

## Diferencias de los ingresos laborales entre los puestos de trabajo asalariado y los puestos de trabajo autogenerados en el Perú, 2007–2011

### *Labor Income Gaps Between Wage-Earner and Self-Employed Jobs in Peru, 2007–2011*

José S. Rodríguez<sup>1</sup>

DOI: 10.13043/DYS.71.4

#### Resumen

El objetivo principal de este artículo es someter a prueba la existencia de diferencias de los ingresos laborales entre los puestos de trabajo asalariado y los puestos de trabajo autogenerados ocupados por personas que son iguales en los mercados laborales del Perú. Para este fin, se utiliza la información de las encuestas de hogares de los años 2007–2011 y se emplean métodos semiparamétricos para hacer las comparaciones (propensity score matching, doble diferencia y doble diferencia con emparejamiento). Los resultados adquieren robustez mediante el uso de tres métodos semiparamétricos o no paramétricos distintos, junto con un conjunto amplio de datos que combina muestras de corte transversal con muestras panel. Los hallazgos sugieren la existencia de segmentación entre estos dos tipos de puestos de trabajo. Los cálculos fueron hechos para diferentes muestras: nacional, diferenciando por sexo y diferenciando por área de residencia. A escala nacional, los premios estimados se

---

1 El autor agradece los comentarios de Albert Berry, Adolfo Figueroa, Gabriel Rodríguez y Alan Sánchez. En etapas iniciales de este trabajo, el apoyo de Minoru Higa y Rodrigo Montes fue muy importante. Más recientemente, para la elaboración de la versión final del trabajo el apoyo de Álvaro Hopkins fue invaluable. El autor es del Departamento de Economía de la Pontificia Universidad Católica del Perú. Correo electrónico: jrodrig@pucp.edu.pe.

Este artículo fue recibido el 2 de marzo de 2012; revisado el 19 de abril de 2013 y, finalmente, aceptado el 24 de abril de 2013.

ubican en el rango de 35% a 53%, dependiendo del método empleado. Entre las mujeres y los residentes de las áreas rurales, los premios son mayores.

*Palabras clave:* Perú, mercados de trabajo segmentados, diferencias de ingresos laborales.

*Clasificación JEL:* J21, J31, J42.

## Abstract

The main objective of the work is to test the existence of differences in earnings between salaried jobs and self-generated jobs in the labor markets of Peru. To this purpose it uses data from household surveys from 2007 to 2011. Also it uses non-parametric and semi-parametric methods for premium estimates (propensity score matching, double difference and double difference with matching). The use of 3 different methods and the use of several datasets that combines cross-section and panel samples, provide robustness to the results. The findings suggest the existence of segmentation between these two types of jobs. The calculations were made for different population samples: national, by sex and by area of residence. At national level premiums are in the range of 35-53% depending on the method used. Among women and residents of rural areas the premiums are larger.

*Key words:* Peru, labor markets segmentation, labor income gaps.

*JEL classification:* J21, J31, J42.

## Introducción

Dos características importantes resaltan en los mercados laborales del Perú. Por un lado, las altas y persistentes proporciones de informalidad de su fuerza laboral (Díaz, 2009; Gasparini y Tornarolli, 2009; Morales, Rodríguez, Higa y Montes, 2010; Perry *et al.*, 2007) y, por otro, los también altos y persistentes porcentajes de la fuerza laboral autoempleada, es decir, que genera sus propios puestos de trabajo. Según recientes estimaciones del Instituto Nacional de Estadística e Informática (INEI, 2012), en el año 2011 la informali-

dad nacional representó poco más del 75% de la población económicamente activa ocupada.

Respecto a la importancia relativa de los puestos de trabajo autogenerados (incluidos aquí los trabajadores familiares no remunerados, TFNR, en adelante), se observa que estos representan poco más de la mitad de la población económicamente activa ocupada nacional. En 2011 poco menos de la mitad de los ocupados eran asalariados (49%), mientras que la diferencia se distribuía entre autoempleados (35%), TFNR (13%) y el resto constituido mayormente por los trabajadores domésticos (3,1%)<sup>2</sup>. En comparación con el año 2002, el empleo asalariado incrementó su participación en cuatro puntos porcentuales, a costa de la de los autoempleados.

Como puede apreciarse en el gráfico 1 con información de mediados de la década pasada, el Perú compartía con Paraguay y Bolivia (países con menor PIB per cápita) proporciones de autoempleados y TFNR cercanas al 50% del empleo. Se observa, además, una relación negativa entre el PIB per cápita y la proporción del empleo autogenerado. En las economías desarrolladas, en promedio solo el 13% de la fuerza de trabajo corresponde a este tipo de ocupaciones y más de la mitad de estas economías tienen porcentajes menores o iguales al 10%<sup>3</sup>. Esto sugiere que el desarrollo económico está asociado con la expansión de relaciones laborales de dependencia (*i. e.*, formas de trabajo asalariado) y la disminución de los empleos autogenerados.

Respecto a las condiciones de trabajo en el Perú, se observa que cuando se comparan los promedios de los ingresos —sin controlar por las características personales— de los asalariados privados con los de los trabajadores independientes, los ingresos de los asalariados en la Administración gubernamental son mayores que los de los asalariados en las empresas privadas en su conjunto. A su vez, los ingresos de los asalariados privados son mayores que los de los independientes. Y aún más notable, los ingresos de los asalariados en empresas privadas del estrato más pequeño (*i. e.*, de hasta nueve trabajadores) son

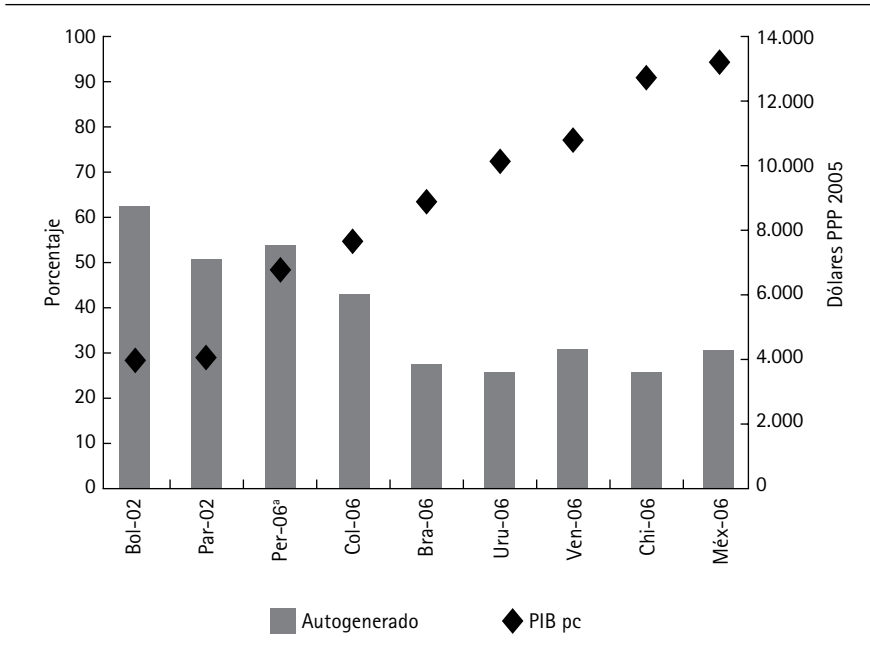
---

2 En las estadísticas laborales oficiales del Perú se excluyen de la población económicamente activa ocupada los TFNR que trabajan menos de quince horas por semana. Para el resto de las categorías ocupacionales, basta con que los individuos trabajen al menos una hora a la semana para ser considerados ocupados.

3 Este porcentaje está casi exclusivamente vinculado a los trabajadores independientes, pues la cantidad de TFNR es muy pequeña o incluso inexistente.

mayores que los ingresos de los independientes<sup>4</sup>. Es cierto, sin embargo, que estas diferencias podrían disminuir e incluso desaparecer, luego de controlar por los atributos productivos del trabajador.

**Gráfico 1.** Autoempleo y trabajo familiar no remunerado, y PIB per cápita en algunos países de América Latina (circa 2006)<sup>a</sup>



<sup>a</sup> El dato de Perú fue estimado por el autor a partir de la ENAHO 2006.

Fuente: *Indicadores de los objetivos del desarrollo del milenio* (<http://mdgs.un.org/unsd/mdg/Data.aspx>) y *World Development Indicators* (<http://data.worldbank.org/data-catalog/world-development-indicators>). Consultadoos 19 de noviembre del 2010).

Por otro lado, entre los trabajadores independientes y los TFNR es en quienes menos se cumplen la protección social y las obligaciones formales. Por ejemplo, más del 95% de los independientes, estén o no registrados ante la Superintendencia Nacional de Aduanas y de Administración Tributaria (SUNAT), no llevan sistema de contabilidad alguno. Entre los TFNR la cobertura de salud o pensiones es inexistente. Mientras tanto, en el grupo de los asalariados en empresas privadas también hay un importante incumplimiento de las normas, pero no en

4 Véase en el portal del Ministerio de Trabajo del Perú: <http://www.mintra.gob.pe/mostrarContenido.php?id=93&tip=54>. Consultado el 10 de abril de 2013.

la magnitud que se observa entre las ocupaciones autogeneradas. Este incumplimiento es notablemente mayor entre las empresas pequeñas (hasta nueve trabajadores) que entre las medianas y grandes (cincuenta y más trabajadores)<sup>5</sup>.

En síntesis, con respecto a su PIB per cápita, es muy grande la fuerza de trabajo del Perú que está trabajando en ocupaciones autogeneradas. El tamaño relativo de estas, en el conjunto del empleo, ha disminuido en la última década, pero muy poco respecto de lo que se esperaba luego de las reformas estructurales de los años noventa y al rápido ritmo de crecimiento que ha mostrado la economía peruana en los últimos años. Los perfiles de los ingresos promedio de los trabajadores independientes y los asalariados muestran que los primeros ganan bastante menos que los segundos, diferencias que se observan aun luego de controlar por el tamaño de los establecimientos en donde trabajan los asalariados y el área de residencia.

En la literatura teórica se pueden encontrar diferentes formas de explicar la existencia de segmentación en los mercados laborales y, asociado a ello, las diferencias de ingresos para trabajadores igualmente productivos. Desde la teoría del dualismo se postula que factores de orden institucional —como la presencia de los sindicatos y los salarios mínimos— están en la base de tales diferencias (Harris y Todaro, 1970; Lewis, 1954). Posteriormente, los modelos de salario de eficiencia (Akerloff y Yellen, 1986) y, en particular, el modelo de rotación laboral (Stiglitz, 1974) han apuntado en la misma dirección, pero haciendo que la determinación de los salarios sea endógena. Figueroa (2011) utiliza el concepto de salarios de eficiencia para modelar el equilibrio en los mercados laborales en economías sobrepobladas, de manera que el modelo replique la existencia de diferencias de ingresos, así como la de subempleo y desempleo.

La investigación aplicada en los últimos años sobre el Perú se ha concentrado más en someter a prueba la existencia de segmentación laboral entre trabajadores formales e informales. Los resultados de estas investigaciones son mixtos (véanse Barco y Vargas, 2010; Saavedra y Chong, 1999; Yamada, 1996). Este no es un resultado particular para el Perú. Otros países de América Latina muestran algo semejante (véanse, para Brasil, Ulyseas, 2006; para México, Duval-Hernández (2006); y para Argentina, Pratap y Quintin, 2006).

---

5 El 90% de los empleadores no están registrados y el 80% no llevan sistema de contabilidad. Además, el 88% de los trabajadores del hogar no tienen seguro de salud (Rodríguez e Higa, 2010).

El presente trabajo busca analizar si las diferencias de ingresos entre los trabajadores autoempleados y los asalariados responden a diferencias en los atributos productivos de la mano de obra o, por el contrario, si son el resultado de alguna forma de segmentación de mercados. Para ello y a diferencia de lo realizado en trabajos anteriores para el Perú, se utilizan metodologías de evaluación de impacto, cuya pieza central es que las muestras utilizadas para medir las diferencias sean comparables. Se implementaron los métodos de propensity score matching (PSM) para emparejar a los asalariados y los autoempleados, el de doble diferencia (DD) y el de doble diferencia con emparejamiento mediante el PSM (DD-PSM)<sup>6</sup>. El primero empareja las poblaciones con base en los atributos observables, mientras que el segundo supera el problema de la comparabilidad utilizando muestras de panel de individuos, con lo que tanto los atributos observables como los no observables son controlados. La literatura de evaluación de impacto sugiere el uso del emparejamiento para robustecer los resultados de la técnica de DD<sup>7</sup>.

Los resultados obtenidos pueden ser considerados como evidencia de la existencia de segmentación de mercados. Con los tres métodos se encuentran diferencias importantes y significativas en los ingresos a favor de los trabajadores que ocupan puestos asalariados, para todos los años analizados con datos de corte transversal y, en el caso de los datos de panel, para todos los subpaneles entre 2007 y 2010. Más aún, la existencia de movilidad de ida y vuelta entre estos dos tipos de puestos entre las mismas personas sugiere que sí es posible que los trabajadores que son autoempleados puedan acceder a puestos asalariados y, además, que tengan un incremento importante en sus ingresos laborales.

El documento consta de esta introducción y cuatro secciones. En la primera se hace una breve revisión de la literatura teórica y aplicada para el caso del Perú en particular. En la segunda sección se presentan a grandes rasgos las metodologías empleadas, así como las bases de datos utilizadas. En la tercera

---

6 Prata y Quintin (2006) utilizan los puntajes del PSM para emparejar y analizar los mercados laborales en Argentina. En cambio, Ñopo (2008) y Barco y Vargas (2010) emparejan con base en las características de los individuos, para analizar la discriminación por sexo y la segmentación en los mercados laborales peruanos.

7 El trabajo de Blundell y Costa Dias (2000) es uno de los primeros en sugerir la aplicación de DD-PSM. Véanse también Gertler, Martínez, Premand, Rawlings y Vermeersch (2011).

sección aparecen los resultados de las estimaciones y la cuarta contiene las conclusiones.

## I. Breve revisión de la literatura

En la literatura teórica se pueden encontrar diferentes formas de explicar la coexistencia de segmentos de mercados de trabajo estructuralmente distintos y de diferencias en los ingresos laborales para trabajadores también productivos. Así mismo, hay investigaciones que cuestionan el hecho de que los distintos sectores o estratos de los mercados laborales implican la existencia de segmentación de mercados y que las diferencias en los ingresos laborales reflejan tal segmentación. Y lo que es más interesante, hay trabajos relativamente recientes que cuestionan la posibilidad de verificar de manera empírica la existencia de segmentación.

Entre los que defienden su existencia, es imprescindible mencionar la teoría del dualismo, la cual postula que la baja dotación de capital con relación a la mano de obra está en la base de un dualismo tecnológico que da origen a un sector de subsistencia (inicialmente asociado con la agricultura tradicional) y un sector capitalista (asociado con la industria en las áreas urbanas). En los términos de Lewis (1954), este exceso de mano obra constituye una oferta ilimitada. Por otro lado, el dualismo tecnológico, que configura a su vez un dualismo económico en los mercados laborales (Ranis, 2006), determina que la remuneración del trabajo no es la misma en ambos sectores. Lewis (1954, 1979) menciona varias razones para ello, pero en particular la presencia de los sindicatos en su negociación y el establecimiento de los salarios en el sector capitalista. Para Todaro (1969) y Harris y Todaro (1970), los salarios mínimos son otra razón para la existencia de diferencias en las remuneraciones del trabajo, dependiendo del sector en donde se labora. Una pieza clave en los modelos de Lewis y de Harris y Todaro es el rol de los salarios en el sector capitalista, en donde estos son exógenos al modelo, bien sea por los sindicatos o por los salarios mínimos. Este modelo dual se extendió más adelante para incorporar al menos un tercer sector: las unidades de producción de subsistencia en las áreas urbanas o el sector urbano informal (Fields, 1975).

La literatura de economía laboral ha desarrollado posteriormente otros argumentos por los que habría segmentación en los mercados laborales, entre los

cuales se emplea la denominación de "salarios de eficiencia" (Akerloff y Yellen, 1986). Por ejemplo, el modelo de costos de rotación laboral de Stiglitz (1974) levanta el supuesto de que el salario en el sector urbano sea exógeno. Este modelo busca explicar simultáneamente la diferencia de los ingresos urbanos con respecto a los rurales, el desempleo urbano y la asignación de la mano de obra entre los sectores urbano y rural. Para ello supone que cuando un trabajador deja un puesto, reemplazarlo implica costos de contratación y entrenamiento relacionado con el puesto. En consecuencia, cada puesto de trabajo implica un costo en la planilla del trabajador más el costo de rotación asociado a ese puesto. En este contexto, las firmas eligen, además de la cantidad de trabajadores a contratar, el salario que han de pagar, como una manera de minimizar el costo laboral por empleado que contratan. Claramente, el salario que se paga a los trabajadores será mayor y la cantidad de empleo menor a los que se observarían en ausencia de costos de rotación.

Figueroa (2011) desarrolla un modelo que busca explicar la existencia de desempleo y de diferencias de ingresos entre sectores en economías sobrepobladas. Un elemento central de este modelo es que en el sector capitalista los salarios se establecen por encima del costo de oportunidad (que son los ingresos del sector de subsistencia) para disciplinar la fuerza de trabajo y disminuir la tendencia a eludir las responsabilidades de los trabajadores. Por otro lado, la mano de obra en el sector de subsistencia se autoemplea hasta que el ingreso en el margen se iguala al salario esperado, que es una fracción del salario capitalista. Gracias a estos supuestos en la economía, hay subempleo (en el sector de subsistencia con bajos ingresos), salarios más altos que los ingresos de los autoempleados (porque es el mecanismo que disciplina la fuerza de trabajo asalariada) y desempleo (dado el racionamiento de empleos en el sector capitalista y el salario esperado como regulador de la asignación de empleo en el sector de subsistencia).

La crítica a los modelos dualistas abarca diferentes aspectos. Por un lado, se argumenta que la informalidad (con la que, en gran medida, se identifica el empleo en el sector de subsistencia) tiene su principal causa en los altos costos de la formalización (De Soto, 1986; Loayza, 1996, 2007; Perry *et al.*, 2007). Por otro lado, se sostiene también que las microempresas informales urbanas son resultado de decisiones voluntarias. Esto se debe en parte a la incapacidad del Estado de supervisar y hacer cumplir la legislación que regula las actividades económicas, entre ellas las empresariales y laborales (Maloney, 2004).



Finalmente, los trabajadores de las empresas informales pueden apreciar ciertas características de los puestos de trabajo, como por ejemplo, la flexibilidad en los horarios, la duración de la jornada de trabajo y el recibir un salario neto mayor en comparación con el que devengarían en términos netos en un puesto formal (Perry *et al.*, 2007). Entre la evidencia que se muestra para sostener estos argumentos están los hechos de que una parte importante de los trabajadores que se vuelven informales lo hacen voluntariamente y que una proporción importante de estos prefieren este tipo de puesto de trabajo que el de un asalariado formal (Maloney, 1999, 2004; Perry *et al.*, 2007).

Los trabajos de Ulyssea (2010) y Ulyssea y Paes de Barros (2010) modelan el funcionamiento de los mercados laborales completamente integrados (en donde no hay segmentación), así como mercados completamente segmentados. Los equilibrios en cada uno de estos tipos de mercados son tales que no es posible afirmar que los hechos estilizados observados en Brasil y, en general, en economías en desarrollo con una importante presencia de empleo informal pueden ser explicados por cualquiera de los dos modelos. Dicho de otra forma, no es posible afirmar que los hechos estilizados son necesariamente evidencia de la existencia de segmentación.

En la escasa literatura específica del Perú, se han estudiado las diferencias de ingresos desde distintas perspectivas. Por un lado, desde el enfoque de discriminación los trabajos han enfatizado las diferencias de ingresos por sexo y, más recientemente, por raza. Por otro lado, los enfoques de segmentación en los mercados laborales han explorado las diferencias de ingresos entre trabajadores formales e informales. En algunos casos se hace la distinción, además, entre los autoempleados y los asalariados. Yamada y Ramos (1996)<sup>8</sup>, con información de Lima Metropolitana de fines de los años ochenta y principios de los noventa, concluyen que en promedio un trabajador autoempleado informal genera más ingreso que como asalariado. Un aspecto central del modelo es que el producto en el sector de servicios depende de la capacidad empresarial. Siendo esta una habilidad que no se distribuye homogéneamente entre la fuerza de trabajo, es de esperarse que quienes tienen más de ella busquen tener su propio negocio, pues ello les proveerá un premio adicional. Los resul-

---

8 Véase también Yamada (1996).

tados indican que los autoempleados ganan más que como asalariados, entre 16% y 123%, dependiendo del año de la estimación<sup>9</sup>.

Saavedra y Chong (1999) examinan la existencia de segmentación en los mercados laborales en las áreas urbanas del Perú para el año 1994. La segmentación que ellos someten a prueba distingue cuatro grupos de trabajadores: asalariados formales, asalariados informales, independientes formales e independientes informales. Los resultados muestran que las diferencias de ingresos entre independientes formales e informales se reducen hasta prácticamente desaparecer. También se reducen, pero sin desaparecer, con relación a los asalariados formales, quienes después de los controles mantienen una ventaja entre 5% y 6% respecto a los independientes formales o informales. Los asalariados informales mantienen diferencias entre 25% y 31% con relación a los otros tres segmentos. Un aspecto que debe resaltarse es que la diferencia entre los asalariados formales e informales, que era inicialmente de seis puntos porcentuales a favor de los formales, crece hasta 31 puntos porcentuales luego de los controles econométricos.

Barco y Vargas (2010) descomponen la diferencia de los ingresos por hora de los trabajadores formales e informales, utilizando el método no paramétrico propuesto por Ñopo (2008). Este método empareja a individuos de los segmentos formal e informal mediante las características observables de aquellos. Utilizan microdatos de 2007 para las áreas urbanas peruanas y hacen la evaluación por separado para los trabajadores asalariados y los independientes. Los resultados muestran que los ingresos de los trabajadores formales, tanto entre los asalariados como entre los independientes, representan aproximadamente 2,8 veces los correspondientes a los de los informales. Más de la mitad de esa diferencia (entre 56% y 61%) se debe a que hay trabajadores que no forman parte del soporte común. La diferencia restante, 41% en el caso de los asalariados y 69% en el caso de los independientes, representa o refleja la segmentación laboral.

En este trabajo se busca aportar a la discusión sobre la existencia o no de segmentación, mediante el uso de métodos novedosos para este campo de investigación, que tienen por virtud aprovechar mejor la información de la distribución de

---

9 En Rodríguez (2011), véase una revisión crítica de Yamada y Ramos (1996) y otros autores para el caso del Perú.

las características de los individuos. Para ello, se emplean tres procedimientos para comparar y hallar las diferencias de ingresos entre los puestos de trabajo asalariado y los autogenerados. Además, se retoma la idea de que la segmentación está asociada a rasgos estructurales (*i.e.*, poco capital con relación a la población) y no solamente a los aspectos institucionales y los costos asociados a ellos. En este sentido, se postula que no es necesariamente lo mismo analizar la segmentación desde el punto de vista institucional (en donde la informalidad puede distinguir segmentos o estratos en las diferentes ocupaciones) que analizarla a partir de la incapacidad o limitada capacidad del sector capitalista de generar suficientes puestos de trabajo para absorber una mayor proporción de mano de obra bajo modalidades de trabajo asalariado. Por último, vale la pena mencionar que se hizo un tratamiento muy cuidadoso y exhaustivo de la información disponible —la cual se ha hecho particularmente rica en los últimos años—, con el propósito de minimizar algunos problemas que se considera han tenido anteriores estudios. En particular, se ha procurado hacer un tratamiento más cuidadoso de los trabajadores autoempleados, quienes por sus particulares formas de producción suelen utilizar a TFNR para producir, con lo cual se tiende a sobrestimar los ingresos que pueden ser atribuidos al que está a cargo de organizar la producción.

## II. Metodología

La pregunta central que se busca responder en este trabajo es si trabajadores idénticos en sus atributos productivos perciben idénticos ingresos laborales, sin importar en dónde se encuentra el puesto laboral que ocupan. En particular, interesa evaluar si existen diferencias en dichos ingresos entre los puestos asalariados y los autogenerados. Nótese que, de acuerdo con lo visto en la revisión de la literatura teórica, la existencia de diferencias de ingresos no es un resultado que se espera observar si el mercado de trabajo es perfectamente competitivo y se encuentra en equilibrio walrasiano. En cambio, los modelos que recogen la existencia de economías sobrepobladas sí son consistentes con la existencia de diferenciales de ingresos entre individuos idénticos, dependiendo del sector o segmento en donde se encuentra el puesto de trabajo.

A continuación se presenta la estrategia metodológica utilizada para tratar de responder la pregunta principal del trabajo. Luego se describirán los diferentes datos empleados en las estimaciones empíricas, así como el tratamiento que

se hizo de la base de datos hasta conseguir el conjunto de información que finalmente se empleó en el trabajo empírico.

## A. Estrategia metodológica

Asumamos que toda persona que forma parte de la población en edad de trabajar tiene solamente dos opciones para hacerlo: como asalariada o en un puesto autogenerado. Si la asignación a los puestos de trabajo fuera aleatoria, entonces la distribución de las características o atributos productivos de los individuos sería semejante en ambas subpoblaciones (en los puestos asalariados y en los autogenerados), sean estos atributos observables o no. En consecuencia, si existiera algún diferencial de ingresos entre ambas subpoblaciones este podría ser atribuido enteramente a los puestos de trabajo y no a las características inherentes de los trabajadores.

Evidentemente, no parece razonable asumir que la asignación sea aleatoria y si se desea evaluar si hay diferencias en los ingresos debidas a los puestos de trabajo, entonces, se debe buscar la manera de que ambas poblaciones sean comparables y equivalentes, es decir, homogéneas en los atributos productivos inherentes a las personas. Empleando los conceptos de la literatura relativa a la evaluación de impacto, este problema podría ser formulado de la siguiente manera. Si lo que se quiere saber es cuál es el premio que recibe un individuo por trabajar en un puesto asalariado en vez de uno autogenerado, el problema es que en un momento determinado solo se observa al individuo en uno de los dos tipos de puestos<sup>10</sup>. Por ejemplo, si el trabajador es asalariado y, por ende, solo se observa su ingreso laboral como tal, el ingreso que recibiría en un puesto autogenerado es un contrafactual. Así, el asunto es cómo estimar el ingreso no observado.

Al menos tres métodos distintos pueden ser empleados para estimar el contrafactual<sup>11</sup>. Un método asume que se sabe de qué factores depende la asignación

---

10 En la literatura, a quienes trabajan en puestos autogenerados se les denomina de diferentes maneras: autoempleados, trabajadores por cuenta propia, trabajadores independientes. Aquí nos referiremos a los puestos autogenerados y a los autoempleados indistintamente.

11 Algunos documentos muy útiles sobre diferentes métodos que se pueden utilizar para las evaluaciones de impacto, y que pueden adaptarse para temas como los que se abordan en este trabajo, son los de Khandker, Koolwal y Samad (2010) y Bernal y Peña (2011).

de las personas a cada uno de los dos tipos de puestos y que dichos factores constituyen variables observables. Si fuera así, se puede implementar el método denominado propensity score matching (PSM) para hacer comparables los dos subconjuntos de población. Es decir, si el "tratamiento" es equivalente a trabajar en un puesto asalariado, la idea es encontrar un grupo de control (*i. e.*, que trabaja en un puesto autogenerado) que sea comparable con el de tratamiento. Esta equiparación se hace utilizando la propensión a ser parte del grupo de tratamiento con base en un conjunto de variables que se asume determinan la asignación a los puestos asalariados. Las propensiones son básicamente las probabilidades estimadas de ser asalariado. Con esta información se emparejan, bajo diferentes algoritmos, personas de ambas subpoblaciones, para con ello estimar la diferencia de ingresos o el premio promedio<sup>12</sup>. Esta metodología se puede emplear cuando se dispone únicamente de información de corte transversal, lo que implica que efectivamente solo se observa a cada individuo en uno o en el otro puesto. La siguiente expresión representa la forma general como puede ser estimado el efecto del tratamiento utilizando PSM:

$$\hat{\alpha}_{PSM} = \sum_{i \in T} \left( Y_i - \sum_{j \in C} W_{ij} Y_j \right) w_i$$

en donde  $Y_i$  y  $Y_j$  representan los ingresos laborales de individuos  $i$  y  $j$ ;  $T$  y  $C$  representan los grupos de tratamiento y control, respectivamente;  $W_{ij}$  es un ponderador de la observación de control  $j$  para el individuo de tratamiento  $i$ ; y  $w_i$  es el ponderador que ajusta la distribución del resultado a la muestra de tratados. La definición de estos ponderadores depende del método de emparejamiento utilizado. Una de las principales limitaciones de este método es que asume (o requiere) que la regla de asignación entre tipos de puestos depende exclusivamente de atributos observables y que estos son medibles a partir de información existente en la base de datos.

Un segundo método permite relajar el supuesto de que la asignación depende exclusivamente de atributos observables o, lo que es equivalente, que depende tanto de atributos observables como de no observables. Si un mismo individuo es

---

12 Cada uno de los individuos que pertenecen al grupo de tratamiento se pueden comparar con: a) el o los vecinos más cercanos, b) los vecinos más cercanos dentro de cierta distancia máxima (definida por un radio) y c) un individuo ficticio construido a partir de la información de todos los individuos que pertenecen al grupo de control (kernel). Véanse Khandker *et al.* (2010) y Bernal y Peña (2011).

observado en dos momentos distintos, siendo en el primero autoempleado y en el segundo asalariado, entonces la diferencia en los ingresos en ambos puestos solo reflejará diferencias en el puesto, pues al ser el mismo individuo, tanto las características observables como las no observables serán las mismas en ambos momentos. Sin embargo, es posible que aún haya un sesgo de selección, pues los que pasan de autoempleados a asalariados pueden ser distintos de los que permanecen como autoempleados en ambos períodos. Luego, la manera de controlar por este sesgo es obtener la diferencia de las diferencias (o doble diferencia), es decir, restar de la diferencia de los ingresos, de los que pasan de ser autoempleados a asalariados, la diferencia de los que permanecen como autoempleados.

El estimador del premio empleando DD es el siguiente:

$$\hat{\alpha}_{DD} = \left( \bar{Y}_{t_1}^T - \bar{Y}_{t_0}^T \right) - \left( \bar{Y}_{t_1}^C - \bar{Y}_{t_0}^C \right)$$

en donde  $\bar{Y}^T$  y  $\bar{Y}^C$  representan los promedios de los ingresos de los grupos de tratamiento y control, respectivamente;  $t_0$  y  $t_1$  son los dos momentos pre y postratamiento en los que es observado cada individuo en el panel. Un requisito para emplear e interpretar correctamente los resultados de este método es que las tendencias de los ingresos de ambas poblaciones sean paralelas.

Un tercer método combina los dos métodos anteriores. Esto es, se calcula la diferencia de las diferencias, pero solo teniendo en cuenta las dos poblaciones emparejadas. La población de tratamiento es la que pasa de ser autoempleada en  $t_0$  a ser asalariada en  $t_1$ . La otra población es la que permanece como autoempleada en ambos momentos. Este método, denominado diferencia de las diferencias con emparejamiento, en opinión de algunos autores es mejor y más robusto, pues controla en los observables y los no observables y guarda la comparabilidad de las poblaciones (Bernal y Peña, 2011; Blundell y Costa Dias, 2000).

La estimación del premio a través de este método se hace de acuerdo con la siguiente expresión:

$$\hat{\alpha}_{PSM-DD} = \sum_{i \in T} \left( \left( Y_{it_1} - Y_{it_0} \right) - \sum_{j \in C} W_{ij} \left( Y_{jt_1} - Y_{jt_0} \right) \right) w_i$$

Es interesante resaltar cómo esta expresión recoge elementos de las dos anteriores para PSM y DD. Por un lado, la primera expresión en paréntesis del lado derecho recoge la diferencia de los ingresos del grupo de tratamiento entre el período pretratamiento  $t_0$  y el período postratamiento  $t_1$ . Por otro lado, la segunda expresión representa la diferencia de los ingresos en tales períodos entre la población que forma parte del tratamiento pero que comparte las características iniciales de la población de tratamiento, pues eran autoempleados en  $t_0$  y son comparables, gracias al emparejamiento, con los que sí cambiaron de puesto de trabajo. Sin embargo, al igual que el estimador DD, se requiere que se cumpla el requisito de que las tendencias de los ingresos de ambos grupos sean paralelas.

## B. Base de datos

Se utilizaron los microdatos de los años 2007 al 2011 de la Encuesta Nacional de Hogares (ENAH) del Perú. Para estos años en particular, esta encuesta incluyó un diseño de muestra tanto de corte transversal como panel. Para la muestra transversal, el número efectivo de viviendas en cada año osciló entre 21.000 y 22.000<sup>13</sup>. En el diseño de muestra original, de cada una de las muestras anuales aproximadamente la tercera parte de las viviendas formaron parte de la muestra panel (alrededor de 7.000 viviendas). Pero como la muestra panel fue rotativa, en dos años consecutivos solo la cuarta parte de la muestra total de viviendas estuvo presente (cerca de 5.600 viviendas)<sup>14</sup>.

La ENAH es aplicada a una muestra continua de viviendas a lo largo de los doce meses de cada año y a escala nacional, y su capacidad de representación permite reportar resultados para todo el país, tanto en las áreas urbanas como rurales, para cada una de las veinticuatro regiones político-administrativas en las que se divide el territorio<sup>15</sup>. Durante la mayor parte de los años aquí con-

---

13 Excepto en 2011, cuando se incrementa a más de 24.000 el número efectivo de viviendas. En el diseño original de la muestra, dicho número se estableció en 22.600.

14 Poco más de 1.800 viviendas deberían estar presentes en la muestra panel de los cinco años, entre 2007 y 2011.

15 También se pueden reportar resultados a escala nacional y en cada una de las áreas para cada uno de los cuatro trimestres del año calendario.

siderados (2007-2011), las aproximadamente 22.000 viviendas incluidas permiten identificar a unas 95.000 personas en cada muestra anual<sup>16</sup>.

Se utilizaron dos tipos de criterios para la exclusión de los individuos de la base que finalmente fue empleada en los análisis cuantitativos. Por un lado, fueron eliminados los casos en los que faltaba información en las variables que se utilizan en los análisis estadísticos y econométricos y, por otro, se excluyeron los individuos cuyas características no corresponden a las de la población que se busca analizar. Es importante señalar que la pérdida de casos por falta de información no es muy importante. Mucho más importante es la disminución en el número de casos debido a la delimitación de la población de referencia del estudio (asalariados y autoempleados) y por la homogenización de esta población más específica por criterios que se presentan y explican a continuación.

De la muestra total de personas registradas en las ENAHO de 2007 a 2011 (alrededor de 95.000 para cada año), el volumen se reduce en aproximadamente 33% cuando solo se considera la población en edad de trabajar (de catorce años y más según las normas peruanas). A su vez, dicha población se reduce a la mitad cuando únicamente se consideran los asalariados y los autoempleados. Es decir, el punto de partida para la muestra corte es de unas 33.000 personas en los años 2007 al 2010 y de 38.000 en 2011 (véase la parte superior del cuadro 1). La muestra panel está constituida por cerca de 3.800 personas (véase la parte inferior del cuadro 1).

### **C. Población de referencia y la implementación de las variables para el estudio**

A la población de referencia (asalariados y autoempleados con información completa) se le aplicaron filtros adicionales para incrementar el grado de confiabilidad de la información y para hacer más homogénea y, por lo tanto, comparable la población entre los dos tipos de puestos de trabajo. Así, como filtro para mejorar la confiabilidad de la información se impuso como requisito que la persona entrevistada fuera miembro del hogar y que la encuesta la respondiera directamente él o ella. Estos dos filtros en conjunto redujeron en aproximadamente 15% el tamaño de la muestra inicial de asalariados y

---

16 Personas de cualquier edad, sean o no miembros del hogar.



Cuadro 1. ENAHO corte transversal y panel. Perú 2007-2011

		Número de casos y su reducción como consecuencia de la imposición de filtros											
		2007		2008		2009		2010		2011			
		Personas	%	Personas	%	Personas	%	Personas	%	Personas	%		
<b>Muestra transversal</b>													
Después de la limpieza pero antes de los filtros <sup>a</sup>		32.947	100,0	32.591	100,0	33.344	100,0	33.283	100,0	38.298	100,0		
Filtros <sup>b</sup>													
Miembros y residentes del hogar		32.185	97,7	31.692	97,2	32.450	97,3	32.389	97,3	37.337	97,5		
Informante directo		28.707	87,1	28.533	87,5	29.013	87,0	28.908	86,9	33.028	86,2		
Excluye discapacitados		28.699	87,1	28.521	87,5	29.003	87,0	28.896	86,8	33.016	86,2		
Excluye asalariados del Estado		25.420	77,2	25.225	77,4	25.711	77,1	25.541	76,7	29.269	76,4		
Excluye autoempleados con TFNR		17.581	53,4	17.356	53,3	17.504	52,5	17.331	52,1	19.928	52,0		
A lo sumo secundaria completa		13.252	40,2	12.955	39,8	12.880	38,6	12.781	38,4	14.632	38,2		
Entre 25 y 65 años de edad		9.082	27,6	8.784	27,0	8.650	25,9	8.576	25,8	9.914	25,9		
Trabajan entre 24 y 72 horas		6.129	18,6	5.793	17,8	5.666	17,0	5.609	16,9	6.631	17,3		
<b>Muestra panel</b>													
Después de la limpieza pero antes de los filtros <sup>c</sup>		3.869	100,0	3.828	100,0	3.912	100,0	3.835	100,0				
Filtros <sup>b</sup>													
Miembros y residentes del hogar		3.457	89,4	3.415	89,2	3.487	89,1	3.411	88,9				
Informante directo		3.444	89,0	3.390	88,6	3.466	88,6	3.391	88,4				
Excluye discapacitados		3.441	88,9	3.389	88,5	3.465	88,6	3.390	88,4				
Excluye asalariados del Estado		3.441	88,9	3.389	88,5	3.465	88,6	3.390	88,4				
Excluye autoempleados con TFNR		1.697	43,9	1.667	43,5	1.704	43,6	1.617	42,2				
A lo sumo secundaria completa		1.402	36,2	1.399	36,5	1.405	35,9	1.299	33,9				
Entre 25 y 65 años de edad		1.109	28,7	1.095	28,6	1.051	26,9	976	25,4				
Trabajan entre 24 y 72 horas		682	17,6	645	16,8	587	15,0	566	14,8				

Notas: <sup>a</sup> La limpieza inicial implica eliminar los casos a los que les falta alguna información en las variables básicas que se emplean a lo largo del trabajo. Solamente incluye a asalariados y autoempleados. <sup>b</sup> La aplicación de los filtros es acumulativa, por ello la magnitud del efecto de cierto filtro es relativa y depende del orden en el que fueron aplicados. <sup>c</sup> La limpieza inicial implica eliminar los casos a los que les falta alguna información en las variables básicas que se emplean a lo largo del trabajo. La muestra se compone únicamente de aquellos trabajadores que hacen la transición en dos años consecutivos de autoempleados a asalariados y de autoempleados a autoempleados.

Fuente: ENAHO 2007 al 2011. Elaboración propia.

autoempleados. Luego se consideraron otros cinco criterios para disminuir la heterogeneidad de las subpoblaciones: a) se excluyeron las personas que trabajaban para el Estado, b) se excluyeron los autoempleados que contaban con el apoyo de TFNR, c) se excluyó la población con educación postsecundaria; d) se excluyeron los que tenían menos de 25 años, o más de 65 años y e) se excluyeron quienes trabajaban semanalmente menos de 24 horas o más de 72 horas (véase cuadro 1).

La exclusión de los asalariados estatales se hizo para evitar los sesgos que podría introducir el hecho de que el Estado no fuera un "tomador de precios" en los mercados de trabajo. El problema con los autoempleados apoyados por TFNR en sus unidades de producción es que la presencia de estos familiares puede sesgar hacia arriba los ingresos del autoempleado, pues no hay cómo restar de los ingresos que estos declaran la parte que corresponde al aporte de los TFNR. Es cierto, sin embargo, que aun después de esta exclusión puede subsistir una sobrestimación de los ingresos en la medida en que no se descuentan de los ingresos declarados la parte que corresponde a otros factores (v. g., infraestructura física y equipos). La restricción de incluir solo a la población con secundaria completa o menos deja por fuera un tipo de mano de obra más especializada y heterogénea, debido a que la educación postsecundaria incluye al menos dos variantes, la superior no universitaria y la universitaria, y en ellas hay mucha heterogeneidad entre las distintas especialidades. Se sabe que en la práctica la participación de personas con estudios postsecundarios es relativamente más importante entre los asalariados que entre los independientes. La cohorte de 25 a 65 años permite concentrar la atención en la población en edad de trabajar con las más altas tasas de participación y que en su mayor parte ha concluido la adquisición de escolaridad básica (i. e., primaria y secundaria). Finalmente, limitar al conjunto de trabajadores que laboran no menos de 24 horas (aproximadamente la mitad de una jornada completa de 48 horas por semana) ni más de 72 (jornada y media) elimina, por un lado, las ocupaciones marginales (pocas horas a la semana) y los valores extremos, por el otro.

Respecto a la forma como se definieron las variables que son objeto del análisis, se debe mencionar que con relación a las variables laborales la atención se concentró en la información relativa a la ocupación principal, independientemente de la existencia, o no, de ocupaciones secundarias.

Los ingresos son aquellos reportados por las propias personas e incluyen, en el caso de los asalariados, los ingresos monetarios más los pagos en especie y los ingresos extraordinarios<sup>17</sup>. En el caso de los autoempleados, los ingresos son las ganancias netas más el valor del autoconsumo. Las magnitudes monetarias han sido todas deflactadas temporal y espacialmente, utilizando los deflatores correspondientes proporcionados por el INEI para estos fines. Los valores reales, por lo tanto, corresponden al promedio del año 2009 de Lima Metropolitana. Los montos de los ingresos reales fueron estandarizados por las horas trabajadas, es decir, fueron divididos por la cantidad de horas efectivamente trabajadas en la ocupación principal durante el período de referencia (la semana anterior a la realización de la encuesta). Es importante mencionar que la base de datos de las ENAHO que pone a disposición del público el INEI incluye dos tipos de imputaciones. Por un lado, la imputación de los valores de ciertas variables clave en caso de que haya una omisión puntual y, por otro, la imputación de la información de todo el registro de una persona (v. g., cuando un individuo se niega a responder o simplemente no es encontrado durante las visitas al hogar). Para el cálculo tanto de la magnitud de los ingresos como de las horas se utilizan las versiones de estas que incluyen las imputaciones.

En el cuadro 2 se reportan por separado algunas estadísticas de las variables empleadas para la muestra transversal y la muestra panel. En la parte superior del cuadro se reportan los valores mínimos y máximos tanto de los promedios como de los desvíos estándar para la muestra transversal, mientras que en la parte inferior se hace lo mismo para la muestra panel. Nótese que para la muestra transversal se trabaja con cinco ENAHO comprendidas entre 2007 y 2011, mientras que para la muestra panel, con cuatro transiciones de dos años consecutivos dentro del conjunto de las ENAHO de 2007 a 2011. De esta forma, los mínimos y máximos son los promedios y desvíos estándar de los cinco años o las cuatro transiciones.

---

17 En los ingresos extraordinarios se incluyen las gratificaciones (fiestas patrias y Navidad), cuando procedan, la bonificación por escolaridad y la compensación por tiempo de servicios. Es importante resaltar, sin embargo, que la encuesta no distingue qué proporción de estos ingresos extraordinarios se debe a la ocupación principal y cuánto a la secundaria. Esto, evidentemente, podría sobrestimar el ingreso de la ocupación principal de los asalariados.

Cuadro 2. Estadísticas descriptivas de las variables empleadas con la muestra transversal y la muestra panel

Variable	Promedio <sup>1</sup>		Des. Est. <sup>2</sup>		Promedio <sup>3</sup>		Des. Est. <sup>3</sup>	
	Mín.	Máx.	Mín.	Máx.	Mín.	Máx.	Mín.	Máx.
<b>Muestra transversal<sup>b</sup></b>								
	<b>Asalariados</b>				<b>Autoempleados</b>			
Edad	39,3	40,8	10,5	10,8	43,6	45,3	11,0	11,4
(Edad) <sup>b</sup>								
Escolaridad	8,0	8,2	3,2	3,3	7,0	7,4	3,4	3,6
(Escolaridad) <sup>b</sup>								
Lengua materna	79,2	80,6	39,5	40,6	72,3	76,3	42,5	44,7
Sexo	72,0	75,2	43,2	44,9	58,6	61,8	48,6	49,3
Área de residencia	71,9	74,4	43,7	45,0	70,1	73,3	44,2	45,8
Gestión de la escuela	2,5	3,4	15,7	18,2	1,9	2,3	13,7	15,0
Jefatura del hogar	59,4	62,5	48,4	49,1	61,5	62,5	48,4	48,7
Estado civil	68,6	71,7	45,0	46,4	65,4	68,2	46,6	47,6
Ingreso laboral	211	270	248	358	139	176	106	185
<b>Muestra panel<sup>b</sup></b>								
	<b>Autoempleado-asalariado<sup>c</sup></b>				<b>Autoempleado-autoempleado<sup>c</sup></b>			
Edad	40,6	42,5	9,7	10,9	44,0	45,8	11,2	11,5
Años de escolaridad	6,4	7,2	3,3	3,6	6,9	7,1	3,4	3,7
Lengua materna	63,6	70,9	45,6	48,3	69,7	75,7	42,9	46,0
Sexo	70,1	83,7	37,1	46,0	56,2	62,8	48,4	49,7
Área de residencia	48,8	56,7	49,8	50,2	66,8	75,7	43,0	47,1
Gestión de la escuela	1,6	2,3	12,5	15,1	1,0	2,3	10,1	14,9
Jefatura del hogar	62,9	76,0	9,7	48,6	63,8	65,3	47,7	48,1
Estado civil	70,1	81,4	39,1	46,0	64,6	66,4	47,3	47,3
Ingreso laboral en $t$	112	165	89	123	121	155	91	133
Ingreso laboral en $t + 1$	168	212	110	136	131	188	104	144

*Notas:* <sup>a</sup> Para cada variable se reporta el menor (Mín) y el mayor (Máx) de los promedios y de los desvíos estándar de las cinco muestras transversales y de las cuatro muestras panel de la ENAHO. <sup>b</sup> Tanto para la muestra transversal como para la panel se considera la población que resulta luego de la "limpieza" y de la aplicación de los filtros. <sup>c</sup> En la muestra panel "autoempleado-asalariado" es aquel autoempleado en  $t$  que es asalariado en  $t + 1$ , mientras que "autoempleado-autoempleado" es quien habiendo sido autoempleado en  $t$  se mantiene como tal en  $t + 1$ .

*Fuente:* ENAHO 2007 al 2011. Elaboración del autor.

Es interesante notar que, a partir de la información de la muestra transversal, los asalariados tienden a ser, en comparación con los autoempleados, más jóvenes, con más años de escolaridad, con mayor presencia relativa de hombres y con una ligera sobrerrepresentación de aquellos cuya lengua materna es el español. Los ingresos laborales estandarizados por horas de trabajo de los asalariados son bastante superiores a los de los autoempleados. La muestra panel presenta algunas coincidencias, pero también interesantes diferencias. También se observa que, frente a aquellos que se mantienen como autoempleados, quienes cambian de autoempleados a asalariados son, desde el punto de vista relativo, más jóvenes pero no necesariamente tienen más años de escolaridad. Predominan los hombres y los de habla vernácula entre los que transitan hacia puestos asalariados. Mientras que en la muestra transversal el área de residencia no muestra diferencias entre asalariados y autoempleados, es claro en la muestra panel entre los que transitan hacia puestos asalariados están relativamente más presentes los que residen en áreas rurales, en comparación con los que se mantienen como autoempleados. Por último, entre los que pasan a ser asalariados hay una evidente presencia relativamente mayor de los casados y convivientes que entre los que se mantienen como autoempleados.

En el caso de los ingresos para el panel, se reportan los ingresos en los dos momentos. Si se consideran los valores mínimos de los promedios, los que transitan del autoempleo a ser asalariados tienen un ingreso relativamente más bajo, pero reportan un mayor incremento en comparación con la correspondiente cifra para los que se mantienen como asalariados. Con los valores máximos también se observa que los primeros incrementan sus ingresos en mayor proporción que los segundos.

Todo esto sugiere que, a pesar de haber aplicado un conjunto de filtros que tienden a homogenizar las poblaciones, estas continúan mostrando diferencias importantes en sus características y que los ingresos laborales estandarizados por las horas trabajadas muestran diferencias brutas en contra de los autoempleados. Sin embargo, una adecuada medición del premio entre asalariados y autoempleados requiere una comparación más cuidadosa y eso es precisamente lo que se hace en la siguiente sección.

### III. Resultados

En esta sección se reportan tres conjuntos de resultados: los que provienen del emparejamiento según las probabilidades de ser asalariado y los resultados de la diferencia de las diferencias de los autoempleados que pasan a ser asalariados, en dos versiones, esto es, con emparejamiento y sin él. Para esto, se hicieron estimaciones para las siguientes poblaciones: a) toda la población, b) hombres y mujeres por separado y c) urbano y rural por separado. Para el primer conjunto de resultados (utilizando solo técnicas de emparejamiento) se utilizaron las bases de datos con las muestras transversales, mientras que para el segundo y el tercer conjunto (diferencia de las diferencias con emparejamiento y sin él) se utilizaron las bases con las muestras panel.

El balanceo de las subpoblaciones, previo al emparejamiento, se hizo utilizando diferentes especificaciones de vectores de variables con las características observables de la persona. Se consideraron hasta veintiséis diferentes especificaciones, en donde tres variables siempre estuvieron presentes: edad, escolaridad y el cuadrado de la edad (véase cuadro 3). Los resultados que se presentan en esta sección se seleccionaron con el siguiente criterio: el conjunto de variables debe ser tal que permita el balanceo de las subpoblaciones en todos los casos de la muestra panel y en no menos de cuatro de los cinco casos de la muestra transversal.

Se encontró que para toda la población cinco especificaciones cumplen con el criterio arriba mencionado. En uno de los casos siempre balancea, y en los otros cuatro falla el balanceo en una especificación en las muestras transversales. Así, en el cuadro 3, en la columna "siempre" en "global", G-1 representa la especificación que siempre balancea, mientras que las especificaciones G-3, G-6, G-16x y G-17x casi siempre balancean. En el caso de mujeres y hombres, por separado, solo una especificación balancea siempre (S-11) para las mujeres, y hasta dos especificaciones para los hombres (S-11 y S-17). Casi siempre balancean tres especificaciones para las mujeres y cuatro para los hombres. Finalmente, en el caso de las áreas de residencia, cuatro especificaciones siempre balancean en el área rural (A-3, A-4, A-5 y A-17), pero ninguna en el área urbana. Casi siempre balancean cinco especificaciones en el área rural y una en la urbana<sup>18</sup>.

---

18 Para el emparejamiento de las subpoblaciones, se usó el procedimiento `psmatch2` de Stata (Leuven y Sianesi, 2012).

**Cuadro 3.** Resultados de las pruebas de balanceo de las diferentes especificaciones del logit

Núm. Espec.	Variables adicionales <sup>a</sup>	Global		Mujer		Hombre		Rural		Urbana	
		Siempre <sup>b</sup> siempre <sup>c</sup>	Casi siempre <sup>c</sup>	Siempre <sup>b</sup> siempre <sup>c</sup>	Casi siempre <sup>c</sup>	Siempre <sup>b</sup> siempre <sup>c</sup>	Casi siempre <sup>c</sup>	Siempre <sup>b</sup> siempre <sup>c</sup>	Casi siempre <sup>c</sup>	Siempre <sup>b</sup> siempre <sup>c</sup>	Casi siempre <sup>c</sup>
1	--	G-1				S-1					
2	civil										
3	castellano	G-3		S-3		S-3		A-3			
4	sexo										
5	urbano										
6	dedupri	G-6				S-5		A-5			
7	dedupri, sexo										
8	dedupri, civil										
9	dedupri, urbano										
10	dedupri, urbano, sexo										
11	dedupri, urbano, jefehh										
12	dedupri, urbano, civil										
13	dedupri, urbano, civil, sexo										
14	jefehh			S-11		S-11					
15	sexo, urbano,										
16	sexo, urbano, civil										
17	sexo, urbano, jefehh										
18	sexo, urbano, castellano										
19	sexo, urbano, civil, castellano										
20	sexo, urbano, jefehh, castellano										
21	sexo, dedupri, castellano										
22	dedupri, castellano	G-16x									
23	dedupri, castellano, aeduc2	G-17x		S-17		S-17		A-17			
24	castellano, aeduc2										
25	castellano, aeduc2, jefehh										
26	castellano, aeduc2, civil										

Notas: <sup>a</sup> Las variables que se incluyen en todas las especificaciones son las siguientes: edad, edad2 y aeduc.

<sup>b</sup> Se obtuvo balanceo tanto en las cinco muestras de corte transversal (ENAH0 2007-2011) como en los cuatro paneles consecutivos de las muestras panel (ENAH0 panel 2007-2008, 2008-2009, 2009-2010 y 2010-2011). <sup>c</sup> Se obtuvo balanceo en al menos cuatro de las cinco muestras de corte transversal (ENAH0 2007-2011) y en los cuatro paneles consecutivos de las muestras panel (ENAH0 panel 2007-2008, 2008-2009, 2009-2010 y 2010-2011).

\* En tanto A-8 y A-10 se usen para la desagregación en áreas rural y urbana, no se considera en estos casos la variable "urbano" en el logit.

## A. Los resultados con PSM<sup>19</sup>

Como se mencionó arriba, en esta sección solamente se reporta un resultado para cada población, sea aquel en el que siempre hay balanceo o, en caso de que esto no suceda, la especificación que casi siempre balancea. Cuando se divide por sexo y por área de residencia, se escogió aquella especificación que es común a ambos grupos (hombre y mujer, y rural y urbano). En el gráfico 2 y el cuadro 4 se reportan los resultados de los cálculos del premio en los ingresos laborales de ocupar un puesto asalariado en vez de uno autoempleado para: a) la población global, b) mujeres, c) hombres, d) residentes en áreas rurales y e) residentes en áreas urbanas. Dado que se dispone de resultados para cada uno de los cinco años de la ENAHO y que se utilizaron cinco algoritmos distintos para calcular los premios, para cada año se reportan el promedio simple, el valor máximo, el valor mínimo de los premios expresados en razón relativa al grupo de control y la diferencia inicial en los ingresos entre ambos tipos de puestos de trabajo. En cada uno de los recuadros del gráfico 2 el rectángulo de la caja de cada año tiene como límites superior e inferior los valores máximo y mínimo, respectivamente. Nótese que la escala en la que se presentan los resultados (eje vertical) es la diferencia de los logaritmos de los ingresos. Es importante mencionar que todos los resultados son estadísticamente significativos al 99%.

Por ejemplo, si se observan los resultados para toda la población ("Perú: nacional", en el gráfico 2 y el cuadro 4) en el año 2007, el premio mínimo es 46,5%, el máximo 50,7% y el promedio 47,8%. La diferencia inicial, antes del emparejamiento es 54,9%. Al mirar el conjunto de los resultados a escala global, se aprecia que: a) en cuatro de los cinco años el premio sin emparejamiento sobrestima el premio si se comparan adecuadamente las poblaciones, b) el tamaño de la distancia entre valores extremos va de cuatro a diez puntos porcentuales, lo que sugiere una importante sensibilidad en la estimación de los premios, dependiendo del algoritmo de estimación que se utilice para el emparejamiento y c) si se utilizan los promedios para medir el tamaño del premio, se podría decir que este no fue de menos de 44% ni más de 53%, lo que implica que el premio en una ocupación asalariada es bastante importante con relación a un puesto autogenerado.

---

19 Dado el gran volumen de cuadros que contienen los resultados de todas las especificaciones consideradas, estos no se incluyen en el texto ni en los anexos. El lector interesado puede solicitarlos al autor.



**Cuadro 4.** Premio en los ingresos laborales por hora del trabajador asalariado con relación al autoempleado. Método: PSM. Perú: 2007-2011

	2007		2008		2009		2010		2011	
	Dif <sup>a</sup>	S <sup>b</sup>	Dif <sup>a</sup>	S <sup>b</sup>	Dif <sup>a</sup>	S <sup>b</sup>	Dif <sup>a</sup>	S <sup>b</sup>	Dif <sup>a</sup>	S <sup>b</sup>
<b>Global (Espec. 1)<sup>d</sup></b>										
Diferencia inicial <sup>c</sup>	0,549	***	0,550	***	0,567	***	0,541	***	0,522	***
Premio con emparejamiento										
One to one	0,507	***	0,423	***	0,615	***	0,487	***	0,433	***
Radious	0,471	***	0,482	***	0,493	***	0,491	***	0,451	***
Kernel	0,465	***	0,486	***	0,493	***	0,487	***	0,460	***
NN(2)	0,478	***	0,484	***	0,533	***	0,484	***	0,423	***
NN(5)	0,469	***	0,516	***	0,500	***	0,518	***	0,444	***
<b>Mujeres (Espec. 11)<sup>e</sup></b>										
Diferencia inicial <sup>c</sup>	0,570	***	0,591	***	0,565	***	0,560	***	0,579	***
Premio con emparejamiento										
One to one	0,482	***	0,463	***	0,626	***	0,561	***	0,454	***
Radious	0,485	***	0,485	***	0,467	***	0,492	***	0,508	***
Kernel	0,479	***	0,493	***	0,459	***	0,499	***	0,509	***
NN(2)	0,464	***	0,494	***	0,556	***	0,559	***	0,422	***
NN(5)	0,449	***	0,503	***	0,469	***	0,495	***	0,461	***
<b>Hombres (Espec. 11)<sup>e</sup></b>										
Diferencia inicial <sup>c</sup>	0,413	***	0,422	***	0,446	***	0,429	***	0,408	***
Premio con emparejamiento										
One to one	0,394	***	0,445	***	0,446	***	0,452	***	0,329	***
Radious	0,382	***	0,414	***	0,429	***	0,414	***	0,386	***
Kernel	0,371	***	0,409	***	0,425	***	0,410	***	0,386	***
NN(2)	0,384	***	0,377	***	0,406	***	0,424	***	0,349	***
NN(5)	0,341	***	0,362	***	0,392	***	0,377	***	0,352	***
<b>Rural (Espec. 19)<sup>f</sup></b>										
Diferencia inicial <sup>c</sup>	0,700	***	0,715	***	0,732	***	0,639	***	0,681	***
Premio con emparejamiento										
One to one	0,548	***	0,652	***	0,665	***	0,527	***	0,551	***
Radious	0,550	***	0,635	***	0,635	***	0,540	***	0,599	***
Kernel	0,551	***	0,632	***	0,656	***	0,551	***	0,592	***
NN(2)	0,542	***	0,604	***	0,643	***	0,489	***	0,552	***
NN(5)	0,528	***	0,627	***	0,624	***	0,507	***	0,560	***
<b>Urbano (Espec. 19)<sup>f</sup></b>										
Diferencia inicial <sup>c</sup>	0,472	***	0,474	***	0,497	***	0,512	***	0,460	***

**Cuadro 4.** Premio en los ingresos laborales por hora del trabajador asalariado con relación al autoempleado. Método: PSM. Perú: 2007-2011 (continuación)

	2007		2008		2009		2010		2011	
	Dif <sup>a</sup>	S <sup>b</sup>	Dif <sup>a</sup>	S <sup>b</sup>	Dif <sup>a</sup>	S <sup>b</sup>	Dif <sup>a</sup>	S <sup>b</sup>	Dif <sup>a</sup>	S <sup>b</sup>
Premio con emparejamiento										
One to one	0,306	***	0,330	***	0,430	***	0,517	***	0,376	***
Radious	0,398	***	0,389	***	0,421	***	0,455	***	0,388	***
Kernel	0,394	***	0,397	***	0,426	***	0,456	***	0,390	***
NN(2)	0,381	***	0,374	***	0,441	***	0,512	***	0,373	***
NN(5)	0,371	***	0,373	***	0,427	***	0,445	***	0,375	***

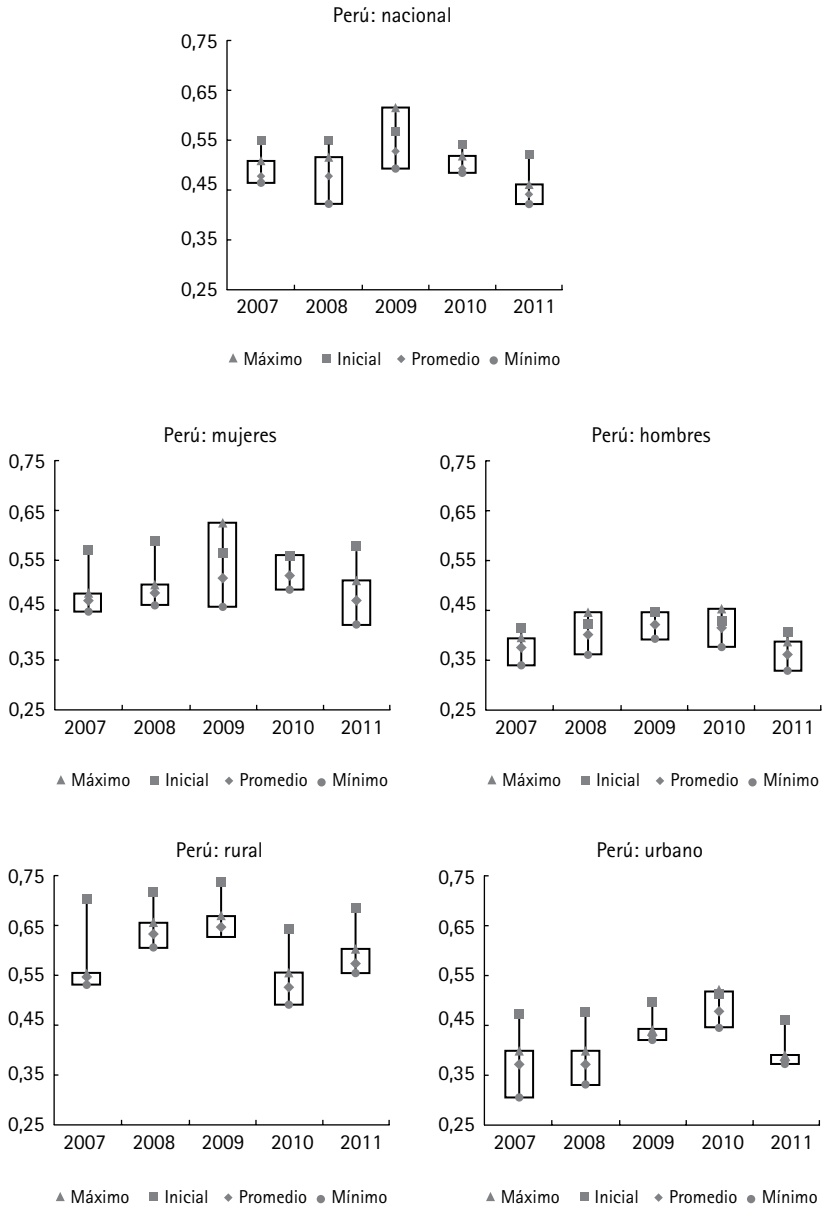
Notas: <sup>a</sup> Diferencia de los promedios de los logaritmos del ingreso laboral por hora de los asalariados y los autoempleados. <sup>b</sup> Nivel de significancia: \*\*\*: 99%, \*\*: 95%. <sup>c</sup> La diferencia inicial, medida en logaritmos, representa la diferencia antes del emparejamiento, pero después de haber aplicado los filtros de "limpieza" y los filtros para homogenizar las poblaciones. <sup>d</sup> La especificación 1 incluye edad, edad al cuadrado y años de escolaridad. <sup>e</sup> La especificación 11 incluye, además de las variables de la especificación 1, si es jefe o jefa del hogar. <sup>f</sup> La especificación 19 incluye, además de las variables de la especificación 1, lengua materna, cuadrado de años de escolaridad y si es jefe o jefa del hogar. Para el estrato urbano no se consiguió alguna especificación que balancee todas las muestras transversales y todas las muestras panel. En todas las especificaciones, NN(2) y NN(5) corresponden a la estimación con los dos y cinco vecinos más cercanos, respectivamente.

*Fuente:* Microdatos de las ENAHO muestra transversal. Elaboración propia.

Los resultados según sexo y, más adelante, según área de residencia muestran una importante heterogeneidad en los resultados. Respecto al sexo, un primer aspecto que resalta es que el premio entre las mujeres es más grande que entre los hombres: los promedios se ubican en el intervalo de 47,1% a 52,1% para las mujeres, frente al 36% a 42% para los hombres. En segundo lugar, mientras que entre las mujeres la diferencia inicial está fuera y por encima de los límites de los premios en tres de las cinco muestras, en el caso de los hombres esto solo ocurre en dos de las cinco muestras y sin estar muy lejos del valor máximo. En tercer lugar, las estimaciones del premio para las mujeres se muestran relativamente más sensibles al método de estimación que para el caso de los hombres. La distancia entre el mínimo y el máximo valor alcanza diecisiete puntos porcentuales entre las mujeres y siete entre los hombres.

Según el área de residencia, lo primero que resalta es la diferencia en la magnitud del premio entre áreas, que es mucho mayor en las rurales que en las urbanas. Segundo, es muy claro que la diferencia inicial (antes de emparejamiento) sobrestima el premio del puesto asalariado en comparación con el autogenerado, pues en las cinco muestras rurales y en cuatro de las urbanas la diferencia inicial está fuera y por encima del rango que contiene los cálculos mínimos y máximos de los premios. Tercero, la sensibilidad del premio al método de estimación es relativamente menos importante en las áreas rurales que en las urbanas.

**Gráfico 2.** Perú: 2007-2011. Premio en los ingresos laborales por hora del trabajador asalariado con relación al trabajador autoempleado



Método: Propensity Score Matching.

Fuente: Microdatos de la ENAHO-muestra transversal. Elaboración propia.

La mirada del conjunto de estos resultados sugiere que, asumiendo que la asignación a los puestos de trabajo asalariado y los autoempleados se basa en las variables observables consideradas, existe un premio a favor de los asalariados atribuible al puesto de trabajo, frente a los que autogeneran sus puestos o, dicho de otra forma, sugiere que hay evidencia de la existencia de segmentación en los mercados laborales.

## B. Resultados al emplear la doble diferencia

De la misma manera como se hizo para el cálculo del premio entre puestos asalariados y autoempleados al utilizar el PSM, para la implementación de la DD se acotó la muestra para obtener una población equivalente a la utilizada en las estimaciones anteriores. También, se eliminaron los casos con inconsistencias en la información de sexo, edad y años de escolaridad, a partir de la comparación de los dos años consecutivos del panel<sup>20</sup>. Se debe tener en cuenta que el número de casos se reduce además debido a que solo se consideran los trabajadores autoempleados en el momento  $t_0$  y no todos los autoempleados y asalariados.

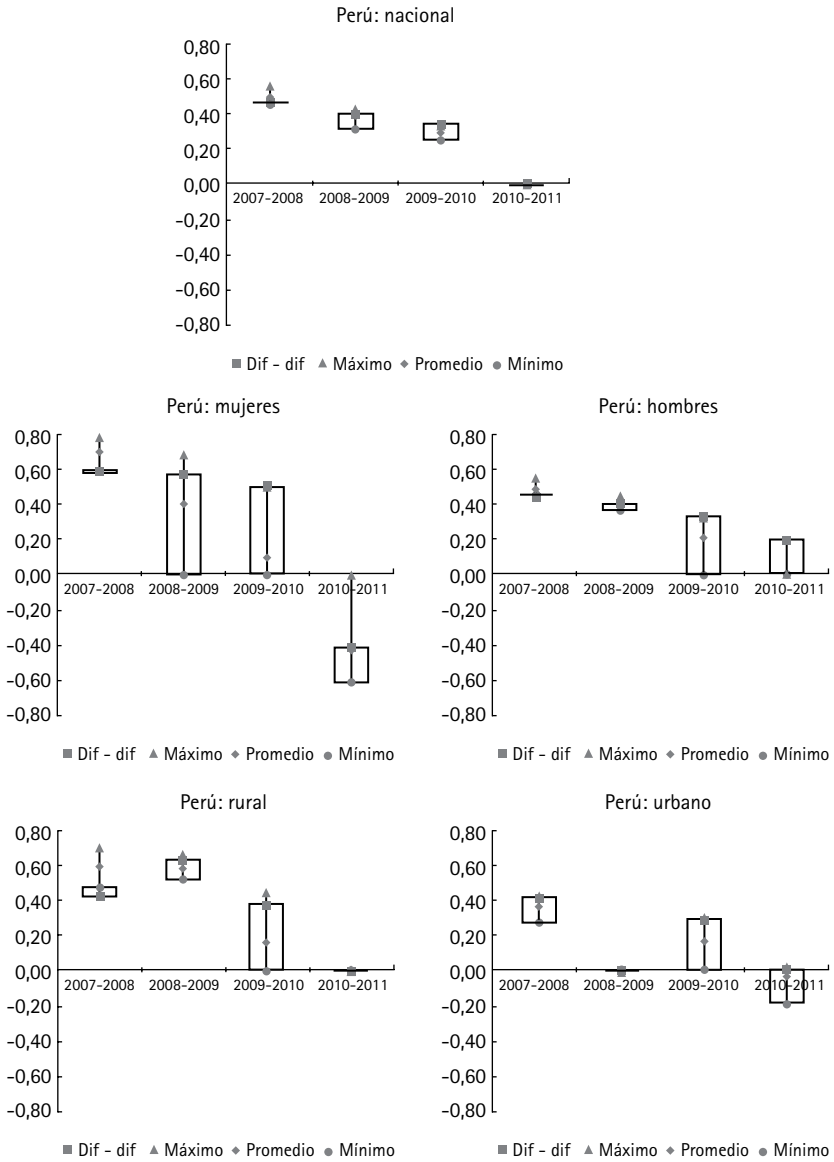
Como se mencionó antes, es requisito para una adecuada interpretación de los resultados que se cumpla el supuesto de tendencias paralelas. Para ello, se implementó un test de diferencia de medias de la variable dependiente entre los individuos tratados y los no tratados en el año base de cada una de las cuatro transiciones. Los resultados, reportados en el cuadro 5, indican que el supuesto se cumple para tres de las cuatro transiciones en cada uno de los estratos considerados (global o nacional, mujeres, hombres, rural y urbano) y que en tres de los cinco casos en que hay evidencia de violación, esto sucede en la última transición (2010-2011). Se volverá a estos resultados más adelante, en el análisis de los resultados por estrato.

En el gráfico 3 y en el cuadro 6 se reportan los resultados de los premios. Recuérdese que el estimador de DD toma en cuenta la brecha de ingresos entre las personas que habiendo sido autoempleadas en el año  $t_0$ , pasan a

---

20 Las inconsistencias surgen de la comparación de los datos en los años  $t_0$  y  $t_1$ . Se hizo exclusión en los siguientes casos: a) diferente sexo, b) diferencia en edad superior en dos años (para más o para menos) y c) años de escolaridad en  $t_1$  inferiores a los de  $t_0$ . En general, no fueron muchos los casos que presentaban inconsistencias.

**Gráfico 3.** Perú: 2007-2011. Premio en el ingreso laboral por hora del trabajador autoempleado en  $t$  que pasa a ser asalariado en  $t + 1$



Método: doble diferencia con emparejamiento y sin él (D-D sin emparejamiento es representado por "dif-dif" en la leyenda de los gráficos, mientras que D-D con emparejamiento son reportados con los valores "máximo", "promedio" y "mínimo").

Fuente: Microdatos de la ENAHO - muestra panel 2007 al 2011. Elaboración del autor.

ocupar un puesto asalariado en el año  $t_1$ , y a esta se le descuenta la diferencia en los ingresos de quienes se mantienen como autoempleados entre  $t_0$  y  $t_1$ . Esta segunda diferencia busca capturar el comportamiento tendencial de los ingresos entre los períodos en consideración. Es muy importante resaltar que, como se puede observar en el cuadro 6 y a diferencia de lo hallado con la metodología anterior, aquí los premios estimados presentan mayor variabilidad respecto a la significación estadística. Con el PSM todos los premios resultaron significativos al 99%. Con la DD este no es el caso. Por ejemplo, en el panel 2007-2008, todos los premios reportados en el cuadro 6 son estadísticamente significativos al 95%; en el panel 2008-2009 un par de resultados no son significativos (en particular para las mujeres y bajo los algoritmos del vecino más cercano). En el siguiente panel, 2009-2010, es más frecuente hallar resultados no significativos, mientras que en el panel 2010-2011 es más bien raro encontrar resultados estadísticamente significativos. Se volverá sobre esto más adelante.

Cuadro 5. Resultado del test de diferencia de medias: prueba de tendencias paralelas

Año $t_0$	Global		Mujeres		Hombres		Rural		Urbano	
	Dif <sup>a</sup>	S <sup>b</sup>	Dif <sup>a</sup>	S <sup>b</sup>	Dif <sup>a</sup>	S <sup>b</sup>	Dif <sup>a</sup>	S <sup>b</sup>	Dif <sup>a</sup>	S <sup>b</sup>
2007	-0,080		-0,150		-0,214	***	0,262	**	-0,118	
2008	0,039		-0,040		-0,115		0,020		0,130	
2009	0,013		0,088		-0,135		0,143		0,066	
2010	0,212	**	0,420	**	0,063		0,268		0,256	**

Notas: <sup>a</sup> Diferencia de los promedios de los logaritmos de los ingresos laborales por hora en el período inicial pretratamiento entre los autoempleados que transitan a asalariados y los autoempleados que se mantienen como tales. <sup>b</sup> Nivel de significancia: \*\*\*: 99%, \*\*: 95%.

Fuente: elaboración del autor.

En el gráfico 3 solo se utilizan los resultados que son estadísticamente significativos al menos al 95%. Cuando esto no sucede, para el cálculo de los valores máximo, mínimo y promedio, los resultados no significativos son considerados con premio igual a cero, pues la estadística nos indica que no hay confianza suficiente para considerar distinta de cero la diferencia en los ingresos. En el gráfico se incluye el premio calculado con DD sin emparejamiento, que en el cuadro 6 aparece como premio sin emparejamiento.

**Cuadro 6.** Premio en los ingresos laborales por hora del trabajador asalariado con relación al autoempleado. Métodos: diferencia de las diferencias sin y con emparejamiento. Perú: 2007-2011

	2007-2008		2008-2009		2009-2010		2010-2011	
	Dif <sup>a</sup>	S <sup>b</sup>	Dif <sup>a</sup>	S <sup>b</sup>	Dif <sup>a</sup>	S <sup>b</sup>	Dif <sup>a</sup>	S <sup>b</sup>
<b>Global (Espec. 1)<sup>d</sup></b>								
Premio sin emparejamiento <sup>c</sup>	0,468	***	0,400	***	0,348	***	-0,010	
Premio con emparejamiento								
One to one	0,472	***	0,319	***	0,182		-0,030	
RADIUS	0,470	***	0,416	***	0,295	***	-0,057	
Kernel	0,472	***	0,427	***	0,329	***	-0,048	
NN(2)	0,561	***	0,366	***	0,184		-0,065	
NN(5)	0,522	***	0,430	***	0,257	***	-0,050	
<b>Mujeres (Espec. 11)<sup>e</sup></b>								
Premio sin emparejamiento <sup>c</sup>	0,594	***	0,572	**	0,504	***	-0,418	**
Premio con emparejamiento								
One to one	0,780	**	0,326		0,422		-0,523	**
RADIUS	0,585	**	0,672	***	0,394		-0,416	
Kernel	0,591	**	0,651	**	0,495	**	-0,421	**
NN(2)	0,777	**	0,505		0,461		-0,547	**
NN(5)	0,761	**	0,684	**	0,336		-0,613	**
<b>Hombres (Espec. 11)<sup>e</sup></b>								
Premio sin emparejamiento <sup>c</sup>	0,451	***	0,400	***	0,336	***	0,192	**
Premio con emparejamiento								
One to one	0,497	***	0,393	***	0,146		0,198	
RADIUS	0,462	***	0,369	***	0,231	**	0,164	
Kernel	0,459	***	0,400	***	0,321	***	0,138	
NN(2)	0,547	***	0,448	***	0,200	**	0,215	
NN(5)	0,462	***	0,394	***	0,295	***	0,193	
<b>Rural (Espec. 19)<sup>f</sup></b>								
Premio sin emparejamiento <sup>c</sup>	0,424	***	0,634	***	0,387	***	0,168	
Premio con emparejamiento								
One to one	0,692	***	0,518	***	0,342		0,321	
RADIUS	0,559	***	0,664	***	0,312		0,182	
Kernel	0,475	***	0,618	***	0,444	**	0,218	
NN(2)	0,699	***	0,566	***	0,251		0,178	

**Cuadro 6.** Premio en los ingresos laborales por hora del trabajador asalariado con relación al autoempleado. Métodos: diferencia de las diferencias con emparejamiento y sin él. Perú: 2007-2011 (continuación)

	2007-2008		2008-2009		2009-2010		2010-2011	
	Dif <sup>a</sup>	S <sup>b</sup>	Dif <sup>a</sup>	S <sup>b</sup>	Dif <sup>a</sup>	S <sup>b</sup>	Dif <sup>a</sup>	S <sup>b</sup>
NN(5)	0,546	***	0,541	***	0,359	**	0,108	
<b>Urbano (Espec. 19)<sup>f</sup></b>								
Premio sin emparejamiento <sup>c</sup>	0,410	***	0,196		0,291	**	-0,116	
Premio con emparejamiento								
One to one	0,274	**	0,045		0,211		-0,270	
Radius	0,399	***	0,136		0,234	**	-0,189	
Kernel	0,421	***	0,195		0,271	**	-0,188	**
NN(2)	0,371	***	0,024		0,253		-0,204	
NN(5)	0,355	***	0,111		0,287	**	-0,197	

Notas: <sup>a</sup> Diferencia de los promedios de los logaritmos del ingreso laboral por hora de los asalariados y los autoempleados. <sup>b</sup> Nivel de significancia: \*\*\*: 99%, \*\*: 95%. <sup>c</sup> La diferencia inicial, medida en logaritmos, representa la diferencia antes del emparejamiento, pero después de haber aplicado los filtros de "limpieza" y los filtros para homogenizar las poblaciones. <sup>d</sup> La especificación 1 incluye edad, edad al cuadrado y años de escolaridad. <sup>e</sup> La especificación 11 incluye, además de las variables de la especificación 1, si es jefe o jefa del hogar. <sup>f</sup> La especificación 19 incluye, además de las variables de la especificación 1, lengua materna, cuadrado de años de escolaridad y si es jefe o jefa del hogar. Para el estrato urbano no se consiguió alguna especificación que balancee todas las muestras transversales y todas las muestras panel. En todas las especificaciones, NN(2) y NN(5) corresponden a la estimación con los dos y cinco vecinos más cercanos, respectivamente.

Fuente: Microdatos de las ENAHO panel. Elaboración propia.

A escala global o nacional, el premio sin emparejamiento muestra que en tres de los cuatro paneles hay diferencias significativas que sugieren que pasar de un puesto autoempleado a uno asalariado representa un incremento en los ingresos reales. Es interesante notar que, visto en conjunto y a lo largo del tiempo, el premio que se ubica en 47% en el primer panel disminuye en los dos siguientes, mientras que el último no es significativamente distinto de cero. Aunque es un espacio de tiempo corto, esto sugiere una tendencia a disminuir y finalmente desaparecer del premio. Los resultados del cálculo del premio con emparejamiento son semejantes: significativos y decrecientes en los tres primeros paneles y no significativo en el último<sup>21</sup>. Tomando en consideración el promedio simple de los premios calculados con diferentes algoritmos con emparejamiento, se observa que el premio oscila entre 29% y 50%.

21 Nótese que precisamente para este estrato para esta transición no se estaría cumpliendo el supuesto de tendencias paralelas, lo que no permite concluir nada sobre la existencia o inexistencia de premio.



Si se consideran por separado hombres y mujeres, en general el premio sin y con emparejamiento muestra que este suele ser más grande entre las mujeres (al menos 50% sin emparejamiento) que entre los hombres (no más de 45% sin emparejamiento). El emparejamiento no necesariamente reduce el premio. Por ejemplo, en el caso de las mujeres, sin emparejamiento está cerca pero por debajo del 60%, mientras que el promedio de los premios con emparejamiento está entre 67% y 69% en los dos primeros paneles. Como en el caso del conjunto de la población, en cada uno de los sexos también se observa, aunque de manera menos pronunciada, la disminución de los premios (con emparejamiento y sin él) en los tres primeros paneles, pero en el cuarto panel (2010-2011) mientras que entre los hombres solo subsiste un premio significativo sin emparejamiento, entre las mujeres se torna negativo con ambos métodos<sup>22</sup>.

La distinción según áreas de residencia muestra comportamientos menos sistemáticos de los premios, en particular entre los residentes de áreas urbanas. Aquí solo hay evidencia estadísticamente significativa de la existencia de premios en dos de los cuatro paneles, para los resultados con emparejamiento y sin él. A lo sumo se podría decir que existe alguna evidencia de la existencia de premios a favor de los que transitan de ocupaciones autogeneradas a puestos asalariados en órdenes de magnitud relativa que oscilan entre 30% y 40%. En el caso de los residentes de áreas rurales, se encuentran resultados que indican la existencia significativa de premios en tres de los cuatro paneles (como en los resultados globales). El premio sin emparejamiento oscila entre 39% y 63%, sin una tendencia clara a lo largo del tiempo. El premio con emparejamiento indica niveles iniciales de 58%-59% (paneles 2007-2008 y 2008-2009), para luego caer a 40% (panel 2009-2010) y dejar de ser significativo en el último panel<sup>23</sup>.

### C. Balance del conjunto de los resultados

A escala nacional o global, la mayor parte de la evidencia sugiere la existencia de segmentación del mercado de trabajo según los puestos sean autogenerados, o generados por una empresa y, en consecuencia, que sean puestos asalariados.

---

22 Nuevamente, nótese que se estaría violando el supuesto de tendencias paralelas en la última transición entre las mujeres (2010-2011) y en la primera de los hombres (2007-2008).

23 Para el grupo de residentes en áreas rurales y el de residentes en áreas urbanas son la primera y última transición, respectivamente, en donde no se cumple el supuesto de tendencias paralelas.

Las tres metodologías empleadas muestran que el premio sería de no menos de 30% ni más de 50%. Los premios calculados con PSM lo ubican en niveles relativamente más altos que lo que la DD indica, lo cual sugiere que efectivamente es más apropiado usar DD, en particular con emparejamiento, pues los atributos no observables tienen un rol en la determinación de una parte de la diferencia en los ingresos. Es importante resaltar, además, que el uso de la muestra panel con DD, tanto sin como con emparejamiento, produce resultados distintos en dos aspectos adicionales con relación al PSM. Primero, muestra mucho mayor variabilidad (lo que se manifiesta en intervalos más grandes en los que se acumulan los distintos resultados con distintos algoritmos y para las distintas transiciones) y, segundo, incluso arroja resultados para un año que indican que no hay premio distinto de cero que sea significativo.

Con relación a la distinción por sexo, en el caso de las mujeres se encuentra que el premio sería de no menos de 42% ni más de 70%. Algunos de los resultados, en particular utilizando DD con emparejamiento, indican que el premio puede llegar a ser mayor que la diferencia inicial antes de hacer ningún emparejamiento. En este caso hay un resultado aún más sorprendente que muestra que en el último panel (2010-2011), y solo en ese caso, el premio es más bien a favor de las autoempleadas y en una magnitud muy importante (42%). Respecto a los hombres, los resultados indican que el premio sería no menor a 26% ni mayor a 49%. En este caso, aunque los rangos en los que se ubican los resultados pueden llegar a límites menores en comparación con los que presentan las diferencias iniciales, sus límites máximos son muy parecidos. Además, la comparación de los premios con PSM y con DD indica que las tres metodologías arrojan premios dentro del mismo orden de magnitud. La única excepción se observa en el último panel con DD emparejadas.

Finalmente, con respecto a la distinción según el área de residencia, la población residente en áreas rurales es la que presenta las mayores diferencias iniciales en los ingresos según segmento laboral. Estas diferencias iniciales se ubican en el rango de 64% a 73%. Al usar alguna de las técnicas de emparejamiento, los premios cuando son estadísticamente significativos oscilan entre 48% y 66%. Una vez más, son los resultados de los cálculos de los premios con la muestra panel los que muestran mayor variabilidad. Es más frecuente que los resultados con esta muestra panel no sean significativos, tanto con DD como con DD con emparejamiento. Entre los residentes de áreas urbanas la diferencia inicial se encuentra entre 46% y 50%, esto es, diferencias

menores que las que se observa en las áreas rurales. Con PSM los premios caen un poco y se ubican en el rango comprendido entre 37% y 48%, mientras que con DD y DD-PSM se halla que no son significativamente distintos de cero y a lo sumo ascienden a 37%, aproximadamente.

#### IV. Conclusiones

La evidencia internacional muestra que en economías desarrolladas la mayor parte de la fuerza de trabajo se emplea en puestos de trabajo asalariado. Los datos de las economías en desarrollo con mayor progreso relativo en Latinoamérica también sugieren que a medida que las economías crecen y muestran mayores niveles de PIB per cápita, el autoempleo y el empleo familiar no remunerado decrecen en importancia. El Perú aún no parece haber llegado a un estado en el que la absorción de la mano de obra en forma asalariada sea lo suficientemente dinámica como para empezar a revertir una situación en la que más de la mitad de la fuerza de trabajo no es empleada en formas de producción típicamente capitalistas. Ni el sostenido e importante crecimiento de la economía peruana de la última década parece haber empezado a modificar de manera significativa este rasgo estructural.

Por otra parte, el autoempleo está muy claramente asociado al subempleo por ingresos y, en comparación con los asalariados, en promedio sus ingresos son relativamente menores. Una primera interpretación de estos hechos podría ser que los autoempleados en realidad no compiten en el mercado laboral con los asalariados, pues tienen menores dotaciones de activos productivos relevantes para el mercado laboral. La gran heterogeneidad en las características, como los años de escolaridad de la población en edad de trabajar o incluso la calidad de dicha escolaridad, parece dar sustento a la existencia de particiones del mercado de laboral (e incluso de distintos mercados), en donde en cada una de ellas se determina un equilibrio para las diferentes cantidades de capital humano.

Alternativamente a la interpretación anterior, podría ser que además de la existencia de distintos mercados laborales exista segmentación de mercados, es decir, que por alguna razón los ingresos laborales para trabajadores idénticos son distintos según el segmento del mercado en donde trabajan. La literatura teórica ha brindado diversas razones para explicar la existencia de segmenta-

ción, desde los aspectos institucionales como la existencia de sindicatos y de salarios mínimos, hasta los distintos mecanismos agrupados bajo el concepto de salarios de eficiencia, los costos de la rotación laboral y los mecanismos de disciplina del trabajo.

En la literatura peruana se ha abordado esta discusión, pero la mayor parte de los pocos trabajos realizados al respecto ponen énfasis a la posible segmentación entre informalidad y formalidad. Los resultados de estos trabajos muestran sistemáticamente que entre los asalariados los formales tienen un premio importante y significativo respecto a sus pares en la informalidad. En cambio, entre los autoempleados los resultados son mixtos. Algunos trabajos muestran que los formales ganan más que los informales, pero otros revelan que las diferencias se estrechan o desaparecen, o incluso que se revierten.

En este trabajo, a diferencia de los anteriores, se sometió a prueba la existencia de segmentación del mercado laboral comparando los ingresos de los asalariados y de los autoempleados. Esta perspectiva es distinta a la de la segmentación entre formales e informales, ya que tanto entre los asalariados como entre los autoempleados hay formalidad e informalidad. Nos concentramos en una partición específica del mercado del trabajo en el Perú, esto es, las personas con no más que la secundaria completa. Se les impuso, además, una serie de filtros a las muestras, buscando: a) incrementar el grado de confiabilidad de la información (v. g., solo considerando la información de las personas que directamente respondieron a las encuestas) b) aumentar la precisión de la información (v. g., eliminando entre los autoempleados aquellos que producen con el concurso de TFNR, cuyos costos no pueden ser deducidos del valor de producción) y c) eliminando puestos de trabajo en donde la fijación de los salarios puede no seguir los mecanismos de mercado (v. g., empleo en la Administración estatal). Se utilizaron metodologías de la evaluación de impacto que emplean técnicas semiparamétricas o no paramétricas que permiten controlar las características observables y no observables de los trabajadores.

Los resultados brindan evidencia mixta, pero sugieren preponderantemente que existe segmentación en los mercados laborales peruanos. Trabajadores con características productivas semejantes observables y no observables perciben distintos ingresos laborales según el tipo de puesto de trabajo. Se calcularon premios para diferentes muestras de la población: total, por sexo y por área de residencia. Empleando el PSM se encuentra que el premio por trabajar en

un puesto asalariado puede ser del orden de 36%, hasta el 65%, según el subconjunto particular de la población que se analice. A escala nacional, el premio oscila entre 44% y 53%. Diferenciando por sexo, entre las mujeres es mayor el premio que entre los hombres (47%-52% y 36%-42%, respectivamente). Distinguiendo por área de residencia, en las rurales se encontraron los mayores premios (52%-65% y 37%-48%, correspondientemente).

Utilizando DD con emparejamiento y sin él, los premios estimados muestran resultados semejantes. A escala nacional, el DD sin emparejamiento arroja un premio en el rango de 35% a 47%, mientras que con emparejamiento está entre 29% y 50%. Entre las mujeres, tres de los cuatro paneles muestra un premio a favor de las asalariadas, que oscila entre 50% y 59% (DD sin emparejamiento) y 50% y 70% (DD con emparejamiento). Entre los hombres los premios con DD sin emparejamiento oscilan entre 19% y 40%, mientras que con emparejamiento, entre 0% y 49%. En las áreas rurales el premio oscila entre 38%-63% y 40%-59% con DD con emparejamiento y sin él, respectivamente. En las áreas urbanas los rangos van de 0% a 42% y de 0% a 36%, con emparejamiento y sin él, respectivamente.

Desde el punto de vista del autor del presente artículo, este conjunto de resultados constituye evidencia de que existe segmentación laboral entre los puestos de trabajo asalariado y los autoempleados. Si se tiene en cuenta que los puestos de trabajo asalariado, cuando son formales, incluyen una serie de beneficios, como el acceso a servicios de salud y las pensiones para la jubilación, entre otros, se podría decir que incluso se estaría subestimando el premio.

Los resultados de los premios estimados con las muestras panel, sin embargo, llevan a considerar la posibilidad de que en los últimos años esté ocurriendo un cambio estructural importante en el funcionamiento de los mercados laborales, como lo sugiere el incumplimiento del supuesto de tendencias paralelas, especialmente entre 2010 y 2011. Así lo sugiere el comportamiento decreciente del premio a lo largo de las cuatro transiciones consideradas, que se observa con claridad a escala nacional. El crecimiento económico sostenido por más de una década podría, finalmente, haber empezado a tener efecto sobre los mercados laborales del Perú.

## Referencias

1. AKERLOFF, G. y YELLEN, J. (1986). *Efficiency wage models of the labor market*. Cambridge: Cambridge University Press.
2. BARCO, D. y VARGAS, P. (2010). Brechas salariales entre formales e informales (Documento de Trabajo 003). Banco Central de la Reserva del Perú.
3. BERNAL, R. y PEÑA, X. (2011). *Guía práctica para la evaluación de impacto*. Bogotá: Universidad de los Andes.
4. BLUNDELL, R. y COSTA DIAS, M. (2000). "Evaluation methods for non-experimental data", *Fiscal Studies*, 21(4):427-468.
5. DE SOTO, H. (1986). *El otro sendero: la revolución informal*. Lima: El Barranco.
6. DÍAZ, J. J. (2009). *Employment and earnings trends in Peru, 1997-2007. Evidence from household surveys*. Lima. Mimeo.
7. DUVAL-HERNÁNDEZ, R. (2006). *Informality, segmentation and earnings in urban Mexico*. San Diego, CA. Mimeo.
8. FIELDS, G. (1975). "Rural-urban migration, urban unemployment and underemployment, and search activity in LCD's", *Journal of Development Economics*, 2(2):165-188.
9. FIGUEROA, A. (2011). "A generalized labor market theory: Inequality as labor discipline device", *Investigación Económica*, 70(276):163-185, abril-junio.
10. GASPARINI, L. y TORNAROLLI, L. (2009). "Labor informality in Latin America and the Caribbean: Patterns and trends from household survey microdata", *Desarrollo y Sociedad*, 63:13-80.
11. GERTLER, P., MARTÍNEZ, S., PREMAND, P., RAWLINGS, L. y VERMEERSCH, C. (2011). *La evaluación de impacto en la práctica*. Washington, D. C.: Banco Mundial.

12. HARRIS, J. y TODARO, M. (1970). "Unemployment and development: A two-sector analysis", *The American Economic Review*, 60(1):126-142.
13. INSTITUTO NACIONAL DE ESTADÍSTICA E INFORMÁTICA, INEI. (2012). *Población económicamente activa ocupada informal*. Lima. Autor.
14. KHANDKER, S., KOOLWAL, G. y SAMAD, H. (2010). *Handbook on impact evaluation. Quantitative methods and practices*. Washington, D. C.: The World Bank.
15. LEUVEN, E. y SIANESI, B. (2012). PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalancing testing. Disponible en <http://ideas.repec.org/c/boc/bocode/s432001.html>. 12 de febrero de 2013.
16. LEWIS, A. (1954). "Economic development with unlimited supplies of labour", *The Manchester School*, 22:139-191.
17. LEWIS, A. (1979). "The dual economy revisited", *The Manchester School*, 47:211-229.
18. LOAYZA, N. (1996). "The economics of the informal sector: A simple model and some empirical evidence from Latin America", *Carnegie-Rochester Conference Series on Public Policy*, 45:129-162.
19. LOAYZA, N. (2007). The causes and consequences of informality in Peru (Documento de Trabajo 018). Banco Central de Reserva del Perú.
20. MALONEY, W. (1999). "Does informality imply segmentation in urban labor markets? Evidence from sectoral transitions in Mexico", *The World Bank Economic Review*, 13(2):275-302.
21. MALONEY, W. (2004). "Informality revisited", *World Development*, 32(7):1159-1178.
22. MORALES, R., RODRÍGUEZ, J., HIGA, M. y MONTES, R. (2010). "Transiciones laborales, reformas estructurales y vulnerabilidad laboral en el Perú: 1998-2008", en J. Rodríguez y A. Berry (eds.), *Desafíos laborales*

en América Latina después de dos décadas de reformas estructurales. *Bolivia, Paraguay y Perú (1997-2008)* (p. 641). Lima: IEP-Fondo Editorial de la PUCP.

23. ÑOPO, H. (2008). "Matching as a tool to decompose wage gaps", *The Review of Economics and Statistics*, 90(2):290-299.
24. PERRY, G., MALONEY, W., ARIAS, O., FAJNZYLBER, P., MASON, A. y SAAVEDRA, J. (2007). *Infomality. Exit and exclusion*. Washington, D. C.: The World Bank.
25. PRATAP, S. y QUINTIN, E. (2006). "Are labor markets segmented in developing countries? A semiparametric approach", *European Economic Review*, 50:1817-1841.
26. RANIS, G. (2006). "Is dualism worth revisiting", en A. de Janvry y R. Kanbur (eds.), *Poverty, inequality and development. Essays in honor of Erik Thorbecke* (pp. 371-385). Nueva York: Springer.
27. RODRÍGUEZ, J. S. (2011). Brechas de ingresos laborales entre asalariados y autoempleados en el Perú (Documento de Trabajo 318). Departamento de Economía, PUC del Perú.
28. RODRÍGUEZ, J. e HIGA, M. (2010). "Informalidad, empleo y productividad en el Perú", en J. Rodríguez y A. Berry (eds.), *Desafíos laborales en América Latina dos décadas después de las reformas estructurales. Bolivia, Paraguay y Perú (1997-2008)* (pp. 124-189). Lima: IEP-Fondo Editorial de la PUCP.
29. SAAVEDRA, J. y CHONG, A. (1999). "Structural reform, institutions and earnings: Evidence from the formal and informal sectors in urban Peru", *The Journal of Development Studies*, 35(4):95-116.
30. STIGLITZ, J. (1974). "Alternative theories of wage determination and unemployment in LDCs: The labour turnover model", *Quarterly Journal of Economics*, 88(2):194-227.



31. TODARO, M. (1969). "A model of labor migration and urban unemployment in less development countries", *The American Economic Review*, 59(1):138-148.
32. ULYSSEA, G. (2006). "Informalidade no mercado de trabalho brasileiro: Uma resenha da literatura", *Brazilian Review of Political Economy*, 26(3):596-618.
33. ULYSSEA, G. (2010). "The formal-informal labor market segmentation hypothesis revisited", *Brazilian Review of Econometrics*, 30(2):311-334.
34. ULYSSEA, G. y PAES DE BARROS, R. (2010). "On the empirical content of the formal-informal labor market segmentation hypothesis", *Brazilian Review of Econometrics*, 30(2):289-310.
35. YAMADA, G. (1996). "Urban informal employment and selfemployment in developing countries: Theory and evidence", *Economic Development and Cultural Change*, 44(2):289-314.
36. YAMADA, G. y RAMOS, M. (1996). "El desempeño del autoempleo y la microempresa en Lima Metropolitana", en G. Yamada, G. Felices, M. Ramos y J. Ruiz (eds.), *Caminos entrelazados. La realidad del empleo urbano en el Perú* (p. 455). Lima: Universidad del Pacífico.