

EL PROBLEMA DEL EMPLEO EN UNA SOCIEDAD SIGMA¹

Adolfo Figueroa

Existe una paradoja poco notada en los países de América Latina: la expansión de la educación ha sido notable en las últimas décadas, pero el grado de desigualdad no se ha reducido. Según los datos que publica el Banco Mundial, la tasa neta de matrícula en América Latina aumentó, entre 1980 y 1997, de 85% a 94% para la educación primaria, y de 55% a 66% para la educación secundaria (Banco Mundial 2001, tabla 6, p. 285). Por otro lado, durante el período 1950-1995, América Latina fue la región más desigual del mundo, con un coeficiente de Gini de 0,50 como promedio (comparado con 0,33 de la región del sudeste asiático); en adición, este promedio se mantuvo casi invariable durante todo ese período (Deininger y Squire 1996, cuadro 5; Li, Squire y Zou 1998).

Emerge, entonces, la pregunta: ¿es la educación un sistema igualador? Uno de los principales mecanismos que podrían transformar la mayor educación en mayores ingresos y en una reducción de la desigualdad es el mercado laboral. Por lo tanto, la explicación de la paradoja debería buscarse, de primera intención, en el funcionamiento del mercado laboral.

¹ Este estudio es resultado del semestre de investigación (2005-II) que me otorgó la Dirección Académica de Investigación y del apoyo financiero que recibí de la oficina del Vicerrectorado Académico de la Pontificia Universidad Católica del Perú (proyecto de la Dirección Académica de Investigación 113.0058). Quiero expresar mi más profundo agradecimiento a mi universidad por este apoyo, que, al igual que en otras oportunidades, me ha permitido dedicarme a la investigación básica en la ciencia económica. Quiero también agradecer a Viviana Cruzado por su excelente trabajo de asistente de investigación durante todo el período del estudio. Carlos Loaiza también participó como asistente de investigación en una etapa del estudio y quiero agradecerle por su contribución. También quiero agradecer a mis colegas del proyecto Centre for Research on Inequality, Human Security and Ethnicity (CRISE, Universidad de Oxford) y a Manuel Barrón y José Rodríguez, por sus valiosos comentarios a mis primeros escritos del artículo. Finalmente, quiero agradecer a Claudia Sánchez, quien me ayudó en la preparación de esta versión del artículo haciendo revisiones y recálculos de los datos y de las pruebas estadísticas.

En efecto, según la teoría económica estándar, si la demanda de trabajo se expandiera constantemente (como efecto de la inversión privada y del consecuente crecimiento económico), aumentaría el salario real, pero además el exceso de oferta laboral disminuiría, el nivel de pobreza se reduciría y el grado de desigualdad también. Todo lo que se necesitaría para resolver el problema del empleo en América Latina sería, entonces, crear los incentivos necesarios para atraer la inversión privada. La evidencia empírica de los últimos cincuenta años refuta esta hipótesis: América Latina ha tenido países y períodos de rápido crecimiento económico, y el mercado laboral no ha mostrado, en general, esos resultados. Una razón de esta refutación puede estar en que esta teoría supone implícitamente una oferta laboral compuesta por una masa homogénea de trabajadores.

En este artículo se intenta resolver la paradoja utilizando una teoría en la que se supone que la oferta laboral está compuesta por una masa heterogénea de trabajadores. Se utiliza un modelo de la teoría Sigma para establecer las interrelaciones entre educación, capital humano, salarios e ingresos. En esta relación, el mercado de trabajo jugará un papel central. Sigma es una sociedad abstracta que es capitalista, superpoblada y multiétnica. Las predicciones de la teoría Sigma han mostrado una consistencia aceptable con los datos básicos de la realidad latinoamericana (Figueroa 2003). Aquí se busca contrastar las predicciones de este modelo particular con los datos de la realidad peruana, como una forma de avanzar en la explicación de la paradoja mencionada.

La literatura internacional ha producido muchos estudios empíricos sobre la relación entre educación e ingresos. Pero no ha producido una respuesta satisfactoria a la pregunta planteada líneas antes. Los análisis sobre la relación educación-ingresos para América Latina se han hecho usualmente en el marco de la teoría neoclásica, en la cual se supone que los mercados laborales son walrasianos, al que se le agregan categorías empíricas como los sectores «formal» e «informal» para analizar cuestiones de segmentación en el mercado laboral, pero se hace abstracción del factor étnico (cf. Maloney 2004).

La teoría Sigma es distinta de la teoría neoclásica, pues parte de otro conjunto de supuestos básicos, y por lo tanto, genera predicciones empíricas distintas. Estas dos teorías compiten entre sí, de modo que ambas no pueden ser empíricamente válidas. Por lo tanto, si las pruebas empíricas corroboran una, esto implica la refutación de la otra. Aquí se hará la prueba empírica de las predicciones de un modelo Sigma.

El artículo está organizado de la siguiente manera. La sección 1 presenta la teoría de los mercados laborales en la sociedad Sigma. El proceso de acumulación del capital humano se muestra en las secciones 2 y 3. Las predicciones del modelo teórico están desarrolladas en la sección 4. Estas hipótesis se someterán a la prueba

empírica para el caso peruano. La sección 5 contiene un análisis crítico de la base de datos que se utilizará para someter a la prueba empírica las hipótesis del modelo teórico. La falsación de las hipótesis aparece en la sección 6. La sección final contiene las conclusiones del estudio y también sus consecuencias para el diseño de políticas. El artículo contiene un anexo estadístico que muestra los resultados de las pruebas estadísticas.

1. Un modelo Sigma de los mercados laborales

El mercado laboral en estudio funcionará en una sociedad capitalista particular: la sociedad Sigma. Ese es su marco institucional.

Los supuestos básicos de la teoría Sigma son los siguientes. En esta sociedad, los individuos participan en el proceso económico con una *dotación desigual* tanto de activos económicos como sociales. Los activos económicos incluyen el capital en todas sus formas: físico, humano y financiero. Los activos sociales se refieren a los derechos políticos y culturales, que aquí se reducirán al grado de ciudadanía de las personas. La ciudadanía es un activo social, pues constituye el conjunto de derechos de las personas, incluido el derecho a tener derechos. La presencia de diferentes grados de ciudadanía implica que no todos son iguales ante la ley ni tienen el mismo acceso a los bienes públicos que ofrece el Estado. El supuesto es que esta diferencia está históricamente determinada. Se origina en la dominación de un grupo social por otro, sea como resultado de una conquista, de una colonización o del esclavismo; es decir, como resultado del episodio fundacional de la sociedad. Sigma es, entonces, una sociedad multiétnica y multicultural, en la que diferentes grupos tienen distintos grados de ciudadanía. En Sigma existen no solo clases sociales sino también ciudadanos de distinta categoría.

Para que la teoría del mercado laboral sea refutable empíricamente, se hace necesario introducir supuestos auxiliares para construir un modelo. En Sigma existen tres grupos étnicos: los azules, los rojos y los morados². Estos grupos étnicos están dotados con distintos grados de ciudadanía, de manera jerarquizada. Los rojos son los descendientes de la población aborigen y esclava, quienes constituyen los ciudadanos de última categoría; los azules, el grupo dominante, son los ciudadanos de primera categoría; los morados, que pertenecen al grupo compuesto

² En un artículo en el que analizan la relación entre preferencias y cultura, Akerlof y Kranton (2000) construyen un mundo abstracto en el que habitan dos grupos sociales, los verdes y los rojos, y los verdes constituyen el grupo dominante. Para utilizar colores primarios como metáfora, aquí se consideran dos grupos étnicos, azules y rojos, y se incluye un tercero, el de los morados, quienes constituyen el grupo mestizo de ambas razas. Al igual que en ese artículo, se supone que los individuos no pueden elegir su identidad étnica; esta es exógena.

por los mestizos, son también ciudadanos de primera categoría. Las clases sociales son dos: capitalistas y trabajadores. Los azules concentran la propiedad del *stock* de capital físico de la economía y constituyen la clase capitalista. Los morados y los rojos son los trabajadores, pero los morados están dotados con mayor capital humano que los rojos. Estas son las condiciones iniciales de la economía.

La estructura social de Sigma se compone de grupos étnicos, tipos de ciudadanos y de clases sociales, categorías que están altamente correlacionadas. En este sentido, se puede decir que Sigma es una sociedad altamente correlacionada o jerárquica. Para analizar el papel de las diferencias en las dotaciones iniciales de activos políticos en la reproducción de la desigualdad señalaremos tres grupos sociales que se diferencian por su grado de ciudadanía. Denominaremos *grupo social Z* a los trabajadores que tienen bajas dotaciones de activos políticos; es decir, son ciudadanos de segunda categoría; *grupo social Y* a los trabajadores que son ciudadanos de primera categoría; y *grupo social A* a los capitalistas, quienes son, obviamente, ciudadanos de primera categoría. La estructura social de sigma se puede resumir en la matriz siguiente:

Grupo étnico	Capital físico	Capital humano	Grado de ciudadanía	Nombre del grupo social
Azules	K_b	K_{h1}	C_1	A
Morados	0	K_{h1}	C_1	Y
Rojos	0	K_{h1}	C_0	Z

Esta matriz señala que existe un grado de desigualdad marcada en la dotación inicial de activos entre los individuos, grado de desigualdad que denominaremos mediante el símbolo δ .

Dada esta estructura social, se puede mostrar que la sociedad Sigma opera con mecanismos de inclusión y exclusión a la vez. El mercado y la democracia, las dos instituciones básicas del capitalismo, constituyen mecanismos de integración social pero, al mismo tiempo, también contienen mecanismos de exclusión. La exclusión económica se refiere a que los trabajadores son excluidos (total o parcialmente) de tres *mercados básicos*, que son aquellos que juegan un papel fundamental en la reproducción de la desigualdad. Estos mercados son: laboral, de crédito y de seguros. Aquí estudiaremos el caso del mercado laboral. En la sociedad Sigma existe, además, la exclusión social. Los trabajadores Z son excluidos (total o parcialmente) del acceso a los *bienes públicos básicos*, que son aquellos que juegan un papel fundamental en la reproducción de la desigualdad. Estos son la educación, la salud, la justicia y la protección social. La exclusión puede tomar la forma de discriminaciones o segregaciones contra los grupos

sociales subordinados. Estos supuestos constituyen los fundamentos institucionales de los mercados laborales.

Cuando contratan trabajadores, ¿qué buscan las firmas? Según la teoría estándar de que la motivación que guía las acciones de las firmas es la búsqueda de la ganancia máxima, estas compran en el mercado laboral servicios de trabajo con vistas a ese objetivo. Las firmas compran como servicio laboral el *stock* de conocimientos productivos que se encuentra incorporado en el trabajador. A este *stock* se le llama *capital humano*. La productividad laboral depende, aparte de la tecnología y del *stock* de capital de la firma, del capital humano de los trabajadores: a mayor capital humano, mayor productividad laboral. Por lo tanto, la firma contratará trabajadores hasta que la productividad marginal laboral sea igual a los salarios del mercado, de manera que a salarios más bajos, la firma contratará a mayor cantidad de trabajadores. La curva de demanda de trabajo del mercado laboral es decreciente.

Cada nivel de capital humano implica un mercado laboral diferente. La oferta de trabajo para cada mercado viene de los trabajadores que están dotados con los activos de capital humano en los niveles correspondientes.

¿Cómo se forman los precios y las cantidades en cada mercado? Según la teoría neoclásica, los salarios reales y la cantidad de empleo en cada mercado se forman por la interacción de puras variables reales, resumidas en las curvas demanda y oferta, y donde ambas curvas se crucen. El salario del mercado es aquel que limpia el mercado. Los mercados laborales funcionan, por lo tanto, como mercados walrasianos. Se puede, entonces, estudiar mercados laborales con el método del equilibrio parcial. La existencia y persistencia del desempleo que se observa en los países es un dato empírico que refuta esta teoría.

Según la teoría Sigma, los salarios reales y el empleo no se determinan en cada mercado laboral. Precios relativos y cantidades no se determinan por la interacción de puras variables reales, como productividad y oferta laboral, sino por la interacción de variables reales y variables monetarias. El supuesto es que, como parte de las normas sociales, el salario nominal no puede descender de su nivel histórico; es decir, es inflexible a la baja. Luego, el salario real dependerá del nivel de precios, el cual se determina de manera endógena por la interacción de variables reales y monetarias. En suma, el nivel de precios y el nivel de empleo se determinan simultáneamente por la interacción de los mercados laborales y monetarios, dados los salarios nominales. El funcionamiento del mercado laboral no se puede explicar con el método del equilibrio parcial sino con el método del equilibrio general (Figuroa 2003, cap. 3).

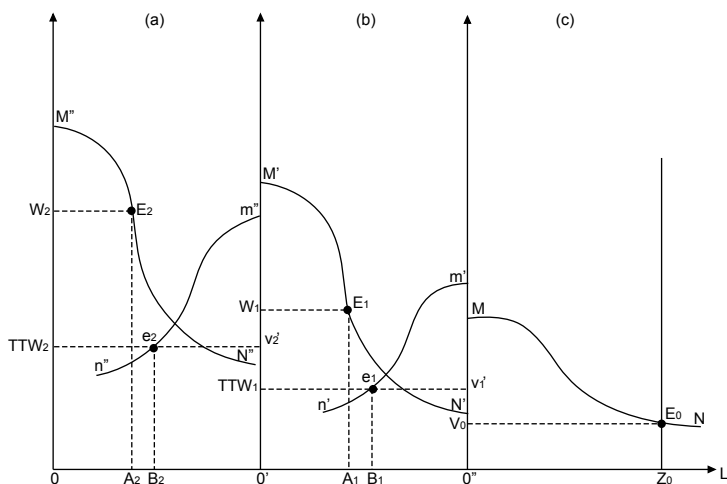
Considérese una economía Sigma con dos mercados laborales que corresponden a dos niveles de capital humano. La figura 1 presenta la situación de

equilibrio en cada uno de los dos mercados laborales. En el panel (a) se presenta el mercado de calificación alta. La cantidad ofrecida está dada y es igual a $00'$, donde la oferta viene de los trabajadores Y. La curva $M''N''$ es la curva de demanda de trabajo. El salario real (w_2) y el nivel de empleo ($0A$) están determinados por las interacciones entre el mercado laboral con otros mercados, incluidos los mercados monetarios. El exceso de oferta resultante es igual al segmento $A20'$. Dada la probabilidad de obtener empleo (π_2), que está exógenamente determinada, el ingreso esperado de los desempleados ($\pi_2 w_2$) queda determinado. La curva $m''n''$ representa la curva de la productividad marginal del trabajo en el sector de autoempleo, medido desde el origen $0'$. La lógica de los trabajadores, quienes buscan la maximización de su ingreso, lleva a que el ingreso esperado del desempleo tenga que ser igual al ingreso marginal del autoempleo (v_2'). Esta igualdad determina la cantidad de autoempleo que es igual al segmento $0'B_2$. Dada la cantidad de exceso de oferta laboral, la cantidad de desempleo también queda determinada (A_2B_2). El tamaño del desempleo se determina de manera secuencial; es decir, el desempleo es residual.

Similar lectura se hace sobre el mercado laboral de calificación baja en el panel (b). La oferta laboral es igual al segmento $0'0''$ y proviene de los trabajadores Z, que pertenecen al grupo llamado Z1. La demanda de trabajo está dada por la curva $M'N'$ y la productividad marginal del trabajo en el sector de autoempleo, por la curva $m'n'$, medida desde el origen $0''$.

Existe otro segmento del grupo social Z, llamado el grupo Z0, con dotaciones demasiado bajas de capital humano para la tecnología que utilizan las firmas. Este

Figura 1. Dos mercados laborales en una economía Sigma



grupo aparece en el panel (c) con un tamaño igual al segmento $0''Z0$. La curva de productividad marginal de este sector de subsistencia está representada por la curva mn. Este grupo está fuera del mercado laboral y solo tiene la opción del autoempleo. Las firmas no tendrán ningún incentivo para invertir en el desarrollo del capital humano (*on the job training*) de este grupo cuando, al mismo tiempo, existe una sobreoferta de trabajadores en los dos mercados laborales.

En este modelo se supone que cada mercado laboral opera como un mercado no walrasiano. Al precio del mercado, existe exceso de oferta laboral. El mercado del trabajo no podría operar con un precio del trabajo que siempre limpie el mercado, pues en este caso los trabajadores no sufrirían ningún costo por su falta de dedicación y esfuerzo en el puesto de trabajo. El mercado del trabajo no puede ser similar al de la papa (como se trata en la teoría neoclásica): mientras la papa no puede modificar su comportamiento según quién sea el consumidor, los trabajadores sí pueden decidir sobre su propio comportamiento, sobre su esfuerzo y productividad, dependiendo de los incentivos que tengan dentro de las firmas.

¿Cuál es, entonces, el incentivo que utilizan las firmas para obtener la productividad laboral deseada en una sociedad Sigma? Dado que existe superpoblación, el mecanismo que asegura la disciplina laboral no podrá ser el desempleo, como sugiere la teoría estándar (cf. Shapiro y Stiglitz 1984), pues sería socialmente inviable tener una economía que operara con 40% o 50% de desempleo. Los trabajadores que forman el exceso de oferta laboral tendrán que generarse su propio ingreso en el sector de autoempleo para hacer viable el sistema. El mecanismo para asegurar la productividad laboral será otro: el salario debe ser superior al costo de oportunidad del trabajador, dado por el ingreso del autoempleo. Las firmas deben pagar una prima a los trabajadores asalariados por encima de su costo de oportunidad. En suma, la predicción del modelo es que el salario medio debe ser mayor que el ingreso medio del autoempleo. En la figura 1, esta relación se expresa por la diferencia que existe entre el salario (w) y el valor de la productividad marginal (v'), lo cual implica que el salario también es mayor que la productividad media (v).

En la teoría Sigma, el mercado laboral es un mercado no walrasiano. El salario real de equilibrio no limpia el mercado laboral; es decir, no es un precio walrasiano. Pero está sujeto a la restricción de que tiene que ser un *salario de eficiencia*. En Sigma, esto significa que el salario tiene que ser mayor que el costo de oportunidad de los asalariados, dado por el ingreso que pueden generarse en el autoempleo. Tiene que existir un premio en el empleo asalariado para que los asalariados tengan el incentivo de desplegar todo su esfuerzo y llegar a la máxima productividad en la firma. Si no lo hacen, perderán su empleo;

y si pierden el empleo, sufrirán un costo económico. Debido a este problema de incentivos para obtener el mayor esfuerzo de los trabajadores, el mercado laboral no podría funcionar como un mercado walrasiano. Las firmas elevarán endógenamente el salario nominal en el caso de que el premio no sea lo suficientemente grande.

El salario de eficiencia implica que cada mercado laboral debe operar con exceso de oferta. Los trabajadores excluidos del empleo asalariado podrán elegir entre buscar empleo asalariado (desempleo) o autoemplearse (subempleo, debido a que el ingreso será inferior al salario del mercado). En el equilibrio, el ingreso esperado del empleo asalariado y el ingreso del autoempleo se igualarán y así se determinará la proporción de desempleo y autoempleo. Existe una elección económica entre desempleo y autoempleo, pero es una elección en una situación de segunda opción, pues la primera es conseguir empleo asalariado. Por lo tanto, existirá un grupo de trabajadores que, con el salario vigente, estaría dispuesto a emplearse como asalariado, pero que no podría lograrlo; es decir, este grupo constituye, por definición, el exceso de oferta laboral, y se compone del desempleo y del subempleo.

Otro supuesto del modelo Sigma de los mercados laborales se refiere al papel de la etnicidad. La demanda de trabajo depende del capital humano del trabajador y no de su origen étnico. Las firmas compran productividad cuando contratan a trabajadores y la productividad se encuentra asociada a la dotación de capital humano de los trabajadores (dado el *stock* de capital físico y la tecnología ahí incorporada). Pero se supone que la dotación inicial de capital humano no es igual entre los trabajadores Y y Z. Como se mostró en la figura 1, el supuesto del modelo sigma es que las firmas contratan a trabajadores de acuerdo con sus dotaciones de capital humano y con los salarios del mercado y no de acuerdo con su etnicidad (la correspondencia entre etnicidad y dotación de capital humano es otro problema).

De la figura 1 se pueden derivar del modelo Sigma las siguientes hipótesis empíricas: (1) en cada mercado laboral, el salario de mercado es mayor que el ingreso medio del autoempleo; (2) en los mercados de mayor capital humano, los salarios serán mayores; es decir, a nivel individual, a mayor dotación de capital humano, mayor salario; (3) a nivel agregado de los trabajadores, a mayor dotación de capital humano, mayor ingreso medio (suma de salarios e ingreso por autoempleo).

Hasta ahora se ha considerado la dotación del capital humano de cada individuo como exógenamente determinada. ¿Por qué existen diferencias en el *stock* de capital humano de los individuos? A resolver esta cuestión teórica se destina la sección siguiente.

2. Acumulación de capital humano en la teoría Sigma

El *stock* de conocimientos productivos que tiene el individuo constituye su *capital humano*. La gente no nace con este *stock* y por eso necesita invertir para adquirirlo y acumularlo y lo hace a través del proceso educativo. Tampoco el *stock* de conocimientos se construye sobre una tabla rasa. Las condiciones iniciales que trae el individuo al proceso educativo en cada nivel son fundamentales para el aprendizaje. Sobre estas dotaciones iniciales, en la literatura de las diferentes ciencias (psicología, biología, neurociencias) se acepta la teoría de que el factor fundamental es la dotación de talentos y que los talentos son múltiples (cf. Gardner 1999).

En esa literatura se acepta, además, la teoría de que si bien las dotaciones iniciales de talentos del individuo, su herencia genética, son exógenas (*nature*), los talentos en el tiempo son endógenos y dependen del medio social (*nurture*), así como de la interacción entre el medio social y la herencia genética de los individuos. Como ha mostrado la literatura moderna de las neurociencias, «[e]l cerebro no es una computadora que simplemente ejecuta programas genéticamente predeterminados. Tampoco es una calabaza pasiva, víctima de las influencias del medio que recaen sobre ella. Genes y ambiente interactúan para modificar continuamente nuestro cerebro, desde el momento en que somos concebidos hasta el momento de nuestra muerte» (Ratey 2002: 17).

En el agregado, los talentos que provienen de la genética se pueden suponer aleatorios, pero no así los que provienen del medio social. La célebre distinción que hizo Rousseau (1755) se refiere justamente a estos factores. Existen, dijo Rousseau, dos tipos de desigualdades entre los individuos: las *naturales*, que son las dotaciones naturales, regalos de la naturaleza, y que son aleatorias; y las *artificiales*, que se originan en el funcionamiento de la sociedad.

2.1 Transformación de educación en capital humano

La teoría del capital humano que se propone aquí tendrá como supuesto básico que los talentos de los individuos son endógenos y que el desarrollo de los talentos depende del nivel socioeconómico de las familias. Las familias ricas construirán talentos superiores en sus hijos en comparación con las familias pobres. La nutrición, la salud y la estimulación intelectual temprana son los principales factores que causan esas diferencias iniciales, las cuales dependen del nivel socioeconómico de las familias.

La nutrición tiene un efecto directo sobre el cerebro y la capacidad cognoscitiva del individuo (Ratey 2002); tiene, además, un efecto indirecto mediante

los episodios de enfermedades. Pero el grado de nutrición depende positivamente del ingreso de las familias.

El estado de salud de los individuos tampoco es neutro con respecto a la desigualdad. Así, en promedio, los episodios de enfermedades serán menos frecuentes en los hijos de familias ricas; además, para su tratamiento, la familia rica podrá asistir a un centro de salud privado, mientras que las familias pobres solo a centros de salud públicos. Pero las diferencias en la calidad del servicio entre los centros de salud privados y públicos son significativas.

Las diferencias en el estado de salud también vienen del lado de la salud ambiental, que tampoco es neutral con respecto a la desigualdad. Las familias ricas pueden evitar los problemas de saneamiento ambiental (calidad de agua, calidad del desagüe y contaminación del aire) mediante la «salida» (*exit*, en el conocido lenguaje *exit-voice* de Albert Hirschman), pues para eso pueden construir barrios residenciales exclusivos. Las familias pobres solo pueden intentar resolver el problema con su «voz» (*voice*): solicitar, protestar, etcétera.

El acceso de la gente a la oferta de los bienes públicos básicos será diferente según el tipo de sociedad de que se trate. En las sociedades tipo Sigma, donde existe la segregación espacial entre los grupos sociales, la salud pública operará no como bienes públicos *universales* sino como bienes públicos *locales*, diferenciados por tipos de ciudadanos. Sea por la provisión privada o sea como bien público, la salud depende de las condiciones socioeconómicas de las familias.

Se supondrá que la estimulación intelectual de los individuos también depende del nivel socioeconómico familiar. La mayor cantidad, calidad y diversidad de bienes y servicios que consumen las familias ricas las induce a descubrir los talentos de sus hijos y a su mayor estimulación intelectual.

Otro factor de desigualdad que se origina en el nivel socioeconómico de las familias es el lenguaje. Existen desigualdades lingüísticas entre individuos. Estas desigualdades se manifiestan en varios aspectos del lenguaje, tales como vocabulario, sintaxis, formas de hablar, capacidades de lectura y escritura. Según la teoría sociolingüística, las desigualdades en el lenguaje se originan más en las experiencias del individuo (su medio social) que en factores genéticos (cf. Hudson 1996: 204).

Las desigualdades en el lenguaje de los niños implican capacidades desiguales de aprendizaje en la escuela. Además, en sociedades heteroglosicas, como es el caso de la sociedad Sigma, las desigualdades iniciales en el lenguaje se mantienen en la edad adulta (la historia cuenta) por efecto de la segregación y la exclusión³.

³ *Heteroglosia* es un concepto que se utiliza en la teoría sociolingüística para referirse a la existencia de varias formas o variaciones de una misma lengua, con una jerarquía entre esas formas, desde las

Se puede, entonces, llegar a decir *Déjame escuchar cómo hablas y te diré quién eres*, pues la manera de hablar del individuo da información sobre su origen social. Y la forma de hablar del individuo puede, entonces, dar señales sobre sus posibles habilidades laborales, positivas o negativas, según la reputación que tenga.

La lengua forma parte del capital humano del individuo adulto. Forma parte del *stock* de conocimientos del individuo y de su competencia laboral. Luego, desigualdades lingüísticas implican desigualdades en capital humano. Una predicción de la teoría sociolingüística sobre las relaciones entre desigualdad social y lingüística es todavía más precisa: «La desigualdad lingüística puede ser vista como una *causa* de la desigualdad social, pero también como una *consecuencia*, debido a que el lenguaje es uno de los medios más importantes a través de los cuales la desigualdad social se reproduce de generación en generación» (Hudson 1996: 205).

Se puede resumir, entonces, la teoría del capital humano propuesta en este estudio: dada la distribución aleatoria de los talentos iniciales (genéticos) en la población, la capacidad de aprendizaje que traen los individuos a la escuela para la acumulación de capital humano depende del medio socioeconómico del cual provienen⁴. La escuela producirá, entonces, diferencias en la acumulación de capital humano entre los hijos de familias ricas y los de familias pobres, aun si las escuelas fuesen homogéneas. A iguales años de educación, los hijos de familias ricas tendrán, en promedio, mayor capital humano que los de las familias pobres.

Hay que analizar ahora los determinantes de la inversión en capital humano. La acumulación de capital humano requiere financiamiento. Los ricos tienen mayor capacidad de financiamiento que los pobres y, por lo tanto, pueden acumular mayor cantidad de capital humano. Luego, el efecto ingreso es positivo: la cantidad demandada de capital humano dependerá positivamente del nivel de ingreso de las familias, el cual llevará a que los individuos que provienen de familias ricas tengan mayor número de años de educación que los que provienen de familias pobres.

Si se abandona el supuesto de que las escuelas son homogéneas, existirá otro factor de diferencia. La escuela privada está más equipada y cuenta con mejores

que se consideran formas superiores hasta las más bajas. «Castellano limeño», «castellano estándar», «castellano costeño», «castellano serrano», «castellano charapa», «castellano motoso» son ejemplos de estas diversas formas del castellano que existen en el Perú. La falta de manejo en el castellano se nota, aparte del acento, en el uso limitado de oraciones, como oraciones en primera persona y verbos en tiempo pasado, presente y futuro, pero no así oraciones que expresen razonamiento abstracto, lo cual requiere el uso de verbos en otros modos, como subjuntivos e impersonales.

⁴ En este estudio, «escuela» se refiere a cualquiera de los niveles educativos: primaria, secundaria, técnica y universitaria.

profesores que la escuela pública. Las escuelas dejan, entonces, de ser homogéneas y más bien se diferencian en calidad y, por lo tanto, en precios. Esta diferencia lleva a que los hijos de los ricos asistan a escuelas privadas y los de los pobres a escuelas públicas. La escuela privada constituye, entonces, un «bien normal» y la escuela pública un «bien inferior». Los hijos de familias ricas tendrán, en promedio, no solo más años de escolaridad sino también un mayor nivel de conocimiento para un nivel de escolaridad dado, comparado con el nivel de los hijos de las familias pobres.

Así como existe una tecnología para producir máquinas, también se puede suponer la existencia de una tecnología para producir capital humano. La cantidad de capital humano que produce una escuela dependerá de los insumos que se pongan en el proceso de producción. Según los supuestos presentados líneas antes, estos insumos son la cantidad y calidad de equipamiento de la escuela, la cantidad y calidad de profesores y la calidad de los estudiantes. La calidad de los estudiantes tiene que ver con sus talentos, los cuales, a su vez, dependen, como se mostró anteriormente, del nivel socioeconómico del cual provienen (dada la distribución inicial de los factores genéticos entre la población, que se supone es aleatoria). Usualmente, se incluye el efecto de las interacciones (externalidades) entre los estudiantes de un grupo social, llamado el «efecto de grupo» (*peer effect*). Ciertamente, dada la segregación que existe en la sociedad Sigma, este efecto será mayor entre los grupos de hijos de familias ricas que entre los de hijos de familias pobres⁵.

La desigualdad tiene un efecto directo sobre la acumulación de capital a través de la calidad del estudiante. Por otro lado, el efecto de la desigualdad también opera mediante la calidad de la escuela. En una sociedad Sigma, los bienes públicos a los que accede la población en cantidad y calidad serán diferenciados y, además, jerarquizados en el orden A-Y-Z. Al igual que en el caso de la salud, el Estado ofrece los bienes públicos en educación bajo la forma de bienes públicos locales.

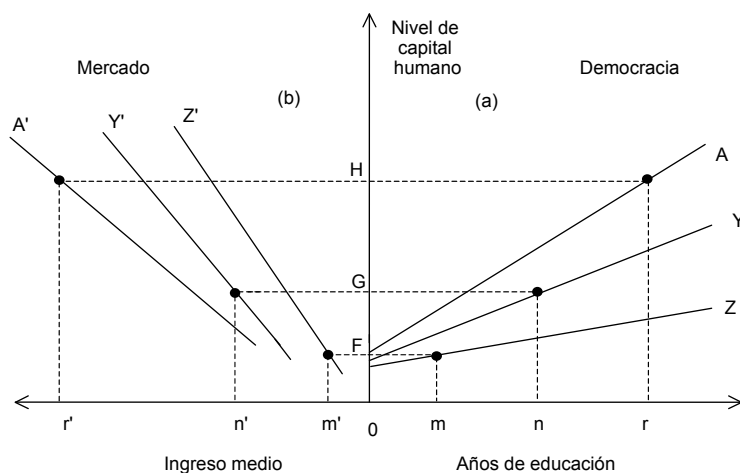
Esta teoría simple de la inversión en capital humano supone, en suma, que educación no es igual que capital humano. En la transformación de la educación en capital humano operan dos tipos de factores: la calidad de la escuela y la calidad del estudiante, y ambos dependen del nivel socioeconómico de las familias. En

⁵ La función de producción del capital humano es un concepto estándar en la teoría neoclásica y se lo presenta como la función de producción de la educación. Usualmente, se supone que el rendimiento educativo depende de dos tipos de variables: las que se refieren a los recursos que utiliza la escuela y las que remiten al origen socioeconómico de las familias. Una revisión reciente de la literatura teórica y empírica referida básicamente al primer mundo se encuentra en Hanushek y Luque (2003) y Todd y Wolpin (2003).

el agregado, entonces, los efectos de estos factores dependen de la desigualdad inicial que exista en la sociedad (la variable δ). En la sociedad Sigma, además, la segregación hace que el grupo Z afronte barreras culturales para acumular capital humano tal como lo requieren las firmas capitalistas (uso del lenguaje, cultura corporativa y otros). A un mismo número de años de educación, el grupo Z obtendrá un nivel más bajo de capital humano que el resto de grupos.

Considérese el modelo simple de la sociedad Sigma construido líneas antes, en el que la estructura social se compone de tres grupos socioeconómicos: A, Y, Z. En esta sociedad existirá una curva de pendiente positiva que relaciona educación (años de escolaridad) con nivel de capital humano. Pero esta relación no será unívoca, sino que tomará formas particulares para cada grupo social. Estas relaciones jerárquicas se muestran en el panel (a) de la figura 2, donde las curvas A, Y, Z corresponden a cada una de las tres poblaciones. Estas relaciones constituyen las ecuaciones estructurales del modelo. Como se muestra en el gráfico, aun si los años de escolaridad fuesen iguales entre los tres grupos, las diferencias en capital humano serían todavía importantes. Esta forma particular de transformación de la educación en capital humano y el orden de las curvas indican también una forma particular de funcionamiento de la democracia: ciudadanos de distinta categoría acceden a los bienes públicos locales de manera diferenciada. La transformación de la educación en capital humano no toma la misma forma para todos los grupos sociales.

Figura 2. Relaciones entre educación, capital humano e ingresos medios



2.2 Transformación de capital humano en ingresos

La transformación de capital humano en ingresos opera a través del sistema del mercado. A mayor capital humano, corresponderá un mayor nivel de ingresos, lo que refleja la rentabilidad positiva de la inversión en capital humano. Esta rentabilidad obedece al efecto positivo que tiene el capital humano sobre la productividad laboral, efecto que opera por dos vías: el de la complementariedad con el capital físico (hace más productivo el uso de la máquina) y el de la adopción de innovaciones tecnológicas (las nuevas máquinas y la nueva tecnología que incorporan requieren un mayor capital humano del trabajador).

Esta transformación también adoptará formas particulares para la sociedad Sigma. A un mismo nivel de capital humano le corresponderá un mayor ingreso medio para los hijos de la clase capitalista (el grupo A) comparado con el de los hijos de los trabajadores (grupos Y, Z). Esta relación se debe a las diferencias en las dotaciones de capital físico y de capital social (redes sociales), lo que implica para el grupo social A un mayor acceso a los mercados básicos, como crédito y seguros, y mayores ventajas económicas por pertenecer a redes sociales de la élite económica. Pero también habrá una diferencia entre los hijos del grupo Y y de Z, pues estos están menos dotados de capital físico y de capital social que aquellos, por lo cual su grado de exclusión de los mercados básicos será mayor.

Esta relación se muestra en la figura 2, panel (b), por las curvas A', Y' y Z'. Estas curvas representan el otro conjunto de ecuaciones estructurales del modelo. A un mismo nivel de capital humano, los individuos que forman el grupo social Z afrontan un mayor grado de exclusión en los mercados básicos (laboral, crédito y de seguros) que los otros grupos. Así, en el mercado laboral los salarios son mayores que el ingreso de autoempleo, y el exceso de oferta laboral es mayor en los trabajadores Z; tal como se muestra en la figura 1, se supone que hay un mayor grado de exclusión laboral en los trabajadores Z que en los trabajadores Y. La falta de acceso al crédito no les permite elevar el retorno a la educación. Finalmente, la falta de acceso al mercado de seguros lleva al trabajador Z a buscar estrategias propias de diversificación de riesgos, asumiendo todos los costos, pues no puede transferir parte de estos costos al mercado de seguros. La consecuencia es que el ingreso medio del grupo Z será menor que el de los otros grupos para cada nivel de capital humano.

3. Transformación de la educación en ingresos

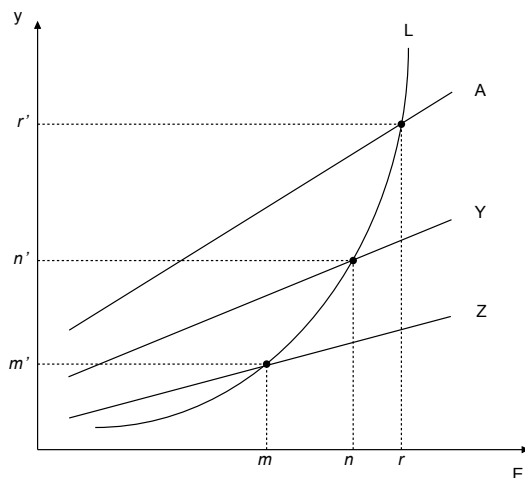
Las ecuaciones estructurales del modelo están representadas en la figura 2. Se necesita ahora derivar la ecuación que muestre la *forma reducida* del modelo. Esta derivación se hará para un modelo estático y otro dinámico.

3.1 Un modelo estático

Consideremos un modelo estático de la teoría Sigma, en el que el nivel de educación de los individuos se determina exógenamente. Como se indica en la figura 2, sean los puntos m , n y r la media de los años de escolaridad de los tres grupos sociales A, Y, Z, y sean los puntos m' , n' y r' los ingresos medios correspondientes. La situación inicial es tal que la clase capitalista tiene el mayor nivel de años de escolaridad, de capital humano y de ingresos, seguida de los otros dos grupos sociales. Ciertamente, más años de escolaridad en el grupo Z se transformarán en mayor capital humano, pero a lo largo de la curva Z; además, este mayor capital humano se transformará en mayor ingreso, pero a lo largo de la curva Z' . Igual proceso se aplica al grupo Y, así como al grupo A.

Si el ingreso depende del capital humano, y este de los años de educación, el ingreso de los individuos dependerá, en última instancia, de sus años de educación, dados los factores que subyacen a las relaciones que aparecen en la figura 2. Así, llegamos a las relaciones que muestran la forma reducida del modelo. Estas relaciones también serán particulares para cada grupo social. Dado que las curvas de la figura 2 son lineales (solo por simplicidad), también las curvas que relacionan años de educación con ingresos serán lineales. Un año adicional de educación aumentará el ingreso en la misma magnitud dentro de cada grupo, pero esta magnitud será mayor para la población A, menor para la población Y y mucho menor para la población Z. Las relaciones que muestran la forma reducida del modelo se muestran en la figura 3.

Figura 3. Hipótesis sobre las relaciones entre educación, ingresos y grupos sociales



Analíticamente, las diferencias de ingresos entre los grupos, como la distancia $m'r'$ entre los grupos Z y A que aparecen en la figura 2 y también en la figura 3, se pueden descomponer en tres efectos. Estos son: la *exclusión cuantitativa de la educación* (menos años de educación), la *exclusión cualitativa de la educación* (menos capital humano acumulado a igualdad de años de educación) y la *exclusión del mercado* (menor ingreso a igualdad de capital humano). De manera similar, se pueden separar estos efectos en las diferencias de ingresos entre los grupos Y y Z.

En el equilibrio inicial de la sociedad Sigma mostrado en la figura 2, se pueden considerar tres casos para analizar el efecto de cambios exógenos de la educación sobre la desigualdad de ingresos futuros. Primero, un año adicional de educación solo en el grupo Z aumentará el ingreso medio del grupo y tendrá el efecto de reducir la desigualdad inicial, pero este efecto será pequeño. Segundo, un año adicional de educación en todos los grupos sociales tendrá un efecto pequeño en la desigualdad; ciertamente, el ingreso medio aumentará en todos los grupos, pero no es claro que los ingresos relativos cambien de manera importante. Tercero, un aumento en años de educación para todos, pero que sea mayor para el grupo Z, tendrá efectos ambiguos y, en todo caso, pequeños.

En suma, el modelo estático predice que una reducción en la desigualdad en los años de educación no implicará una reducción importante en la desigualdad de los ingresos. La razón es que la diferencia en años de educación constituye solo uno de los tres efectos que operan en la generación de diferencias en los ingresos de la gente.

En el caso particular de la sociedad Sigma, por lo tanto, la reducción de la desigualdad en años de educación no será un mecanismo suficiente para aminorar de manera significativa y cierta la desigualdad inicial de ingresos. Otra sería la conclusión si las relaciones entre años de educación e ingresos fueran inseparables; es decir, si existiera una sola curva para todos los grupos, como la curva L de la figura 3. Si las tres curvas se redujeran a una, y a una que fuese lineal (supóngase que la curva L de la figura 3 fuese lineal), un año adicional de educación aumentaría el ingreso en la misma magnitud en los tres grupos sociales; por lo tanto, si las diferencias en años de educación se redujeran, la desigualdad de ingresos sería menor. Si la curva L mostrara rendimientos decrecientes, la desigualdad tendería a reducirse aún más (los ingresos relativos de los más pobres aumentarían); y si mostrara rendimientos crecientes (como muestra la línea L), el cambio en la desigualdad sería ambiguo. La reducción en la desigualdad educativa tendría que ser muy alta para que pueda compensar con holgura las diferencias en las tasas de retorno. El modelo desarrollado aquí predice, precisamente, una curva L de rendimientos crecientes.

Las relaciones que se presentan en la figura 3 se refieren, por lo tanto, a los valores medios de las variables analizadas por grupos sociales (no por individuos). Se pueden dar casos de éxito en el grupo Z, pero la predicción de la teoría es que estos casos serán la excepción y no la regla. La figura 3 muestra la forma reducida del modelo, mientras que las ecuaciones estructurales están representadas por las relaciones que se establecieron en la figura 2.

3.2 Discriminación versus exclusión en el mercado laboral

El modelo Sigma también predice que en el mercado laboral las firmas no pagarán iguales salarios por igual número de años de educación. Las firmas, que actúan con la motivación de la ganancia máxima, no comprarán educación sino capital humano, que es el factor que influye en la productividad y la ganancia. Por lo tanto, la diferencia salarial por años de educación que se observa empíricamente no constituye un caso de discriminación de salarios, como erróneamente se sostiene en la literatura estándar.

La discriminación de salarios propiamente dicha ocurriría si las firmas pagaran salarios diferentes a igualdad de capital humano de los trabajadores. El origen de esta discriminación podría provenir de factores tales como las preferencias de los consumidores, que estarían sesgadas en contra de los bienes producidos por trabajadores Z. La demanda derivada de trabajadores Z sería diferente de la de los trabajadores Y. También se podría originar en el problema de información incompleta que rige en el mercado laboral. Los capitalistas podrían tener poca confianza en los trabajadores Z debido a diferencias culturales y prejuicios étnicos. En este caso, los costos de transacción de emplear a trabajadores Z serían mayores en relación con los trabajadores Y. Como consecuencia, se generaría una segmentación en el mercado laboral: a igualdad de niveles de capital humano, los trabajadores Z obtendrán en el mercado laboral un salario inferior que el de los trabajadores Y. Este efecto está incluido implícitamente en la ecuación estructural que se encuentra representada en la figura 2(b).

Las diferencias salariales entre los grupos de trabajadores Y y Z pueden descomponerse analíticamente en tres efectos: la *exclusión cuantitativa de la educación* (menos años de educación), la *exclusión cualitativa de la educación* (menor acumulación de capital humano a igualdad de años de educación) y la *discriminación de precios en el mercado laboral* (menor salario a igualdad de capital humano). El segmento m'n' de la figura 2(b) se puede descomponer en estos efectos si el eje horizontal se lee como salarios en lugar de ingresos.

¿Cuál de estos efectos tiene mayor importancia en explicar los bajos ingresos relativos del grupo Z? En el contexto general de las relaciones que establece la

teoría Sigma, su predicción es que los bajos ingresos relativos de los trabajadores Z se deben principalmente a la exclusión antes que a la discriminación. La exclusión es el factor esencial y general, mientras que la discriminación se refiere a casos especiales. La diferencia de ingresos entre el grupo Z y el grupo Y no se debe tanto a que los ingenieros del grupo Z reciban un salario inferior que el de los ingenieros del grupo Y, sino a que la proporción de ingenieros dentro del grupo Z es baja comparada con la del grupo Y; además, la proporción de ingenieros con un nivel alto de capital humano es menor en el grupo Z que en el grupo Y. La figura 1 muestra el caso extremo de exclusión: los trabajadores Y están dotados con un mayor *stock* de capital que los trabajadores Z.

3.3 *Un modelo dinámico*

El supuesto de que la cantidad de educación está exógenamente determinada será abandonado ahora. Considérese una situación inicial de equilibrio general de la economía Sigma, donde el ingreso nacional y su distribución están determinados. Con esos valores dados, el número de años de educación de cada grupo social quedará también determinado. La cantidad de educación de los individuos será endógena. Por lo tanto, los valores del capital humano y el ingreso medio de la siguiente generación también quedarán establecidos (puntos m, n, r y puntos m', n', r' en la figura 2).

Aunque la educación se ofrezca como un bien público, existe un componente de inversión privada en la acumulación de capital humano. El ingreso de los padres determina, por eso, el capital humano de los hijos; por lo tanto, el ingreso de los padres determina el ingreso de los hijos, pero dentro de cada grupo social A, Y, Z. Por lo tanto, los padres más ricos y con mayor posición social (grupo A) transmitirán a sus hijos ingresos superiores en comparación con los que transmitirán los padres pobres y de muy baja posición social (grupo Z) por cada año adicional de educación. Pero los ricos invertirán más en capital humano por año de educación y sobre condiciones iniciales para el aprendizaje de sus hijos que son relativamente superiores. Este resultado proviene del mayor ahorro, así como del mayor acceso al mercado de crédito que tienen los ricos. En suma, los grupos sociales acumularán capital humano a lo largo de sus respectivas líneas de aprendizaje, las que son jerárquicas, señaladas en la figura 2.

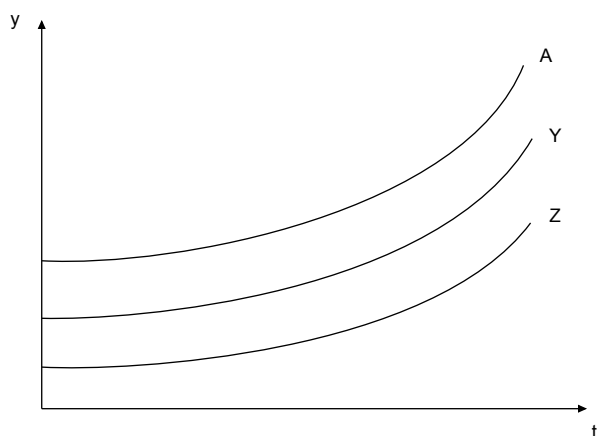
Como resultado, aun si los hijos del grupo Z aumentaran sus años de educación en mayor proporción que los otros grupos, debido a las políticas de expansión de la educación de los gobiernos, los ingresos relativos no cambiarían en una dirección definida en el futuro. Por lo tanto, el equilibrio dinámico del ingreso entre generaciones será tal que tenderá a mantener las desigualdades iniciales de

ingresos. El equilibrio dinámico aparece en la figura 4. La trayectoria de ingresos medios de los tres grupos es creciente en el tiempo, pero no hay tendencia a la convergencia. En el largo plazo, las trayectorias de ingresos de los grupos sociales dependerán de las condiciones iniciales de las que partieron. Este modelo dinámico predice que los hijos heredarán la posición relativa de los padres tanto en cuanto a ingreso como a posición social; es decir, la historia cuenta.

La teoría del capital humano presentada aquí muestra que la desigualdad en los ingresos (flujos) depende esencialmente de las diferencias individuales en las dotaciones iniciales en los activos económicos y sociales (*stocks*). Los capitalistas cuentan con una dotación de activos tal que tienen un mayor *stock* de todos los tipos de capital. Si los trabajadores pudieran igualar a los capitalistas en el número de años de escolaridad, no podrían llegar a acumular el mismo nivel de capital humano, pues los insumos que ponen a la producción del capital humano son diferentes; y aun si la igualdad en capital humano se diera, no podrían igualarlos en el nivel de ingresos, porque no podrían acumular capital físico ni capital social para estar en igualdad de condiciones en el mercado con los capitalistas. El funcionamiento de los mercados básicos, que operan con exclusiones, impide tal resultado.

En suma, el modelo Sigma predice que, dada la pronunciada desigualdad inicial en la dotación de activos entre los grupos sociales, la educación no constituye un mecanismo que conduzca a la reducción significativa de esa desigualdad. La educación no es un sistema igualador. Se puede dar la movilidad educativa pero no la movilidad económico-social. Los trabajadores Z pueden llegar a tener

Figura 4. Hipótesis sobre las trayectorias no convergentes de ingresos por grupos sociales



mayor nivel en años de educación, pero no pueden llegar a igualar la posición de los trabajadores Y, mucho menos la de los capitalistas. En términos intergeneracionales, los «hijos» de un grupo social tenderán a heredar el estatus económico de los «padres» de ese mismo grupo social. La teoría predice que la tendencia a la igualdad intergeneracional entre hijos y padres biológicos, que enunció el biólogo Galton como la «ley de la regresión hacia la media» para el caso de las tallas, no se dará en el caso de los ingresos de los distintos grupos sociales de una sociedad Sigma.

4. Hipótesis empíricas

Las predicciones empíricas que se derivan del modelo Sigma son las siguientes:

H1. Hipótesis de la exclusión

Existe una correspondencia entre el *stock* de capital y el grado de ciudadanía de los grupos étnicos. Debido a que el capital humano depende de los años de educación, las dotaciones diferenciadas en capital humano también se pueden medir en años de educación. Debido a las diferencias en la calidad de la educación entre A-Y-Z, las brechas en años de educación subestiman las verdaderas diferencias en capital humano.

H2. Hipótesis de la separabilidad y jerarquía en las relaciones entre años de educación e ingresos totales

Existe una relación positiva entre años de educación e ingresos. Esta relación es separable para cada grupo social y es jerarquizada por grupos étnicos, en el orden A-Y-Z (tal como se indica en la figura 3)⁶.

⁶ Esta hipótesis se deriva de las siguientes ecuaciones:

(1) $Kh_j = u(E_j, X)$, $j \in X$, $X = Z, Y, A$, $u_i > 0$

(2) $Y_j = f(Kh_j, X) = F(E_j, X)$, $j \in X$, $X = Z, Y, A$, $F_i > 0$

La primera ecuación muestra que el capital humano del *individuo j* depende positivamente de sus años de educación y del grupo étnico X al que pertenece. La segunda ecuación muestra que el ingreso del *individuo j* depende positivamente de su capital humano y del grupo étnico X al que pertenece. La función F es la forma reducida del modelo y constituye H2.

H3. Hipótesis de la separabilidad y jerarquía en las relaciones entre años de educación e ingresos en el mercado laboral

Para cada mercado laboral, existe una relación positiva entre años de educación y precios del trabajo (sueldos o salarios). La relación es separable por grupos étnicos. A igualdad de años de educación, los sueldos y salarios del mercado son jerárquicos por grupos étnicos, en el orden A-Y-Z (tal como se indica en la figura 3, si en el eje vertical se miden salarios). Las diferencias individuales de sueldos o salarios observadas en el mercado laboral no implican discriminación de precios. Estas diferencias reflejan, más bien, el hecho de que a igualdad de años de educación no les corresponden iguales cantidades de capital humano a todos los grupos sociales. (A igualdad de capital humano le corresponderán precios del trabajo más uniformes).

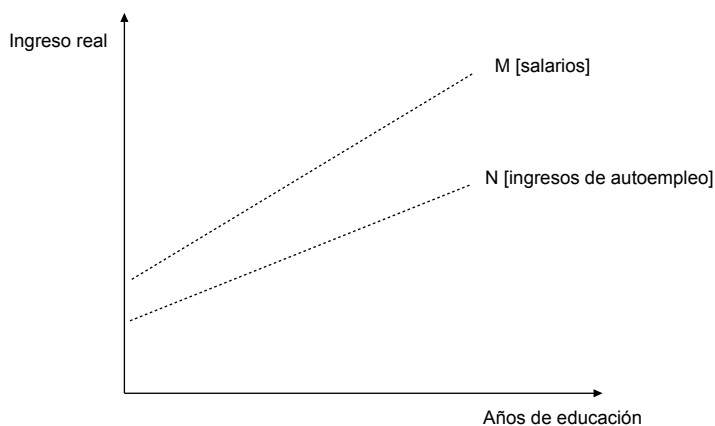
H4. Hipótesis del salario de eficiencia

Para un nivel educativo dado, el precio del trabajo es mayor que el ingreso medio de los trabajadores autoempleados en cada mercado laboral. Esta relación se muestra en la figura 5.

H5. Hipótesis de la exclusión cuantitativa en el mercado laboral

El exceso de oferta laboral depende del mercado laboral: es mayor en los mercados de baja calificación (donde predominan los trabajadores Z) que en los de

Figura 5. Hipótesis sobre la relación entre salarios e ingresos de autoempleo por años de educación para un grupo social dado (A, Y o Z)



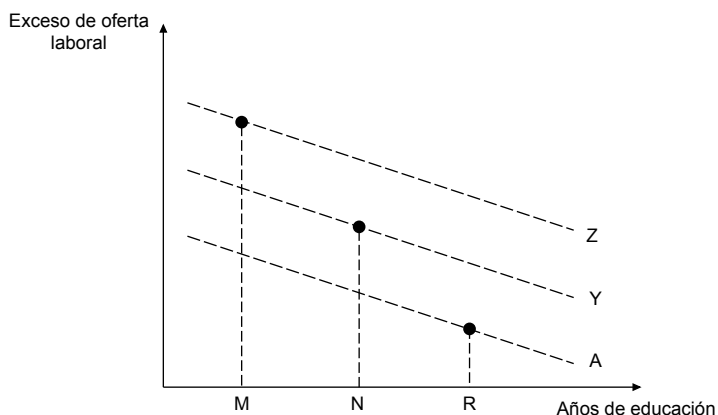
alta calificación (donde predominan los trabajadores Y), para años de educación dados. Esta relación se muestra en la figura 6.

H6. Hipótesis de la persistencia intergeneracional en la desigualdad

En términos de años de educación, existe una tendencia a la convergencia entre grupos sociales. Las diferencias de años de educación serán menores entre generaciones más jóvenes que entre generaciones más viejas. En términos de ingresos medios, la tendencia hacia la convergencia no es significativa, como se indicó en la figura 4. Dadas las desigualdades iniciales, el grado de desigualdad δ , las razones de la falta de convergencia en los ingresos son dos:

- a) Sobre el proceso de transformación de educación en capital humano, la mayor igualdad en años de educación no implica mayor igualdad en capital humano, pues los grupos sociales acumulan capital humano a lo largo de diferentes trayectorias.
- b) Sobre la transformación de capital humano en ingresos, una parte del mayor capital humano se va al empleo asalariado y otra al autoempleo y desempleo; no hay nada que asegure que la primera porción crecerá o cuando menos se mantendrá fija. Utilizando el método de la matriz de transición, esta hipótesis implica una matriz poco correlacionada para la distribución de la educación entre generaciones («padres» e «hijos» socialmente hablando, no en el sentido biológico) y una matriz altamente correlacionada para la distribución del ingreso entre generaciones.

Figura 6. Hipótesis sobre la relación entre exceso de oferta laboral (desempleo y subempleo), años de educación y grupo social



Ciertamente, este conjunto de hipótesis empíricas no podría haber sido derivado de la teoría neoclásica. Los supuestos básicos de la teoría neoclásica son distintos de los de la teoría Sigma. La teoría neoclásica supone que los mercados laborales son walrasianos y sin superpoblación. También hace abstracción del efecto que tienen las desigualdades iniciales en los activos sociales.

La mayoría de los modelos neoclásicos suponen implícitamente una sola curva que transforma educación en capital humano, con lo cual la curva Y coincide con la curva A y tendría, además, una pendiente de 45° en el panel (a) de la figura 2. Mientras que en la teoría Sigma se hace una distinción entre educación y capital humano como una cuestión esencial, en estos modelos neoclásicos se supone implícitamente que educación es igual que capital humano. En general, los modelos neoclásicos predicen una sola relación entre ingresos y años de educación (la línea L en la figura 3), sin tomar en cuenta el efecto de la etnicidad, mientras que el modelo Sigma estudiado aquí predice que la etnicidad de las personas sí importa.

Así, Lucas (1990) introdujo el capital humano en la teoría del crecimiento económico, pero lo hizo suponiendo una economía que es socialmente homogénea y que la educación es igual que capital humano. Los trabajos empíricos sobre determinantes del crecimiento económico se han hecho aceptando estos supuestos (cf. Barro y Sala-i-Martin 2005). Es cierto que se han construido modelos neoclásicos para explicar diferencias étnicas en ingresos, especialmente entre negros y blancos en Estados Unidos (cf. Becker y Murphy 2001); sin embargo, usualmente se emplea el método de equilibrio microeconómico y, por lo tanto, se hace abstracción de las implicancias de la desigualdad en activos sociales en el equilibrio general. En todo caso, estos estudios empíricos sirven usualmente para interpretar la realidad a la luz de la teoría neoclásica, pero no buscan someter la teoría a la prueba de la falsación.

En este estudio, se utilizarán datos peruanos para someter a la prueba empírica las predicciones del modelo Sigma. Sin embargo, dadas las diferencias en los supuestos básicos con la teoría neoclásica, los resultados permitirán contrastar ambas teorías a la vez. La pregunta específica que se busca responder aquí es si la economía peruana se parece a la sociedad Sigma o a la sociedad neoclásica.

5. La base de datos

La base de datos para llevar a cabo el proceso de falsación de los modelos teóricos presentados líneas antes está constituida por la Encuesta Nacional de Hogares (ENAHOG) de 2002 y 2003 del Perú, producida por el Instituto Nacional

de Estadística e Informática (INEI). Estas bases de datos están disponibles en Internet, en la siguiente dirección: <<http://www.inei.gob.pe>>.

La muestra para cada encuesta es representativa de todo el país; aún más, su diseño particular es tal que la muestra es también representativa de varios segmentos importantes de la economía peruana: urbano-rural, y regiones naturales: costa-sierra-selva. La unidad de muestreo es el hogar. Además de los datos referidos al hogar, las encuestas recogieron datos individuales sobre educación, empleo e ingresos, así como datos sobre lugar de nacimiento y lugar de residencia, que son todos importantes para las pruebas empíricas del modelo.

El tamaño de la muestra fue de cerca de 18 mil viviendas en el 2002 y de 20 mil en el 2003. La submuestra utilizada en este estudio se compone de los individuos de 25 o más años de edad, cerca de 38 mil observaciones en el 2002 y de 39 mil en el 2003.

Un análisis crítico de las encuestas, así como las diferentes agregaciones realizadas para construir variables consistentes con las que aparecen en el modelo teórico y las tabulaciones básicas, se encuentra en el anexo metodológico de este artículo (Figueroa, Cruzado y Sánchez 2007). El análisis crítico aplicado a ambas encuestas arrojó los siguientes resultados. La prueba de consistencia lógica interna de los datos es aceptable. La prueba de representatividad, excluidos los casos de «sin información», también fue aceptable. Aunque se encontró una proporción alta de datos «sin información» en diversas variables claves, un análisis de distribuciones bivariadas de los datos permitió distinguir los datos «sin información» de los datos que debieron consignarse como «cero» o «sin significación» al momento de la codificación. Hecha esta corrección, los problemas de datos «sin información» se redujeron a un rango aceptable.

Los hogares cuentan usualmente con negocios familiares, como una bodega en el medio urbano o una finca en el medio rural. Estos negocios generan ingresos colectivos que son resultado de la actividad de los miembros de la familia. ¿Cómo transformar el ingreso colectivo en ingreso individual? Existen tres opciones. Primero, atribuir este ingreso a un miembro particular de la familia; por ejemplo, al jefe de familia. Este método crea sesgos, pues sobreestima el ingreso de esta persona y aumenta el retorno a su educación. Segundo, asignar este ingreso a todos los miembros, lo cual crea el problema de utilizar criterios apropiados para hacerlo. Tercero, no atribuir este ingreso a ningún miembro y mantenerlo como ingreso colectivo, lo cual tiene el problema de que reduce el retorno a la educación de todos los individuos.

Entre estas opciones, INEI, al parecer, ha utilizado la primera, pues no aparece esta distinción en los cuestionarios. La suma de los ingresos individuales es, entonces, igual a la suma de los ingresos familiares.

El INEI publica los datos de ingresos corregidos por la inflación de precios durante el período de la muestra. El período de referencia de la primera encuesta fue de octubre a diciembre de 2002, y el de la segunda, de mayo de 2003 a abril de 2004. Los ingresos nominales tienen distintas unidades de tiempo, pues los entrevistados declararon sus ingresos por los períodos habituales de pago o de generación de ingresos. El INEI llevó estos ingresos a una unidad de tiempo estandarizada, que fue el trimestre. Los ingresos reales se refieren a precios de noviembre de 2002 y de julio de 2003. Hay que señalar que en este período la tasa de inflación *anual* en el Perú fue baja: 1,5% en 2002 y 2,5% en 2003. La corrección por inflación de los datos de ingresos debió de modificar poco los datos recolectados sobre ingresos nominales.

En cuanto a los ingresos reales, se hicieron dos modificaciones. La primera es que se utiliza como unidad de tiempo estandarizada el mes. La segunda es una modificación que se hizo a todos los datos de ingresos, que consistió en corregirlos por las diferencias en los niveles de precios que existen entre Lima y las ciudades del interior del país. Por lo tanto, los datos de ingresos reales para el Perú que se utilizan en este estudio están medidos en nuevos soles de Lima de noviembre de 2002 para la primera encuesta, y en nuevos soles de Lima de julio de 2003 para la segunda.

En el conjunto de tabulaciones básicas que se hicieron de ambas encuestas, se encontró que los estadísticos sobre tendencias centrales para las variables claves del estudio eran muy similares. Para dos encuestas aplicadas en años contiguos, este resultado era de esperar. Este resultado se aplica también al caso de las variables de ingresos (salarios, sueldos, ingresos de autoempleados, ingresos totales). Las tabulaciones y el análisis estadístico se han hecho para las dos encuestas y los resultados son consistentes. Aquí se presentan solo los resultados basados en la encuesta del año 2003, y por ninguna otra razón que la de ser la más reciente.

Las principales variables de los modelos teóricos en estudio incluyen: grupos sociales (A, Y, Z), categoría laboral, años de educación, nivel de capital humano (nivel de conocimientos y edad), ingresos totales, sueldos, salarios, ingresos netos de autoempleo, empleo dependiente (empleado y obrero), autoempleo, desempleo y subempleo.

5.1 Estimación empírica de grupos étnicos

La estimación empírica del tamaño de los grupos étnicos no es tarea sencilla. Se pueden considerar hasta cuatro criterios, que son los siguientes:

- c) El criterio de la lengua materna tiene el problema de que reduce la etnicidad a la lengua, de manera que los hijos que ya no hablan la lengua aborigen de sus padres dejan de ser indígenas. Poblaciones indígenas que perdieron su lengua con el paso del tiempo dejarán de ser indígenas.
- d) El criterio de la autoidentificación subestimaré, en una sociedad jerárquica, el tamaño de las poblaciones que son social y racialmente subordinadas.
- e) El criterio de la región de residencia tiene el problema de que la gente cambiará su etnicidad por el simple hecho de cambiar su residencia a otra región, o del campo a la ciudad.
- f) El criterio de la región de nacimiento tiene el problema de que los hijos que nacen en un lugar distinto del que nació el padre cambiarán de etnicidad.

En el caso del Perú, se optó por el criterio (d), pues parece el menos imperfecto. Este criterio, además, retiene las condiciones históricas y geográficas de la formación de la sociedad peruana actual, en la que los descendientes de poblaciones aborígenes se pueden encontrar en las tres regiones naturales, incluida la costa, y no solo en la sierra y en la selva (Figueroa y Barrón 2005).

En todos los criterios se tiende a subestimar el tamaño de la población indígena, pero este criterio lo hace posiblemente en menor medida. Sobre otras etnias inmigrantes, como los descendientes de la población africana y asiática, los datos son menos confiables, pero todo indica que se trata de grupos étnicos realmente minoritarios.

El grupo social Z (que corresponde empíricamente a la población indígena) se definió como la población nacida en los distritos rurales del país en las tres regiones naturales: costa, sierra y selva. El grupo social A (que corresponde empíricamente a la población blanca) se definió como la población que nació en los 11 distritos más residenciales de Lima (Barranco, Jesús María, La Molina, Lince, Magdalena, Miraflores, Pueblo Libre, San Borja, San Isidro, San Miguel y Surco). El grupo social Y (que corresponde empíricamente a la población mestiza) se calculó por diferencia e incluye a la población nacida en los distritos de Lima que están fuera del núcleo residencial y en los distritos que son capitales de las provincias que, a su vez, son capitales de departamento.

Las proporciones de la población adulta (con 25 o más años de edad) que resulta de este cálculo son: 3,5% como grupo A, 27,5% como grupo Y, 69% como grupo Z. Si se toma una definición más restringida de grupo Z, que incluya a los nacidos en los distritos rurales de la sierra sur y de la selva (donde

se hablan lenguas aborígenes), la proporción se reduce a 28%; si se restringe aún más para incluir solo la sierra sur, la proporción baja a 21% (Figuerola *et al.* 2007, cuadro 54)⁷.

En este estudio, se utilizará el concepto amplio de grupo Z. Ciertamente, es de esperar que cuanto más restringida sea la definición de este grupo, las pruebas de hipótesis tenderán a ser más robustas. Este análisis de sensibilidad de los resultados se ha realizado y será reportado aquí solo en el caso en que se encuentre un patrón diferente.

6. Falsación de las hipótesis

Las predicciones empíricas que se han derivado del modelo Sigma constituyen hipótesis empíricas que tienen la propiedad de ser refutables. En esta sección estas hipótesis se someten a la prueba estadística. Se utilizan pruebas paramétricas y no paramétricas en la prueba de las medias, pues la primera prueba supone homocedasticidad en las distribuciones de frecuencias, mientras que la segunda no lo hace. Ambas pruebas aparecen el anexo estadístico. Los datos se refieren al Perú y provienen de la ENAHO de 2003.

H1: Hipótesis de la exclusión

Como no tenemos mediciones de capital humano, y el capital humano depende de los años de educación, mediremos su concentración por los años de educación. En primer lugar, se puede examinar la relación empírica entre años de educación por grupos sociales. El cuadro 1 muestra esta relación. Para la población de 25 años o más, en términos de nivel educativo, las dos terceras partes de la población se ubican en los niveles primario y secundario; en términos de grupos sociales, la muestra arroja que 3% de la población pertenece al grupo A, 26% al grupo Y y 71% al Z.

⁷ La ENAHO de 2001 preguntó al jefe de familia sobre su autoidentificación étnica. El resultado fue que 43% de los jefes de familia se declararon de origen indígena, mientras que 34% declararon como su lengua un idioma aborigen (Hall y Patrinos 2005, tablas 2.1, 2.7, 7.1). Los estimados sobre el tamaño del grupo indígena son variados, pero muestran en todos los casos que el orden de magnitud es significativo. No se trata de un grupo minoritario.

Cuadro 1. Perú: Nivel educativo por grupo social, 2003
(miles de personas y porcentajes)

Nivel educativo	Grupos sociales						Total	
	Z		Y		A			
	N	%	N	%	N	%	N	%
Ninguno	1.283	14,1	110*	3,3	0,7*	0,2	1.394	10,9
Primario	3.903	43,0	556	16,5	10,1*	2,8	4.470	34,9
Secundario	2.611	28,7	1.475	43,8	94,9	26,9	4.180	32,6
Técnico	725*	8,0	530	15,7	73,4	20,8	1.328	10,4
Universitario	562*	6,2	697	20,7	174,4	49,3	1.433	11,2
Total N	9.083	100,0	3.368	100,0	353,4	100,0	12.805	100,0
Porcentaje (%) horizontal	70,9		26,3		2,8		100,0	
Años de educación:								
Media	6,9		10,8		13,7		8,2	
Mediana	5,0		11,0		14,0		9,0	

Notas:

N: Población de 25 años a más (expansión muestral al universo)

A: Población nacida en los 11 distritos más residenciales de Lima; Y: población nacida en el resto de los distritos de Lima o en distritos residenciales de provincias; Z: población nacida en el resto del Perú.

* Población relativamente pequeña, 10% o menos de su grupo.

Fuente: ENAHO 2003. El método de cálculo aparece en Figueroa *et al.* (2007), cuadro 1B

La media de años de educación varía por grupos sociales: 14 años para el grupo A, 11 años para Y y 7 años para Z. La mediana también muestra la misma relación: 14 años para el grupo A, 11 años para Y, pero solo 6 años para Z. Estos datos muestran una desigualdad pronunciada. El grupo A tiene una media de años de educación que es el doble de la que tiene el grupo Z; en términos de la mediana, la diferencia es de 2,3 veces. Las pruebas estadísticas, tanto paramétricas como no paramétricas, muestran que las diferencias de medias observadas son significativas, tal como se muestra en el anexo de pruebas estadísticas, al final del artículo.

Utilizando los niveles de educación (primaria, secundaria, técnica y universitaria) en lugar de los años de educación, la desigualdad aparece también pronunciada. En el grupo A, cerca de 70% tiene educación postsecundaria, mientras que en el grupo Y esta proporción llega solo a 36%, y en el grupo Z, apenas a 14%. La exclusión del grupo Z de la educación postsecundaria es notoria.

Debe quedar muy en claro que estos hallazgos empíricos no provienen de comparar a poblaciones «indígenas» que viven *actualmente* en zonas rurales con

poblaciones que viven en las zonas residenciales de Lima. La comparación se refiere a poblaciones «indígenas» contra poblaciones mestizas y blancas *independientemente* del lugar de residencia. Como ejemplo, se puede decir que a los que nacieron en las zonas rurales del departamento de Huancavelica se los define como «indígenas» y su nivel de educación medio logrado se mide independientemente del lugar en que residan, que puede ser Huancavelica, Huancayo o Lima. El efecto de la emigración está incluido en las diferencias mencionadas. Si se hace el cálculo solo con los miembros del grupo Z que permanecen en sus lugares de origen y los que residen fuera de su lugar de origen, las medias de años de educación son de cerca de cinco años y nueve años, respectivamente, lo que da como promedio los siete años de promedio para todo el grupo, tal como se mencionó antes. Para el grupo Z, la migración implica cuatro años más de educación, en promedio. La separación teórica que se ha hecho del grupo Z en dos grupos de trabajadores, con diferentes niveles de capital humano, y que podría corresponder a diferentes áreas de residencia, uno urbano y otro rural, tiene, así, una justificación empírica.

Existen algunos estudios sobre los determinantes del nivel de conocimientos que se adquiere en la escuela en el Perú, como una aproximación a los determinantes de la acumulación de capital humano. Así, se ha mostrado que el nivel de conocimientos de los estudiantes para un mismo año de escolaridad varía según el tipo de escuela (privada-pública) y según su ubicación (rural-urbana). El nivel de conocimientos adquiridos es mayor en la escuela privada que en la escuela estatal, y mayor en la escuela urbana que en la rural (Rivera 1979; Cueto, Jacoby y Pollit 1997; Ministerio de Educación del Perú 2005). Los dos factores de la función de producción del capital humano (recursos de la escuela y condiciones socioeconómicas de las familias) parecen diferenciarse bien entre esos tipos de escuelas.

En términos de los factores socioeconómicos, existe alguna evidencia sobre el papel negativo que tiene la desnutrición sobre el aprendizaje escolar. Los Estudios Conjuntos de Integración Económica Latinoamericana (ECIEL), en particular, indican que hay una correlación negativa, que es estadísticamente significativa y cuantitativamente importante, entre rendimiento escolar y grado de desnutrición en una muestra de poblaciones escolares de Lima, Puno-ciudad y Puno-rural (Rivera 1979).

Los datos del cuadro 1 muestran claramente la desigualdad en la dotación de años de educación entre los grupos sociales en el año 2003. Estos datos no refutan la hipótesis, más bien la corroboran. Después de 180 años de vida republicana, y a pesar de la expansión del sistema educativo, en el Perú todavía se observa una exclusión educativa significativa.

H2: Hipótesis de la separabilidad y jerarquía en la relación educación-ingresos

El cuadro 2 presenta datos sobre ingresos totales por nivel educativo y por grupo social. La población se refiere a la población económicamente activa (PEA) que se encuentra empleada (se excluye a la desempleada). La media de los años de educación aumenta en un año y la mediana en dos años respecto a los promedios del cuadro 1. El promedio también sube para los grupos sociales, en especial para el grupo Z.

Cuadro 2. Perú: Ingresos medios por nivel educativo y grupo social, 2003
(miles de personas, soles/mes y porcentajes)

Nivel educativo	Grupos sociales						Total		
	Z		Y		A				
	N (%)	IM	N (%)	IM	N (%)	IM	N	%	IM
Ninguno	9,6	171*	1,9	297*	--	--	604	7,1	180
Primario	41,0	330	13,1	479	1,0	137*	2.704	32,0	346
Secundario	32,7	592	44,0	710	22,7	798	2.998	35,4	637
Técnico	9,6	844*	17,2	912	15,8	965	1.001	11,9	876
Universitario	7,1	1.529*	23,8	2.005	60,5	2.721	1.148	13,6	1.943
Total	100,0	535	100,0	1.015	100,0	1.981	8.464	100,0	717
Total N	5.841		2.330		293		8.464		
%	69,0		27,5		3,5		100,0		
Años de educación:									
Media	7,6		11,4		14,2		9,0		
Mediana	8,0		11,0		15,0		11,0		

Notas:

N: Población de 25 años a más (expansión muestral al universo) que incluye solo a la PEA empleada y con ingresos declarados.

IM: Ingreso medio mensual medido a precios de Lima, de octubre de 2003.

-- No existe población en esta categoría.

* Corresponde a poblaciones relativamente pequeñas, 10% o menos de su grupo social.

Fuente: ENAHO 2003. El método de cálculo aparece en Figueroa et al (2007), cuadro 4B.

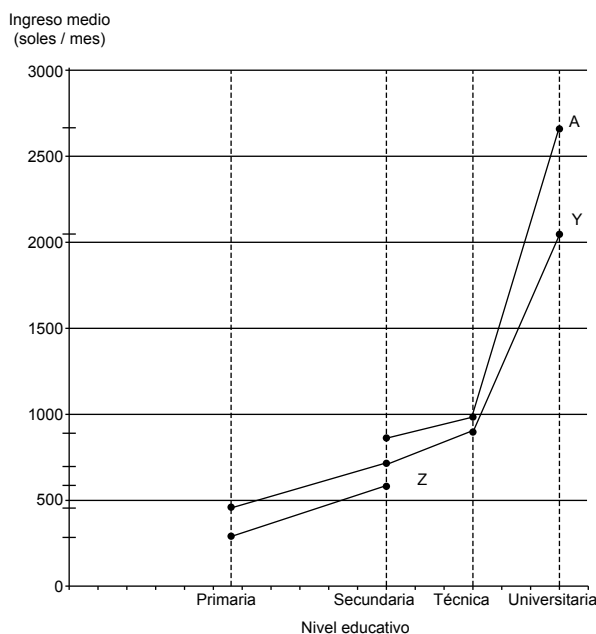
Este cuadro muestra que existe una relación positiva entre la media de ingresos y el nivel de educación para toda la población. Esta relación positiva también se observa para cada uno de los grupos sociales.

La pregunta es si las relaciones entre ingresos y educación son separables estadísticamente por grupos sociales y si dan cuenta de una jerarquía entre A-Y-Z, como sostiene la hipótesis. En general, en el cuadro 2 se observa que para cada nivel de educación, el ingreso total medio es mayor en el grupo A, y este es mayor que el del grupo Y, y este, a su vez, es mayor que el del grupo Z. Se observa la jerarquía en los ingresos que propone la hipótesis. La prueba estadística para evaluar la significancia de estas diferencias debe tomar en cuenta los resultados de la hipótesis H1. Dada la exclusión educativa corroborada en esta hipótesis, esta prueba tiene que hacerse solo entre niveles de educación *comparables*. La figura 7 muestra la naturaleza de la relación entre ingresos y educación. En el grupo Z se encuentra que las personas están distribuidas principalmente solo en los primeros niveles educativos, con tamaños muy pequeños en la postsecundaria, menos de 10%. En el grupo Y, la gran mayoría tiene educación postprimaria, y en el grupo A, la gran mayoría tiene educación postsecundaria.

Los niveles educativos y los grupos sociales comparables son, entonces, los siguientes: en el nivel de primaria, los grupos Y-Z; en el nivel de secundaria, los tres grupos, A-Y-Z; y en los niveles de técnica y universitaria, los grupos A-Y. Las pruebas estadísticas paramétricas y no paramétricas muestran que las diferencias son significativas en los niveles primaria y universitaria. Para los otros niveles, en la medida en que las pruebas paramétricas no pasan la prueba de la igualdad de varianzas, como criterio de evaluación solo quedan las pruebas no paramétricas. Las diferencias no son significativas en el nivel técnico; mientras que en la secundaria, las diferencias son significativas, excepto por los grupos A-Y. Se puede decir que, en general (en cuatro de los seis casos), la hipótesis no es refutada (véase el anexo).

Las pendientes de las líneas que relacionan los ingresos medios con los niveles educativos en la figura 7 son crecientes en los tres grupos sociales. La pendiente de la curva A parece crecer más rápidamente que la de Y, y esta más rápidamente que la de Z. El método estándar de regresiones podría aplicarse a estos datos para tener una prueba econométrica sobre estas relaciones entre las pendientes, utilizando los grupos sociales A, Y, Z como variables ficticias (*dummy variables*). Sin embargo, los datos hacen que se pueda comparar solo el nivel de secundaria. Un supuesto implícito del modelo estándar de regresiones es que las variables independientes deben estar distribuidas de manera uniforme en todo el rango de las variables dependientes de la muestra; si existen exclusiones, el método es inaplicable. En realidad, los modelos econométricos suponen implícitamente ausencia de exclusiones en las relaciones entre variables. Los fenómenos de exclusiones no pueden, entonces, ser analizados con este método. Un análisis de regresiones

Figura 7. Perú: relación empírica entre ingresos medios y niveles educativos por grupos sociales



por separado para cada grupo social mostró que, en efecto, las pendientes son diferenciadas y estas diferencias son estadísticamente significativas. En suma, los datos no refutan la hipótesis H2.

H3: Hipótesis de la separabilidad y jerarquía en las relaciones educación-salarios

En el estudio sobre las relaciones entre niveles educativos y remuneraciones laborales se tienen que distinguir dos categorías laborales: empleados y obreros. Estas dos categorías se diferencian por el nivel de capital humano; es decir, por el nivel de educación, pero también se diferencian cualitativamente, pues la categoría empleados tiene mayor estatus social que la de obreros. Constituyen, entonces, dos mercados separados.

Los cuadros 3a y 3b presentan los datos sobre educación y remuneraciones para los dos mercados, el de obreros y el de empleados. La media de años de educación es de 13,7 años para los empleados y de 8,3 para los obreros; los valores de las medianas son de 14 y 9. Los niveles educativos aparecen, así, como un factor de diferencia entre estos dos grupos. La figura 8 muestra los resultados de manera gráfica.

Cuadro 3a. Perú: Salarios medios por nivel educativo y grupos sociales, 2003
(miles de personas, soles/mes y porcentajes)

Nivel educativo	Grupos sociales						Total	
	Z		Y		A			
	N (%)	IM	N (%)	IM	N (%)	IM	N %	IM
Ninguno	6,0	306*	1,8	349*	--	--	4,7	310
Primario	37,5	489	21,6	518	--	--	32,4	490
Secundario	46,7	616	59,4	651	78,2	569*	50,8	627
Técnico	8,1	695*	13,8	712	21,8	743*	10,0	703
Universitario	1,7	763*	3,4	796*	--	--	2,1	778
Total	100,0	557	100,0	630	100,0	606*	100,0	579
Total N	1.163		482		23		1.668	
%	69,7		28,9		1,4		100,0	
% de PEA (del cuadro 2)	19,9		20,7		7,8*		19,7	
Años de educación:								
Media	8,0		9,7		11,3		8,6	
Mediana	9,0		11,0		11,0		10,0	

Notas:

N: Población de trabajadores obreros (expansión muestral al universo).

IM: Ingreso mensual medido en soles de Lima, de octubre de 2003.

-- No existe población en esta categoría.

* Corresponde a poblaciones muy pequeñas, 10% o menos de su grupo social en este cuadro o con respecto a la PEA empleada del cuadro 2.

Fuente: ENAHO 2003. El método de cálculo aparece en Figueroa *et al.* (2007), cuadro 10B.

Cuadro 3b. Perú: Sueldos medios por nivel educativo y grupo social, 2003
(miles de personas, soles/mes y porcentajes)

Nivel educativo	Grupos sociales						Total	
	Z		Y		A			
	N (%)	IM	N (%)	IM	N (%)	IM	N %	IM
Ninguno	0,1	357*	0,5	431*	--	--	0,3	415
Primario	4,6	703*	1,8	641*	--	--	2,8	685
Secundario	29,2	871	28,1	707	14,2	763	27,2	788
Técnico	32,9	1.045	26,2	1.009	17,7	1.102	28,4	1.033
Universitario	33,2	1.780	43,4	2.203	68,1	3.182	41,3	2.212
Total	100,0	1.222	100,0	1.433	100,0	2.469	100,0	1.442
Total N	807		833		181		1.821	
%	44,3		45,7		10,0		100,0	
% de la PEA (del cuadro 2)	13,8		35,7		61,8		21,5	
Años de educación:								
Media	13,5		13,7		14,9		13,7	
Mediana	14,0		14,0		16,0		14,0	

Notas:

-- No existe población en esta categoría.

* Corresponde a poblaciones muy pequeñas, 10% o menos de su grupo social.

Fuente: ENAHO 2003. Para el método de cálculo, véase Figueroa *et al.* (2007), cuadro 19-B.

El cuadro 3a presenta los datos para el grupo de los obreros. La gran mayoría de este grupo tiene educación primaria o secundaria. Los analfabetos no participan en este mercado, tampoco los que tienen postsecundaria. Por lo tanto, el grupo social A no es parte de la oferta laboral en este mercado. Los casos relevantes son los grupos Y-Z en los niveles primaria y secundaria. Dentro del grupo de obreros también se observa que los que tienen educación secundaria obtienen un salario mayor.

Esta relación también es separable para cada grupo social. Dado el nivel de primaria, el grupo social Y recibe un salario mayor que el de Z; similar resultado se observa en el nivel secundario. Sin embargo, y utilizando solo el test no paramétrico (debido a la falta de homocedasticidad en el test paramétrico), estas diferencias son estadísticamente significativas en el caso del nivel secundario, pero no lo son en el caso del nivel primario (véase el anexo). La comparación es mucho más relevante en el nivel secundario, pues ahí se encuentra la mayoría de ambos grupos: 60% de los obreros del grupo Y y 47% de los obreros del grupo Z.

El cuadro 3b presenta datos sobre niveles de educación y sueldos medios de los empleados. La media y la mediana de años de educación de los empleados muestra muy poca variación entre grupos sociales. Los empleados parecen, por lo tanto, constituir un grupo relativamente homogéneo en cuanto a dotación de años de educación, pues todos tienen educación postprimaria y cerca de 70% cuentan con educación postsecundaria.

Dentro del grupo de empleados, se corrobora la relación de que a mayor educación, los sueldos son mayores. Igual relación positiva entre sueldos y educación se observa para cada uno de los grupos sociales. La relación que establece esta hipótesis es que, para un nivel dado de educación, los sueldos deben mostrar una jerarquía por grupos sociales A-Y-Z. Los datos muestran que, en efecto, existe esta jerarquía en el nivel universitario; sin embargo, las diferencias que se observan en los niveles de secundaria y técnica son pequeñas.

Las pruebas estadísticas, tanto paramétricas como no paramétricas, muestran que las diferencias son significativas para el nivel universitario. Para los otros dos niveles, las pruebas paramétricas muestran que la prueba de la igualdad entre varianzas falla; quedan, entonces, las pruebas no paramétricas como el criterio de evaluación, las cuales muestran que las diferencias no son significativas en los casos de los grupos Y-A en secundaria y en los grupos Z-A y A-Y en técnica. El resultado es que la hipótesis se corrobora solo en el nivel universitario, que es tal vez el caso relevante de comparación, pues la gran mayoría de cada grupo se encuentra ahí: 68% del grupo A, 43% de Y y 33% de Z, haciendo que la comparación en los otros niveles sea menos relevante.

La evidencia empírica presentada hasta aquí muestra que existe una relación estadísticamente positiva entre los ingresos del trabajo dependiente y el nivel educativo del trabajador. Esta relación también existe tanto dentro del grupo de empleados como en el de obreros. Como el nivel educativo de los empleados es mayor que el de los obreros, también se observa que el sueldo medio es mayor que el salario medio. Finalmente, para un nivel educativo dado y para los casos que son comparables, los sueldos y salarios tienden, en general, a variar por grupos sociales en el orden que indica la hipótesis. La relación estadística es significativa cuanto mayores son las proporciones de los grupos sociales en un nivel educativo dado. En suma, los datos empíricos no parecen refutar la hipótesis H3.

Las diferencias de sueldos y salarios que se observan por nivel educativo no implican, por lo tanto, discriminación salarial. Si, como supone el modelo teórico, las firmas pagan igual remuneración por igual capital humano, el salario del grupo Z debería ser menor que el que obtiene el grupo Y para igual número de años de escolaridad; de manera similar, para años de educación dados, el sueldo del grupo Y (profesional egresado de una universidad pública y de provincia)

debería ser inferior que el del grupo A (profesional egresado de una universidad privada). Estos son, precisamente, los resultados empíricos encontrados aquí. Las diferencias en remuneraciones muestran, entonces, exclusión en la calidad educativa (el orden de la dotación de capital humano es A-Y-Z) antes que discriminación salarial.

H4: Hipótesis del salario de eficiencia

El cuadro 4 presenta datos sobre ingresos medios de los trabajadores dependientes y autoempleados, por grupo social. La hipótesis del precio del trabajo de eficiencia se refiere a cada grupo social; por lo tanto, examinaremos las relaciones por grupo social.

Cuadro 4. Perú: Ingresos medios por trabajo dependiente o autoempleado, según grupos sociales en 2003 (soles/mes)

Grupo Social	Nivel educativo					
	Ninguno	Primario	Secundario	Técnico	Universitario	Total
Grupo Z						
Sueldo	357*	703*	871*	1.045*	1.780*	1.222
Salario	306*	483	618	695*	763*	557
Ingreso de autoempleo	136	260	451	568*	818*	327
Grupo Y						
Sueldo	431*	641*	707	1.009	2.203	1.433
Salario	349*	518*	651	712*	796*	630
Ingreso de autoempleo	221*	384*	550	696	1.180	621
Grupo A						
Sueldo	--	--	763	1.102	3.182	2.469
Salario	--	--	568*	743*	--	608
Ingreso de autoempleo	--	137*	1.032	508*	1.135	1.013

Notas:

N: Población de trabajadores empleados (expansión muestral al universo)

IM: Ingreso mensual medido en soles de Lima, de octubre de 2003.

-- No existe población en esta categoría.

* Corresponde a poblaciones muy pequeñas, 10% o menos de su grupo social en este cuadro o con respecto a la PEA empleada en el cuadro 2.

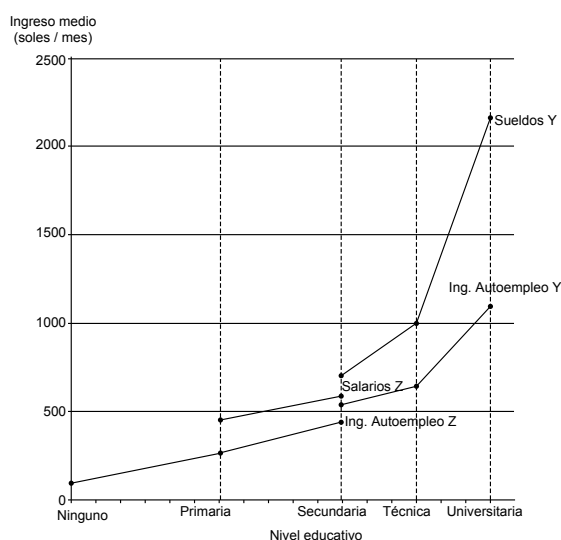
Fuente: Enaho 2003. El método de cálculo aparece en Figueroa *et al.* (2007), cuadro 7B.

En el grupo social Z, se observa que la mayoría de los trabajadores son o autoempleados u obreros. La hipótesis establece que los ingresos medios de los primeros deben ser inferiores que los de los segundos. Esta relación se cumple para el agregado. Se observa que esta relación se cumple también para cada nivel educativo, excepto para el universitario. Sin embargo, se debe notar que existe una cantidad relativamente pequeña del grupo Z en el nivel universitario, lo cual hace la comparación inadecuada en términos estadísticos. La prueba estadística muestra que las diferencias observadas son todas significativas en los niveles relevantes, que son: primaria, secundaria y técnica, así como para el agregado (véase el anexo).

En el grupo social Y, la mayoría de los trabajadores son o autoempleados o empleados. Se observa que el ingreso de autoempleo es menor que el sueldo medio para cada nivel educativo y también para el agregado. Se observa que existen muy pocos casos de trabajadores en el nivel *ninguno*. La prueba estadística muestra que las diferencias observadas son todas significativas en los niveles relevantes, que son: secundaria, técnica y universitaria, así como para el agregado (véase el anexo).

En el grupo A, la mayoría son o autoempleados o empleados. En este grupo, los niveles educativos *ninguno* y *primaria* son irrelevantes. En el resto de los niveles, la relación que establece la hipótesis se cumple para todos los casos, excepto en el nivel secundario. En el agregado, el ingreso medio del autoempleo es menor que el sueldo medio. Dado el tamaño pequeño de la muestra, estas diferencias no fueron puestas a la prueba estadística.

Figura 8: Perú. Relaciones empíricas entre salarios-ingresos de autoempleo y entre sueldos-ingresos de autoempleo por grupos sociales



Es importante también anotar que en el cuadro 4 se observa una relación positiva generalizada entre ingresos y niveles de educación; es decir, en cada columna del cuadro, independientemente del grupo social y de la categoría laboral, se observa esta relación. La figura 9 presenta los resultados de manera resumida.

En suma, los datos empíricos no parecen refutar la hipótesis H4. Los mercados laborales parecen operar con sueldos y salarios que son de eficiencia, como dice el modelo de la teoría Sigma, en los cuales los autoempleados constituyen parte del exceso de oferta laboral.

H5: Hipótesis de la exclusión diferenciada del mercado laboral

El cuadro 5 muestra los datos sobre desempleo y subempleo por grupos sociales. El grupo de los *desempleados* incluye a los trabajadores que se encuentran buscando empleo activamente. El grupo de los *subempleados* se define como el de aquellos que están autoempleados y que tienen ingresos menores que el ingreso medio de los dependientes, para un nivel de educación dado.

Cuadro 5. Perú: Tasa de exceso de oferta laboral (desempleo y subempleo) por niveles de educación por grupos sociales en 2003 (porcentajes)

Nivel educativo	Grupos sociales			Total
	Z	Y	A	
Ninguno	88,1 (11,7)	73,7* (12,3)	--	87,1 (11,7)
Primario	79,7 (6,5)	61,8 (7,2)	100,0* 0,0	77,7 (6,6)
Secundario	57,6 (5,4)	49,8 (8,2)	39,3 (11,4)	54,4 (6,5)
Técnico	35,0* (5,6)	31,1 (7,3)	25,7 (12,4)	32,9 (6,6)
Universitario	29,3* (7,0)	32,7 (8,7)	28,8 (2,7)	31,0 (7,2)
Total	65,6 (6,6)	44,5 (8,1)	31,5 (6,4)	58,4 (7,0)

Notas:

-- No existe población en esta categoría.

* Corresponde a poblaciones muy pequeñas, 10% o menos de su grupo social.

En paréntesis y cursiva: tasa de desempleo. La diferencia con el total es la tasa de subempleo.

Fuente: ENAHO 2003. Para el método de cálculo, véase Figueroa *et al.* (2007), cuadro 38-B.

En el agregado, la tasa de exceso de oferta laboral es de 59% para el Perú en el año 2003. Esta tasa se compone de la tasa de desempleo, que es de 8%, y de la tasa de subempleo, que es de 51%.

En cuanto a las diferencias por niveles educativos, los datos muestran que la tasa de exceso de oferta laboral está correlacionada de manera negativa con el nivel educativo, como predice el modelo teórico: la tasa es menor cuanto mayor es el nivel educativo de los trabajadores. Las tasas de desempleo también tienden a mostrar una relación negativa por niveles educativos. (El modelo teórico no tiene predicciones precisas sobre esta relación).

En cuanto a las diferencias por grupos sociales, los resultados indican que estas son muy pronunciadas: 66% en el grupo Z, 46% en Y, y 33% en A. Las diferencias en la tasa de desempleo señalan que la tasa más alta se encuentra en el grupo Y (10%) y la menor en el grupo Z (7%). Dada la pequeña magnitud de las tasas de desempleo, la diferencia en las tasas de exceso de oferta laboral se debe a las altas tasas de subempleo.

En el grupo social Z se puede hacer la distinción entre los dos grupos de trabajadores que se planteó en el modelo teórico. Si se define empíricamente como grupo Z_0 a los que no llegaron a la educación secundaria, este grupo es de cerca de 2,9 millones de personas, que representa 35% de la fuerza laboral nacional (que es de 8,4 millones). Luego se podría descomponer el exceso de oferta laboral entre 8% de desempleo, 16% de subempleo (exclusión parcial del mercado laboral) y 35% de exclusión total del mercado laboral (los trabajadores Z_0).

La prueba crucial de la hipótesis es que la tasa de exclusión laboral es menor en los grupos sociales ordenados jerárquicamente, para un nivel dado de educación. Para el nivel educativo *ninguno* no existe relación entre grupos, pues su tamaño es muy pequeño en el grupo social Y y en A. En el nivel primario, los grupos comparables son Z e Y: las tasas de exclusión son 80% y 62%. En el nivel secundario, se pueden comparar los tres grupos y las tasas resultantes son 58%, 50% y 39%, conforme con la hipótesis. Para el nivel técnico, los grupos comparables son Y y A, y las tasas observadas son 31% y 26%. En el nivel universitario, los grupos comparables son también Y y A, con tasas de 33% y 29%. Las pruebas estadísticas indican que, en efecto, las diferencias observadas en estas tasas son estadísticamente significativas (véase el anexo). En suma, los datos no refutan la hipótesis H5.

H6: Persistencia de la desigualdad entre generaciones

Los cuadros 6a y 6b presentan las razones de educación e ingresos por grupos de edad (una medida de diferentes generaciones) y por grupos sociales. En el cuadro 6a se puede apreciar un aumento relativamente importante en los años de educación entre generaciones del grupo social Z: los jóvenes («hijos») tienen casi el doble de educación que los viejos («padres»). Los aumentos que se han

producido en los otros grupos sociales son menos pronunciados. En el grupo social A, parece haberse llegado a un techo educativo de 14 años. La consecuencia es que las brechas educativas entre el grupo Z y los otros grupos se han acortado considerablemente.

Cuadro 6a. Perú: Diferencias intergeneracionales en educación por grupos sociales, 2003 (media de años de educación)

	Grupo generacional												Total		
	25-34			35-44			45-54			55-65					
Grupo social	N		e	N		e	N		e	N		e	N		e
Z	1.551		11	1.636		8	1.260		7	850		5	5.297		8
Y	912		12	670		12	449		11	218		9	2.249		12
A	146		14	111		14	29		14	6		12	291		14
Total	2.609		12	2.417		10	1.738		8	1.074		6	7.837		9
Ratio Z/A		0,8			0,6			0,5			0,4			0,6	
Ratio W/A		0,9			0,8			0,8			0,7			1,2	

Notas:

N: población de PEA ocupada (millones de personas).

e: media de años de educación.

Fuente: ENAHO 2003. Para el método de cálculo, véase Figueroa *et al.* (2007), cuadro 41 B.

Cuadro 6b. Perú: Diferencias intergeneracionales en educación por grupos sociales, 2003 (media de ingresos en soles/mes)

	Grupo generacional												Total		
	25-34			35-44			45-54			55-65					
Grupo social	N		e	N		e	N		e	N		e	N		IM
Z	1.551		530	1.636		596	1.260		623	850		473	5.297		563
Y	912		771	670		987	449		1.195	218		1.626	2.249		1.003
A	146		1.456	111		2.455	29		2.982	6		1.648	291		1.991
Total	2.609		666	2.417		790	1.738		810	1.074		713	7.837		743
Ratio Z/A		0,4			0,2			0,5			0,4			0,6	
Ratio W/A		0,5			0,4			0,8			0,7			1,2	

Notas:

N: población de PEA ocupada.

IM: media de ingreso mensual.

Fuente: ENAHO 2003. Para el método de cálculo, véase Figueroa *et al.* (2007), cuadro 43B.

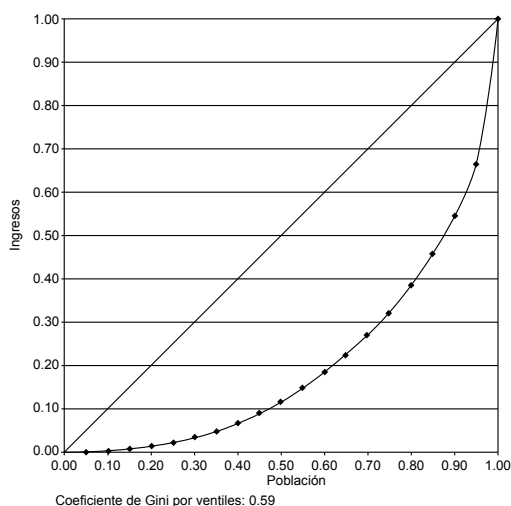
En el cuadro 6b se puede apreciar el cambio en las brechas de ingresos. En el grupo social Z, se observa que, en general, el ingreso medio del grupo de los «padres» (55 a 65 años) es inferior al de los hijos. Similar tendencia se observa en el grupo A. En el grupo Y, en cambio, la tendencia es en dirección contraria: los ingresos medios son mayores para las generaciones mayores.

En términos de ingresos relativos, los «padres» del grupo Z tienen un ingreso relativo al del grupo A que es de 30%, y los «hijos», de 40%. Este pequeño cambio contrasta con los grandes cambios ocurridos en la reducción en las brechas de años de educación que se mostró en el cuadro anterior. La situación del grupo social Y parece tener una trayectoria distinta cuando se la compara con la del grupo A. El ingreso relativo de «padres» e «hijos» es casi constante, en el rango de 40% y 50%, a excepción del grupo más viejo, en el que esta cifra es 100%. Esta última cifra es ciertamente extraña. Podría ser que el pequeño tamaño de la muestra en el grupo A sea un factor que explique parte de este resultado.

La conclusión que se debe destacar de estos cuadros sobre la convergencia entre generaciones y entre grupos sociales es que si bien existe una tendencia clara a la disminución de las diferencias en los años de educación, no existe una tendencia similar hacia la disminución de la desigualdad en los ingresos entre generaciones. Este resultado es consistente con la hipótesis H6.

De la base de datos utilizada en este estudio se ha hecho una estimación del grado de desigualdad. La distribución del ingreso por ventiles se presenta en el cuadro 7. La correspondiente curva de Lorenz se muestra en la figura 9. El grado de desigualdad es bastante pronunciado. Así, el tercio más pobre de la población

Figura 9: Perú: Curva de Lorenz, por ventiles, 2003



Fuente: Cuadro 7.

Cuadro 7. Distribución de ingresos de la PEA ocupada
(población de 25 o más años), 2003

	Total		Z	Y	A
Ventiles	Fi	Qi	Fi	Fi	Fi
1	0.05	0.01	0.07	0.01	0.01
2	0.10	0.01	0.13	0.03	0.01
3	0.15	0.01	0.20	0.05	0.02
4	0.20	0.01	0.26	0.07	0.03
5	0.25	0.02	0.32	0.10	0.04
6	0.30	0.03	0.38	0.14	0.06
7	0.35	0.05	0.43	0.18	0.08
8	0.40	0.07	0.49	0.22	0.09
9	0.45	0.09	0.54	0.27	0.12
10	0.50	0.12	0.59	0.32	0.13
11	0.55	0.15	0.64	0.38	0.15
12	0.60	0.18	0.68	0.45	0.18
13	0.65	0.22	0.73	0.52	0.27
14	0.70	0.27	0.77	0.58	0.32
15	0.75	0.32	0.81	0.64	0.40
16	0.80	0.39	0.85	0.71	0.48
17	0.85	0.46	0.90	0.78	0.54
18	0.90	0.55	0.94	0.84	0.61
19	0.95	0.67	0.98	0.92	0.74
20	1.00	1.00	1.00	1.00	1.00
Gini (ventiles)	0.59				

Fi: frecuencia acumulada.

Qi: participación de ingreso por ventil acumulada

1/Coeficiente de Gini calculado en base a ventiles

Fuente: Cálculos propios basados en datos de ENAHO 2003 y Cruzado 2006.

recibe apenas 5% del ingreso total, mientras que el decil más rico recibe 45% del ingreso total. Por grupos sociales, el cuadro 7 muestra que la incidencia de la pobreza en el grupo Z es la más alta: 43% se encuentra en el tercio más pobre de la población; por otro lado, la incidencia de la riqueza es la más alta en el grupo A: 40% se encuentra en el decil superior.

El coeficiente de Gini, calculado por ventiles, es de aproximadamente 0,60 para el Perú de 2003. Esta cifra es relativamente alta para los estándares internacionales, pero debe de ser una cifra subestimada. Ciertamente, como es usual en las encuestas de hogares en todos los países, la muestra no incluye a la élite económica peruana. En realidad, el grupo social A que se ha utilizado en este estudio corresponde a la llamada clase media (administradores, ejecutivos,

profesionales de alto nivel) antes que a la clase capitalista. En una sociedad de dos clases, como es el modelo teórico construido, la clase media usualmente se incluye en la categoría de la clase capitalista, pues se supone que ella se encuentra más ligada a esta clase (Wright 1997). Este es el procedimiento que también se ha seguido en este estudio.

Cálculos anteriores sobre el coeficiente de Gini para el Perú llegaron a valores de 0,62 (Figueroa 2003, cuadro 6.2). El Perú es un país desigual. La educación no parece ser, en efecto, un sistema igualador. Este resultado es consistente con la hipótesis H6.

7. Conclusiones e implicaciones para las políticas

Las relaciones estadísticas presentadas no refutan las predicciones del modelo de la teoría Sigma. En las relaciones entre educación e ingresos en el Perú, la etnicidad de la gente sí importa, tal como predice el modelo Sigma. En suma, el Perú se parece a una sociedad Sigma y las predicciones del modelo teórico Sigma son consistentes con los datos observados.

El papel de la etnicidad en la relación educación-ingresos no ha sido muy estudiado en la literatura internacional, como se muestra en la revisión de la literatura realizada por Loaiza (2006). Existen, sin embargo, algunos trabajos empíricos, sobre todo para el primer mundo, que permiten hacer una comparación con los resultados obtenidos en este estudio. Sobre la hipótesis H3, los trabajos de Shapiro (1984) y Trejo (1997) para Estados Unidos, y Chiswick (1980) para Inglaterra, muestran que, en efecto, en la relación entre educación y salarios la etnicidad sí importa. Sobre el Perú, el estudio de Ñopo *et al.* (2004) encontró que en las diferencias salariales urbanas la etnicidad sí cuenta.

Sobre la hipótesis de la persistencia en las desigualdades entre generaciones, los estudios de Couch y Dunn (1997) para Estados Unidos y Alemania, y los de Solon (1992) y Zimmerman (1992) para Estados Unidos, encuentran que la movilidad intergeneracional es muy débil, que los hijos tienden a heredar la posición relativa de sus padres en años de educación, en ocupación y en ingresos. El estudio de Borjas (1992) sobre Estados Unidos introduce la variable étnica y muestra que esta tiene poder explicativo. Con métodos menos cuantitativos, Bowles, Gintis y Groves (2005) llegan a ese mismo resultado. Para América Latina, los trabajos de Dahan y Gaviria (1999) y de Behrman, Gaviria y Székeley (2001) muestran resultados poco concluyentes sobre la movilidad intergeneracional, pero en ellos la variable étnica es ignorada. En suma, los resultados obtenidos para el Perú no son contradictorios con los obtenidos en otras partes del mundo.

Es probable que estos resultados se puedan hacer extensivos a la mayoría de los países de América Latina. Por lo tanto, no hay razón alguna para rechazar la teoría Sigma en esta etapa de la investigación. Este resultado implica que la teoría neoclásica, que es la que compite, no podría explicar esos hechos.

La teoría Sigma, una vez aceptada por su consistencia empírica, puede servir de guía para discutir las políticas de empleo. Cuanto más homogénea sea la masa trabajadora, el mercado laboral podrá generar salarios más homogéneos y menor grado de desigualdad en la sociedad. Pero una sociedad Sigma no presenta estas características. Por el contrario, como muestra el caso peruano, la oferta laboral se compone de una masa heterogénea de trabajadores, donde existe una correspondencia entre dotación baja de capital humano y ciudadanía de segunda clase.

Una sociedad capitalista será más desarrollada cuando menos desigualdad exista en sus condiciones iniciales. Ciertamente, la historia no se puede cambiar, pero su legado se puede romper. Un objetivo claro de política de largo plazo consiste en reducir las desigualdades en capital humano entre grupos étnicos.

Se ha escrito mucho en la literatura del desarrollo económico sobre la necesidad de implementar políticas de *igualdad de oportunidades*, pero sin mucha claridad sobre su contenido. Con ayuda de la figura 2 se puede dar contenido analítico a esta propuesta. En una sociedad Sigma, igualdad de oportunidades significaría reducir las tres curvas a una sola. También queda claro que no se puede llamar igualdad de oportunidades al acceso de los trabajadores Z a más años de educación solamente, pues eso les significaría moverse a lo largo del sendero dado por la línea Z en el proceso de acumulación de capital humano, cuando lo que constituiría igualdad de oportunidades es que se muevan de la curva Z a la curva Y o, aún mejor, a la curva A, según aparecen estas curvas en la figura 2(a). Esto último es lo que sería tener un sistema educativo igualador.

La otra política es la *acción afirmativa*. Esta política consiste en aplicar medidas de discriminación (positiva) a favor de los pobres para que sus hijos tengan acceso a la escuela. Entregar un subsidio a la familia pobre, sujeto a la condición de que sus hijos no abandonen la escuela, es una política que se ha extendido en América Latina. Nuevamente, esta política implica que los trabajadores se mueven a lo largo de distintos senderos en el proceso de acumulación de capital humano, sea a lo largo de la curva Z o de la curva Y, pero en ningún caso a lo largo de la curva A. La acción afirmativa como una política que llevará al desplazamiento de la curva Z hacia la curva Y, y de ambas hacia la curva A, implicaría un paquete diferente de medidas. Se tendrían que modificar los parámetros que subyacen a las relaciones establecidas en la figura 1.

Los servicios de salud constituyen otro factor esencial en la acumulación de capital humano. En este estudio, los servicios de salud juegan, en efecto, un papel muy importante en ese proceso. Uno de los factores que subyacen a las relaciones mostradas en el panel (a) de la figura 1 es la oferta de servicios de salud. Las distinciones analíticas hechas anteriormente sobre las políticas de igualdad de oportunidades y de acción afirmativa también se aplican al caso de la oferta de los servicios de salud. Una manera de desplazar las curvas de acumulación de capital es elevando la cantidad y calidad de los servicios de salud que se ofrecen a la población Z.

En el corto plazo, un objetivo central de la política económica tendría que estar dirigido a reducir el problema del exceso de oferta laboral. Para el caso del Perú, la base de datos utilizada revela un nivel de exceso de oferta laboral muy grande, cercano a 60% de la fuerza laboral. Esta tasa se compone de 8% de desempleo (principalmente población mestiza), 16% de subempleo, que es la exclusión parcial del mercado laboral (principalmente población mestiza), y 35% de exclusión total del mercado laboral, excepto por empleos asalariados temporales (principalmente población indígena).

Frente a esta estructura en el exceso de oferta laboral, la política de empleo no puede dirigirse al grupo de desempleados solamente, como es la práctica usual. Incluir en los grupos objetivo al segundo grupo de subempleados es igualmente importante. En todo caso, las demandas de empleo provienen principalmente de estos dos grupos, que en términos étnicos corresponden sobre todo a los mestizos. Se trata de incluir en las políticas de empleo también a los trabajadores indígenas, tanto del campo como de la ciudad, que en el caso del Perú son los agricultores, artesanos y comerciantes de pequeña escala. Ellos, debido a su exclusión política, no tienen voz en las decisiones de las políticas públicas; sin embargo, ellos también son parte del exceso de oferta laboral.

Es probable que estas conclusiones se puedan extender a la mayoría de los países de América Latina. El problema del empleo se refiere a la existencia de un enorme exceso de oferta laboral que varía según los grupos étnicos. En consecuencia, las políticas públicas tendrían por objetivo la reducción de este exceso de oferta laboral. Definido así «el problema del empleo», los objetivos de las políticas públicas y los instrumentos que se utilicen tendrían que ser muy distintos de los actuales. Se dice que entender un problema ya constituye la mitad de su solución. Este estudio ha buscado, por eso, contribuir a entender la naturaleza del problema del empleo.

Anexo: Pruebas estadísticas

Hipótesis 1: Diferencias de medias en años de educación
Prueba F de diferencia de medias

Grupos sociales	Test de igualdad de varianzas		Test de diferencia de medias				
	Estadístico T	p-value	Diferencia	Intervalo de confianza		Estadístico F	p-value
Z - Y	354.56	0.00	-396	[-4.10	-3.81]	2024.40	0.00
Z - A	150.21	0.00	-6.87	[-7.20	-6.55]	1620.48	0.00
Y - A	53.21	0.00	-2.92	[-3.26	-2.58]	258.30	0.00

Test U de Mann Withney de contraste de distribuciones

Grupos sociales	Test de contraste de distribución			
	Suma de rangos		Estadístico Z	p-value
	Z	Y		
Z - Y	490,300,000	218,000,000	-61.49	0.00
Z - A	420,400,000	10,058,291	-24.62	0.00
Y - A	38,984,586	2,479,585	-12.57	0.00

Realizamos una prueba F de contraste de medias (comparando el promedio de los años de educación entre los tres grupos sociales), expandiendo los datos al universo y corrigiéndolos según el diseño muestral. Vemos que la diferencia es significativa entre los tres grupos. Luego realizamos la prueba U de Mann Whitney, que contrasta la distribución de una variable en distintos grupos y vemos, nuevamente, que la diferencia entre los grupos sociales es estadísticamente significativa.

Hipótesis 2: Diferencias de medias en ingresos
Prueba F de diferencia de medias

Nivel educativo	Grupos sociales	Test de igualdad de varianzas		Test de diferencia de medias				
		Estadístico T	p-value	Diferencia	Intervalo de confianza		Estadístico F	p-value
Primaria	Z - Y	4.24	0.04	-126.41	[-164.86	-87.97]	33.04	0.00
Secundaria	Z - Y	0.63	0.43	-102.02	[-162.35	-41.68]	6.36	0.01
	Z - A	0.15	0.70	-217.56	[-513.60	78.47]	2.11	0.15
	Y - A	0.07	0.80	-115.54	[-416.41	185.32]	0.36	0.55
Superior técnica	Y - A	0.96	0.33	-41.85	[-307.86	224.17]	0.15	0.69
Superior universitaria	Y - A	21.48	0.00	-938.12	[-1677.40	-198.83]	0.30	0.07

Test U de Mann Withney de contraste de distribuciones

Nivel educativo	Grupos sociales	Test de contraste de distribución			
		Suma de rangos		Estadístico Z	p-value
		Z	Y		
Primaria	Z - Y	34,224,589	4,808,441	-9.83	0.00
Secundaria	Z - Y	23,098,507	10,714,470	-7.61	0.00
	Z - A	17,017,958	289,828	-2.76	0.01
	Y - A	3,033,103	113,184	-1.20	0.23
Superior técnica	Y - A	553,451	31,370	-0.72	0.47
Superior universitaria	Y - A	785,177	103,934	-0.66	0.00

Para corroborar la segunda hipótesis, realizamos nuevamente un test de medias, pero lo que comparamos esta vez son los ingresos promedio de los distintos grupos sociales. Las pruebas estadísticas paramétricas y no paramétricas muestran que las diferencias son significativas en los niveles primaria y universitaria. Para los otros niveles, las pruebas paramétricas no pasan la prueba de la igualdad de varianzas; por lo tanto, como criterio de evaluación solo quedan las pruebas no paramétricas. Las diferencias no son significativas en el nivel técnico, mientras que en la secundaria, las diferencias son significativas, excepto para los grupos Y-A. Se puede concluir que, en general (en cuatro de los seis casos), la hipótesis no es refutada.

*Hipótesis 3: Diferencias de medias en remuneraciones en el mercado laboral**Empleados**Prueba F de diferencia de medias*

Nivel educativo	Grupos sociales	Test de igualdad de varianzas		Test de diferencia de medias				
		Estadístico T	p-value	Diferencia	Intervalo de confianza		Estadístico F	p-value
Secundaria	Z - Y	8.73	0.00	164.01	[93.99	234.03]	21.08	0.00
	Z - A	1.71	0.19	107.76	[-143.60	358.52]	0.71	0.40
	Y - A	0.39	0.53	-56.25	[-307.71	195.21]	0.19	0.66
Superior técnica	Z - Y	4.75	0.03	35.85	[-67.13	138.83]	0.47	0.50
	Z - A	0.70	0.40	-56.94	[-378.31	264.43]	0.12	0.73
	Y - A	0.00	0.97	-92.79	[-421.73	236.15]	0.31	0.58
Superior universitaria	Z - Y	21.64	0.00	-422.79	[-1011.39	165.82]	1.98	0.16
	Z - A	60.13	0.00	-1402.16	[-2212.04	-592.28]	11.52	0.00
	Y - A	16.44	0.00	-979.37	[-1822.26	-136.48]	5.19	0.02

Test U de Mann Withney de contraste de distribuciones

Nivel educativo	Grupos sociales	Test de contraste de distribución			
		Suma de rangos		Estadístico Z	p-value
		Z	Y		
Secundaria	Z - Y	476,879	271,598	4.27	0.00
	Z - A	284,144	8,852	1.63	0.10
	Y - A	125,362	6,994	0.28	0.78
Superior técnica	Z - Y	667,938	391,302	3.39	0.00
	Z - A	405,470	13,601	1.30	0.19
	Y - A	175,132	10,004	0.35	0.73
Superior universitaria	Z - Y	646,406	660,130	-1.85	0.06
	Z - A	363,848	56,139	-5.15	0.00
	Y - A	344,454	52,041	-4.10	0.00

Las pruebas estadísticas, tanto paramétricas como no paramétricas, muestran que las diferencias son significativas para el nivel universitario. Para los otros dos niveles, las pruebas paramétricas muestran que la prueba de la igualdad entre varianzas falla; como criterio de evaluación quedan, entonces, las pruebas no paramétricas, las cuales muestran que las diferencias no son significativas en los casos de los grupos Y-A en secundaria y de los grupos Z-A e Y-A en técnica. El resultado es que la hipótesis es corroborada solo en el nivel universitario.

*Obreros**Prueba F de diferencia de medias*

Nivel educativo	Grupos sociales	Test de igualdad de varianzas		Test de diferencia de medias				
		Estadístico T	p-value	Diferencia	Intervalo de confianza		Estadístico F	p-value
Primaria	Z - Y	0.38	0.54	-35.08	[-115.23	45.07]	0.74	0.39
Secundaria	Z - Y	4.01	0.05	-34.34	[-91.07	22.40]	1.41	0.24

Test U de Mann Withney de contraste de distribuciones

Nivel educativo	Grupos sociales	Test de contraste de distribución			
		Suma de rangos		Estadístico Z	p-value
		Z	Y		
Primaria	Z - Y	1,139,292	212,899	-0.93	0.35
Secundaria	Z - Y	1,653,321	774,385	-2.71	0.01

Las comparaciones relevantes son para el nivel primario y secundario, y solo para los grupos Z y Y. Utilizando solo el test no paramétrico (debido a la falta de homocedasticidad en el test paramétrico), las diferencias de salarios son estadísticamente significativas en el caso del nivel secundario, pero no así en el caso del nivel primario.

***Hipótesis 4: Diferencias de medias entre remuneraciones
e ingresos de autoempleo***

Grupo social Z

Prueba F de diferencia de medias

Nivel educativo	Test de igualdad de varianzas		Test de diferencia de medias				
	Estadístico T	p-value	Diferencia	Intervalo de confianza		Estadístico F	p-value
Primaria	72,56	0,00	443,13	[287,90	598,36]	31,31	0,00
Secundaria	24,64	0,00	420,17	[365,23	475,12]	224,64	0,00
Total	495,77	0,00	894,33	[757,69	1.030, 97]	164,58	0,00

Test U de Mann Withney de contraste de distribuciones

Nivel educativo	Test U de Mann Withney de contraste de distribuciones			
	Suma de rangos		Estadístico Z	p-value
	Autoempleados	Empleados		
Primaria	16.417.007	603.688	-11,43	0,00
Secundaria	4.943.900	2.010.686	-24,30	0,00

Los niveles de educación relevantes para la comparación son el primario, el secundario y el técnico. Las pruebas muestran que las diferencias son estadísticamente significativas en todos los casos.

Grupo social Y

Prueba F de diferencia de medias

Nivel educativo	Test de igualdad de varianzas		Test de diferencia de medias				
	Estadístico T	p-value	Diferencia	Intervalo de confianza		Estadístico F	p-value
Secundaria	0,42	0,52	157,72	[80,83	234,61]	16,17	0,00
Superior técnica	0,21	0,65	313,21	[70,70	555,73]	6,41	0,01
Universitaria	20,77	0,00	1.022,85	[502,82	1.542,88]	14,86	0,00
Total	140,96	0,00	812,48	[586,29	1.038,68]	49,57	0,00

Test U de Mann Withney de contraste de distribuciones

Nivel educativo	Test U de Mann Withney de contraste de distribuciones			
	Suma de rangos		Estadístico Z	p-value
	Autoempleados	Empleados		
Secundaria	772.113	492.733	-12,58	0,00
Superior técnica	64.839	285.864	-13,97	0,00
Superior universitaria	89.954	479.825	-12,06	0,00

Para los niveles educativos relevantes, secundario, técnico y universitario, el test paramétrico señala que las diferencias son significativas en el nivel universitario y en el agregado; en los otros casos, no hay resultados porque el test de igualdad de varianzas falla. El test no paramétrico señala que las diferencias son significativas en todos los casos.

Hipótesis 5: Diferencias de medias en tasas de exclusión del mercado laboral
Prueba chi cuadrado de contraste de proporciones

Nivel educativo	Grupos sociales	Chi-cuadrado	p-value
Primaria	Z - Y	6,855	0.00
Secundaria	Z - Y - A	4,855	0.00
Superior técnica	Z - Y	1,034	0.00
Superior universitaria	Z - Y	819	0.00

Finalmente, realizamos una prueba chi cuadrado de contraste de proporciones para comprobar la significancia estadística de las diferencias en las tasas de exclusión del mercado laboral por grupos sociales. Podemos observar que dicha diferencia es significativa para todos los niveles educativos y los grupos sociales relevantes.

Referencias bibliográficas

- AKERLOF, George y Rachel Kranton (2000) «Economics and Identity». *Quarterly Journal of Economics*, vol. 115, n.º 3, pp. 715-733.
- BANCO MUNDIAL (2001) *World Development Report*. Washington D. C.
- BARRO, Robert y Xavier SALA-I-MARTIN (1995) *Economic Growth*. McGraw Hill.
- BECKER, Gary y Kevin MURPHY (2001) *Social Economics: Market Behavior in a Social Environment*. Harvard University Press.
- BEHRMAN, Jere, Alejandro GAVIRIA y Miguel SZÉKELY (2001) «Intergenerational Mobility in Latin America». Working Paper 452. Washington D. C.: Banco Interamericano de Desarrollo.

- BORJAS, George (1992) «Ethnic Capital and Intergenerational Mobility». *Quarterly Journal of Economics*, vol. 107, n.º 1, pp. 123-150.
- BOWLES, Samuel, Herbert GINTIS y Melissa GROVES (eds.) (2005) *Unequal Chances: Family Background and Economic Success*. Princeton University Press.
- CHISWICK, Barry (1980) «The Earnings of White and Coloured Males Immigrants in Britain». *Económica*, febrero, pp. 81-87.
- COUCH, Kenneth y Thomas DUNN (1997) «Intergenerational Correlation in Labor Markets Status: A Comparison of the United States». *Journal of Human Resources*, vol. 32, n.º 1, invierno, pp. 210-232.
- CUETO, Santiago, Enrique JACOBY y Ernesto POLLIT (1997) «Rendimiento escolar de niños y niñas en zonas rurales y urbanas del Perú». *Revista de Psicología de la Universidad Católica del Perú*, vol. XV, n.º 1, pp. 115-133.
- DAHAN, Momi y Alejandro GAVIRIA (1999) «Sibling Correlations and Social Mobility in Latin America». Working Paper 395, Office of the Chief Economist. Washington D. C.: Banco Interamericano de Desarrollo.
- DEININGER, Klaus y Lyn SQUIRE (1996) «A New Data Set Measuring Inequality». *The World Bank Economic Review*, vol. 10, n.º 3, setiembre, pp. 565-591.
- FIGUEROA, Adolfo (2003) *La sociedad sigma. Una teoría del desarrollo económico*. Lima-México D. F.: Fondo Editorial de la Pontificia Universidad Católica del Perú-Fondo de Cultura Económica.
- FIGUEROA, Adolfo y Manuel BARRÓN (2005) «Inequality, Ethnicity and Social Disorder in Peru». Working Paper 8. CRISE, Universidad de Oxford.
- FIGUEROA, Adolfo, Viviana CRUZADO y Claudia SÁNCHEZ (2007) «El problema del empleo en una sociedad sigma: anexo metodológico y base de datos». Documento de trabajo. Lima: Departamento de Economía, Pontificia Universidad Católica del Perú, Lima.
- GARDNER, Howard (1999) *Intelligence Reframed: Multiple Intelligences for the 21st Century*. Basic Books.
- HALL, Gillette and Anthony PATRINOS (2005) *Indigenous Peoples, Poverty, and Human Development*. Palgrave.
- HANUSHEK, Eric y Javier LUQUE (2003) «Efficiency and Equity in Schools around the World». National Bureau of Economic Research, Working Paper 8949. Cambridge, Massachusetts.
- HUDSON, R. A. (1996) *Sociolinguistics*. Cambridge University Press.
- LI, Hongyi, Lyn SQUIRE y Heng-fu ZOU (1998) «Explaining International and Inter-temporal Variations in Income Inequality». *Economic Journal* 108 (1), enero, pp. 26-43.
- LOAIZA, Carlos (2006) «Revisión de la literatura internacional sobre educación y mercados laborales». Lima: Departamento de Economía, Pontificia Universidad Católica del Perú (manuscrito).

- LUCAS, Robert (1990) «Why Doesn't Capital Flow from Rich to Poor Countries». *American Economic Review*, vol. 80, n.º 2, pp. 92-96.
- MALONEY, William (2004) «Informality Revisited». *World Development*, vol. 22, n.º 7, julio, pp. 1159-1178.
- MINISTERIO DE EDUCACIÓN DEL PERÚ (2005) Evaluación nacional 2004. Lima [conferencia de prensa del 28/10/2005].
- ÑOPO, Hugo, Jaime SAAVEDRA y Máximo TORERO (2004) «Ethnicity and Earnings in Urban Peru». Discussion Paper 980. Institute for the Study of Labor, IZA, Bonn.
- RATEY, John (2002) *A User's Guide to Brain*. Vintage Books.
- RIVERA, Iván (1979) «Los determinantes de la calidad de la educación en el Perú. Un estudio ECIEL». Publicaciones del Centro de Investigaciones Sociológicas, Económicas, Políticas y Antropológicas (Cisepa). Documento de trabajo 44. Lima: Pontificia Universidad Católica del Perú.
- ROUSSEAU, Jean-Jacques (1984 [1755]) *A Discourse on Inequality*. Traducción del francés por Maurice Cranston. Penguin Books.
- SHAPIRO, Carl y Joseph STIGLITZ (1984) «Equilibrium Unemployment as a Worker Discipline Device». *The American Economic Review*, vol. 74, n.º 3, junio, pp. 433-444.
- SHAPIRO, David (1984) «Wage Differentials among Black, Spanic, and White Young Men». *Industrial and Labor Relations Review*, vol. 37, n.º 4, julio, pp. 570-581.
- SOLON, Gary (1992) «Intergenerational Income Mobility in the United States». *The American Economic Review*, vol. 82, n.º 3, junio, pp. 393-408.
- TODD, Petra y Kenneth WOLPIN (2003) «On the Specification and Estimation of the Production Function for Cognitive Achievements». *Economic Journal*, vol. 113, n.º 45, febrero, pp. F3-F33.
- TREJO, Stephen (1997) «Why Mexican American Earn Low Wages?». *Journal of Political Economy*, vol. 105, n.º 6, diciembre, pp. 1235-1268.
- WRIGHT, Erick (1997) *Class Counts. Comparative Studies in Class Analysis*. Cambridge University Press.
- Zimmerman, David (1992) «Regression toward Mediocrity in Economic Stature». *The American Economic Review*, vol. 82, n.º 3, junio, pp. 409-419.