferentes líneas de pobreza. Hay un amplio rango de estimaciones de pobreza en distintos estudios para cada país. SEDLAC incluye un conjunto de estimaciones de pobreza basadas en (i) líneas de pobreza internacionales que utiliza el Banco Mundial (3.00, 4.20 y 8.30 dólares diarios a PPP 2021), (ii) líneas de pobreza nacionales (extrema y moderada), y (iii) líneas de pobreza relativas (50% del ingreso mediano). Usar un rango de líneas es particularmente relevante dada la arbitrariedad en la definición que mencionáramos anteriormente. Mientras que la medición de pobreza con líneas nacionales toma en cuenta la consideración de que las sociedades difieren en el criterio utilizado para identificar a los pobres, las líneas internacionales son instrumentos inevitables para comparar niveles de pobreza absolutos y tendencias entre países, y proveer números de pobreza a nivel regional y mundial.

### Líneas internacionales

La línea de pobreza de 1 dólar diario a precios PPP es una línea de pobreza internacional cuyo sentido es definir una norma internacional para capturar la incapacidad de satisfacer

las necesidades alimentarias. Esta línea fue propuesta en Ravallion *et al.* (1991) y utilizada por el Banco Mundial (1990). Originalmente, era un valor medido en precios internacionales de 1985 y ajustado a la moneda local usando la paridad del poder de compra (PPP) para tener en cuenta los precios locales. El estándar de un dólar diario fue elegido para ser representativo de las líneas de pobreza nacionales entre países de ingresos bajos. La línea fue recalculada en 1993 a términos PPP en \$1.0763 dólares por día (Chen y Ravallion, 2001). Posteriormente, en 2008, esta línea básica fue fijada en \$1.25 dólares por día a PPP de 2005 (Ravallion, Chen y Sangraula, 2008). En 2015, la línea se volvió a actualizar, fijándose su valor en \$1.90 dólares diarios a PPP de 2011, mientras que en 2022 subió a \$2.15 dólares por día a PPP de 2017. Finalmente, en 2025 la línea se estableció en \$3.00 dólares diarios a PPP de 2021. Los factores PPP de 2021 que se utilizan a partir de 2025 fueron computados en la ronda del International Comparison Program del 2024 (World Bank, 2024).

Además de la línea de \$3.00, la cual parece relativamente baja para la mayoría de los países de ALC que son de ingresos medios, el Banco Mundial utiliza otras 2 líneas de pobreza internacionales, las de \$4.20 y \$8.30 dólares por día. Ambas líneas son incluidas en las estimaciones de pobreza de SEDLAC, ya que se considera que las mismas son más apropiadas para capturar la situación socioeconómica de ALC. Debe tenerse en cuenta que para expresar los valores de las líneas de pobreza en términos mensuales el valor diario es multiplicado por 30.42.

Aunque las líneas en dólares han sido criticadas, su simplicidad y la falta de alternativas razonables y fáciles de implementar han hecho de ellas el principal instrumento en las comparaciones internacionales de pobreza.<sup>8</sup> Por ejemplo, el primero de los Objetivos de Desarrollo Sostenible de las Naciones Unidas – poner fin a la pobreza en todas las formas para 2030 – fue establecido en términos de la línea de pobreza de \$2.15 dólares por día.

Siguiendo a Deaton (2003) y WDI (2004) computamos las líneas de pobreza para 2021 en unidades de moneda local usando el ajuste por PPP, y luego la actualizamos al valor del mes en que se realizó la encuesta usando el índice de precios al consumidor nacional del país. Aplicamos estas líneas a nuestra definición homogénea del ingreso per cápita familiar entre países/años, que incluye todas las fuentes ordinarias de ingreso y la estimación de la renta implícita de la vivienda propia (ver sección 2). Por supuesto, aun cuando sigamos el mismo procedimiento, debido a que las encuestas de hogares difieren entre países podemos obtener variables no estrictamente comparables.

Es ampliamente reconocido que el consumo familiar es una mejor proxy del bienestar que el ingreso familiar.<sup>9</sup> A pesar de esta dominancia, prácticamente todos los estudios

---

<sup>8</sup> Ver Srinivasan (2004), Kakwani (2004) y Ravallion (2004) para una discusión sobre las ventajas y desventajas de la línea de 1 dólar diario.

<sup>9</sup> Ver, por ejemplo, Deaton y Zaidi (2003).

comparativos de pobreza en ALC usan el ingreso como indicador de bienestar. Una simple razón justifica esta práctica: pocos países en la región realizan rutinariamente encuestas de hogares nacionales con cuestionarios de consumo/gasto, mientras que todos incluyen preguntas sobre el ingreso individual y familiar. Algunos autores ajustan el ingreso medio para que coincida con los datos de consumo de cuentas nacionales para estimar pobreza (CEPAL, 2003; Wodon, 2000; WDI, 2002). Sin embargo, no es claro que tal ajuste por consumo incremente la comparabilidad, ya que la confiabilidad de los datos de cuentas nacionales no es necesariamente mayor que la confiabilidad de los datos de las encuestas de hogares. Deaton (2003) argumenta decididamente a favor de la utilización exclusiva de microdatos en el cómputo de la pobreza, ya que los ajustes por cuentas nacionales “tienden a sobreestimar la reducción de la pobreza a través del tiempo, y exageran las diferencias de pobreza entre países”. En esta base de datos no realizamos ningún ajuste para computar la pobreza a partir de las encuestas de hogares. WDI (2003) reporta que las medidas de pobreza basadas en consumo y aquellas basadas en ingreso sin ajuste no difieren significativamente, debido a dos efectos que tienden a cancelarse mutuamente: el ingreso medio es mayor que el consumo medio, mientras que la desigualdad en el ingreso es mayor que la desigualdad en el consumo. Por esta razón, desde 2003 los WDI reportan estadísticas de pobreza computadas sin ajustar el ingreso para que esté de acuerdo con el consumo de cuentas nacionales.

#### *Líneas nacionales*

La mayoría de los países de ALC tienen líneas de pobreza extrema nacionales normalmente basadas en el costo de una canasta básica de alimentos, y líneas de pobreza moderadas computadas a partir de las líneas extremas usando el ratio de Engel/Orshansky para gastos en alimentos. Esta metodología es también usada por CEPAL, que en algunos casos ayuda a los gobiernos a implementar líneas de pobreza nacionales. Si bien existen algunas similitudes, las metodologías usadas para estimar la pobreza nacional difieren sustancialmente entre países. Por ejemplo, algunos países usan gastos como variable de bienestar (e.g. Perú, Guatemala, Nicaragua), aunque la mayoría de los países elige como variable de bienestar alguna variante del ingreso del hogar (e.g. Argentina, Chile, Paraguay, Uruguay). Incluso entre los países que eligen el ingreso familiar como variable para calcular la pobreza existen diferencias significativas. Una de ellas es que algunos países incluyen en el agregado de ingreso la renta implícita de la vivienda propia, mientras que otros no ajustan los ingresos por dicha alternativa.

#### *Líneas relativas*

Algunos países (e.g. aquellos pertenecientes a la Unión Europea) utilizan medidas relativas de pobreza en lugar de medidas absolutas. Según esta visión, ya que las

percepciones sociales de pobreza cambian cuando el país se desarrolla y los niveles de vida crecen, las líneas de pobreza deberían incrementarse junto con el crecimiento económico. Probablemente la línea de pobreza relativa más popular es aquella fijada como el 50% de la mediana de la distribución del ingreso per cápita familiar. Cuando la economía crece esta línea se incrementa, y la pobreza posiblemente crezca más que con líneas de pobreza fijas.

## Agregación

La segunda etapa en la medición de la pobreza –la agregación– es relativamente más simple. En este proyecto para cada línea de pobreza computamos tres de los indicadores de pobreza más usados: la tasa de pobreza, la brecha de pobreza y el FGT (2).<sup>10</sup> En todos los casos calculamos la pobreza basándonos en la distribución de individuos, no de hogares.

## Necesidades Básicas (NBI)

Las medidas de pobreza por ingreso tienen dos importantes limitaciones. Primero, dado que se usa el ingreso *mensual* como indicador de bienestar, algunas personas serán consideradas incorrectamente como pobres si tienen un shock negativo temporario o debido a razones de estacionalidad. Segundo, existen argumentos convincentes como para considerar a la pobreza como un asunto multidimensional.<sup>11</sup> La insuficiencia de ingresos es sólo una de las manifestaciones de un fenómeno más complejo.

Esas dos razones llevan a la búsqueda de otras variables para medir la pobreza. Investigadores, agencias y ONEs han usado diferentes medidas de educación, salud, vivienda, empleo y acceso a servicios sociales para definir la pobreza *estructural*. Dada la práctica corriente en algunos países, y la disponibilidad de información en todos los países de ALC, construimos un indicador de pobreza de acuerdo a las siguientes condiciones:

- (i) Más de cuatro personas por habitación
- (ii) La familia vive en un lugar “pobre” (e.g. en la calle, en una villa)
- (iii) La vivienda es de materiales de baja calidad o precarios
- (iv) La vivienda no tiene acceso a agua
- (v) La vivienda no tiene un baño sanitario

---

<sup>10</sup> Ver Foster, Greer y Thornbecke (1984) para más referencias.

<sup>11</sup> Bourguignon (2003) discute la necesidad y el problema que supone pasar de pobreza de ingresos a un enfoque multidimensional de dotaciones. Attanasio y Székely (eds.) (2001) muestran evidencia de pobreza como falta de ciertos activos para países de LAC.



- (vi) En la familia hay niños de entre siete y once años que no asisten a la escuela
- (vii) El jefe de hogar tiene educación inferior a primaria completa
- (viii) El jefe de hogar tiene educación inferior a secundaria completa y en el hogar hay más de cuatro miembros por perceptor de ingreso

Todas las personas en un hogar son consideradas pobres si ellas cumplen al menos una de las condiciones mencionadas arriba. Este indicador es similar al popular indicador de NBI. También combinamos este enfoque con el de pobreza de ingresos, utilizando simultáneamente el indicador de NBI con la medida de pobreza de 4.20 dólares diarios: sólo si un individuo es pobre bajo ambos criterios es considerado como “crónicamente” pobre.

Restringimos el análisis a áreas urbanas, ya que de otro modo las condiciones para el indicador de NBI deberían diferir entre zonas rurales y urbanas (e.g. acceso a sanidad).

### Variabilidad muestral

Las medidas de las diferentes dimensiones de la distribución están sujetas a problemas de variabilidad muestral, ya que provienen de datos de encuestas, no de datos censales. Supongamos que nuestro tamaño muestral fueran sólo dos individuos, en ese caso los indicadores seguramente variarían ampliamente en el tiempo, aún si la población permaneciera completamente sin cambios, en la medida que seleccionamos aleatoriamente dos individuos diferentes cada año. Este problema puede reducirse pero no eliminarse completamente usando muestras más grandes. En el proyecto tenemos en cuenta este problema y calculamos errores estándar e intervalos de confianza para algunos indicadores de pobreza y desigualdad. Los errores estándar se calculan con técnicas de *bootstrap*, la cual provee estimaciones de intervalos de confianza y medidas de dispersión para los indicadores distributivos en una forma sencilla y eficiente.<sup>12</sup> Para un indicador dado computamos los errores estándar, el coeficiente de variación y el correspondiente intervalo de confianza al 95% de significatividad.

---

<sup>12</sup> La implementación del método de bootstrap seguirá a Sosa Escudero y Gasparini (2000). Para más referencias teóricas en el tema ver Biewen (2002), Davidson y Duclos (2000) y Mills y Zandvakili (1997).

## 4. DESIGUALDAD

Esta sección contiene un archivo ([AÑO\\_inequality\\_LAC.xls](#)) con varias tablas:

### *deciles pci*

- Participación de deciles y ratios de ingresos. Distribución del ingreso per cápita familiar.

### *indices pci*

- Indicadores de desigualdad. Distribución del ingreso per cápita familiar.

### *intervals pci*

- Estimación puntual, error estándar, coeficiente de variación e intervalo de confianza del 95% para el coeficiente de Gini de la distribución del ingreso per cápita familiar. Computado por *bootstrap* con 100 reposiciones.

### *deciles ei*

- Participación de deciles y ratios de ingresos. Distribución del ingreso familiar equivalente.

### *indices ei*

- Indicadores de desigualdad. Distribución del ingreso familiar equivalente.

### *intervals ei*

- Estimación puntual, error estándar, coeficiente de variación e intervalo de confianza del 95% para el coeficiente de Gini de la distribución del ingreso familiar equivalente. Computado por *bootstrap* con 100 reposiciones.

### *deciles lmi*

- Participación de deciles y ratios de ingresos. Distribución del ingreso monetario laboral familiar equivalente en áreas urbanas.

### *indices lmi*

- Indicadores de desigualdad. Distribución del ingreso monetario laboral familiar equivalente en áreas urbanas.

### *intervals lmi*

- Estimación puntual, error estándar, coeficiente de variación e intervalo de confianza del 95% para el coeficiente de Gini de la distribución del ingreso monetario laboral familiar equivalente en áreas urbanas. Computado por *bootstrap* con 100 reposiciones.

### *deciles ni*

- Participación de deciles y ratios de ingresos. Distribución de la variable de ingreso/consumo usada para estimaciones de pobreza con líneas nacionales.

### *indices ni*

- Indicadores de desigualdad. Distribución de la variable de ingreso/consumo usada para estimaciones de pobreza con líneas nacionales.

### *intervals ni*

- Estimación puntual, error estándar, coeficiente de variación e intervalo de confianza del 95% para el coeficiente de Gini de la distribución de la variable de ingreso/consumo usada en las estimaciones de pobreza con líneas nacionales. Computado por bootstrap con 100 reposiciones.

### *gini1*

- Coeficiente de Gini. Distribución entre individuos de diferentes variables de ingreso familiar:
  - Ingreso per cápita.
  - Ingreso equivalente (diferentes escalas).
  - Ingreso total familiar.
  - Ingreso equivalente por grupos etarios.

### *gini2*

- Coeficiente de Gini. Distribución entre individuos de diferentes variables de ingreso familiar:
  - Ingreso per cápita familiar. Áreas urbanas.
  - Ingreso per cápita familiar. Áreas rurales.
  - Ingreso familiar equivalente. Áreas urbanas.
  - Ingreso familiar equivalente. Áreas rurales.
  - Ingreso per cápita familiar. Ingreso de fuentes laborales.
  - Ingreso per cápita familiar. Ingreso monetario.
  - Ingreso per cápita familiar. Ingreso monetario de fuentes laborales.
  - Ingreso per cápita familiar. Ingreso monetario de fuentes laborales, áreas urbanas.

### *gini3*

- Coeficiente de Gini. Distribución entre individuos de diferentes variables de ingreso familiar. Incluyendo y excluyendo los ceros de las estimaciones.

### **polarization**

- Indicadores de bipolarización computados sobre la distribución del ingreso per cápita familiar, ingreso familiar equivalente y la variable de consumo/ingreso utilizada para las estimaciones de pobreza con líneas nacionales. Índices de Esteban, Gradín y Ray y Wolfson.

## **Cuestiones metodológicas**

### **Deciles**

Cada decil contiene al 10% de la población, ordenada de menor a mayor de acuerdo a la variable de ingreso considerada. Los deciles tienen un número igual de individuos (no hogares).

### **Ratios de ingresos**

10/1: ingreso medio del decil 10 / ingreso medio del decil 1

90/10: ingreso medio del percentil 90 / ingreso medio del percentil 10

95/80: ingreso medio del percentil 95 / ingreso medio del percentil 80

95/5: ingreso medio del percentil 95 / ingreso medio del percentil 5

95/50: ingreso medio del percentil 95 / ingreso medio del percentil 50

50/5: ingreso medio del percentil 50 / ingreso medio del percentil 5

### **Índices**

Las tablas muestran estadísticas de los siguientes indicadores de desigualdad (por referencias ver: Lambert, 1993 o Cowell, 1995), para varias medidas de ingresos:

- Coeficiente de Gini
- Índice de Theil
- Coeficiente de variación
- Índice de Atkinson con parámetros 0.5, 1 y 2
- Índice de entropía generalizada con parámetros 0 y 2 (el Índice de Theil es el de entropía con parámetro 1).

## Ingreso equivalente

El ingreso familiar equivalente se obtienen dividiendo el ingreso total familiar por  $(A + \alpha_1 \cdot K_1 + \alpha_2 \cdot K_2)^\theta$ , donde  $A$  es el número de adultos,  $K_1$  el número de niños menores a 5 años y  $K_2$  el número de niños entre 6 y 14 años. Como referencia tomamos el caso en que  $\alpha_1 = 0.50$ ,  $\alpha_2 = 0.75$  y  $\theta = 0.90$ .

En la hoja **gini1** computamos el ingreso equivalente con distintas escalas:

**A:**  $\theta = 0.90$ ,  $\alpha_1 = 0.50$ ,  $\alpha_2 = 0.75$  (caso de referencia)

**B:**  $\theta = 0.75$ ,  $\alpha_1 = 0.50$ ,  $\alpha_2 = 0.75$

**C:**  $\theta = 0.90$ ,  $\alpha_1 = 0.30$ ,  $\alpha_2 = 0.50$

**D:**  $\theta = 0.75$ ,  $\alpha_1 = 0.30$ ,  $\alpha_2 = 0.50$

**E:** escala de Ámsterdam ( $\theta = 1$ , hombres entre 14 y 17 años = 0.98, mujeres mayores de 14 años = 0.90 y niños menores de 14 años = 0.52)

## Polarización

Computamos para cada país/año dos índices de bipolarización: el Índice de Wolfson, el cual corta a la distribución en el ingreso mediano, el Índice DER y el Índice EGR, el cual busca el ingreso óptimo de corte.<sup>13</sup> La tabla **polarization** muestral los resultados para esas medidas de bipolarización para la distribución del ingreso per cápita familiar y la distribución de la variable de ingreso/consumo utilizada para las estimaciones de pobreza con líneas nacionales.

---

<sup>13</sup> EGR se refiere a Esteban, Gradín y Ray (1999).

---

## **BOX 1: Agrupando la población**

Las siguientes secciones muestran estadísticas sobre un amplio rango de variables. Dichas estadísticas son usualmente presentadas dividiendo la población en grupos. A continuación listamos los principales grupos utilizados en dichas tablas:

### *Quintiles de ingreso equivalente*

Todos los individuos son ordenados de acuerdo a su ingreso familiar equivalente (i.e. ingreso total familiar ajustado por factores demográficos – ver el caso de referencia arriba por detalles) y divididos en cinco grupos de igual tamaño (cada uno con 20% de la población). El quintil 1 incluye el 20% más pobre de la población. Sólo individuos con ingreso familiar equivalente válido son incluidos en las tablas.

### *Educación*

Se forman tres grupos de acuerdo a los años de educación formal. Baja = 0 a 8 años, Media = 9 a 13 años, y Alta = más de 13 años.

### *Área*

La población es dividida en urbano y rural, de acuerdo a la clasificación hecha por la ONE e incluida en la base de datos original de la encuesta de hogares.

También presentamos la información agrupada por edad y género. Las definiciones son claras.

---

## 5. DEMOGRAFÍA

Esta sección tiene dos archivos: [AÑO\\_ActX\\_demographics\\_LAC.xls](#) y [AÑO\\_ActX\\_regions\\_migrations\\_LAC.xls](#).

El archivo [AÑO\\_ActX\\_demographics\\_LAC.xls](#) tiene las siguientes hojas:

### *hh size*

- Tamaño del hogar: número de personas por hogar, excluyendo miembros “externos” (pensionistas y empleados domésticos y sus familias).
- Tablas por quintiles de ingreso equivalente y por nivel educativo del jefe de hogar.

### *children*

- Número de niños menores de 12 años por hogar.
- Tablas por quintiles de ingreso de los padres y por nivel educativo del jefe de hogar.
- “Ingreso de los padres” es el ingreso total percibido por el/la jefe/a de hogar y su esposa/o. En la mayoría de los casos (pero no en todos) ellos son los padres de los niños.
- Niños de miembros externos del hogar (pensionistas, empleados domésticos) son excluidos.

### *dependency*

- Tasa de dependencia: tamaño del hogar sobre número de perceptores de ingreso en el hogar.
- Tablas por quintiles de ingreso equivalente y por nivel educativo del jefe de hogar.

### *age structure*

- Proporción de la población total en cada grupo de edad: [0,14], [15,24], [25,40], [41,64] y [65+).

### *age*

- Edad promedio por quintil de ingreso equivalente.

### *mating*

- Indicadores de *assortative mating*: coeficiente de correlación lineal entre parejas en años de educación y horas de trabajo.
- Análisis restringido a jefes de hogar y sus parejas.
- Las últimas dos columnas en la tabla reportan el coeficiente de correlación para el total de horas trabajadas (en todas las actividades). El coeficiente de correlación

en la última columna es computado para todas las parejas en las cuales el total de horas trabajadas es positivo.

El archivo *AÑO\_ActX\_regions\_migrations\_LAC.xls* tiene las siguientes hojas:

**area**

- Porcentaje de la población en áreas urbanas y rurales. La información proviene de datos de encuesta de hogares, no de datos censales. Los porcentajes son obtenidos usando ponderadores (como el resto de las estadísticas en la base de datos). Las definiciones de urbano y rural son las establecidas por las diferentes ONEs.

**regions**

- Porcentaje de la población (cubierta en la encuesta) en cada región. La información proviene de datos de encuesta de hogares, no de datos censales. Los porcentajes son obtenidos usando ponderadores. Las regiones son aquellas usadas por las ONEs.

**migration**

- Porcentaje de migrantes en la población. En la mayoría de los países los migrantes se definen como aquellos individuos que no nacieron en la ciudad en la que viven actualmente. Porcentaje de migrantes recientes y migrantes extranjeros. Caracterización de la población migrante por ingreso, educación y área.



## 6. EDUCACIÓN

Esta sección contiene cuatro archivos con información sobre años de educación, alfabetismo, matriculación en la escuela y movilidad educativa.

### Archivo 1: *AÑO\_ActX\_years\_edu\_LAC.xls*

Incluye varias hojas con información sobre años de educación formal.

#### *structure*

- **Estructura educativa:** porcentaje de adultos (entre 25 y 65 años) en cada grupo de calificación (definidos en términos de años de educación).
- **Grupos de calificación:** Baja = 0 a 8 años de educación formal, Media = 9 a 13 años y Alta = más de 13 años.

Tablas para todos los adultos, hombres, mujeres, trabajadores hombres y trabajadoras mujeres.

#### *years*

- Distribución de los años de educación formal en la población adulta (25-65).

#### *age\_gender*

- Años de educación formal promedio por grupos de edad y género.

#### *income*

- Años de educación formal promedio de la población adulta (25-65) por quintiles de ingreso equivalente.

#### *income\_age*

- Años de educación formal promedio por grupos de edad y quintiles de ingreso equivalente.

#### *area*

- Años de educación formal promedio por área (urbana-rural). Tablas para toda la población, adultos (25-65), hombres adultos y mujeres adultas.

#### *Gini\_years*

- Coeficiente de Gini para la distribución de años de educación formal. Tablas por grupos de edad.

### **Archivo 2:** *AÑO\_ActX\_literacy\_LAC.xls*

Incluye tres hojas. Consideramos a un individuo como analfabeto si el mismo reporta no saber leer o escribir (ver abajo).

#### *age\_gender*

- Tasas de alfabetización por edad y género.

#### *income*

- Tasas de alfabetización por quintiles de ingreso equivalente. Jóvenes (15-24) y adultos (25-65).

#### *area*

- Tasas de alfabetización por área (urbana-rural) y grupos de edad.

### **Archivo 3:** *AÑO\_ActX\_enrollment\_LAC.xls*

Este archivo contiene varias hojas con información sobre tasas brutas y netas de matriculación escolar. Las tasas brutas se refieren al porcentaje de una población dada asistiendo a cualquier nivel educativo. Las tasas netas se refieren al porcentaje de individuos en edad de asistir a un cierto nivel educativo que realmente asisten a ese nivel.

#### *gender*

- Tasas de matriculación bruta por grupos de edad y género.

#### *income*

- Tasas de matriculación bruta por grupos de edad y quintiles de ingreso equivalente.

#### *area*

- Tasas de matriculación bruta por grupos de edad y área (urbana-rural).

#### *primary*

- Tasas de matriculación neta en educación primaria por género, quintiles de ingreso equivalente y área.

#### *secondary*

- Tasas de matriculación neta en educación secundaria por género, quintiles de ingreso equivalente y área.

### *tertiary*

- Tasas de matriculación neta en educación superior por género, quintiles de ingreso equivalente y área.

### *primary completion*

- Tasas de graduación en educación primaria, *i.e.* porcentaje de jóvenes de entre 15 y 24 años con educación primaria completa. Tablas por género, quintiles de ingreso equivalente y área.

### *public*

- Porcentaje de estudiantes en escuelas públicas. Por nivel educativo. Tablas por quintiles de ingreso equivalente y área.

## **Archivo 4: *AÑO\_ActX\_mobility\_LAC.xls***

### *mobility*

- Índice de Movilidad Educativa para adolescentes (entre 13 y 19 años) y jóvenes adultos (entre 20 y 25 años) (ver abajo por detalles).

## ***Cuestiones metodológicas***

### ***Años de educación***

La mayoría de las encuestas reportan años de educación formal. En aquellos casos donde esta variable no existe, es posible aproximarla a partir del máximo nivel educativo alcanzado por una persona y de su edad. Para algunos países la distribución está truncada en su cola superior, ya que las encuestas no preguntan sobre años de educación posteriores a la graduación universitaria (e.g. Argentina).

### ***Grupos educativos***

Dividimos a la población adulta en tres grupos de acuerdo a sus años de educación: Baja (menos de 9 años de educación), Media (entre 9 y 13 años de educación) y Alta (más de 13 años de educación). Dado que los sistemas educativos difieren entre países y en el tiempo, podemos estar incluyendo gente con diferentes niveles educativos (*e.g.* secundario completo e incompleto) dentro de la misma categoría. Consideramos que este es un costo aceptable a pagar a fin de presentar estadísticas simples de niveles educativos en un formato similar para todos los países/años.

## Coeficientes de Gini

Computamos el coeficiente de Gini sobre la distribución de los años de educación, como una medida de desigualdad en educación. El uso de Ginis educacionales se ha incrementado en los últimos años.<sup>14</sup> Sin embargo, no es obvio que la desigualdad en educación debiera ser medida por un índice de diferencias *relativas* más que por uno de diferencias *absolutas* entre individuos.<sup>15</sup>

## Alfabetismo

Prácticamente todas las encuestas en ALC reportan si un individuo sabe leer y escribir. Consideramos a un individuo como analfabeto si reporta no saber leer o escribir (o ambas). En caso de que no exista la pregunta específica sobre alfabetismo en una encuesta, consideramos como analfabeto a toda persona con menos de dos años de educación formal.

## Niveles educativos

Un sistema educativo formal latinoamericano comienza típicamente a los 6 años de edad. Habitualmente comprende 7 años de educación primaria y 5 años de educación secundaria. La duración de la educación superior depende del tipo de institución y de la carrera elegida, variando normalmente entre 3 y 6 años. La mayoría de los países se apartan de este estándar en distintas direcciones. En algunos países la educación formal comienza a los 5 años. En otros la escuela primaria es más larga y la secundaria más corta. En esos casos la escuela primaria usualmente se divide en tres secciones.

Cuando tomamos decisiones sobre la asignación de individuos a niveles educativos preferimos mantener la comparabilidad dentro del mismo país al costo de perder comparabilidad entre países. Si el sistema educativo de un país ha cambiado, hacemos el esfuerzo de utilizar la clasificación educativa presente en la encuesta más reciente disponible.

## Tasas de matriculación

Las tasas de matriculación escolar pueden definirse en formas alternativas. Una definición simple de matriculación sería computar el porcentaje de personas dentro de un grupo etario que están asistiendo a la escuela, sin considerar el nivel educativo al que asisten.

---

<sup>14</sup> Por ejemplo, Thomas, Wang y Fan (2002) calculan coeficientes de Gini sobre la distribución de los años de educación para 140 países en el período 1960-2000.

<sup>15</sup> El coeficiente de Gini, como la mayoría de los índices de desigualdad, es invariante a la escala (ver Lambert, 1993).

Aunque imperfecto, este indicador de matriculación escolar *bruta* provee información útil sobre aquellos que están en el sistema educativo formal. Estadísticas de tasas de matriculación bruta para niños y jóvenes desde 3 a 23 años se presentan en las tablas *gender*, *income* y *area* en [AÑO\\_ActX\\_enrollment\\_LAC.xls](#).

Otra posibilidad es reportar el porcentaje de personas dentro de un cierto grupo etario que asisten al nivel educativo correspondiente a su edad (matriculación *neta*). Esta alternativa enfrenta varios problemas, ya que no hay una única forma de relacionar edades y niveles educativos (e.g. en Argentina los jóvenes de 18 años podrían estar en la universidad o en escuelas técnicas secundarias). Adicionalmente, como fue mencionado arriba, los sistemas educativos difieren entre países, inclusive a veces son diferentes dentro de un mismo país. Por ejemplo, para el período 1996-2010 un joven argentino de 14 años debería haber estado en la escuela secundaria si hubiera vivido en la ciudad de Buenos Aires, pero en la EGB (un programa de 9 años de duración que comienza a la edad de 6) si viviera en la provincia de Buenos Aires. A pesar de esos problemas intentamos computar medidas de matriculación neta en escuelas primarias y secundarias con la información de las encuestas de hogares. Las hojas *primary*, *secondary* y *tertiary* muestran dichas estadísticas.

## Movilidad educacional

Las estadísticas de movilidad educacional son computadas siguiendo la metodología desarrollada en Andersen (2001). La variable dependiente es la brecha educacional, definida como la diferencia entre (i) años de educación que un niño habría completado si hubiese entrado en la escuela a la edad indicada y hubiese avanzado un grado por año, y (ii) los años de educación que completó en realidad. En otras palabras, la brecha educacional mide los años de educación repetidos o fallados. El Índice de Movilidad Educacional es definido como 1 menos la proporción de la varianza de la brecha educacional que es explicada por los antecedentes familiares.<sup>16</sup> En una economía con muy baja movilidad, los antecedentes familiares serían importantes y, entonces, el índice sería cercano a cero.<sup>17</sup>

---

<sup>16</sup> La regresión, que incluye varios controles, se corre para jóvenes que viven con sus padres (ellos son los únicos con información de la educación y el ingreso de sus padres).

<sup>17</sup> Por detalles técnicos ver Andersen (2001).

## 7. EMPLEO

Esta sección contiene tres archivos: [AÑO\\_ActX\\_employment\\_LAC.xls](#), [AÑO\\_ActX\\_wages\\_hours\\_LAC.xls](#) y [AÑO\\_ActX\\_labor\\_benefits\\_LAC.xls](#).

El archivo [AÑO\\_ActX\\_employment\\_LAC.xls](#) tiene las siguientes hojas:

### *labor force*

- Participación en la fuerza laboral por género, edad, educación y área.

### *employment*

- Porcentaje de adultos empleados por género, edad, educación y área.

### *unemployment*

- Tasas de desempleo por género, edad, educación y área.

### *duration*

- Duración del desempleo (en meses) por género, edad, educación y área.

### *change*

- Porcentaje de trabajadores empleados con deseos de cambiar de empleo y/o incrementar sus horas de trabajo por género, edad, educación y área.

### *structure*

- Distribución de trabajadores por género, edad, educación y área.

### *stru\_region*

- Distribución de trabajadores por región.

### *stru\_type*

- Distribución de trabajadores por relación laboral, tipo de firma y categoría laboral.

### *stru\_sector*

- Distribución de trabajadores por sector económico.

### *informal\_1*

- Porcentaje de trabajadores en trabajos informales.  
Definición 1: Informal = trabajadores asalariados en firmas pequeñas, cuentapropistas no profesionales y trabajadores sin ingreso.

### ***informal\_2***

- Porcentaje de asalariados en trabajos informales.  
Definición 2: Ausencia de derechos de seguridad social.

### ***child***

- Porcentaje de niños entre 10 y 14 años que están empleados.

### ***MDG***

- Porcentaje de mujeres en empleos no agrícolas.

El archivo *AÑO\_ActX\_wages\_hours\_LAC.xls* tiene las siguientes hojas:

### ***wage\_1***

- Salarios horarios en la actividad principal en moneda local nominal por género, edad, educación y área.

### ***wage\_2***

- Salarios horarios en la actividad principal en moneda local nominal por tipo de trabajo.

### ***wage\_3***

- Salarios horarios en la actividad principal en moneda local nominal por status de informalidad.

### ***wage\_4***

- Salarios horarios en la actividad principal en moneda local nominal por sector.

### ***hours\_1***

- Horas trabajadas semanales por género, edad, educación y área.

### ***hours\_2***

- Horas trabajadas semanales por tipo de trabajo.

### ***hours\_3***

- Horas trabajadas semanales por status de informalidad.

### ***hours\_4***

- Horas trabajadas semanales por sector.

#### *li\_1*

- Ingreso laboral mensual en moneda local nominal por género, edad, educación y área.

#### *li\_2*

- Ingreso laboral mensual en moneda local nominal por tipo de trabajo.

#### *li\_3*

- Ingreso laboral mensual en moneda local nominal por status de informalidad.

#### *li\_4*

- Ingreso laboral mensual en moneda local nominal por sector.

#### *Shares*

- Distribución de los ingresos laborales por relación laboral

#### *Ginis*

- Coeficientes de Gini para la distribución de salarios e ingresos.

#### *correlation*

- Coeficientes de correlación lineal entre horas de trabajo y salarios horarios.

#### *wagegaps*

- Ratios de salarios horarios por grupos educacionales. Hombres en edad primaria.

#### *Mincer*

- Coeficientes de las *dummies* educacionales en las ecuaciones de Mincer. Hombres y mujeres.

El archivo *AÑO\_ActX\_labor\_benefits\_LAC.xls* tiene las siguientes hojas:

#### *contract*

- Porcentaje de trabajadores asalariados con contratos laborales por género, edad, educación y área.

#### *pensions*

- Porcentaje de trabajadores asalariados con derecho a pensión al retirarse por género, edad, educación y área.



### **health**

- Porcentaje de trabajadores asalariados con derecho a seguro de salud vinculado al empleo por género, edad, educación y área.

### **others**

- Porcentaje de trabajadores asalariados con acceso a aguinaldo y vacaciones.
- Porcentaje de trabajadores asalariados con un empleo permanente
- Porcentaje de trabajadores asalariados en un programa de empleo
- Porcentaje de trabajadores asalariados sindicalizados.

## **Cuestiones metodológicas**

### **Tipo de firma**

Los trabajadores son clasificados en tres grupos de acuerdo a si ellos trabajan en firmas privadas pequeñas, firmas privadas grandes o en el sector público. El último grupo incluye trabajadores en firmas estatales, escuelas, hospitales y otros servicios públicos y en la administración pública.

### **Informalidad**

Existen como mínimo dos conceptos diferentes a los que se refiere el término *informalidad laboral*. La definición “productiva” clasifica como trabajadores informales a aquellos en trabajos de baja productividad, marginales y de baja calificación, mientras que la definición “legal” o de “protección social” subraya la falta de protección laboral y de beneficios de la seguridad social. La definición productiva se relaciona con el tipo de trabajo (e.g. asalariados vs. cuentapropistas, firmas pequeñas vs. firmas grandes), mientras la definición legal se vincula con el cumplimiento de ciertas leyes en la relación laboral, principalmente referidas a la protección laboral.

La implementación empírica de la noción productiva de informalidad ha sido vinculada a (i) el tipo de trabajo (asalariado, cuentapropistas), (ii) el tipo de unidad económica (firma grande o pequeña o sector público), (iii) y la calificación del trabajador. Siguiendo esta práctica dividimos la población trabajadora en 7 grupos:

1. Empresarios (*patrones*)
2. Trabajadores asalariados en firmas privadas grandes
3. Trabajadores asalariados en el sector público
4. Trabajadores asalariados en firmas privadas pequeñas
5. Cuentapropistas calificados

6. Cuentapropistas no calificados

7. Trabajadores con ingreso cero

Para implementar esta clasificación incluimos como *no calificados* a todos los individuos sin un título de educación terciaria o superior y definimos como *pequeñas* a todas las firmas con cinco o menos empleados.<sup>18</sup> Dado que un individuo puede tener más de un trabajo, aplicamos esta definición sólo a la ocupación principal. Implementamos entonces la siguiente definición de informalidad laboral:

**Definición 1 (definición productiva):** *Un individuo es considerado como trabajador informal si pertenece a una de las siguientes categorías: (i) cuentapropistas no calificados, (ii) asalariados en firmas pequeñas, (iii) trabajadores con ingreso cero.*

Una segunda corriente de literatura se concentra en la noción de informalidad “legal” o de “protección social”. Firms informales son aquellas que no cumplen con las normas en términos de contratos laborales, impuestos laborales y regulaciones laborales, de modo que sus trabajadores no tienen derechos a protección laboral o beneficios sociales vinculados al empleo.

El derecho a recibir una pensión al retirarse es el beneficio de seguridad social más comúnmente incluido entre las preguntas de las encuestas de hogares de ALC. Sin embargo, no todos los países tienen preguntas sobre este tema, y entre aquellos que tienen las preguntas difieren (ver tabla abajo). Adicionalmente, en la mayoría de los países las preguntas se realizan sólo a trabajadores asalariados, dejando de lado a cuentapropistas.

Aplicamos entonces la siguiente definición de informalidad legal/protección social:

**Definición 2 (definición legal o de protección social):** *Un trabajador asalariado es informal si no tiene derecho a una pensión vinculada al empleo una vez que se retire.*

### Ecuaciones de Mincer

Corremos regresiones del logaritmo del salario horario en la ocupación principal para adultos de entre 25 y 55 años. Como regresores incluimos dummies educacionales, edad, edad al cuadrado y dummies de área y región. Estimamos el modelo separadamente para hombres y mujeres por el método de Heckman (1979) de máxima verosimilitud. La ecuación de selección incluye adicionalmente matriculación escolar y número de hijos. En la hoja **Mincer** del archivo *AÑO\_ActX\_wages\_hours\_LAC.xls* reportamos los retornos marginales de completar cada nivel educativo.

---

<sup>18</sup> Dadas las diferencias en las encuestas, el punto de corte no es cinco empleados en todos los países. Ver la tabla arriba.

La ecuación de Mincer (1958, 1974) es también informativa sobre dos importantes factores: el rol de variables inobservables y la brecha salarial entre géneros. El término de error en la regresión de Mincer es interpretado usualmente como capturando el efecto en los salarios horarios de factores que son inobservables en las encuestas de hogares, como ser la habilidad natural, los contactos y la ética laboral. Un incremento en la dispersión del término de error puede reflejar un incremento en los retornos a esos factores inobservables en términos de salarios horarios. Mostramos el desvío estándar del término de error de cada ecuación de Mincer.

Los coeficientes en la regresión de Mincer son normalmente diferentes para hombres y mujeres, indicando que reciben diferentes retornos aun cuando tienen las mismas características observables (educación, edad, lugar de residencia). Para investigar un poco más sobre este punto simulamos el salario contrafactual que ganarían los hombres si se les pagara igual que a las mujeres.

## 8. VIVIENDA

El archivo *AÑO\_ActX\_housing\_LAC.xls* tiene las siguientes hojas:

### *housing*

Estadísticas nacionales por quintiles de ingreso equivalente de:

- Porcentaje de propietarios de la vivienda
- Número de habitaciones en cada vivienda
- Número de personas por habitación en un hogar
- Porcentaje de viviendas pobres (*i.e.* localizadas en vecindarios pobres o lugares inconvenientes –ver abajo)
- Porcentaje de viviendas de materiales precarios (ver abajo)

### *housing urban*

- Las mismas estadísticas de arriba pero sólo para áreas urbanas.

### *house age*

- Las mismas estadísticas de arriba por edad del jefe de hogar.

### *house area*

- Las mismas estadísticas de arriba por área (rural y urbana).

### *house edu*

- Las mismas estadísticas de arriba por educación del jefe de hogar.

## *Cuestiones metodológicas*

### *Propiedad de la vivienda*

La vivienda es probablemente el principal activo que posee la mayoría de la gente. Muchas encuestas de hogares en ALC reportan si la vivienda es propiedad de la familia que vive en ella y algunas también reportan el valor de la renta de la vivienda. Las preguntas de las encuestas usualmente diferencian entre propiedad de la vivienda y propiedad del terreno. Construimos una variable que toma el valor de 1 cuando el hogar es propietario tanto del terreno como de la vivienda.

### *Número de habitaciones*

El tamaño de la vivienda es un indicador de riqueza y una posible proxy para pobreza. Aunque el tamaño en m<sup>2</sup> rara vez se reporta en las encuestas, suele haber información sobre el número de habitaciones. Construimos una variable que captura el número de habitaciones en la vivienda de uso exclusivo del hogar. En un reporte que se incluirá en la página web se especificará como se construye esta variable en cada país. En las tablas de SEDLAC mostramos el número de habitaciones, y también esa variable normalizada por el tamaño del hogar.

### **Vivienda pobre**

Esta variable captura si la vivienda está localizada en una villa (“shanty town”) u otro vecindario identificable claramente como pobre, o si el individuo/hogar mora en un lugar inconveniente (e.g. calle). Las definiciones varían ampliamente entre países. Proveemos detalles en un reporte metodológico.

### **Calidad de la vivienda**

Esta variable registra la precariedad de los materiales utilizados en las paredes, pisos y techos de la vivienda. Los países difieren significativamente en los materiales utilizados para construir las viviendas y en la cobertura de esta información en encuestas de hogares. Materiales que pueden ser un claro indicador de pobreza en un país (o región) pueden no relacionarse con la pobreza en otro país. Las comparaciones que se basen en estas variables deben hacerse con mucha precaución y preferiblemente entre años de un mismo país. Estos indicadores pueden ser muy específicos de cada país. El reporte metodológico tiene más información al respecto.

## 9. INFRAESTRUCTURA

El archivo *AÑO\_ActX\_infrastructure\_LAC.xls* tiene las siguientes hojas:

### *infrastructure*

- **Agua:** Porcentaje de hogares con acceso a una fuente de agua en la vivienda o en el terreno.
- **Baño sanitario:** Porcentaje de hogares con baño conectado al sistema de cloacas o a pozo séptico.
- **Cloacas:** porcentaje de hogares con una vivienda conectada al sistema público de cloacas.
- **Electricidad:** Porcentaje de hogares con acceso a electricidad en la casa, de cualquier fuente.
- **Teléfono:** Porcentaje de hogares con un teléfono fijo o celular.

Todas las estadísticas por quintiles de ingreso equivalente y promedios nacionales.

### *infrastructure urban*

- Las mismas estadísticas de arriba pero sólo para áreas urbanas.

### *Infrastructure\_areas*

- Las mismas estadísticas de arriba por área (rural y urbana).

## *Cuestiones metodológicas*

### *Agua*

El fácil acceso a una fuente segura de agua es uno de los indicadores fundamentales de desarrollo. La mayoría de las encuestas en ALC no preguntan sobre agua potable, pero si lo hacen sobre la localización de la fuente de agua. Construimos una variable que toma el valor de 1 si el hogar tiene acceso a una fuente de agua (agua segura si se reporta en la encuesta) en la vivienda o en el terreno.

### *Baño sanitario*

Esta variable toma el valor de 1 si el hogar tiene baño conectado al sistema de cloacas o pozo séptico.

### *Cloacas*

Esta variable es 1 cuando la vivienda está conectada al sistema público de cloacas.

### **Electricidad**

El acceso a electricidad es reportado en la mayoría de las encuestas de hogares. Nuestra variable es igual a 1 si el hogar tiene acceso a electricidad, de cualquier fuente.

### **Teléfono**

El acceso a teléfono es incluido en una variable que toma el valor de 1 si el hogar tiene un teléfono fijo o alguien en el hogar tiene un teléfono celular. La segunda posibilidad, cada vez de mayor importancia, se captura en cuestionarios de consumo de muchas encuestas. Esto introduce una diferencia entre aquellos países que tienen preguntas presupuestarias y capturan gastos en teléfono celular y aquellos países que no incluyen preguntas sobre gasto del hogar.

## Referencias

- Andersen, L. (2001).** Social mobility in Latin America: links with adolescent schooling. *IADB Research Network Working Paper #R-433*.
- Attanasio, O. and Székely, M. (eds.) (2001).** Portrait of the poor. An assets-based approach. IADB.
- Baulch, B. and McCulloch, N. (2000).** Tracking pro-poor growth. ID21 insights 31, Sussex, Institute of Development studies.
- Beccaria, L. and Gluzmann, P. A. (2013).** Medición de los Ingresos y la Pobreza Oficial en América Latina y el Caribe. Working Papers 0148, CEDLAS, Universidad Nacional de La Plata.
- Biewen, M. (2002).** Bootstrap inference for inequality, mobility and poverty measurement. *Journal of Econometrics* 108, 317-342.
- Bourguignon, F. (2003).** From income to endowments: the difficult task of expanding the income poverty paradigm. Delta WP 2003-03.
- Buhmann, B., Rainwater, G., Schmaus, G. and Smeeding, T. (1988).** Equivalence scales, well being, inequality and poverty: sensitivity estimates across ten countries using the Luxembourg Income Study database. *Review of Income and Wealth* 34, 115-142.
- Chen, S. and Ravallion, M (1997).** How did the world's poor fare in the 1990s? *The World Bank Economic Review*.
- Chen, S. and Ravallion., M. (2001).** How did the world's poorest fare in the 1990s? World Bank working paper.
- Cowell, F. (1995).** *Measuring Inequality*. LSE Handbook in Economic Series. Prentice Hall.
- Davidson, R. and Duclos, J.Y (2000).** Statistical inference for stochastic dominance and for the measurement of poverty and inequality. *Econometrica* 68 (6) nov, 1435-1464.
- Deaton, A. (1997).** *The analysis of household surveys. Microeconomic analysis for development policy*. Washington D.C.: The World Bank.
- Deaton, A. (2003).** How to monitor poverty for the Millennium Development Goals. Research Program in Development Studies, Princeton University.
- Deaton, A. and Zaidi, S. (2003).** Guidelines for constructing consumption aggregates for welfare analysis. *LSMS Working Paper* 135.



- Esteban, J., Gradin, C. and Ray, D. (1999).** Extension of a Measure of Polarization, with an Application to the Income Distribution of Five OECD Countries. El Instituto de Estudios Economicos de Galicia Pedro Barrie de la Maza, Papers 24.
- Foster, J., Greer, J. and Thorbecke, E. (1984).** A Class of Decomposable Poverty Measures. *Econometrica* 52(3), 761-766.
- Gasparini L. and Sosa Escudero, W. (2001).** Assessing Aggregate Welfare: Growth and Inequality in Argentina. Cuadernos de Economía, Instituto de Economía. Pontificia Universidad Católica de Chile. 38(113), 49-71.
- Heckman, J. (1979).** Sample selection bias as a specification error. *Econometrica* 47(1), 153-161.
- Kakwani, N. and Pernia, E. (2000).** What is pro-poor growth?. *Asian Development Review* 18, 1-16.
- Kakwani, N., Khandker, S. and Son, H. (2003).** Poverty equivalent growth rate: with applications to Korea and Thailand. Economic Commission for Africa, mimeo.
- Lambert, P. (1993).** *The Distribution and Redistribution of Income: A Mathematical Analysis*. Manchester, UK. Manchester University Press.
- Mills, J. and Zandvakili, S. (1997).** Statistical inference via bootstrapping for measures of inequality. *Journal of Applied Econometrics* 12, 133-150.
- Mincer, J. (1958).** Investment in human capital and personal income distribution. *Journal of Political Economy* 66(4), 281-302.
- Mincer, J. (1974).** *Schooling, experience, and earnings*. New York: NBER Press.
- Piketty, T. and Saez, E. (2003).** Income inequality in the United States, 1913-1998. The Quarterly Journal of Economics, CXVII, 1, February, 1-39.
- Pizzolitto, G. (2005).** Poverty and Inequality in Chile: Methodological Issues and a Literature Review. Working Papers 0020, CEDLAS, Universidad Nacional de La Plata.
- Ravallion, M. and Chen, S. (1997).** What can new survey data tell us about recent changes in distribution and poverty? *The World Bank Economic Review* 11 (2), 357-82.
- Ravallion, M., Datt, G. and van de Walle, D. (1991).** Quantifying absolute poverty in the Developing World. *Review of Income and Wealth* 37, 345-361.
- Ravallion, M. and S. Chen (2003).** Measuring Pro-Poor Growth, *Economic Letters* 78: pp.93-99.
- Ravallion, M. S. Chen, and P. Sangraula (2008).** Dollar a day revisited. World Bank Policy Research Working paper 4620.

- Son, H. (2004).** Pro-poor growth: Asian experience. Mimeo.
- Sosa Escudero, W. and Gasparini, L. (2000).** A note on the statistical significance of changes in inequality. *Económica* XLVI (1). Enero-Junio.
- Székely, M. and Hilgert, M (1999).** What's behind the inequality we measure: an investigation using Latin American data. *IADB Working Paper* # 409.
- Székely, M. Lustig, N., Cumpa, M. and Mejía, J. (2000).** Do we know how much poverty there is?, *IADB Working Paper* # WP-437
- Székely, M. (2001).** The 1990s in Latin America: another decade of persistent inequality, but with somewhat lower poverty. *IADB Working Paper* # 454.
- Thomas, V., Wang, Y. and Fan, X. (2001).** Measuring education inequality - Gini coefficients of education. The World Bank, Policy Research Working Paper Series 2525.
- World Bank (1990).** World Bank's World Development Report 1990: Poverty.
- World Bank (2003).** World Bank's World Development Report 2003: Poverty.
- World Bank (2004).** *Breaking with history? Inequality in Latin America and the Caribbean.*
- Zheng, B. (1993).** An Axiomatic Characterization of the Watts Index. *Economic Letters* 42, 81-86.