

Efectos de los Ingresos No Reportados en el Nivel y Tendencia de la Pobreza Laboral en México

Raymundo M. Campos Vázquez

Colmex

Marzo 27, 2013
El Colegio de México

Pobreza Laboral

- ▶ La mayoría de los estudios en economía laboral así como las mediciones de pobreza en México ignoran a los individuos que reportan trabajar por remuneración pero que deciden no reportar ingresos.
- ▶ En la medición del Índice de la Tendencia Laboral de la Pobreza por el CONEVAL (2010) se incluyen pero como si no tuvieran un ingreso
- ▶ El supuesto detrás de esas decisiones es que las personas que deciden no reportar ingresos son una muestra aleatoria de los trabajadores o bien que la composición de ese tipo de personas se mantiene constante en el tiempo.
- ▶ En este artículo argumento que esos supuestos dejaron de ser válidos para el caso de México.

Importancia

- ▶ Tener un dato preciso de pobreza es sumamente importante: cambios en política pública, asignaciones presupuestales por entidad federativa.
- ▶ La proporción de individuos que decide no declarar ingresos aun cuando trabaja por remuneración ha aumentado.
- ▶ Por ejemplo, en 2005 la proporción era de 10 por ciento mientras que en 2012 era de 20 por ciento.
- ▶ Más aún, la proporción de ingresos inválidos ha aumentado principalmente entre los individuos con educación superior. Esta proporción era de 20 por ciento en 2005 pero para 2012 era de 38 por ciento.

Métodos

- ▶ En este artículo utilizo métodos de imputación para asignar un ingreso a los trabajadores que no lo reportan.
- ▶ Existe un amplio consenso sobre la conveniencia de corregir datos faltantes con métodos de imputación: Little y Rubin (2003) y Rubin (2004).
- ▶ Imputar con la media es incorrecto: picos en la distribución y subestimación de la varianza.
- ▶ 4 métodos de imputación: Pareamiento de puntajes de propensión, Hot-Deck, Medias por grupo más un error, pareamiento por puntajes predictivos

Resultados

- ▶ Encuentro que la pobreza laboral está sobreestimada: CONEVAL-pobreza $\uparrow 25\%$ para 2005-2012.
- ▶ Imputación: $\uparrow 12\%$.
- ▶ Es decir, para 2012:3 existen 8.1 millones de personas que son declarados como pobres laborales pero, de acuerdo al método de imputación, no lo son.
- ▶ Restricción de este estudio: Pobreza laboral, y no la medición oficial de pobreza con ENIGH.

Datos

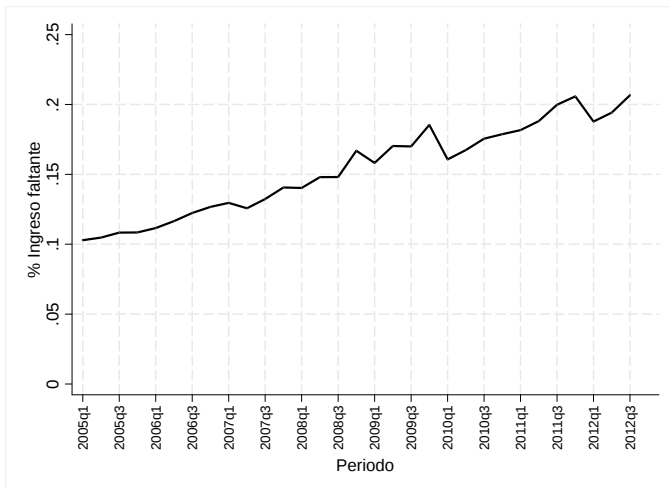
- ▶ ENOE: 2005-2012
- ▶ Imputación nos restringimos a trabajadores con horas positivas pero que no reportan ingresos.
- ▶ Trabajadores sin pago, desempleados, o trabajadores con horas cero NO se les imputa ingreso. En la muestra de imputación no los utilizo, pero los incluyo en el cálculo de pobreza.
- ▶ No hay restricción por edad.
- ▶ Todos los ingresos en pesos constantes del segundo trimestre de 2012.

Problema (I)

- ▶ Las personas que deciden no reportar ingresos son una muestra aleatoria de los trabajadores.
- ▶ O bien que la composición de ese tipo de personas se mantiene constante en el tiempo.
- ▶ Esto no se cumple ya para el caso de México
- ▶ Si las personas que no declaran ingresos son una muestra aleatoria de la población en general el único problema es de eficiencia (mayores errores estándar) sin problema de sesgo.
- ▶ Si las personas con ingresos inválidos no son una muestra aleatoria de la población entonces los estadísticos de interés no son representativos de la población en su conjunto

Problema (II)

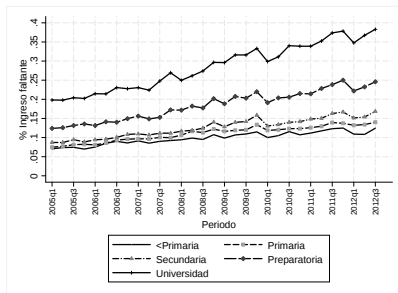
Porcentaje de Trabajadores con ingreso faltante



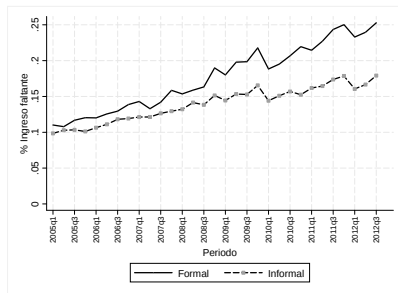
Problema (III)

Porcentaje de Trabajadores con ingreso faltante: Educación y Formalidad

A. Educación



B. Formalidad



Problema (IV)

Estadísticas

	2005				2012				Inválidos	
	Válidos	Inválidos	Diferencia		Válidos	Inválidos	Diferencia		2005-2012	
Edad	36.86	40.25	-3.4	*	37.8	40.8	-2.99	*	-0.54	*
			[0.125]				[0.090]		[0.144]	
Formal	0.41	0.41	0		0.39	0.46	-0.07	*	-0.06	*
			[0.004]				[0.003]		[0.005]	
Rural	0.25	0.19	0.06	*	0.28	0.18	0.1	*	0.01	
			[0.003]				[0.002]		[0.004]	
Preparatoria completa	0.19	0.22	-0.03	*	0.21	0.25	-0.04	*	-0.03	*
			[0.003]				[0.003]		[0.004]	
Universidad completa	0.14	0.27	-0.13	*	0.15	0.31	-0.16	*	-0.04	*
			[0.003]				[0.003]		[0.004]	
N	132849	13977			120491	29344				

Métodos

Explicación General

- ▶ Abundante literatura para tratar datos faltantes: Allison, 2000; Cameron y Trivedi, 2005; Little y Rubin, 2003; Rubin, 2004.
- ▶ Uno de los métodos comúnmente utilizados para resolver el problema de ingresos faltantes es imputar con la media del valor observado.
 1. Erróneo por: distribución de ingreso imputada tendría picos en ciertas partes de la distribución
 2. Erróneo por: subestima la varianza del ingreso sustancialmente.
- ▶ Utilizamos métodos de imputación única y múltiple utilizando el supuesto de Rubin (1976) Missing at Random: $\Pr(Y \text{ faltante} | Y, X) = \Pr(Y \text{ faltante} | X)$.
- ▶ Imputación múltiple reconoce que el resultado final puede variar dependiendo de la aleatoriedad. Con IM el resultado final es el promedio de cada imputación.

Métodos

Pareamiento y Hot-Deck

1. Pareamiento por puntaje de propensión: Rosenbaum y Rubin (1983); DiNardo et al (1996) – PSM estimar puntaje y luego pareamiento con puntaje.
 - 1.1 Método intensivo computacionalmente.
 - 1.2 No es posible extenderlo a IM
2. Hot-Deck: Ampliamente utilizado, por ej: Censo o encuestas en Estados Unidos: Andridge y Little (2010), Rubin (1987).
 - 2.1 Muestra aleatoria con reemplazo (bootstrap) de los ingresos válidos (la muestra se mantiene del mismo tamaño n_0)
 - 2.2 De n_0 , tomar una muestra aleatoria con reemplazo de tamaño n_1 .
 - 2.3 Sustituir el ingreso de la muestra n_1 en el paso 2 como ingreso imputado para aquellos con ingreso inválido.
 - ▶ Ventaja: Toma valores observados, extensión a imputación múltiple, no es computacionalmente intenso.

Métodos

Media más ruido y pareamiento por promedios predictivos

1. Grupo con aleatoriedad: Se calcula la mediana más un "ruido" para evitar el problema de subestimar la varianza.
Desventaja: falta de rigor teórico.
 2. Pareamiento por promedios predictivos: Little (1988)
 - 2.1 Obtener los parámetros OLS de $Y = X\beta + \xi$.
 - 2.2 Obtener los parámetros $(\hat{\beta}, \hat{\sigma}^2)$ de una distribución aleatoria.
La varianza se obtiene de una χ^2 , y $\hat{\beta}$ de una normal multivariada.
 - 2.3 Se obtiene una predicción con los nuevos parámetros, y se hace el pareamiento y se sustituye el ingreso inválido con el verdadero observado.
- Es computacionalmente intenso.

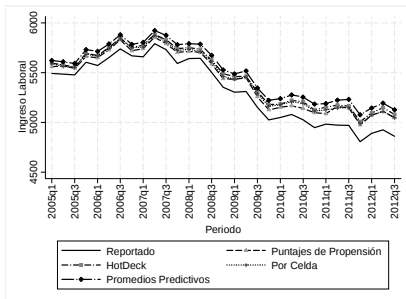
Métodos

Estimación y restricción de datos

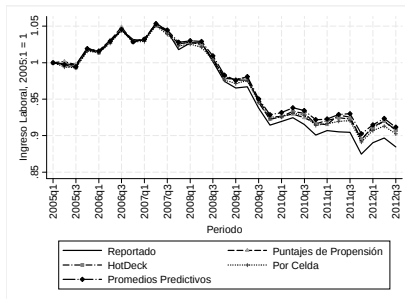
1. Se tiene información de ingreso en términos de los múltiplos de salario mínimo. Para esos individuos siempre se busca un ingreso dentro de ese intervalo para la imputación $\approx 50\%$
2. PSM: polinomio edad, variables dummy de educación, estado civil casado, mujer, trabajador formal, vive en sector rural, trabajador de tiempo completo, geográficas e interacciones entre algunas variables. El pareamiento se realiza diferenciado por género.
3. Hot-Deck: utiliza variables dicotómicas de edad (12 grupos), área metropolitana, nivel educativo, trabajador formal, vive en sector rural, trabajador de tiempo completo. Si no hay match, se realiza eliminando formal y tiempo completo.
4. Mediana más ruido: Igual que hot deck.
5. Pareamiento por promedios predictivos: Igual que PSM

Resultados: Ingreso Individual

A. Promedio de ingreso laboral

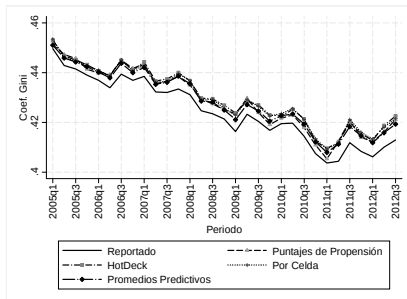


B. Razón de ingreso laboral

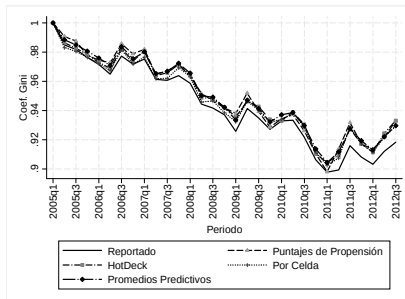


Resultados: Desigualdad Ingreso Individual

A. Desigualdad de ingreso



B. Razón de desigualdad

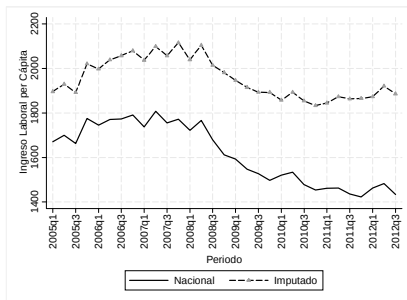


Consecuencias en Ingreso Familiar

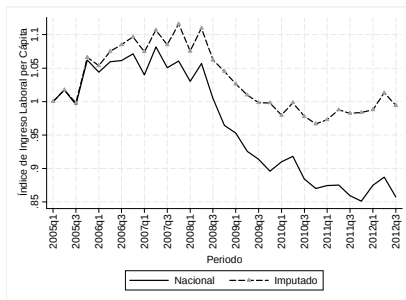
- ▶ Los resultados anteriores se refieren al individuo
- ▶ Si hogares con mayores ingresos y niveles de educación no reportan, al calcular ingresos por hogar se puede aumentar el problema.
- ▶ Para lo siguiente se incluye a todos los hogares en la ENOE.
- ▶ Trabajadores sin pago, desempleados o con horas cero tienen un ingreso cero.

Resultados: Ingreso Familiar per Cápita

A. Promedio

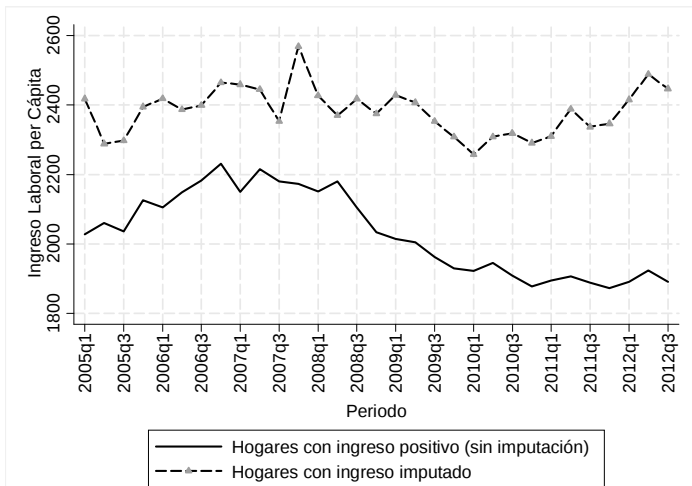


B. Razón



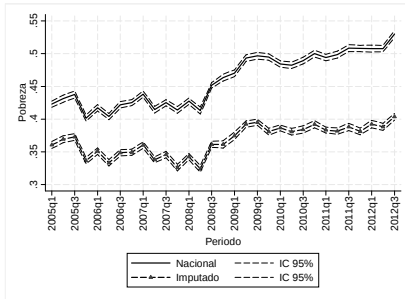
Resultados: Ingreso Familiar per Cápita

No es un problema de los ingresos cero únicamente

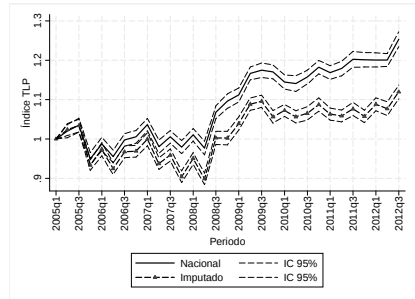


Resultados: Pobreza Laboral

A. Pobreza

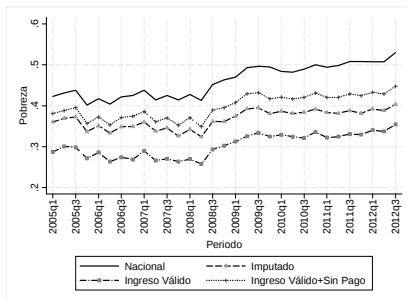


B. ITLP

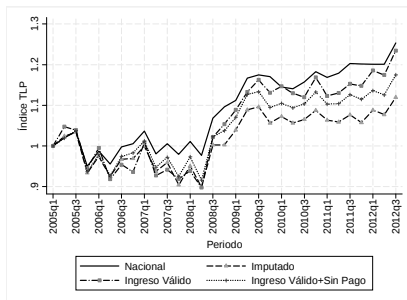


Robustez de resultados

A. Pobreza



B. ITLP



Implicaciones en Pobreza Laboral

- ▶ Encuentro que la pobreza laboral está sobreestimada: CONEVAL-pobreza $\uparrow$ 25% para 2005-2012.
- ▶ Imputación: $\uparrow$ 12%.
- ▶ Tomando la población total de INEGI y los % pobreza tenemos que: la medición oficial implica un aumento de la pobreza laboral de 17.3 millones de personas en el periodo 2005 a 2012, pero utilizando el ingreso imputado se tiene un incremento de 9.2 millones de personas en ese periodo.
- ▶ Es decir, para 2012:3 existen 8.1 millones de personas que son declarados como pobres laborales pero, de acuerdo al método de imputación, no lo son.
- ▶ Son número no triviales.

Conclusiones

Resumen

La proporción de trabajadores que no declara ingresos ha ido en aumento, y su composición está cambiando de tal manera que cada vez más trabajadores con educación superior evitan dar información de ingreso.

1. Para el año 2012 el ingreso individual está subestimado en aproximadamente 4%.
2. La caída en la desigualdad de ingresos individuales está sobreestimada.
3. El ingreso familiar está subestimado en 36%. En particular, se encuentra que el ingreso familiar per cápita verdadero tiene el mismo valor real en el año 2005 primer trimestre y el año 2012 tercer trimestre. Mientras que el ingreso reportado en ese mismo periodo ha caído cerca de 15 por ciento en términos reales.
4. La pobreza laboral como la calcula CONEVAL está sobreestimada. Tanto la subestimación de ingresos familiares como la sobreestimación de la pobreza laboral se deben principalmente a que se toman como ingresos cero los ingresos de personas que sí tienen un ingreso positivo (pero deciden no reportarlo).

Conclusiones

Implicación de Política Pública

- ▶ Esta información es importante, tanto que no tomarla en cuenta sesga sustancialmente las cifras oficiales de pobreza y del ingreso familiar per cápita.
- ▶ Existe una seria sobreestimación en la pobreza laboral del país, y en la tendencia observada.
- ▶ Los resultados anteriores implican que las instituciones oficiales en México deben de considerar imputar los ingresos inválidos siguiendo la misma metodología que en este artículo, o alguna cercana. Suma importancia si se considera asignar recursos con esa medición.
- ▶ Limitación importante y trabajo futuro: ENIGH vs ENOE. No tan sencillo pues ENIGH diferentes fuentes de ingreso.