

RELACIÓN DE KUZNETS EN AMÉRICA LATINA. EXPLORANDO MÁS ALLÁ DE LA MEDIA CONDICIONAL

JAVIER ALEJO

RESUMEN

Este trabajo estudia la relación de Kuznets para un conjunto de países latinoamericanos con el objetivo de caracterizar el cambio en la tendencia de los indicadores de desigualdad. Los cambios distributivos están vinculados con múltiples factores observables asociados al desarrollo, tales como el crecimiento económico o el capital humano. Sin embargo existen muchos otros que no son directamente observables tales como la historia o la idiosincrasia de cada región. La incorporación de los cuantiles condicionales al análisis de la relación de Kuznets puede ayudar a estudiar si éste conjunto de factores inobservables puede generar senderos de desarrollo y equidad heterogéneos.

Clasificación JEL: C15, D31, I38

Palabras Clave: Hipótesis de Kuznets, Desarrollo económico, Desigualdad, América Latina, Datos en panel, Regresión por cuantiles.

ABSTRACT

This paper studies the Kuznets relationship for a set of Latin American countries with the aim of characterizing the change in the evolution trend of the regional inequality indicators. Distributional changes are related with a lot of observable factors linked to development variables such as economic growth or human capital stock. However, there are many non-measurable inequality determinants like history or regional idiosyncrasies. The incorporation of conditional quantiles to the Kuznets relationship analysis can help to explain whether the unobservable factors can generate alternative paths for development and inequality.

JEL Classification: C15, D31, I38

Keywords: Kuznets hypothesis, Economic development, Inequality, Latin America, Panel data, Quantile regression.

RELACIÓN DE KUZNETS EN AMÉRICA LATINA. EXPLORANDO MÁS ALLÁ DE LA MEDIA CONDICIONAL¹

JAVIER ALEJO²

I. Introducción

Un hecho estilizado para los países de América Latina es que los indicadores de desigualdad han cambiado su tendencia en la última década. Varios documentos publicados por distintos organismos de investigación registran este cambio distributivo en la mayoría de los países de la región conjuntamente con un proceso de recuperación económica que comienza en 2003 (Gasparini et al. 2010). Mejoras en las condiciones externas conjuntamente con cambios en las agendas de política económica en varios de los países latinoamericanos favorecen a este nuevo panorama social latinoamericano (CEPAL 2006). Tratar de entender el trasfondo de este hecho es relevante desde varias perspectivas: por un lado es un fenómeno interesante *per se* ya que los países de Latinoamérica son los de mayor desigualdad en comparación con el resto de las regiones globales y por lo tanto el estudio de sus determinantes ayuda a la caracterización del fenómeno, al menos en su historia reciente. Por otro lado, este cambio en las agendas de política económica sugiere que una mejora en la distribución de los ingresos presupone un progreso en términos de bienestar social. Por lo tanto, es necesario para los hacedores de política contar con herramientas que le sirvan de guía en el proceso de desarrollo. Si bien son múltiples las causas que podrían asociarse a esta mejora distributiva reciente, este trabajo trata de cuantificar el protagonismo que ha tenido la recuperación económica de la región durante este proceso redistributivo.

¹ Este trabajo es parte del plan de tesis del Doctorado en Economía de la UNLP. Agradezco a mi director Walter Sosa Escudero, a Leonardo Gasparini, Guillermo Cruces, Pablo Gluzmann y Marcelo Bérngolo por sus valiosos comentarios y observaciones. También a todas las sugerencias relevantes recibidas durante los Seminarios del Doctorado y a dos árbitros anónimos que aportaron diversos comentarios. Esta investigación se benefició del financiamiento de las Becas Internas de Postgrado otorgadas por el CONICET. Cualquier error es de mi exclusiva responsabilidad.

² Centro de Estudios Distributivos, Laborales y Sociales, Facultad de Ciencias Económicas, Universidad Nacional de La Plata. javier.alejo@depeco.econo.unlp.edu.ar.

Desde que Simon Kuznets presentara su conjetura sobre una relación de U invertida entre la desigualdad de ingresos y el crecimiento económico, las distintas fuentes de interrelación entre la inequidad y el desarrollo han ocupado un lugar cada vez más importante dentro de la literatura académica (Moran, 2005). Como era de esperar, este tipo de relación tuvo un gran impacto tanto en la visión de los *policy makers* sobre las economías capitalistas como en las opiniones dentro del ámbito académico. En el primer caso la conjetura implicaba que la política económica sólo debía preocuparse por el crecimiento económico dado que en algún momento el desarrollo llevaría hacia una mejora en la distribución del ingreso. En el contexto académico, la conjetura de Kuznets planteó un hecho estilizado que debía ser explicado. Si bien el campo abordado por la literatura para estudiar este problema ha sido en su mayor parte empírico, han surgido diversos conceptos teóricos importantes durante el proceso de búsqueda que confluyen hacia un conjunto de explicaciones posibles para la relación conjeturada y que ha dado sus frutos en las distintas teorías del desarrollo y el crecimiento económico. Por lo tanto, entender las distintas etapas distributivas del proceso de desarrollo puede ayudar a establecer las fuerzas principales que actúan sobre el cambio observado en la evolución reciente de la desigualdad para los países Latinoamericanos.

En lo que respecta al campo empírico, la complejidad para caracterizar correctamente el proceso de desarrollo no es menor. Se han hecho notables esfuerzos para tratar de encontrar evidencia a favor o refutar definitivamente la conjetura de Kuznets mediante distintas metodologías y estructura de los datos. La gran mayoría de esos trabajos se concentran en el análisis de la relación a través de la media condicional, es decir estimando una relación promedio entre desigualdad y desarrollo. Sin embargo, extender el análisis hacia otras partes de la distribución condicional puede ser interesante desde el punto de vista analítico y conceptual al considerar la posibilidad de que existan trayectorias de desarrollo heterogéneas. En lo que respecta a los datos utilizados para obtener las estimaciones, los trabajos empíricos más destacados presentan como un punto de relevancia la buena calidad en la selección y confiabilidad de los datos a utilizar. Como se verá, éste aspecto es ampliamente considerado por este estudio.

El principal resultado del trabajo es que, aún luego de controlar por los efectos idiosincráticos de cada región y por algunos de los determinantes del desarrollo, las estimaciones encuentran una relación de Kuznets para las regiones de América Latina. Sin embargo, al analizar la relevancia económica

del resultado se observa que una vez que se incorpora al análisis las distintas dimensiones del desarrollo, la relación de Kuznets se torna ilusoria debido a la incompatibilidad con los datos de la región.

El trabajo se ordena de la siguiente manera: la sección II hace una breve reseña de la literatura sobre la desigualdad y el desarrollo económico; la sección III ofrece una descripción de los datos a utilizar, la sección IV explica la metodología econométrica empleada, la sección V muestra los resultados obtenidos y finalmente en la sección VI se presentan las conclusiones del trabajo.

II. Desigualdad y crecimiento

II.1. Aspectos teóricos

La conjetura de Kuznets propone que existe una relación de U invertida entre el nivel de desigualdad y el desarrollo económico. Es decir, plantea que durante proceso económico existe una primera etapa de subdesarrollo en la cual la desigualdad de ingresos es baja, luego una etapa posterior de crecimiento en donde surge una mayor tensión distributiva para finalmente pasar al desarrollo económico con mejoras en la equidad. Simon Kuznets (1955) ejemplificó esta regularidad empírica con datos para el Reino Unido, los Estados Unidos, Puerto Rico, India y Ceylan.

Varios autores modelaron la relación entre desigualdad y el nivel de los ingresos bajo distintos supuestos y mecanismos de transmisión. El mismo Kuznets (1955) trató de explicar a la relación de U invertida con un modelo migratorio sencillo en donde la mano de obra, en búsqueda de mejores salarios, se muda desde el ámbito rural hacia el área urbana. Este modelo sería adaptado a versiones más sofisticadas de cambios de empleo entre sectores como consecuencia de los diferenciales de ingresos, como por ejemplo la migración desde un sector con baja tecnología (productiva o financiera) hacia otro con técnicas más modernas (Greenwood y Jovanovic, 1990).

Otras explicaciones teóricas que vinculan el desarrollo con el crecimiento provienen de distintas ramas de la economía: la teoría del crecimiento endógeno (Romer, 1986), las teorías de votación y conflicto social (Saint Paul, 2000; Perotti, 1996), modelos con imperfecciones del mercado de capitales

(Banerjee y Newman, 1993), entre otras explicaciones.³ Es menester aclarar que gran parte de estas teorías hacen referencia al crecimiento económico más que al nivel de ingresos. Ambos conceptos están claramente relacionados y es por ello que esta breve reseña trata de reflejar la multiplicidad de conceptos detrás del complejo mecanismo que determina la desigualdad y el desarrollo.

La teoría del crecimiento endógeno vincula al crecimiento de la economía con su dotación de capital humano a través de las ganancias en la productividad laboral y sus externalidades positivas. A su vez, la asignación del capital humano determina los diferentes ingresos de la economía por medio de los retornos a la educación. En forma sintética, esta teoría predice una complementariedad entre el crecimiento y la equidad, por lo tanto se esperaría observar altos niveles de desarrollo junto con bajos niveles de desigualdad.

El canal establecido por las teorías de votación para encadenar crecimiento y desigualdad es a través del sistema electoral y la influencia del esquema impositivo sobre la inversión privada. Las preferencias de los votantes acerca de la progresividad de los impuestos dependen de su posicionamiento en la distribución del ingreso personal. Por lo tanto, una mayor desigualdad implicará que se vote a favor de impuestos más redistributivos y en consecuencia la mayor presión tributaria desincentiva a la inversión y el crecimiento a largo plazo. También dentro de esta literatura se encuentran las argumentaciones vinculadas al conflicto social entre los agentes económicos: una elevada desigualdad puede generar sectores insatisfechos y crear un ambiente de inestabilidad política y amenazas contra los derechos de propiedad; en consecuencia, el crecimiento se detiene por el efecto del pesimismo sobre las decisiones de inversión a largo plazo. Las imperfecciones del mercado de capitales son otro factor clave en el proceso de desarrollo.

Las restricciones al crédito hacen que las decisiones de inversión en capital físico y humano dependan de los ingresos familiares y por lo tanto, desde la perspectiva de una economía con movilidad perfecta de los capitales, las decisiones productivas serán sub-óptimas. Es decir, el stock de capital será escaso para algunas familias (pobres) y ocioso para otras (ricos). El efecto de ésta distorsión en el mercado crediticio será mayor cuanto más elevada sea la

³ Es extensa la lista de autores que componen las distintas teorías del desarrollo, por nombrar sólo algunos: Lewis (1954), Kaldor (1956), Persson y Tabellini (1994), Galor y Zeira (1993), Barro (1999), Rodrik (1999), Barro y Becker (1988); Becker, Murphy y Tamura (1990), entre otros.

desigualdad de los recursos y por lo tanto la teoría predice una relación negativa entre la desigualdad y el desarrollo.

Cabe destacar que si bien cada una de estas teorías explica alguna dimensión de la relación entre el desarrollo económico y la desigualdad, no existe una teoría unificadora que englobe a todos estos aspectos en una forma compacta y elegante. De hecho, se han descrito otros canales entre desigualdad y desarrollo basados en teorías demográficas como la de fecundidad endógena (Becker et al., 1990) o segregación residencial (Ortega, 2003), entre otros. Es por eso que la literatura ha utilizado fundamentalmente la investigación empírica tratando de encontrar evidencia a favor o en contra de la conjetura planteada por Kuznets.

II.2. Antecedentes empíricos

La literatura empírica sobre la relación entre desigualdad y crecimiento es extensa y los resultados obtenidos no siempre llegan a las mismas conclusiones. El indicador de desigualdad más empleado es el coeficiente de Gini dado que es fácil de interpretar y no hay diferencias cualitativas demasiado importantes por usar otros indicadores. Por lo general se utilizan modelos de regresión en los cuales el indicador de desigualdad es una función del ingreso per cápita del país. La especificación puede ser sencilla o incorporar otras variables consideradas como determinantes y/o relacionadas con la desigualdad, según las distintas teorías comentadas anteriormente. Para capturar la forma de U invertida es usual utilizar una forma funcional polinómica de segundo grado en el ingreso per cápita, medido en escala logarítmica. La prueba de hipótesis usual para evaluar la relación de Kuznets es analizar la significatividad estadística del coeficiente asociado a la variable cuadrática. Sin embargo, como veremos en la sección de resultados, se debe ser cuidadoso al interpretar el ajuste cuadrático (cóncavo) como una evidencia directa a favor de la curva de Kuznets.

Los primeros trabajos cuantitativos sobre la relación de Kuznets utilizaban datos de corte transversal (*cross-section*) entre países para realizar las estimaciones: Kuznets (1955), Fields (1980), Anand y Cambur (1993); aunque hay trabajos recientes como Gasparini y Gluzmann (2012) que emplean ésta estructura de los datos provenientes de la encuesta Gallup. Una buena parte de la evidencia a favor de la relación de tipo Kuznets se obtenía por la presencia

de América Latina dentro de los países de la muestra, debido a que la región se caracteriza por economías de ingreso per cápita medio y elevada desigualdad (Fields, 2001). Es decir, se sospechaba que en lugar de ser un hecho estilizado, la relación de U invertida era producto de la estructura de los datos, marcada por las características idiosincráticas de cada observación.

La mejora en la recopilación de estadísticas permitió que los trabajos empíricos más recientes utilicen el formato de datos en panel para realizar las estimaciones, y por lo tanto se abrió un camino para corregir el efecto del corte transversal, problema latente en los trabajos anteriores. Con ésta nueva disposición en la configuración de los datos los resultados cambiaron sustancialmente ya que, al depurar los efectos fijos por país en las estimaciones, la hipótesis de U invertida se vio debilitada. Algunos ejemplos son: Deininger y Squire (1996) y Angeles (2010), entre otros.

En un meta-análisis, Dominics et al. (2008) utiliza como insumo las estimaciones publicadas en la literatura y observa que las conclusiones de los papers empíricos sobre la conjetura de Kuznets difieren en base a las siguientes dimensiones:

- Estructura de los datos: los trabajos que utilizan un *cross-section* encuentran una relación más fuerte entre desigualdad y crecimiento.
- Método de estimación: el uso de efectos fijos tiende a debilitar la evidencia a favor de la relación de Kuznets mientras que el método de efectos aleatorios, variables instrumentales y momentos generalizados no difieren sistemáticamente del estimador de OLS.
- Características de los datos: utilizar bases de datos de calidad baja afecta significativamente las conclusiones de los papers.
- Nivel geográfico: por lo general, los trabajos que analizan desigualdad y desarrollo comparando las regiones dentro de un país tienden a encontrar una relación más débil que aquellos que utilizan países como unidad de análisis.

Es probable que las variables involucradas en el análisis empírico de la curva de Kuznets se determinen conjuntamente, es decir, puede existir un

problema de endogeneidad en la estimación. Es por eso que se habla de una relación entre el desarrollo y la desigualdad en lugar de un efecto causal de uno sobre el otro. Cómo se verá en la sección metodológica, si la fuente de endogeneidad es invariante en el tiempo, la misma puede solucionarse mediante la inclusión de efectos fijos en la estimación. Sin embargo, si aún persiste la fuente de endogeneidad en el término de error remanente, se debe emplear algún otro método. Dentro de las conclusiones del meta-análisis de Dominics et al. (2008) se observa que las estrategias de estimación que consideran correcciones de endogeneidad tales como VI o GMM, no difieren sistemáticamente de las estimaciones de OLS. Sin embargo, si bien esta evidencia es interesante, se debe ser cauto con respecto a las conclusiones de causalidad en este tipo de trabajo.

II.3. Enfoque de cuantiles condicionales

La literatura empírica sobre la Curva de Kuznets tradicionalmente ha girado en torno a la relación entre desigualdad y desarrollo vista como un patrón promedio, es decir como una única relación que describiría el comportamiento de un país o región promedio en ausencia de otras perturbaciones. Sin embargo, es posible que la Curva de Kuznets pueda presentar trayectorias heterogéneas, dependiendo de los distintos factores que afectan a la desigualdad y el desarrollo. Considérese una ilustración numérica basada en una economía con dos sectores $S = R, U$ (rural y urbano, por ejemplo) similar a la presentada por Kuznets (1955). Supóngase que dentro de cada sector S la mitad gana un ingreso bajo $w_B(S)$ y la otra mitad un ingreso alto $w_A(S)$, luego el ingreso promedio en cada sector es $\mu_w(S) = [w_A(S) + w_B(S)]/2$ y la desigualdad de ingresos dentro del sector (medida por el índice de Gini) es $G_w(S) = 0.25[w_A(S) - w_B(S)]/\mu_w(S)^4$. Sin pérdida de generalidad, considérese que $w_A(R) < w_B(U)$, asegurando que el ingreso promedio del área urbana sea mayor que el rural. La proporción de personas que trabajan en la ciudad es ψ y por lo tanto el ingreso promedio de toda la economía es $\mu_w = (1 - \psi)\mu_w(R) + \psi\mu_w(U)$. Según la lógica planteada por Kuznets, en la etapa inicial del desarrollo todos trabajan en el sector rural, es decir $\psi = 0$ y por lo tanto el ingreso promedio de

⁴ El índice de Gini se calcula a partir de la fórmula $2Cov[w, F(w)]/E(w)$ aplicado a este caso particular donde $F(w_B) = 0.5$ y $F(w_A) = 1$. Ver en el Apéndice para más detalles.

la economía coincide con el del sector rural. A medida que la población se muda a la ciudad en busca de mejores ingresos, el ingreso promedio de la economía aumenta lentamente hasta llegar, en el otro extremo ($\psi = 1$), a coincidir con el promedio del sector urbano. ¿Qué ocurre con la desigualdad? El Gráfico 1 muestra ejemplos numéricos del índice de Gini en cada momento de la transición para economías con distintas configuraciones en la desigualdad al interior de cada región (Tabla 1).⁵ En todas las economías ejemplificadas, el ingreso promedio del sector rural es 10 mientras que en el sector urbano es 90, por lo tanto la economía en su conjunto oscilará entre estos dos valores de ingreso. En todos los casos aparece un patrón de U invertida como producto de la transición de un sector al otro. Los casos *A*, *B* y *C* son economías en donde la desigualdad entre ellas difiere pero al interior de cada una de las mismas presenta homogeneidad en la equidad sectorial. El caso *B* representa a una economía con desigualdad crónicamente alta, mientras que el caso opuesto es ilustrado por *C*. Las tres curvas de Kuznets son paralelas y por lo tanto, más allá del posicionamiento vertical de la misma, la forma de U invertida es bien capturada por el caso intermedio (economía *A*). En cambio, en otros casos donde la desigualdad entre sectores difiere, la transición genera senderos de desarrollo y desigualdad heterogéneos. Los casos *D* y *E* ejemplifican esa idea: si bien ambos tienen la forma de U invertida, el primero es un caso en donde la desigualdad al final del desarrollo es mayor que la inicial, mientras que el segundo es el caso opuesto.

Si bien este simple ejercicio es solo una visión muy acotada de la economía, muestra que los distintos factores subyacentes detrás del proceso de desarrollo y desigualdad pueden generar senderos heterogéneos. Muchos de esos factores son cuantificables empíricamente y pueden ser considerados en las estimaciones mediante un análisis condicional, sin embargo, hay otra multiplicidad de factores que pueden afectar al sendero de desarrollo y de los que no se dispone de una observación directa. Éste es un aspecto claramente interesante para las economías de América Latina cuya historia económica y social (aspecto que claramente es de difícil medición) ha estado marcada por innumerables inequidades. Cómo se explicará en la sección IV, el estudio de la relación de Kuznets a través de la metodología de cuantiles condicionales permite estudiar el efecto de estos factores no observables sobre la conjetura

⁵ La fórmula del Gini agregado así como una explicación más detallada del ejemplo numérico se muestran en el Apéndice.

de Kuznets. La evidencia que surja de este tipo de enfoque es interesante tanto desde el punto de vista académico (abriendo nuevos interrogantes) como para la política económica (como herramienta de análisis). En consecuencia, este trabajo empírico extiende a sus antecesores al estudiar la conjetura de Kuznets a través de la metodología de regresiones por cuantiles.

III. Datos

Como en todo análisis empírico, es deseable contar con una base de datos de calidad aceptable que permita realizar comparaciones de las estadísticas de desigualdad y desarrollo tanto entre los distintos países como en el tiempo. Por lo general, los datos utilizados en la literatura provienen de distintas fuentes: institutos nacionales de estadística, organismos internacionales, ministerios de hacienda, entre otros. Si bien cada una de estas instituciones probablemente mantenga una coherencia de criterios en la generación de estadísticas, el procesamiento de datos así como el cálculo de indicadores contienen una amplia cantidad de decisiones metodológicas que no siempre son explícitas en la información publicada y puede llevar a realizar comparaciones incorrectas. El caso de los índices de desigualdad es un claro ejemplo de éste problema. Algunos países computan el indicador de Gini utilizando el ingreso equivalente del hogar como variable de bienestar, mientras que otros lo calculan en base al consumo o al ingreso total familiar.⁶ El trabajo de Deininger y Squire (1996) propone tres requisitos sobre la aptitud de los datos para que sean considerados como información de buena calidad: (i) que estén calculados con encuestas de hogares en lugar de estadísticas basadas en cuentas nacionales, (ii) que tengan una amplia cobertura de todas las fuentes de ingresos y (iii) deben ser descriptivas de la población nacional más que ser representativas únicamente de las áreas urbanas o rurales. Si bien estos requisitos permiten algún grado de coherencia en la comparación de los datos, aún distan de garantizar un buen grado de precisión y comparabilidad debido a las diferencias metodológicas en la generación de estadísticas por parte de cada institución.

⁶ El ingreso equivalente es el ingreso total del hogar ajustado por factores demográficos de acuerdo a la composición de sus miembros. En general se utilizan escalas de necesidades de kilo-calorías por edad y género, pero la metodología para su cómputo no siempre es aclarada por los institutos de estadísticas.

En este trabajo se utiliza la base SEDLAC que cuenta con información a nivel de microdatos que provienen de las encuestas de hogares, para distintos países de América Latina y que fueron procesados con una metodología común. La base SEDLAC es un proyecto conjunto entre el CEDLAS y el grupo de Pobreza y Género para América Latina del Banco Mundial.⁷ Si bien no es posible homogeneizar completamente los aspectos relacionados con el diseño de las encuestas (cuestionarios, período de recolección, etc.), el acceso a los microdatos hace posible construir una serie de indicadores de desigualdad y desarrollo con criterios uniformes, otorgando un buen nivel de comparabilidad entre países y en el tiempo. Por lo tanto, éste tratamiento estandarizado de la base SEDLAC permite trabajar con una calidad de información superior a la utilizada por Deininger y Squire, al menos para algunos países de América Latina.

Dado que el trabajo está acotado a los países latinoamericanos, la dimensión *cross-section* es reducida. La estrategia seguida es generar un panel de indicadores de desigualdad y desarrollo a nivel de regiones sub-nacionales, definidas según el instituto de estadística de cada país (ver Tabla 2 con el listado de regiones consideradas dentro de cada país). Por otro lado, en la construcción del panel existe un *trade-off* en la elección de la dimensión de corte transversal *versus* la dimensión tiempo. Para la determinación del número de periodos a considerar se siguieron dos criterios alternativos:

- *Panel corto*: se compone de 61 regiones que pertenecen a 10 países (Argentina, Brasil, Costa Rica, El Salvador, Honduras, México, Paraguay, Perú, Uruguay y Venezuela) para 6 períodos bianuales que van desde 1995 hasta 2006.
- *Panel largo*: comprende solamente 27 regiones de Argentina (1992 a 2009), Brasil (1990 a 2008), Costa Rica (1989 a 2008), Honduras (1991 a 2009) y Uruguay (1989 a 2009). La dimensión tiempo representa entonces más de 17 períodos anuales.

⁷ Para más referencias sobre la base SEDLAC ver metodología en:
<http://sedlac.econo.unlp.edu.ar>

El indicador de desigualdad utilizado es el índice de Gini, calculado en base al ingreso per cápita familiar.⁸ La Tabla 3 muestra las estadísticas básicas de desigualdad e ingreso per cápita mensual (medido en PPA de 2005) para cada una de las regiones utilizadas. Claramente, la disponibilidad de datos por regiones provee una gran variabilidad en los datos y observaciones. En el Gráfico 2 se muestra la evolución del índice del Gini, calculado para los países del segundo panel. Como se observa, a excepción de Brasil que muestra una evolución que decrece suavemente, la tendencia de la desigualdad en todos los países ha sido creciente hasta la primera mitad de la década de los 2000. Luego, el valor del índice comienza a disminuir (Argentina, Uruguay y Honduras) o a permanecer sin cambios (Costa Rica). Es decir, hay un quiebre en la evolución de la desigualdad en los primeros años de los 2000.

Para la construcción del ingreso per cápita se utilizaron los ingresos reportados en las encuestas de hogares ya que esto permite tener la variabilidad muestral necesaria para desagregar los indicadores a niveles subnacionales. Deben mencionarse dos aspectos importantes del uso de encuestas de hogares: por un lado los ingresos provenientes de las encuestas presentan diferencias con los que surgen del sistema de cuentas nacionales, lo cual podría presentar algún problema de error de medición; por otro lado, la sensibilidad en el reporte de los ingresos laborales a la estabilidad en el mercado de empleo puede ser una fuente de ruido adicional en el cálculo del nivel económico de cada región.

Con respecto al primer punto, son múltiples las causas de la discrepancia entre las cuentas nacionales y las encuestas de hogares: subdeclaración de los ingresos de capital, economía informal, ingresos de firmas dentro del país que envían sus utilidades al exterior, tratamientos de los ingresos nulos, entre otras. Adicionalmente, los países utilizan distintas metodologías para medir el poder adquisitivo (típicamente, consumos, ingresos, gastos) con claras consecuencias en la comparabilidad de los datos. Actualmente, la literatura cuenta con dos líneas principales en cuanto al ajuste por cuentas nacionales. Por un lado, no es claro que el ejercicio de ajustar los ingresos de las encuestas por cuentas nacionales sea la mejor alternativa ya que si bien se esperarían que los países sigan los procedimientos estandarizados de la División de Estadísticas de las

⁸ Para una lectura más sencilla de las estimaciones de regresión se utiliza el índice de desigualdad multiplicado por 100, por lo tanto el valor 0 indica perfecta igualdad y 100 representa la máxima desigualdad.

Naciones Unidas, en la práctica, el alineamiento a estas normas contables por parte de los países en desarrollo son, en general, bastante dispares (Chen y Ravallion, 2008). Por otro lado, la alternativa de usar microdatos para reconstruir los ingresos tampoco está exenta de fallas, dado que el diseño de las encuestas realizadas dentro de cada país (cuestionarios, muestreo, etc.) varían de un país a otro. Si bien es posible aplicar reglas homogéneas al procesamiento de los microdatos, tal como las aplicadas en la base SEDLAC, esto sería solo una solución parcial a este problema debido a la discrepancia en las encuestas antes mencionada es *ex ante*. Por lo tanto, no es obvio en la literatura evaluar qué alternativa es la mejor, dadas las discrepancias en ambas mediciones.⁹ Por lo general, la literatura ha utilizado los ingresos de cuentas nacionales para estudiar la relación de Kuznets con datos de todo el mundo. Sin embargo, dentro del marco de este trabajo que se focaliza en el estudio de América Latina, la disponibilidad de los datos es un problema no menor. Es aquí donde el uso de microdatos permite contar con un indicador del nivel de ingresos de los habitantes para cada región latinoamericana con un grado de comparabilidad interna aceptable. Naturalmente, se debe ser cauto al comparar los resultados que se desprenden de este trabajo con los de sus antecesores.

En segundo lugar, la estabilidad en el mercado de empleos tiene otro efecto adicional sobre la medición de los ingresos mediante las encuestas de hogares. Si existe una amplia movilidad entre ocupados y desocupados, el salario reportado en las encuestas puede reflejar tan solo una situación temporal del individuo que no es representativa de su capacidad real para generar recursos. Para contemplar de alguna manera este efecto, se incluye dentro de los modelos de regresión a la tasa de desempleo como variable de control.

En síntesis, a partir de la base SEDLAC se construyen las siguientes variables de desarrollo a nivel de regiones sub-nacionales: promedio de años de educación, tasa de desempleo, porcentaje de población adulta y la participación del sector industrial y del sector público en el empleo.¹⁰ También se incluyen otras variables de desarrollo a nivel nacional que provienen de los

⁹ Se remite a Gasparini y Gluzmann (2012) para una discusión más detallada sobre las diferencias en el uso de cuentas nacionales y encuestas de hogares.

¹⁰ Esta última variable también intenta morigerar el problema de comparabilidad regional de los ingresos como consecuencia de las políticas redistributivas a través del empleo público.

World Development Indicators (WDI 2010, Banco Mundial): grado de apertura, tasa de inflación y el porcentaje de población urbana.¹¹

IV. Metodología de estimación

Como fue mencionado, los trabajos recientes sobre desigualdad y desarrollo utilizan por lo general la estructura de datos en panel para realizar sus estimaciones dado que, además de representar un mayor número de observaciones para el cálculo de estimadores, la importancia fundamental reside en el hecho de que la hipótesis de Kuznets está planteada en un contexto dinámico. Al utilizar datos de corte transversal, el supuesto implícito es que los países más pobres se comportarán como los más avanzados una vez que recorran todo el sendero del desarrollo. Este supuesto puede ser razonable en cierto contexto, pero existen factores idiosincráticos a cada país que están relacionados con el desarrollo y que puede arrojar estimaciones incorrectas. Algunos ejemplos son los niveles de corrupción, la historia particular de cada país, los juicios de valor sobre la equidad que cada sociedad considera como aceptable, entre otros. Es por eso que los datos en panel permiten un control por tales características inherentes a cada región, y por lo tanto, una estimación de la relación entre desigualdad y desarrollo más confiable.

IV.1. Media condicional

El estudio de la relación entre la desigualdad y el desarrollo dentro de la literatura empírica se ha basado principalmente en la media condicional. Algunos ejemplos son los trabajos de Li y Zou (1998), Barro (1999), Forbes (2000), Banerjee y Duflo (2000b), Mbabazi et al. (2001), Panizza (2002), Castelló (2004) y Benjamin et al. (2006). Siguiendo esta línea de investigación se procedió a estimar dos modelos de regresión para la media condicional:

¹¹ Si bien es cierto que en la mayoría de las encuestas de hogares puede estimarse la proporción de población rural, la encuesta de Argentina es realizada en aglomerados urbanos. No obstante, las estadísticas de desigualdad para el área urbana calculadas con la EPH no distan demasiado de las calculadas para población total con otras encuestas de cobertura nacional (Gasparini y Cruces, 2008).

$$g_{rt} = \beta_1 y_{rt} + \beta_2 y_{rt}^2 + x'_{rt} \delta + u_r + \varepsilon_{rt} \quad (1)$$

$$g_{rt} = f(y_{rt}) + x'_{rt} \delta + u_r + \varepsilon_{rt} \quad (2)$$

para todo $r = 1, \dots, R$ y $t = 1, \dots, T$, donde:

- g_{rt} : coeficiente de Gini multiplicado por 100 de la región r en el periodo t ,
- y_{rt} : ingreso promedio (en logaritmos) de la región r en el periodo t ,
- x_{rt} : conjunto de otros indicadores de desarrollo para la región r en el periodo t ,
- u_r : efecto fijo idiosincrático (no observado) en la región r .
- ε_{rt} : error aleatorio que satisface $E(\varepsilon_{rt}) = 0$ y $V(\varepsilon_{rt}) = \sigma^2$.

El modelo (1) es la típica especificación cuadrática para poner a prueba la hipótesis de Kuznets, mientras que el modelo (2) es menos usual en la literatura. Este último es un modelo semi-paramétrico en donde $f(y_{rt})$ es una función desconocida, es decir, no se asume ninguna forma funcional específica para la relación entre el ingreso y la esperanza condicional de la desigualdad. El modelo (2) será utilizado para analizar la bondad de ajuste del modelo (1).

Con respecto a la forma funcional paramétrica, se esperaría que si la curva de Kuznets es verdadera esto se vea reflejado en un ajuste cuadrático cóncavo de la variable y_{rt} , es decir $\beta_2 < 0$. Sin embargo, se debe ser cuidadoso en interpretar un buen ajuste cuadrático con la aceptación inmediata de la hipótesis de Kuznets dado que el mismo puede que tan sólo sea el reflejo de una relación cóncava sin que dé lugar a una curva con forma de U invertida. Este aspecto se discutirá nuevamente en la sección de resultados.

De la misma manera que en el análisis *cross-section*, la idea implícita en ambos modelos es que la relación entre el desarrollo y la desigualdad es un patrón común entre países: tanto los coeficientes β_1 y β_2 como la función $f(\cdot)$ es la misma, independientemente de la región y/o el periodo de tiempo que se considere. Este supuesto es conveniente puesto que, dado un momento en el tiempo, sólo se observa a países en distintas etapas del desarrollo. Si bien el supuesto de que los países pobres al desarrollarse se comportarán de la misma forma que aquellos contemporáneamente ricos no es necesariamente válido, ésta idea ayuda a acortar el camino para el entendimiento del proceso de

desarrollo en su totalidad. Claramente, considerar cada país por separado puede modificar los resultados del análisis, pero este supuesto permite buscar un patrón común a todos los países. La clave en ésta metodología de paneles es que permite la depuración de los aspectos que son permanentes (efectos fijos) dentro de cada región y por lo tanto permite un análisis que va más allá de sus idiosincrasias. También es importante que el período de tiempo sea lo suficientemente largo de forma tal que el proceso del desarrollo sea observable en su totalidad. Esto tal vez nunca suceda para algunos países y es por eso que el supuesto de considerar un mismo conjunto de parámetros para todas las regiones permite que el desarrollo se observe, al menos en el agregado. Por lo tanto, es aquí donde el supuesto anterior ayuda a sortear ése obstáculo.

Desde el punto de vista descriptivo, la curva de Kuznets vista como una media condicional es una relación promedio entre la desigualdad y el desarrollo, en otras palabras, es un resumen de la relación. En consecuencia, dado un conjunto de regiones con un determinado nivel de desarrollo, la desigualdad observada en cada uno de ellos puede diferir de aquella predicha por la curva de Kuznets. Sin embargo, esto es producto de que se está estimando un patrón de desarrollo que sintetice aquel que se debería observar para cada región.

La metodología de estimación implementada para la ecuación (1) es una regresión estándar de paneles con efectos fijos (Baltagi, 1999) mientras que para la estimación del modelo (2) se utilizó una adaptación sencilla del estimador semi-paramétrico en diferencias (Yatchew, 2000) en donde simplemente se incorporan variables binarias por región para capturar los efectos fijos sobre la desigualdad. En términos más formales, la ecuación (2) puede escribirse de la siguiente manera:

$$g_{rt} = f(y_{rt}) + x'_{rt} \delta + z' \gamma + \varepsilon_{rt} \quad (2')$$

donde se ha reemplazado al término u_r con un vector z que contiene variables binarias que indican a la región r (vale 1 en la r -ésima posición y 0 en otro lado) multiplicado con el vector de coeficientes γ que contiene a los efectos fijos de cada región. El método en diferencias consiste en los siguientes pasos:

- *Paso 1:* ordenar los datos de acuerdo a la variable y , es decir las observaciones quedan ordenadas según $y_{(1)} < y_{(2)} < \dots < y_{(i)} < \dots < y_{(RT)}$.¹²
- *Paso 2:* estimar los parámetros δ y γ por OLS utilizando el siguiente modelo en diferencias¹³:

$$g_{(i)} - g_{(i-1)} = [x'_{(i)} - x'_{(i-1)}]\delta + [z'_{(i)} - z'_{(i-1)}]\gamma + \varepsilon_{(i)} - \varepsilon_{(i-1)}$$
- *Paso 3:* estimar por algún método no paramétrico a $f(y)$ utilizando la variable transformada $\tilde{g}_{rt} = g_{rt} - x'_{rt}\hat{\delta} - z'_{rt}\hat{\gamma}$ como variable dependiente y a y como explicativa. Para este trabajo se utilizó el método de suavizado conocido como LOWESS.

Tanto para el estimador de efectos fijos como para el procedimiento basado en diferencias ordenadas, los autores demuestran que ambos métodos arrojan estimadores que son consistentes y asintóticamente normales. Como fue comentado, un problema latente dentro de la literatura de Kuznets es la endogeneidad. Económicamente, esto significa que dentro del error de estimación (entendido como $u_r + \varepsilon_{rt}$) existe algún componente que está correlacionado con las variables explicativas resultando en un sesgo de estimación. En el caso de que la fuente de endogeneidad sea invariante en el tiempo, el hecho de incluir efectos fijos por región u_r permite corregir el problema eliminando la fuente de sesgo, dado que el error residual ε_{rt} es ortogonal con las variables explicativas. Sin embargo, es probable que la endogeneidad aún persista y por lo tanto se necesitaría de algún otro método de estimación tal como variables instrumentales o el método de momentos generalizados. Más allá de la evidencia poco concluyente sobre la relevancia práctica de corregir o no por dichos métodos, es necesario ser cauto al analizar los resultados sobre la curva de Kuznets, interpretándola como una relación.

¹² El subíndice entre paréntesis indica la posición en la que se encuentra cada observación una vez que fueron ordenadas.

¹³ Este método también admite mayores órdenes de diferenciación, requiriendo un esquema óptimo de ponderadores para lograr mejoras de eficiencia asintótica en las estimaciones. Ver Yatchew (2000) para más detalles.

IV.2. Cuantil condicional

La relación entre la desigualdad y el desarrollo ha sido menos explorada fuera de la media condicional. El análisis de una relación promedio es razonable si se tienen en mente casos en donde los senderos representados por la curva de Kuznets son homogéneos. El estudio de patrones heterogéneos en la relación de U invertida puede ser abordado por el método de cuantiles condicionales, dado que analiza a la variable dependiente no solo en el centro de la distribución condicional sino en otros puntos de la misma. En este sentido, el método de cuantiles condicionales ofrece un complemento al análisis brindado por la media condicional dado que permite evaluar el alcance del mismo como forma de resumir el patrón de U invertida de Kuznets.

En este trabajo se estiman dos modelos alternativos para el cuantil τ -ésimo:

$$Q_{\tau}(g_{rt} | y_{rt}, x_{rt}) = \beta_1(\tau)y_{rt} + \beta_2(\tau)y_{rt}^2 + x'_{rt} \delta(\tau) \quad (3)$$

$$Q_{\tau}(g_{rt} | y_{rt}, x_{rt}) = \beta_1(\tau)y_{rt} + \beta_2(\tau)y_{rt}^2 + x'_{rt} \delta(\tau) + \alpha_r(\tau) \quad (4)$$

para todo $r = 1, \dots, N$ y $t = 1, \dots, T$, donde:

- g_{rt} : coeficiente de Gini multiplicado por 100 de la región r en el periodo t ,
- y_{rt} : ingreso promedio (en logaritmos) de la región r en el periodo t ,
- x_{rt} : conjunto de otros indicadores de desarrollo para la región r en el periodo t ,
- $\alpha_r(\tau)$: parámetro que representa un efecto fijo para la región r .

La interpretación de los cuantiles condicionales no difiere drásticamente de la realizada para la esperanza condicional, sólo requiere una leve intuición sobre qué es lo que representa cada cuantil. Si se toma a las regiones con un mismo nivel de desarrollo lo más probable es que las mismas difieran en sus niveles de desigualdad. Lógicamente, existen otros factores que no son capturados por el mero crecimiento de los ingresos y que afectan a la desigualdad; es decir, aún condicionando por el desarrollo económico las regiones pueden ser ordenadas en un ranking de equidad. Éste ranking de desigualdad condicional se origina en otros factores que afectan a la

distribución de los ingresos pero que usualmente no son fácilmente cuantificables en una investigación empírica, y por ende no pueden ser incluidos dentro de un análisis de regresión. En este sentido, el cuantil condicional actúa como un indicador de cómo afectan esos factores inobservables en la determinación del ranking, una vez que se ha controlado por todos los factores observables del desarrollo. Utilizando ésta interpretación, las curvas representadas por los distintos cuantiles condicionales muestran cómo evolucionaría cada una de las posiciones del ranking de desigualdad durante las distintas etapas del desarrollo, bajo el supuesto de que el conjunto de los factores inobservables permanece inalterado.¹⁴ Los cuantiles superiores modelarían curvas de Kuznets en donde la desigualdad es crónicamente alta, mientras que lo opuesto ocurre con los cuantiles inferiores. Por lo tanto, un resultado posible e interesante es que los patrones de desarrollo pueden diferir entre cuantiles llevando a que la hipótesis de Kuznets tal vez sea válida bajo ciertos contextos, pero totalmente irrelevante en otros. Una forma de tener alguna noción de cómo afectan estos factores inobservables a este patrón es a través de los efectos marginales del ingreso sobre cada cuantil condicional τ , medido por los coeficientes $\beta_1(\tau)$ y $\beta_2(\tau)$: si los mismos difieren entre cuantiles indicaría que existen factores que no fueron incluidos en la regresión que afectan al sendero de desarrollo.

Para la estimación de la ecuación (3) se utilizó el estimador tradicional de Koenker y Basset (1978) con el pool de datos, es decir que el vector $[\hat{\beta}_1(\tau), \hat{\beta}_2(\tau), \hat{\delta}(\tau)]$ resuelve el siguiente problema de optimización:

$$\min_{(\beta_1, \beta_2, \delta)} \sum_{r=1}^R \sum_{t=1}^T \rho_{\tau}(g_{rt} - \beta_1 y_{rt} - \beta_2 y_{rt}^2 - x'_{rt} \delta) \text{ con } \rho_{\tau}(h) = \begin{cases} \tau h & \text{si } h \geq 0 \\ (\tau - 1)h & \text{si } h < 0 \end{cases} \quad (5)$$

En la estimación de la ecuación (4) es necesaria una metodología de cuantiles para datos en panel, literatura que es novedosa y que ha avanzado rápidamente en los últimos años. Los trabajos de Koenker (2004, 2005) y Lamarche (2010) se basan en una analogía de una propiedad de los

¹⁴ El hecho de que una misma región no necesariamente se mantendrá siempre sobre el mismo cuantil condicional no es un argumento en contra de ésta interpretación puesto que el mismo también es válido contra la interpretación de la media condicional, ampliamente aceptada en la literatura.

estimadores de la media condicional en donde el estimador de efectos aleatorios puede describirse como un estimador de OLS con efectos fijos penalizados. En consecuencia, los autores proponen un estimador de cuantiles condicionales para datos en panel utilizando un problema de optimización similar a (5) pero incluyendo una penalización al tamaño de los efectos fijos. Por otro lado, el trabajo de Canay (2010) plantea un estimador de efectos fijos para paneles en dos etapas basándose en el supuesto adicional de que $\alpha_r(\tau) = \alpha_r$ para todo τ (modelo *location shift*). Intuitivamente este supuesto implica que los factores idiosincráticos de la región r afectan de la misma manera en los distintos cuantiles condicionales. Si bien el estimador de Koenker (2005) es más general y además es óptimo (Lamarche, 2010), el problema de optimización es computacionalmente complejo, incluso cuando el número de observaciones no es demasiado grande. Por tal motivo, en este trabajo se utiliza el estimador en etapas que consta de los siguientes pasos:

- *Paso 1:* estimar consistentemente α_r con un estadístico de paneles con efectos fijos, tal como se hace para la ecuación (1).
- *Paso 2:* estimar $\beta(\tau)$ por la regresión por cuantiles estándar de Koenker y Basset (1978) utilizando a $\hat{g}_{ir} = g_{ir} - \hat{\alpha}_r$ como variable dependiente y a y_{rt} (junto con su cuadrado) y a x_{rt} como regresores.

Canay (2010) demuestra que éste estimador en etapas es consistente y asintóticamente normal cuando T es grande. Los errores estándar son calculados mediante el método de *bootstrap* en donde la dimensión de remuestreo es r dado que con una serie de experimentos de Monte Carlo, el trabajo de Kato, Galvao y Montes-Rojas (2010) muestra que, para el caso de un modelo de *location shift*, los intervalos de confianza construidos con este procedimiento tienen un buen desempeño, medido por la tasa de rechazo de la hipótesis nula. Los mismos comentarios generales sobre los efectos de la endogeneidad hechos para la media condicional son válidos para el cuantil condicional, pero a diferencia del anterior, no hay evidencia que muestre la relevancia práctica de su corrección en la literatura de desigualdad y desarrollo. Nuevamente, se debe ser cuidadoso en las interpretaciones causales de los resultados.

V. Resultados empíricos

V.1. Relación simple

Como primer paso se considera a la relación simple entre desigualdad y desarrollo, tal como fuese planteado por Kuznets. En otras palabras, se asume que en los modelos (1) a (4) los parámetros $\delta(\tau)$ (y por lo tanto δ también) son nulos. Los Gráficos 3y 4 muestran que ajustar por efectos fijos por región lleva a que la relación entre desigualdad y desarrollo sea menos difusa en comparación al caso en que se contempla al pool de datos. A su vez, al comparar los gráficos que surgen de ambos paneles se observa que utilizar una mayor cantidad de años y menos países muestra una relación más definida entre la desigualdad y el nivel de ingresos.

En el caso de la media condicional (Gráfico 3), aun permitiendo la estimación de una forma funcional flexible, el pool de datos no presenta señales de una relación de Kuznets en ambos paneles. Al incorporarle al modelo los efectos fijos por región (bloques (b) y (d) del Gráfico 3) aparecen algunas relaciones entre desigualdad y desarrollo más definidas: en el caso del panel corto la pendiente de la media condicional se vuelve levemente negativa mientras que para el panel largo aparece una marcada relación de Kuznets. Este resultado claramente difiere de los resultados encontrados en la literatura según el cual al incorporar los efectos por región la relación entre ingresos y desigualdad se volvía débil.

Al considerar el comportamiento de los cuantiles (Gráfico 4) las conclusiones difieren según el panel de datos utilizado. En el caso del panel corto (bloque (i)), la estimación con el pool de datos muestra un comportamiento muy heterogéneo en los cuantiles condicionales indicando que el sendero de desarrollo de una región depende de la posición en la que se ubique en el ranking de desigualdad condicional. Tal como muestra el bloque a del Gráfico 4, si la región se caracteriza por una baja desigualdad al inicio del desarrollo seguirá un sendero a la Kuznets, mientras que no ocurre lo mismo en el resto de los cuantiles. Al depurar los efectos del corte transversal con el estimador de efectos fijos, las curvas por cuantiles vuelven a ser homogéneas y por lo tanto cuentan una historia de desarrollo y desigualdad similar, independientemente del nivel de desigualdad observado en la etapa inicial. Sin embargo, se debe ser cauteloso al analizar los resultados del bloque

(i) ya que el estimador Canay (2010) es más confiable a mayor número de períodos considerados en la regresión. Es por ello que las curvas que se muestran en el bloque (ii) del Gráfico 4 (panel largo) son estimaciones más precisas. Como se observa, el pool de datos sigue presentando una cierta heterogeneidad en el comportamiento de los cuantiles, aunque menos marcada que en el caso anterior. Nuevamente, al controlarse los efectos fijos por región, el comportamiento de los cuantiles se vuelve homogéneo y con forma de U invertida en cada uno de ellos. Este comportamiento de las estimaciones con ambos paneles pareciera indicar que la heterogeneidad en las distintas curvas de Kuznets se debe fundamentalmente a los efectos idiosincráticos de cada región, dado que al ser depurados por el estimador de EF, la relación se vuelve marcadamente homogénea.

V.2. Relación multivariada

Como fuese mencionado en la revisión teórica, son múltiples los factores que generan una asociación entre el nivel de ingresos y la desigualdad y por lo tanto es interesante ver cómo se alteran los resultados anteriores al considerarlos dentro del análisis de regresión. En otras palabras, el patrón de Kuznets que se observa en una relación simple como la anterior puede que sea el resultado de la interacción con otros factores tal como el grado de urbanización, ejemplo planteado por el mismo Kuznets. Son múltiples las dimensiones adicionales a considerar: el stock de capital humano, la composición etaria, el nivel de apertura económica, la estabilidad monetaria y laboral de cada país, la estructura sectorial del empleo, la desigualdad del factor tierra, el respeto por las leyes propiedad, entre otros. Para controlar algunos de estos aspectos se estimaron las relaciones modeladas en las ecuaciones (1) a (4) en donde el vector x_{rt} incluye las variables de desarrollo mencionadas.

Además del ingreso per cápita, de las encuestas de hogares se calcularon los años de educación promedio, la participación del sector público y del sector industrial en el empleo, el porcentaje de población adulta y la tasa de desempleo regional. La tasa de inflación anual, el nivel de apertura y el porcentaje de población urbana provienen de la *World Development Indicators*

(WDI 2010, Banco Mundial) y son a nivel país¹⁵. Las estimaciones de todos los modelos con ambos paneles se presentan en las tablas 4 y 5. En todos los casos, se muestra la estimación que se considera controlada por los efectos fijos (FE) de cada región.

Centrando la atención en la esperanza condicional, el panel corto (Tabla 4) muestra que en el modelo paramétrico estimado los coeficientes del logaritmo del ingreso per cápita resultaron ser no significativos, tal como sucede en los trabajos empíricos. En otras palabras, al mantener constante el resto de los factores observables del desarrollo, la relación entre el ingreso y la desigualdad resultó ser estadísticamente no significativa. En las últimas filas de la Tabla 4 se incluyen dos test estadísticos para la forma funcional de $f(y)$ basados en Yatchew (2010). El primero compara la estimación no paramétrica con una forma lineal, mientras que el segundo lo hace contra la especificación cuadrática. Claramente no es posible rechazar la hipótesis nula de que la forma funcional semi-paramétrica sea compatible con un modelo lineal o cuadrático.¹⁶ Por lo tanto, el análisis del resto de las variables de desarrollo con este panel puede hacerse tanto con las estimaciones del modelo flexible como con las del paramétrico: las únicas variables que no aparecen estadísticamente relacionadas con la desigualdad son la estabilidad de precios y el empleo en el sector público. Por otro lado, analizando el panel que considera una mayor cantidad de períodos (Tabla 5) las estimaciones muestran el ajuste de una relación cuadrática en forma estadísticamente significativa. También cobran poder de explicación la estabilidad de precios mientras que sólo el nivel educativo, el grado de urbanización y el empleo en el sector público parecen no correlacionados con la desigualdad.¹⁷ El Gráfico 5 muestra el ajuste de la esperanza condicional en ambos paneles. Como fue mencionado al analizar el test estadístico sobre la forma funcional de $f(y)$, en el caso del panel corto la forma de U invertida pareciera levemente forzada por la forma funcional del modelo ya que la misma no es demasiado diferente (estadísticamente) de una recta horizontal. Al utilizar el panel largo aparece una relación similar a la

¹⁵ El indicador del nivel de apertura es la suma de exportaciones e importaciones sobre el PBI, todo en moneda constante.

¹⁶ Para más detalles, ver la sección A.1 del Apéndice, al final del trabajo, donde se comenta brevemente el procedimiento del test.

¹⁷ Para el caso de la educación, resulta ser levemente significativo en el caso flexible. Sin embargo, estadísticamente no puede distinguirse del modelo paramétrico y por lo tanto la conclusión no es obvia.

curva de Kuznets (concavidad), aun cuando se controla por efectos fijos y no se restringe la forma funcional de la media condicional. Nótese que el test de hipótesis (últimas filas de Tabla 5) sugiere que ésta forma funcional es compatible tanto con una ecuación lineal como cuadrática. Por lo tanto, este resultado del panel largo muestra que, dado el resto de los factores observables del desarrollo, la relación entre la desigualdad y el nivel de ingresos es similar a una relación de Kuznets.

Al analizar los coeficientes de los cuantiles condicionales sucede algo similar a lo que ocurre con la media condicional, pero en este caso las estimaciones más confiables son las del panel largo. Nuevamente, al considerar un panel con mayor cantidad de períodos, el ajuste polinómico muestra un buen ajuste para casi todos los cuantiles. En base a estas estimaciones se puede decir que, además del nivel de ingreso per cápita, la estabilidad del nivel general de precios, la participación de la industria en el empleo y la tasa de desempleo impactan sobre la desigualdad en casi todas las posiciones del ranking de desigualdad condicional. Distinto es el caso del porcentaje de población adulta, en donde el impacto pareciera estar concentrado sólo en la mediana condicional, mientras que la apertura comercial tiene efecto en los cuantiles superiores de la desigualdad condicional. La relación entre ingreso y desigualdad, dado el resto de las variables de desarrollo, pareciera ser homogénea dado que los cuantiles condicionales tienen una curvatura similar para distintos valores de z . Por lo tanto, se esperaría que la media condicional ofrezca un buen resumen de la dependencia entre desigualdad y desarrollo. Esto es más claro en el Gráfico 6, donde se representa a las estimaciones de la curva de Kuznets por cuantiles para una región con factores de desarrollo promedio. Para poner a prueba este comportamiento de los cuantiles condicionales se realizó un test de Wald, en donde la hipótesis nula es la igualdad de todos los coeficientes asociados a la variable de ingreso per cápita. Intuitivamente, la hipótesis nula considera que el patrón de desarrollo no difiere entre los distintos cuantiles condicionales.¹⁸ Los resultados del test se muestran en la Tabla 6. En ambos paneles la hipótesis de homogeneidad en el sendero de desarrollo parece poco sostenible (a un nivel de significación del 2%) cuando la estimación no depura al efecto idiosincrático de cada región sobre la desigualdad (columna *pool QR*). Cuando se controla a los efectos fijos (columna *FE QR*) parece no encontrarse

¹⁸ Ver la sección A.2 del Apéndice para una explicación más formal del test de Wald.

evidencia significativa en contra de la hipótesis de homogeneidad. Por lo tanto, el análisis de cuantiles condicionales muestra que, aun controlando por otras variables de desarrollo, son las características idiosincráticas de cada región las que generan un sendero de crecimiento y desarrollo diferente, dependiendo de la posición inicial en el ranking de desigualdad. Una vez que se limpia el efecto particular de cada región controlando por efectos fijos, entonces la media condicional se vuelve un buen resumen del patrón de desigualdad y desarrollo para el agregado de regiones.

Como fuese mencionado, los resultados difieren con una gran parte de la literatura empírica que muestra que la incorporación de efectos fijos al modelo empeora el ajuste cuadrático. Esto puede ocurrir básicamente porque la estructura de datos utilizada en este trabajo difiere de los anteriores, en el sentido de que por un lado sólo se está considerando a los países de América Latina y por otro, la unidad de estudio es una región sub-nacional en lugar de un país. La Tabla 7 muestra las estimaciones realizadas con los mismos países y años utilizados en el panel largo, pero con un nivel de agregación nacional. Como se observa, en la regresión para la media condicional la relación de Kuznets sigue apareciendo pero con una menor significatividad, y en el caso de los cuantiles el ajuste cuadrático del ingreso per cápita se pierde casi por completo. Lamentablemente, no se puede dar una conclusión definitiva a estas estimaciones con estos menores niveles de potencia estadística. En otras palabras, o bien es válido el resultado encontrado previamente en la literatura empírica o la pérdida de poder explicativo sea una consecuencia de la menor cantidad de observaciones. Para tomar una dimensión de la pérdida de grados de libertad, la Tabla 8 muestra una descomposición de la variabilidad muestral del índice de Gini para las distintas muestras utilizadas. Como se observa, en ambos paneles la unidad del *cross-section*, ya sea definida como región sub-nacional o por país, representa más del 82% de la variabilidad total, mientras que el número de observaciones es mucho más acotado en el caso en donde se consideren solamente al país como la unidad de análisis. Por otro lado, si se considera a cada país por separado con sus respectivas regiones, además de observaciones también se pierde variabilidad. La Tabla 9 muestra las estimaciones de los coeficientes del ingreso per cápita (en logaritmos) y su cuadrado, considerando esta estructura de datos. A excepción de Uruguay, las regresiones muestran que dentro de cada país no es posible respaldar a la conjetura de Kuznets al considerarse a los países por separado. Las figuras del Gráfico 7 muestran un comportamiento dispar entre países: Brasil y Honduras

parecieran estar aún en una etapa inicial en la cual el crecimiento económico, dado todo lo demás, está asociado a incrementos en la inequidad, mientras en Uruguay esto es relativo al nivel del ingreso per cápita, tal como sugiere la conjetura de Kuznets. Argentina ha mostrado sus mayores niveles de desigualdad en los ciclos de bajo perfil económico mientras que en Costa Rica la desigualdad pareciera estar relacionada a otros factores distintos que al nivel de los ingresos. Claramente, el supuesto mencionado en la sección metodológica de que la relación entre el desarrollo y la desigualdad es un patrón común entre países es necesario si se quiere obtener alguna conclusión sobre la conjetura de Kuznets. La poca disponibilidad de datos completos para países de América Latina es claramente un obstáculo que afortunadamente irá reduciendo su influencia sobre este tipo de trabajos con el correr del tiempo.

V.3. Umbrales de desarrollo

Es importante destacar que el ajuste de los datos a una forma cuadrática cóncava para la relación entre la desigualdad y el desarrollo es una condición necesaria para aportar evidencia a favor de la conjetura de Kuznets, pero no suficiente para que ello dé lugar a una forma de U invertida que sea compatible con los datos. La relevancia de evaluar este aspecto radica en el carácter práctico de la curva de Kuznets para ofrecer una guía de política económica que muestre en qué momento finalizaría el *trade-off* entre desigualdad y desarrollo. Por lo tanto, es importante tener una idea de cuál es nivel de ingreso que representa a ese umbral del desarrollo, es decir, estimar cuál sería el punto de quiebre de la curva de Kuznets en el cual la desigualdad llega a su valor máximo. Con esta información se puede saber si es razonable esperar que junto con el mayor crecimiento económico se observe un cambio en la desigualdad hacia una distribución del ingreso más equitativa. En la Tabla 10 se muestran los umbrales de ingreso per cápita estimados para cada regresión realizada con la muestra del panel largo.¹⁹ Dado que la media condicional es un buen resumen de toda la relación condicional entre desigualdad y nivel de ingresos, el análisis de los umbrales de desarrollo estará basado en tales regresiones.

¹⁹ En la sección A.3 del Apéndice se da más detalles sobre el procedimiento utilizado para calcular los umbrales de ingreso y los intervalos de confianza.

En el caso del análisis simple de la relación entre la desigualdad y el desarrollo, tal como fuera planteada por Kuznets, se encuentra que el umbral ronda los 272 dólares mensuales per cápita, medido en PPA de 2005, que al compararlo con los valores promedio de las regiones de Argentina, Brasil y Uruguay (Tabla 3) es relativamente bajo, mientras que las regiones que pertenecen a los dos países de América Central incluidos en la muestra (Honduras y Costa Rica) tienen valores de ingreso per cápita que en promedio están muy por debajo del umbral estimado. La excepción a esto es la región central de Costa Rica con un ingreso per cápita superior al umbral estimado. De la misma manera, las regiones del norte y nordeste de Brasil y noroeste de Argentina también entrarían en esta categoría de regiones rezagadas. Un análisis apresurado de estos resultados concluiría que los países del Cono Sur han superado el umbral del desarrollo. Por lo tanto, dado que estos países han experimentado una etapa de crecimiento económico durante los últimos años, sería razonable que cambie la tendencia de la desigualdad de ingresos.

Sin embargo, nótese que la historia contada en el párrafo anterior es construida utilizando los resultados del análisis no condicionado. Conjuntamente con el nivel de desarrollo económico, existen otros factores que afectan a la desigualdad y que se han ido modificando a la largo de la última década. Por lo tanto, el umbral estimado puede ser ficticio si lo que se desea es conocer la importancia que tuvo exclusivamente el crecimiento sobre la desigualdad, sin verse afectado por el resto de los determinantes del desarrollo. Recalculando nuevamente con las estimaciones que surgen del análisis multivariado (bloque *b* de la Tabla 8) se obtiene que estos umbrales son mucho mayores cuando se considera que el resto de los factores de desarrollo permanecen constantes. Es decir, si no se alterara el stock de capital humano, ni el nivel general de precios, ni la apertura comercial con el resto de los países, ni el desempleo, etc., el cambio en la tendencia de la desigualdad se esperaría en aquellas regiones que superen los 544 dólares mensuales per cápita, medidos en PPA 2005. Al comparar estos valores con los ingresos per cápita de la muestra (panel corto) se ve claramente que casi ninguna región del Cono Sur supera ese umbral excepto la región conformada por Montevideo y su periferia en Uruguay (Gran Montevideo). Aun considerando un umbral un poco más laxo de 443 dólares PPA, sólo se incluiría a la región patagónica de Argentina.

Esto arroja dudas sobre la relación de U invertida en el análisis condicionado, dado que el umbral estimado sería una extrapolación del modelo

que no está respaldada por el rango de ingresos de la muestra. Para evaluar este punto se realizó un test propuesto por Lind y Mehlum (2010) para verificar si la función estimada alcanza un máximo dentro del rango del soporte de estimación.²⁰ Los valores del test se reportan en la Tabla 11: para el caso del análisis simple, el comportamiento de la función en los valores extremos del soporte son estadísticamente compatibles con una forma de U invertida dentro del rango de ingresos del panel, mientras que para el caso del análisis multivariado el comportamiento de la media condicional en el ingreso máximo no es estadísticamente distinto de una función estrictamente creciente. Por lo tanto, la razón por la cual el ajuste polinómico es bueno al utilizar el panel largo es simplemente por la presencia de una relación convexa creciente, pero que no avala a la Conjetura de Kuznets. Ambas estimaciones utilizadas en este test controlan por los efectos fijos, por lo tanto vimos que al depurarse los efectos idiosincráticos de cada región la media condicional es representativa del comportamiento de relación entre la desigualdad y el desarrollo. Por lo tanto, la forma de U invertida sólo aparece en el análisis simple, es decir, cuando el crecimiento no interactúa con el resto de los factores observables de desarrollo (educación, estabilidad de precios, apertura, etc.). Sin embargo, una vez que se controla por el efecto que los mismos tienen sobre la desigualdad (análisis multivariado) la forma de U desaparece, quedando solamente una relación cóncava creciente. Es por eso que estos resultados muestran que es poco probable que el crecimiento reciente de América Latina haya sido el protagonista exclusivo del cambio en la tendencia distributiva de la región.

VI. Conclusiones

En este trabajo se ha hecho un análisis empírico exhaustivo sobre la relación entre la desigualdad y el nivel de ingresos de América Latina, en la búsqueda de una explicación del cambio en la tendencia en los indicadores de desigualdad que la región viene presentando durante casi una década, conjuntamente con una buena *performance* en términos de crecimiento económico. Basándose en las distintas teorías del desarrollo económico se intentó cuantificar la relevancia de la hipótesis planteada por Kuznets, según la cual se esperaría que luego de un período de crecimiento se debiera concluir

²⁰ En la sección A.4 del Apéndice se explica la intuición del test.

con una etapa de mejora distributiva. El trabajo expande los análisis precedentes en dos dimensiones: (i) mediante la utilización de datos de una calidad superior a los utilizados en la literatura empírica previa; y (ii) analizando la relación conjeturada desde distintas perspectivas mediante metodologías econométricas complementarias que se consideraron idóneas. En particular, el trabajo evalúa la posibilidad de senderos de desarrollo heterogéneos mediante la aplicación del método de cuantiles condicionales. Los resultados del trabajo arrojan varias conclusiones, algunas de las cuales desafían los resultados encontrados en la literatura empírica.

En primer lugar, se refuerza la relevancia de utilizar la estructura de datos de paneles de larga duración para analizar un proceso dinámico como el desarrollo económico. Al considerar un panel más largo la calidad de los resultados mejora, en el sentido de que la relación entre desigualdad y desarrollo se vuelve más definida.

En segundo lugar, la utilización de los efectos fijos para capturar el componente idiosincrático de cada región mejora aún más el ajuste cuadrático de las estimaciones. Según el análisis de Dominics et al. (2008), este resultado contradice otros trabajos empíricos en los cuales la utilización de los efectos fijos debilitaría la idea de una relación de Kuznets. Sin embargo, dado el contexto peculiar de este trabajo que se enfoca en una sub-muestra de países Latinoamericanos y con una estructura de datos más desagregados, existe la posibilidad de que no contradiga del todo las conclusiones de la literatura, puesto que al considerar la relación a nivel país la correlación de U invertida se debilita. Lamentablemente, la falta de grados de libertad también es otra explicación posible para esta pérdida de poder explicativo. Adicionalmente, el resto de los trabajos empíricos considera al ingreso per cápita que surge de cuentas nacionales, mientras que este trabajo considera a los ingresos de encuestas, de forma tal de ganar observaciones manteniendo un nivel de variabilidad muestral aceptable. Por lo tanto, es muy importante ser cauteloso al comparar estos resultados con la literatura empírica previa ya que, por el momento, el análisis de la relación entre desigualdad y crecimiento en Latinoamérica presenta una gran restricción en la disponibilidad de los datos.

Por otro lado, al incorporar el análisis de cuantiles en la curva de Kuznets para caracterizar el comportamiento del ranking de desigualdad entre países con un mismo nivel de desarrollo, se llega a la conclusión de que las características particulares no observables de cada región generan senderos de

desarrollo y desigualdad heterogéneos. Si se depura la influencia de los mismos a la relación entre la inequidad y el desarrollo controlando las estimaciones por efectos fijos, entonces cada posición de la curva de Kuznets se comporta de manera muy similar al promedio. En consecuencia, un análisis de la media condicional sería suficiente para analizar la relación entre crecimiento y desigualdad de América Latina en su conjunto siempre y cuando tales efectos regionales sean controlados. A su vez, la conclusión de estimar a la media condicional utilizando un método semi-paramétrico como alternativa a la especificación cuadrática indica que en la mayoría de los casos el ajuste paramétrico funcionó relativamente bien.

Por último, si bien el análisis multivariado muestra que aún después de controlar por distintos factores observables e inobservables que influyen sobre la desigualdad y el crecimiento se encuentra una forma funcional que coincide con la conjetura de Kuznets, al analizar la relevancia práctica de la misma se descubre que el umbral de desarrollo estimado queda casi fuera del soporte de ingresos de la muestra utilizada. En particular, sólo dos regiones de la muestra pueden superar esa brecha. Por lo tanto, lo que se encuentra realmente es una relación creciente y cóncava entre la desigualdad y el nivel de ingresos. Ésta conclusión es apoyada por un test de hipótesis que evalúa la forma de U invertida. Consecuentemente, hablar de un umbral de máxima desigualdad o de una relación a la Kuznets con este ajuste cuadrático de la ecuación con los datos analizados se convierte en una extrapolación del modelo que puede ser ficticia. Además, el hecho de que la hipótesis de U invertida sea aceptada sólo en el caso en cual no se controla por el resto de los factores observables del desarrollo (capital humano, porcentaje de población rural, etc.) indica que es la interacción del nivel de ingreso con esas variables las que generan un patrón a la Kuznets.

En conclusión, asignar al crecimiento económico como principal motivo del cambio en la tendencia de América Latina hacia la igualdad de ingresos no parece ser creíble. Si bien es difícil hacer recomendaciones de políticas con este tipo de relaciones, desde ese punto de vista la conclusión sería que, abstrayéndose de las particularidades de cada región, la mejora distributiva no debería buscarse en tratar de impulsar indefinidamente el crecimiento, sino que pareciera más efectivo poner el foco en otras dimensiones del desarrollo tales como la calidad del capital humano y los retornos a la educación, las políticas de empleo, la cobertura social y la estabilidad económica. Además, dado que

las particularidades idiosincráticas generan senderos de desarrollo y desigualdad heterogéneos, no debe faltar en este tipo de evaluaciones la consideración de la realidad social de cada uno de los países latinoamericanos.

Referencias

Anand, S. and R. Kanbur (1993a). "The Kuznets Process and the Inequality. Development Relationship." *Journal of Development Economics*, Vol. 40: 25-52.

Anand, S. and R. Kanbur (1993b). "Inequality and Development: A Critique." *Journal of Development Economics*, Vol. 41: 19-43.

Angeles, L. (2010). "An alternative test of Kuznets' hypothesis." *Journal of Economic Inequality*, Vol. 8: 463-473.

Baltagi, B. (1999). *Econometric Analysis of Panel Data*. 2nd edition, Wiley, New York.

Banerjee, A. V. and Duflo, E. (2003). "Inequality and growth: what can the data say?" *Journal of Economic Growth*, Vol. 8: 267-299.

Barro R. (1999). "Inequality, growth and investment." Working Paper N° 7038, National Bureau of Economic Research.

Barro R. y Becker G. (1988). "A reformulation of the economic theory of fertility." *Quarterly Journal of Economics*, Vol. 103:1-25.

Barro, R. J. (2000). "Inequality and growth in a panel of countries." *Journal of Economic Growth*, Vol. 5: 5-32.

Beccaria, L. (2006). "Notas sobre la evolución de la distribución de las remuneraciones en la Argentina." Estudios del Trabajo N° 32. Asociación Argentina de Especialistas en Estudios del Trabajo, Buenos Aires.

Becker G., Murphy K. y Tamura R. (1990). "Human capital, fertility and economic growth." *Journal of Political Economy*, Vol. 98:S12-S37.

Benjamin, D., Brandt, L. And Giles, J. (2006). "Inequality and growth in rural China: does higher inequality impede growth?" IZA Discussion Paper Series no. 2344.

Cameron, C. and P. Trivedi (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press, New York.

Canay, I. (2001). "A Note on Quantile Regression for Panel Data Models." Department of Economics, Northwestern University.

Castelló, A. (2004). "A reassessment of the relationship between inequality and growth: what human capital inequality data say?" IVIE Working Paper no. 15.

CEPAL (2006). *Panorama Social de América Latina*.

Chen, S. y Ravallion, M. (2008). "The Developing World Is Poorer Than We Thought, But No Less Successful in the Fight against Poverty." Policy Research Working Paper 4703.

Deininger, K. y Squire, L. (1996). "Measuring Income Inequality: A New Data-Base." Paper no. 537, Harvard - Institute for International Development.

Dominics, L., Florax, R. and de Groot, H. (2008). "A meta-analysis on the relationship between income inequality and economic growth." *Scottish Journal of Political Economy*, Vol. 55(5).

Fields G. (2001). *Distribution and development*. Russell Sage Foundation, New York, The MIT Press, Cambridge and London.

Flores C., Flores-Lagunes A. y Kapetanakis D. (2009). "Lesson from Quantile Panel Estimation of the Environmental Kuznets Curve." 15th International Conference on Panel Data at the University of Bonn.

Forbes, K. J. (2000). "A reassessment of the relationship between inequality and growth." *American Economic Review*, Vol. 90(4): 869-87.

Galor O. y Zeira J. (1993). "Income distribution and macroeconomics." *Review of Economic Studies*, Vol. 60: 35-52.

Gasparini L. y Gluzmann P. (2012). "Estimating Income Poverty and Inequality from the Gallup World Poll: The Case of Latin America and the Caribbean." *Journal of Income Distribution*, Vol. 21(1): 3-27.

Gasparini L., Cruces G., Tornarolli L. y Marchionni M. (2009). "A Turning Point? Recent Developments on Inequality in Latin America and the Caribbean." Documento de Trabajo CEDLAS N° 81, UNLP, Argentina. Publicado en En prensa, Economía 2010.

Gasparini, L. y Cruces, G. (2008). "A Distribution in Motion: The Case of Argentina." Documento de Trabajo CEDLAS N° 78, UNLP, Argentina.

Greenwood, J. y Jovanovic, B. (1990). "Financial development, growth and the distribution of income." *Journal of Political Economy*, Vol. 98: 1076-1107.

Kaldor N. (1956). "Alternative theories of distribution." *Review of Economic Studies*, Vol. 23.

Kengo y Montes Rojas (2010). "Asymptotics and bootstrap inference for panel QR models with FE." Department of Economics, Northwestern University.

Koenker, R. (2004). "Quantile Regression for longitudinal data." *Journal of Multivariate Analysis*, Vol. 91: 74-89.

Koenker, R. (2005). *Quantile Regression*. Cambridge University Press.

Koenker R. y Bassett G. Jr. (1982). "Robust Tests for Heteroscedasticity Based on Regression Quantiles." *Econometrica*.

Koenker, R. y Basset, G. (1978). "Regression quantiles." *Econometrica*.

Kuznets S. (1955). "Economic growth and income inequality." *American Economic Review*, Vol. 45(1).

Lamarche, C. (2010). "Robust Penalized Quantile Regression Estimation for Panel Data." *Journal of Econometrics*, forthcoming.

Lewis W. (1954). "Economic development with unlimited supplies of labour." Manchester School 22.

Li, H. y Zou, H. (1998). "Income inequality is not harmful for growth: theory and evidence." *Review of Development Economics*, Vol. 2(3): 318-34.

Lind , Jo Thori y Mehlum , Halvor (2010). "With or Without U? The appropriate test for a U shaped relationship." *Oxford Bulletin Of Economics and Statistics*, Vol. 72(1).

Lopez-Calva and Lustig (eds.) (2010). *Declining Inequality in Latin America: a Decade of Progress?*. Brookings Institution Press and UNDP.

Mbabazi, J., Morrissey, O. And Milner, C. (2001). "Are inequality and trade liberalization influences on growth and poverty?" WIDER Discussion Paper no. 2001/132.

Moran, T. (2005). "Kuznets's Inverted U-Curve Hypothesis. The Rise, Demise, and Continued Relevance of a Socioeconomic Law." *Sociological Forum*, Vol. 20(2): 209-244.

Panizza, U. (2002). "Income inequality and economic growth: evidence from American data." *Journal of Economic Growth*, Vol. 7: 25-41.

Perotti, R. (1996). "Growth, income distribution, and democracy: what data say." *Journal of Economic Growth*, Vol. 1: 149-87.

Persson T. y Tabellini G. (1994). "Is inequality harmful for growth?" *American Economic Review*, Vol. 84 (3): 600-621.

Rodrik D. (1999). "Where did all the growth go? External shocks, social conflict and growth collapses." *Journal of Economic Growth*, Vol. 4: 385-412.

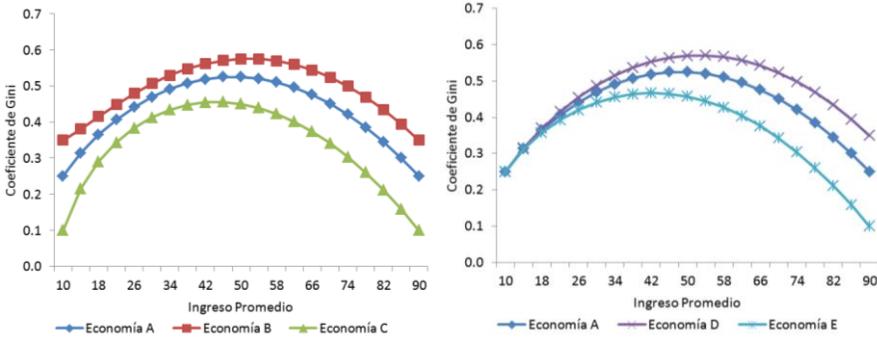
Yatchew A. (2000). "Differencing Methods in Nonparametric Regression: simple techniques for applied econometrician." University of Toronto.

Gráfico 1.

Curvas de Kuznets bajo distintas configuraciones de desigualdad

(a) Senderos homogéneos

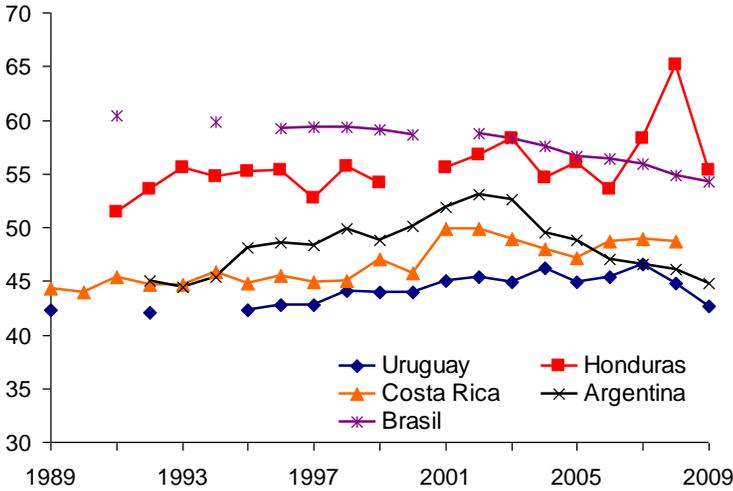
(b) Senderos heterogéneos



Fuente: elaboración propia en base a ejemplo numérico.

Gráfico 2.

Evolución de la desigualdad en América Latina



Fuente: elaboración propia en base a SEDLAC (CEDLAS - Banco Mundial)

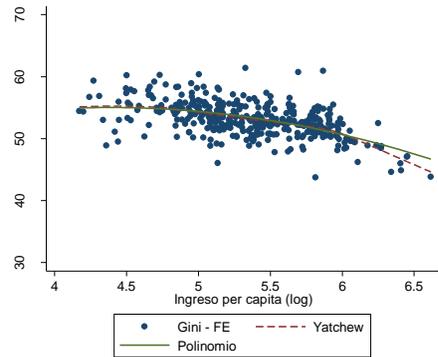
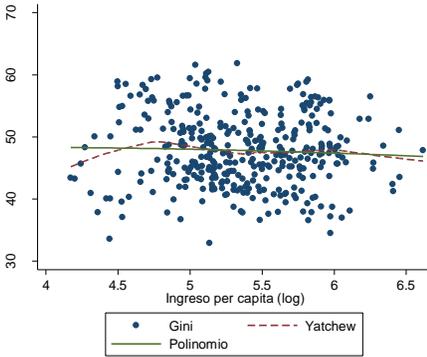
Gráfico 3.

Desigualdad e ingreso per cápita – relación simple. Esperanza condicional

(i) Panel corto

(a) Pool de datos

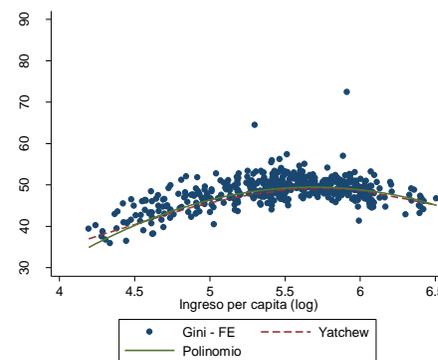
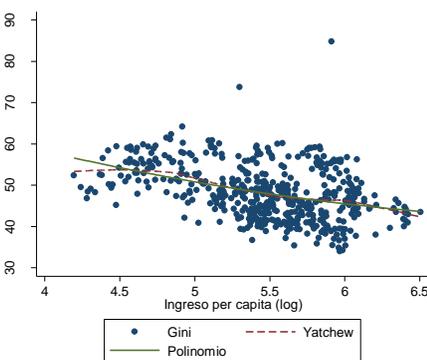
(b) Efectos fijos



(ii) Panel largo

(a) Pool de datos

(b) Efectos fijos

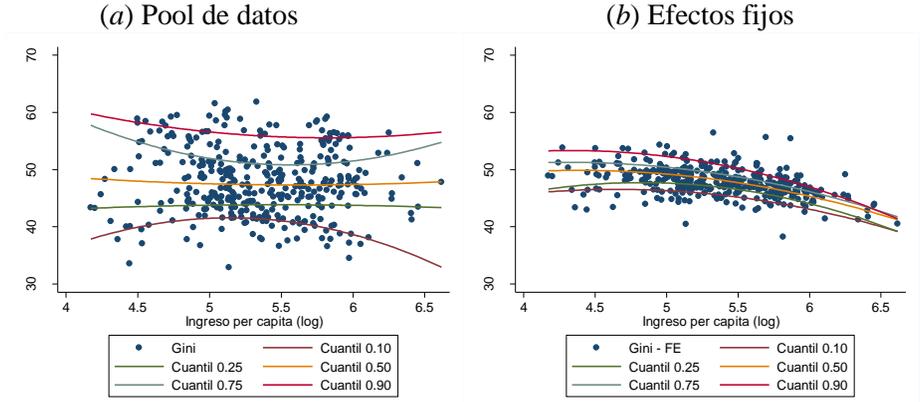


Fuente: elaboración propia en base a SEDLAC (CEDLAS - Banco Mundial)

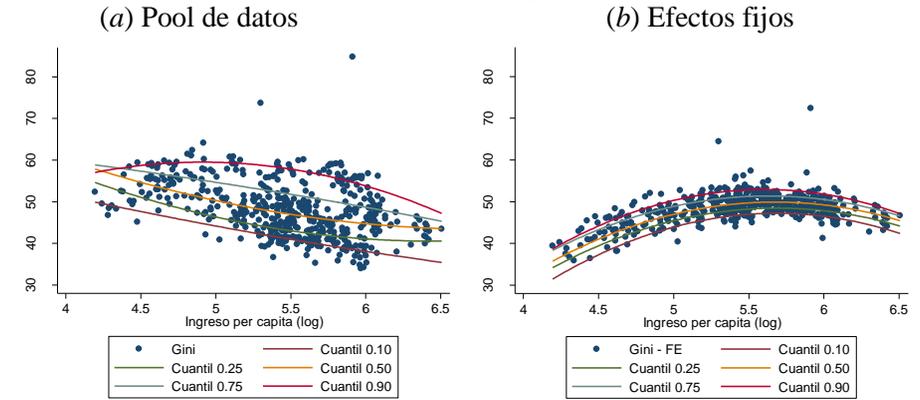
Nota: el eje vertical de los gráficos (b) y (d) representa al coeficiente de Gini sin los efectos fijos, es decir $g_r - \alpha_r$.

Gráfico 4.
Desigualdad e ingreso per cápita. Relación simple. Cuantiles condicionales

(i) Panel corto



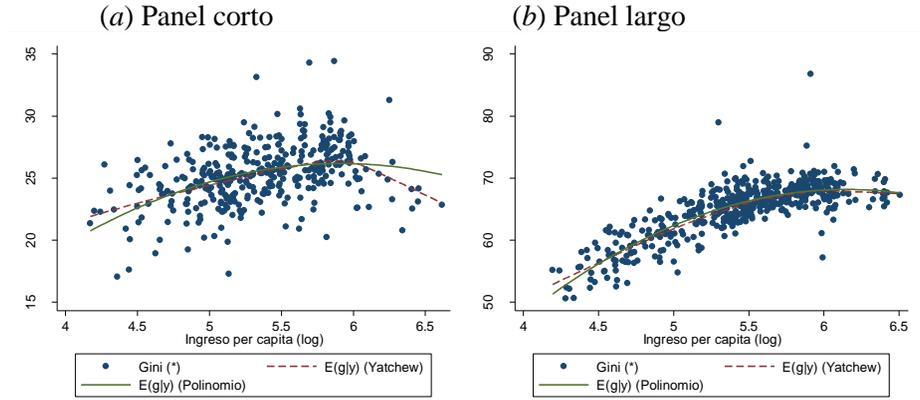
(i) Panel largo



Fuente: elaboración propia en base a SEDLAC (CEDLAS - Banco Mundial)

Nota: el eje vertical de los gráficos (b) y (d) representa al coeficiente de Gini sin los efectos fijos, es decir $g_r - \alpha_r...$

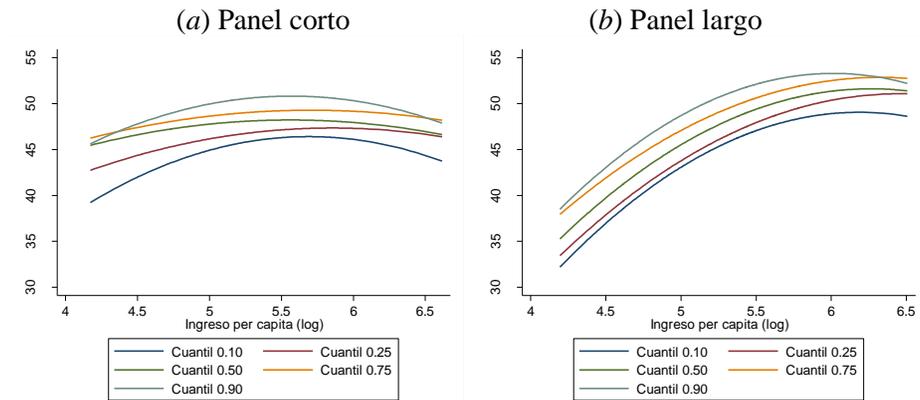
Gráfico 5.
Desigualdad e ingresos per cápita. Análisis multivariado
Esperanza condicional



Fuente: elaboración propia en base a SEDLAC (CEDLAS - Banco Mundial)

Nota: (*) indica que en el eje vertical representa a $g_{rt} - x_{rt}^T \hat{\delta}$, es decir al coeficiente de Gini sin el efecto del resto de los regresores.

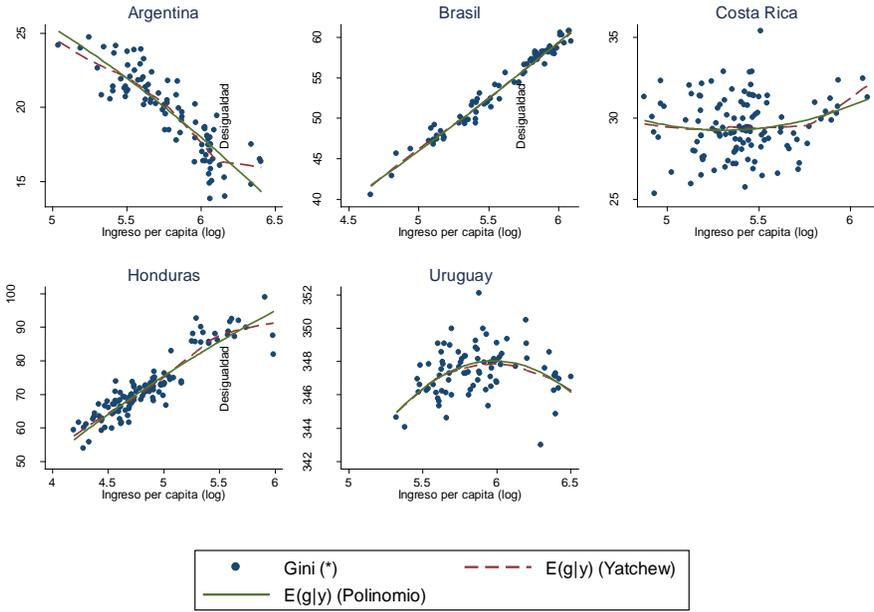
Gráfico 6.
Desigualdad e ingreso per cápita. Análisis multivariado
Cuantiles condicionales



Fuente: elaboración propia en base a SEDLAC (CEDLAS - Banco Mundial)

Nota: cada cuantil condicional tiene las variables explicativas evaluadas en sus medias muestrales, a excepción del ingreso per cápita.

Gráfico 7.
Desigualdad e ingreso per cápita por países de LAC (panel largo)



Fuente: elaboración propia en base a SEDLAC (CEDLAS - Banco Mundial)

Nota: (*) indica que en el eje vertical representa a $g_{rt} - x_{rt}^T \hat{\delta}$, es decir al coeficiente de Gini sin el efecto del resto de los regresores.

Tabla 1.
Configuración de parámetros en sectores urbano-rural

Economía	A	B	C	D	E
Rural					
w_B	5	3	8	5	5
w_A	15	17	12	15	15
$\mu_w(R)$	10	10	10	10	10
$Gini(R)$	0.25	0.35	0.10	0.25	0.25
Urbano					
w_B	45	27	72	27	72
w_A	135	153	108	153	108
$\mu_w(U)$	90	90	90	90	90
$Gini(U)$	0.25	0.35	0.10	0.35	0.10

Fuente: elaboración propia en base a SEDLAC (CEDLAS - Banco Mundial)

Tabla 2.
Listado de regiones por país

Argentina	Brasil	Costa Rica	El Salvador	Honduras
Gran Buenos Aires	Norte	Central	Occidental	Area metropolitana
Pampeana	Nordeste	Chorotega	Central 1	Norte
Cuyo	Sudeste	Pacífico Central	Central 2	Occidente
Noroeste	Sur	Brunca	Oriental	Sur
Patagonia	Centro-Oeste	Huetar Atlántica	AMSS	Oriente
		Huetar Norte		Central
México	Paraguay	Perú	Uruguay	Venezuela
Noroeste	Costa urbana	Costa urbana	Gran Montevideo	Capital
Norte	Sierra urbana	Sierra urbana	Interior Norte	Central
Noreste	Selva urbana	Selva urbana	Interior Centro Norte	Insular
Centro-Occidente	Costa rural	Costa rural	Interior Centro Sur	Nororiental
Centro-Este	Sierra rural	Sierra rural	Interior Sur	Guayana
Sur		Selva rural		Centro Occidental
Oriente		Lima Metropolitana		Zuliana
Península de Yucatan				Andes
				Los Llanos

Fuente: elaboración propia en base a SEDLAC (CEDLAS - Banco Mundial)

Tabla 3.
Estadísticas básicas por región

	Desigualdad (Gini)				Ingreso per cápita			
	Media	Desv.	Min	Max	Media	Desv.	Min	Max
Argentina (1992 - 2009)								
Gran Buenos Aires	48.6	3.1	44.2	54.4	409.0	49.4	269.2	464.7
Pampeana	44.5	2.1	41.5	48.3	332.6	54.1	201.7	427.2
Cuyo	45.5	2.4	41.6	49.2	275.2	42.2	180.0	349.3
Noroeste	47.7	2.2	44.5	51.2	235.4	33.5	154.8	296.5
Patagonia	45.5	1.7	42.1	49.0	448.4	82.3	289.9	594.9
Brasil (1990 - 2008)								
Norte	55.3	2.6	50.6	59.0	216.5	27.2	140.9	250.4
Nordeste	58.6	2.2	55.2	62.3	166.7	30.1	105.5	224.4
Sudeste	54.6	2.2	49.2	56.9	364.9	54.8	224.1	436.5
Sur	53.1	2.5	48.7	56.5	342.1	60.0	208.4	442.4
Centro-Oeste	57.6	2.1	51.5	60.2	325.3	59.8	179.1	438.1
Costa Rica (1989 - 2008)								
Central	45.2	2.3	41.6	48.9	327.0	62.7	227.3	447.0
Chorotega	48.7	2.8	45.4	57.1	200.3	43.7	131.4	277.2
Pacífico Central	44.3	2.3	40.9	50.9	227.5	33.5	170.0	296.9
Brunca	47.4	2.3	42.1	51.2	187.6	31.9	137.1	250.6
Huetar Atlántica	41.5	1.4	39.3	44.0	221.0	16.8	196.2	258.8
Huetar Norte	45.8	2.4	40.9	50.4	220.0	39.1	152.2	306.8
El Salvador (1995 - 2006)								
Occidental	49.7	3.3	45.4	53.9	163.2	8.4	153.3	177.1
Central 1	44.2	2.0	41.8	46.8	162.7	11.9	144.1	175.2
Central 2	50.0	2.7	45.8	54.0	139.0	9.9	127.5	153.8
Oriental	51.1	3.3	46.9	56.6	154.3	11.6	139.5	169.0
AMSS	46.6	0.9	45.2	47.7	321.7	21.6	296.1	344.4
Honduras (1991 - 2009)								
Area metropolitana	49.6	2.7	41.7	53.5	258.6	70.0	161.6	405.9
Norte	50.7	3.0	45.4	57.9	150.2	35.5	88.3	234.1
Occidente	56.4	4.7	45.6	63.2	99.6	16.8	66.8	128.8
Sur	54.5	4.1	46.4	65.1	111.1	26.5	69.8	153.3
Oriente	57.7	4.3	48.5	64.8	124.7	30.9	76.5	195.7
Central	55.0	3.9	47.7	60.4	112.6	23.7	79.4	166.3
México (1996 - 2006)								
Noroeste	49.2	2.5	46.1	53.5	332.1	42.7	279.3	391.8
Norte	47.7	2.5	43.3	50.1	251.8	48.2	169.2	302.1
Noreste	47.5	6.1	39.3	56.5	356.6	45.0	280.4	405.9
Centro-Occidente	49.7	4.3	45.3	57.7	257.7	55.7	161.2	304.3
Centro-Este	52.9	1.5	51.0	54.6	294.7	48.5	221.8	340.8
Sur	52.9	3.6	48.7	57.7	143.5	29.4	109.4	183.5
Oriente	50.5	1.8	48.7	53.4	186.9	38.0	140.9	242.5
Península de Yucatan	53.5	5.0	48.3	61.8	235.0	46.9	183.8	306.5
Paraguay (1995 - 2006)								
Costa urbana	51.5	3.3	47.5	56.5	557.1	64.8	482.4	634.8
Sierra urbana	44.1	2.7	40.4	47.4	308.7	52.2	252.7	391.5
Selva urbana	40.2	3.1	36.5	44.0	223.0	47.9	173.1	301.9
Costa rural	53.7	2.2	50.3	56.4	275.9	44.8	226.1	344.1
Sierra rural	58.3	2.3	54.9	61.4	157.1	14.4	142.1	184.0

Fuente: elaboración propia en base a SEDLAC (CEDLAS - Banco Mundial)

Tabla 3: (continuación)

	Desigualdad (Gini)				Ingreso per cápita			
	Media	Desv.	Min	Max	Media	Desv.	Min	Max
Perú (1996 - 2006)								
Costa urbana	40.3	1.7	37.8	41.8	193.6	16.5	173.6	214.2
Sierra urbana	43.7	1.5	42.1	45.8	190.2	9.4	174.6	201.5
Selva urbana	43.0	2.1	39.6	45.0	176.8	11.6	161.7	197.5
Costa rural	40.5	2.8	36.7	44.1	115.0	15.6	97.4	139.4
Sierra rural	43.1	3.8	37.8	48.3	72.4	7.0	64.9	83.2
Selva rural	39.3	3.6	33.6	44.3	86.7	6.9	74.6	92.8
Lima Metropolitana	48.9	2.0	46.1	51.3	360.8	23.9	329.4	392.3
Uruguay (1989 - 2009)								
Gran Montevideo	43.8	2.5	39.5	47.8	543.5	75.6	418.6	668.4
Interior Norte	41.3	2.1	36.9	46.2	272.2	31.6	213.6	332.3
Interior Centro Norte	41.2	2.2	38.1	48.4	307.1	37.8	235.9	360.3
Interior Centro Sur	39.3	1.6	35.6	42.1	347.4	39.4	278.2	412.9
Interior Sur	37.1	2.2	32.7	39.9	373.6	53.7	272.0	470.3
Venezuela (1995 - 2006)								
Capital	45.9	0.8	45.2	47.2	243.7	35.7	182.3	284.5
Central	42.9	1.6	41.5	45.8	173.1	25.2	138.0	203.3
Insular	40.9	4.4	32.9	45.1	163.0	36.8	132.7	229.6
Nororiental	46.1	2.3	43.8	49.0	166.6	20.3	142.0	201.1
Guayana	45.4	1.3	43.4	46.7	166.6	21.0	129.0	191.3
Centro Occidental	44.9	0.8	44.0	46.0	134.8	21.3	113.7	175.8
Zuliana	44.9	1.7	41.7	46.3	154.8	23.7	118.3	185.6
Andes	43.6	0.9	42.8	45.0	150.9	19.9	126.4	186.2
Los Llanos	49.1	3.5	44.3	54.4	147.1	19.4	125.1	171.9

Fuente: elaboración propia en base a SEDLAC (CEDLAS - Banco Mundial)

Tabla 4.
Relación de Kuznets para América Latina (*panel corto*)

Estimaciones para paneles, modelo de efectos fijos

(r = país-región, t = periodo bianual)

	$E(g/y, x)$		$Q_{\theta}(g/y, x)$				
	FE	SPFE	$\theta = 0.10$	$\theta = 0.25$	$\theta = 0.50$	$\theta = 0.75$	$\theta = 0.90$
ln(IPCF)	18.54 (1.48)		35.50 (2.68)*	19.09 (1.60)	15.88 (1.49)	14.78 (1.34)	29.70 (2.01)
ln(IPCF) ²	-1.62 (1.39)		-3.12 (2.58)*	-1.63 (1.45)	-1.43 (1.39)	-1.30 (1.22)	-2.67 (1.92)
Educación (años promedio)	-2.08 (4.06)**	-2.10 (3.57)**	-2.17 (3.4)**	-2.22 (3.75)**	-1.89 (3.47)**	-2.10 (3.63)**	-1.97 (3.22)**
Inflación	0.027 (1.95)	0.02 (1.23)	0.043 (1.32)	0.027 (1.98)*	0.023 (1.41)	0.030 (1.64)	0.023 (1.00)
Apertura	0.024 (2.13)*	0.022 (1.83)	0.015 (1.01)	0.022 (1.78)*	0.027 (2.13)*	0.024 (1.62)	0.032 (1.93)
% Pobl. Urbana	0.6 (3.86)**	0.559 (3.15)**	0.582 (3.2)**	0.606 (3.3)**	0.606 (3.22)**	0.591 (3.32)**	0.610 (3.47)**
Desempleo	0.295 (3.63)**	0.342 (3.88)**	0.297 (2.48)**	0.327 (3.48)**	0.276 (3.13)**	0.274 (2.69)*	0.269 (2.46)
Población Adulta	-0.677 (2.36)*	-0.73 (2.28)*	-0.567 (1.97)*	-0.693 (2.59)*	-0.642 (2.16)*	-0.698 (2.51)*	-0.792 (2.74)*
Empleo Industrial	-0.4 (4.41)**	-0.34 (3.43)**	-0.444 (3.04)**	-0.418 (3.31)**	-0.433 (3.76)**	-0.389 (3)**	-0.420 (3.03)*
Empleo Sector Público	-0.241 (1.45)	-0.212 (1.19)	-0.229 (0.85)	-0.204 (0.82)	-0.250 (1.13)	-0.188 (0.74)	-0.314 (1.15)
Nro. de observaciones	366	365	366	366	366	366	366
Nro. de regiones	61		61	61	61	61	61
R ²	0.21	0.88					
Test sobre $f(y)$: (FE vs SPFE)	V	Valor p					
- Lineal	1.29	0.197					
- Cuadrática	1.43	0.154					

Fuente: estimaciones propias en base a SEDLAC (CEDLAS - Banco Mundial)

Nota: Estadísticos z entre paréntesis (valor absoluto). * indica significatividad estadística al 5%;

** al 1%

Tabla 5.
Relación de Kuznets para América Latina (panel largo)
 Estimaciones para paneles, modelo de efectos fijos
 (r = país-región, t = periodo anual)

	$E(g/y, x)$		$Q_{\theta}(g/y, x)$				
	FE	SPFE	$\theta = 0.10$	$\theta = 0.25$	$\theta = 0.50$	$\theta = 0.75$	$\theta = 0.90$
ln(IPCF)	48.02 (5.57)**		52.67 (2.98)**	44.63 (2.93)**	47.89 (2.88)**	41.49 (2.77)**	53.86 (3.46)**
ln(IPCF) ²	-3.81 (4.71)**		-4.26 (2.39)	-3.46 (2.23)**	-3.82 (2.35)*	-3.28 (2.25)*	-4.48 (2.96)**
Educación (años promedio)	-0.79 (1.82)	-1.14 (2.13)*	-0.78 (1.22)	-0.87 (1.49)	-0.59 (1.06)	-0.69 (1.18)	-0.46 (0.81)
Inflación	0.001 (3.32)**	0.001 (2.91)**	0.002 (0.55)	0.002 (.99)**	0.001 (3.56)**	0.001 (2.44)**	0.001 (.28)**
Apertura	0.045 (3.25)**	0.024 (1.44)	0.037 (1.75)	0.039 (2.10)	0.036 (1.74)	0.045 (1.85)*	0.057 (2.48)*
% Pobl. Urbana	-0.145 (1.27)	-0.038 (0.27)	-0.129 (0.68)	-0.128 (0.70)	-0.150 (0.84)	-0.152 (0.79)	-0.140 (0.67)
Desempleo	0.399 (5.81)**	0.386 (4.60)**	0.435 (3.42)**	0.411 (3.71)**	0.417 (3.25)**	0.368 (2.8)**	0.339 (1.87)*
Población Adulta	-0.525 (2.47)*	-0.549 (2.06)*	-0.516 (1.84)	-0.583 (2.25)	-0.576 (2.31)*	-0.546 (2.14)	-0.544 (1.91)
Empleo Industrial	-0.252 (3.69)**	-0.312 (3.72)**	-0.241 (2.12)*	-0.247 (2.26)	-0.250 (2.65)*	-0.252 (2.37)*	-0.253 (2.14)*
Empleo Sector Público	-0.125 (0.89)	-0.209 (1.21)	-0.175 (0.86)	-0.147 (0.70)	-0.172 (0.87)	-0.102 (0.57)	-0.118 (0.62)
Nro. de observaciones	483	482	483	483	483	483	483
Nro. de regiones	27		27	27	27	27	27
R ²	0.34	0.84					
Test sobre $f(y)$: (FE vs SPFE)	V	Valor p					
- Lineal	-0.448	0.654					
- Cuadrática	0.622	0.534					

Fuente: estimaciones propias en base a SEDLAC (CEDLAS - Banco Mundial)

Nota: Estadísticos z entre paréntesis (valor absoluto). * indica significatividad estadística al 5%;

** al 1%

Tabla 6.
Homogeneidad de los cuantiles condicionales en relación al ingreso per cápita. Análisis multivariado

Test de Wald basado en regresiones por cuantiles (Pool y Efectos Fijos)
(r = país-región, t = periodo anual)

	Panel Corto		Panel Largo	
	Pool QR	FE QR	Pool QR	FE QR
Estadístico de Wald	21.09	12.75	18.16	10.26
Valor P	0.007	0.121	0.020	0.247

Fuente: estimaciones propias en base a SEDLAC (CEDLAS - Banco Mundial)

Nota: varianzas conjuntas calculadas por bootstrap de 500 réplicas.

Tabla 7.
Relación de Kuznets para América Latina (panel largo)

Estimaciones para paneles, modelo de efectos fijos (r = país, t = periodo anual)

	$E(g/y, x)$		$Q_{\theta}(g/y, x)$				
	FE	SPFE	$\theta = 0.10$	$\theta = 0.25$	$\theta = 0.50$	$\theta = 0.75$	$\theta = 0.90$
ln(IPCF)	63.07 (2.87)**		48.55 (0.52)	51.54 (0.58)	82.10 (1.09)	67.98 (1.00)	81.45 (1.09)
ln(IPCF) ²	-5.21 (2.59)*		-3.83 (0.47)	-4.22 (0.54)	-7.12 (1.07)	-5.75 (0.93)	-6.97 (1.03)
Educación (años promedio)	0.36 (0.44)	-0.75 (0.98)	-0.02 (0.02)	0.29 (0.38)	0.43 (0.46)	0.30 (0.30)	0.41 (0.31)
Inflación	0.001 (2.02)*	0.001 (1.89)	-0.002 (0.07)	0.002 (0.06)	0.001 (0.05)	0.001 (0.05)	0.001 (0.02)
Apertura	0.031 (1.18)	0.047 (1.67)	0.023 (0.78)	0.028 (1.26)	0.017 (0.81)	0.047 (1.60)	0.053 (1.02)
% Pobl. Urbana	-0.141 (0.54)	0.004 (0.02)	-0.132 (0.34)	-0.103 (0.33)	-0.163 (0.64)	-0.111 (0.34)	-0.082 (0.20)
Desempleo	0.433 (3.77)**	0.318 (2.56)*	0.426 (1.48)	0.411 (1.62)	0.365 (1.78)*	0.390 (1.45)*	0.443 (1.72)
Población Adulta	-1.58 (3.34)**	-1.17 (2.30)*	-1.73 (1.91)**	-1.64 (2.21)**	-1.35 (2.18)**	-1.56 (2.23)*	-1.61 (1.76)
Empleo Industrial	-0.462 (3.86)**	-0.367 (2.91)**	-0.449 (1.15)	-0.447 (1.27)	-0.582 (1.56)	-0.469 (1.63)	-0.370 (1.31)*
Empleo Sector Público	0.385 (0.91)	-0.063 (0.11)	0.654 (0.70)	0.355 (0.40)	0.464 (0.58)	0.429 (0.49)	0.024 (0.03)
Nro. de observaciones	89	88	89	89	89	89	89
Nro. de regiones	5		5	5	5	5	5
R ²	0.59	0.87					
Test sobre $f(y)$: (FE vs SPFE)	V	Valor p					
- Lineal	2.844	0.004					
- Cuadrática	3.956	0.000					

Fuente: estimaciones propias en base a SEDLAC (CEDLAS - Banco Mundial).

Nota: Estadísticos z entre paréntesis (valor absoluto). * indica significatividad estadística al 5%;

** al 1%

Tabla 8.**Descomposición de la variabilidad muestral del índice de Gini**

Unidad del cross-section de la muestra	Panel Corto		Panel Largo	
	% cross-section en varianza total	Cantidad de observaciones	% cross-section en varianza total	Cantidad de observaciones
<i>País-región</i>	82%	366	84%	483
<i>País</i>	88%	60	87%	89
<i>Regiones por país</i>				
Argentina	49%	30	32%	90
Brasil	62%	30	46%	80
Costa Rica	59%	36	54%	120
Honduras	54%	36	39%	108
Uruguay	71%	30	57%	85
El Salvador	50%	30		
México	19%	48		
Paraguay	87%	30		
Perú	57%	42		
Venezuela	47%	54		

Fuente: estimaciones propias en base a SEDLAC (CEDLAS - Banco Mundial)

Tabla 9.**Relación de Kuznets para América Latina (*panel largo*)**Estimaciones para paneles, modelo de efectos fijos (r = región, t = periodo anual)

	$E(g y, x)$	$Q_{\theta}(g y, x)$				
	FE	$\theta = 0.10$	$\theta = 0.25$	$\theta = 0.50$	$\theta = 0.75$	$\theta = 0.90$
Argentina						
ln(IPCF)	7.12 (0.32)	24.42 (0.30)	8.50 (0.16)	-3.17 (0.08)	-26.54 (0.42)	1.25 (0.02)
ln(IPCF) ²	-1.12 (0.58)	-2.70 (0.38)	-1.27 (0.27)	-0.19 (0.06)	1.79 (0.33)	-0.55 (0.09)
Brasil						
ln(IPCF)	15.25 (1.09)	12.75 (0.22)	23.90 (0.53)	26.57 (0.62)	25.50 (0.57)	29.36 (0.61)
ln(IPCF) ²	-0.101 (0.08)	0.283 (0.05)	-0.772 (0.19)	-1.117 (0.29)	-1.269 (0.32)	-1.355 (0.31)
Costa Rica						
ln(IPCF)	-23.15 (0.75)	-32.86 (0.34)	-76.64 (1.37)	-57.31 (1.04)	31.12 (0.40)	-0.39 (0.00)
ln(IPCF) ²	2.26 (0.79)	3.255 (0.35)	7.148 (1.38)	5.321 (1.06)	-2.891 (0.40)	0.021 (0.00)
Honduras						
ln(IPCF)	57.06 (2.68)**	107.05 (1.25)	87.16 (0.96)	77.01 (0.96)	23.69 (0.33)	19.50 (0.24)
ln(IPCF) ²	-3.985 (1.90)	-10.481 (1.17)	-8.378 (0.87)	-6.692 (0.79)	-0.063 (0.01)	0.155 (0.02)
Uruguay						
ln(IPCF)	91.75 (3.36)**	141.18 (1.59)	72.80 (1.12)	89.82 (1.52)	92.14 (1.32)	135.35 (1.82)
ln(IPCF) ²	-7.58 (3.26)**	-12.17 (1.58)	-6.16 (1.09)	-7.44 (1.46)	-7.36 (1.21)	-11.14 (1.72)

Fuente: estimaciones propias en base a SEDLAC (CEDLAS - Banco Mundial)

Nota: Estadísticos z entre paréntesis (valor absoluto). * indica significatividad estadística al 5%;

** al 1%.

El resto de los regresores no se muestran pero fueron incluidos en las estimaciones.

Tabla 10.**Umbral de ingreso (*panel largo*)**

Ingresos mensuales en moneda constante (U\$S PPA de 2005)

a) Análisis simple

Regresión	Estimado	Interv. Conf. 95%	
Media	272.2	270.7	273.7
Cuantil 0.10	293.3	289.4	297.3
Cuantil 0.25	295.0	291.0	299.1
Cuantil 0.50	291.6	287.4	295.8
Cuantil 0.75	279.2	275.3	283.1
Cuantil 0.90	271.7	268.5	274.9

b) Análisis multivariado

Regresión	Estimado	Interv. Conf. 95%	
Media	543.5	531.9	555.4
Cuantil 0.10	485.7	460.6	512.2
Cuantil 0.25	635.0	592.2	680.8
Cuantil 0.50	524.8	496.6	554.5
Cuantil 0.75	560.3	526.5	596.3
Cuantil 0.90	408.6	392.2	425.6

Fuente: estimaciones propias en base a SEDLAC (CEDLAS - Banco Mundial)

Nota: los intervalos de confianza del 95% de confianza fueron calculados por el método delta.

Tabla 11.**Test de hipótesis sobre la forma de U invertida (*panel largo*)**

Pruebas basadas en las estimaciones por paneles con efectos fijos.

(r = país-región, t = periodo anual)

Análisis de regresión	(a) Simple		(b) Multivariado	
	<i>y_L</i>	<i>y_H</i>	<i>y_L</i>	<i>y_H</i>
<i>- Test individual</i>				
Ingreso per capita (log)	4.19	6.50	4.19	6.50
Primera derivada de $f(y)$	17.89	-11.38	16.04	-1.58
Estadístico <i>t</i>	9.42	-6.08	8.08	-0.73
Valor <i>p</i>	0.000	0.000	0.000	0.231
<i>- Test conjunto</i>				
Estadístico <i>t</i>	6.08		0.73	
Valor <i>p</i>	0.000		0.231	

Fuente: estimaciones propias en base a SEDLAC (CEDLAS - Banco Mundial)

Nota: para más detalles sobre el cómputo de los estadísticos de prueba, ver Lind y Mehlum (2010).

Apéndice

A.1. Test hipótesis sobre para la forma funcional $f(y)$

La prueba de hipótesis implementada en este trabajo contrasta la estimación no paramétrica de $f(y)$ contra dos formas funcionales conocidas como son los polinomios de grado 1 (recta) y 2 (parábola). El test consiste simplemente en comparar las varianzas residuales de cada modelo paramétrico confrontándolo con el ajuste no paramétrico. El estadístico de prueba es:

$$V = \sqrt{mT} (s_P^2 - s_{NP}^2) / s_{NP}^2 \xrightarrow{d} N(0,1)$$

donde el subíndice P hace referencia al modelo paramétrico estimado por OLS con efectos fijos (EF) y NP indica al modelo que no presupone ninguna forma funcional para la media condicional, estimado por método de diferencias con efectos fijos (SPFE).

Para más detalles sobre el test ver Yatchew (2010).

A.2. Test de homocedasticidad basado en el estadístico de Wald

Se trata de un enfoque general para hacer pruebas de hipótesis sobre combinaciones lineales de parámetros. El mismo fue adaptado por Koenker y Bassett (1982) para proponer un test de homocedasticidad robusto utilizando regresión por cuantiles. Siguiendo la notación propuesta en la sección 4, la hipótesis nula plantea lo siguiente:

$$H_0: \quad \beta_1(0.10) = \beta_1(0.25) = \dots = \beta_1(0.90) \\ \beta_2(0.10) = \beta_2(0.25) = \dots = \beta_2(0.90)$$

$$H_1: \quad H_0 \text{ no es verdadera}$$

En otras palabras, bajo la H_0 tanto el coeficiente que multiplica al logaritmo del ingreso per cápita como a su cuadrado son constantes en los distintos cuantiles.

Esta hipótesis nula puede escribirse en forma matricial: $H\zeta = 0$, donde el vector $\zeta \equiv [\beta(\tau_1)', \beta(\tau_2)', \dots, \beta(\tau_5)']'$ y H es una matriz no estocástica que reproduce lo enunciado H_0 . El estadístico de prueba es:

$$W = \hat{\zeta}' H' [H V^{-1} H'] H \hat{\zeta} \xrightarrow{d} \chi_q^2$$

donde q es el rango de la matriz H y $V = \text{Var}(\hat{\zeta})$.

En este trabajo, el test fue aplicado tanto a los coeficientes estimados con el método de regresión por cuantiles considerando a los datos como un *pool* (Koenker y Bassett, 1978) como al estimador de efectos fijos (Canay, 2010). La varianza conjunta V fue estimada por el método de bootstrap con 500 réplicas.

Para más detalles sobre el test remitirse a Koenker y Bassett (1982).

A.3. Cálculo de umbrales de desarrollo e inferencia

El cálculo de los umbrales de desarrollo implica encontrar el nivel de ingreso per cápita que maximiza la desigualdad, según la Curva de Kuznets. Para ello se utilizan las estimaciones paramétricas de las ecuaciones (1) y (4). Analíticamente el problema consiste en igualar a cero las primeras derivadas parciales. Por lo tanto, los umbrales estimados con la ecuación de la media condicional quedan definidos de la siguiente manera:

$$\hat{U}_m = -\frac{\hat{\beta}_1}{2\hat{\beta}_2}$$

mientras que con los cuantiles, los umbrales son:

$$\hat{U}_q(\tau) = -\frac{\hat{\beta}_1(\tau)}{2\hat{\beta}_2(\tau)} \quad \tau = \{0.10, 0.25, 0.50, 0.75, 0.90\}$$

Recordar que los ingresos están expresados en escala logarítmica, por lo tanto se le aplica la función exponencial a estos valores para que queden expresados en US\$ PPA de 2005.

Nótese que estos umbrales son funciones no lineales de los parámetros previamente estimados y para los cuales fueron computados sus respectivas varianzas y covarianzas. Por lo tanto se recurre al uso del método delta para estimar los errores estándar e intervalos de confianza de los umbrales de ingreso, dado que su aplicación es inmediata.

Para más detalles de este método ver Cameron y Trivedi (2005).

A.4. Prueba de hipótesis sobre la forma de U invertida

Este test es propuesto por Lind y Mehlum (2010) y es una forma apropiada para testear si una relación paramétrica estimada a través de una regresión tiene o no forma de U invertida. A diferencia del test explicado en A.1 en donde se evalúa las ganancias del ajuste de una versión paramétrica para la relación entre la desigualdad y el crecimiento económico, este test evalúa el comportamiento de la función en los extremos del soporte sobre la variable explicativa de interés, en este caso el logaritmo del ingreso per cápita.

Utilizando la ecuación (1), las hipótesis del test son las siguientes:

$$H_0 : \beta_1 + 2\beta_2 y_L \leq 0 \quad \vee \quad \beta_1 + 2\beta_2 y_H \geq 0$$

$$H_1 : \beta_1 + 2\beta_2 y_L > 0 \quad \wedge \quad \beta_1 + 2\beta_2 y_H < 0$$

donde y_L es el logaritmo del ingreso per cápita de la región más pobre y_H es el de la región más rica de la muestra. Intuitivamente, el test evalúa la derivada primera en ambos extremos del soporte de la muestra de ingresos. La hipótesis nula plantea la existencia de una relación entre la desigualdad y el crecimiento que puede ser monótona o bien con forma de U; por el contrario, bajo la hipótesis alternativa la relación es estrictamente una U invertida. Para más detalles sobre los estadísticos de prueba y la inferencia ver Lind y Mehlum (2010).

A.5. Ejemplo numérico

El ejemplo consta de una economía de dos sectores R y U en donde, dentro de cada uno de ellos, la mitad obtiene un ingreso alto $w_A(S)$ y otro bajo $w_B(S)$, con $S = U, R$. Para simplificar el ejercicio, se supone que $w_B(U) > w_A(R)$.

Luego, el ingreso promedio de cada sector es simplemente $\mu_w(S) = [w_A(S) + w_B(S)]/2$.

El ingreso promedio de toda la economía es $\mu_w = (1-\psi)\mu_w(R) + \psi\mu_w(U)$ donde ψ es la proporción de personas en el sector U . La desigualdad dentro de cada sector S medida por el índice de Gini se calcula como:

$$Gini(S) = \frac{2Cov(w, F)}{E(w)} = \frac{2[E(wF) - E(w)E(F)]}{E(w)}$$

Considerando que $F[w_B(S)] = 1/2$ y $F[w_A(S)] = 1$ es la distribución acumulada de ingresos, y que $E(w) = \mu_w(S)$, es simple calcular el resto de los valores esperados:

$$E(wF) = (1/2)w_B(S)(1/2) + (1/2)w_A(S)(1) = \mu_w(S) - (1/4)w_B(S)$$

$$E(F) = (1/2)(1/2) + (1/2)(1) = 3/4$$

y sustituyendo cada componente en la fórmula del Gini se obtiene que:

$$Gini(S) = 0.25[w_A(S) - w_B(S)] \quad S = R, U$$

Para calcular la desigualdad agregada G , notar que $w_B(R) < w_A(R) < w_B(U) < w_A(U)$, con distribución acumulada $F = (1 - \psi)/2$, $F = (1 - \psi)$, $F = (1 - \psi/2)$ y $F = 1$, respectivamente.

Nuevamente, se calcula cada componente:

$$E(wF) = w_B(R)(1/4) (1 - \psi)^2 + w_A(R)(1/2) (1 - \psi)^2 + w_B(U)(1 - \psi/2)(\psi/2) + w_A(U)(\psi/2)$$

$$E(F) = (3/4)(1 - \psi)^2 + (1 - \psi/4)\psi$$

$$E(w) = (1 - \psi)\mu_w(R) + \psi\mu_w(U) = \mu_w$$

y se sustituye en la fórmula del coeficiente de Gini.

El ejercicio realizado toma los valores de la Tabla 1 y computa los distintos valores de μ_w y $Gini$ para una grilla de valores de $\psi = 0, 0.05, 0.10, \dots, 0.95, 1$; que representan la transición del R a U .