



# El impacto Social del Earned Income Tax Credit (EITC)

Eduardo Fajnzylber

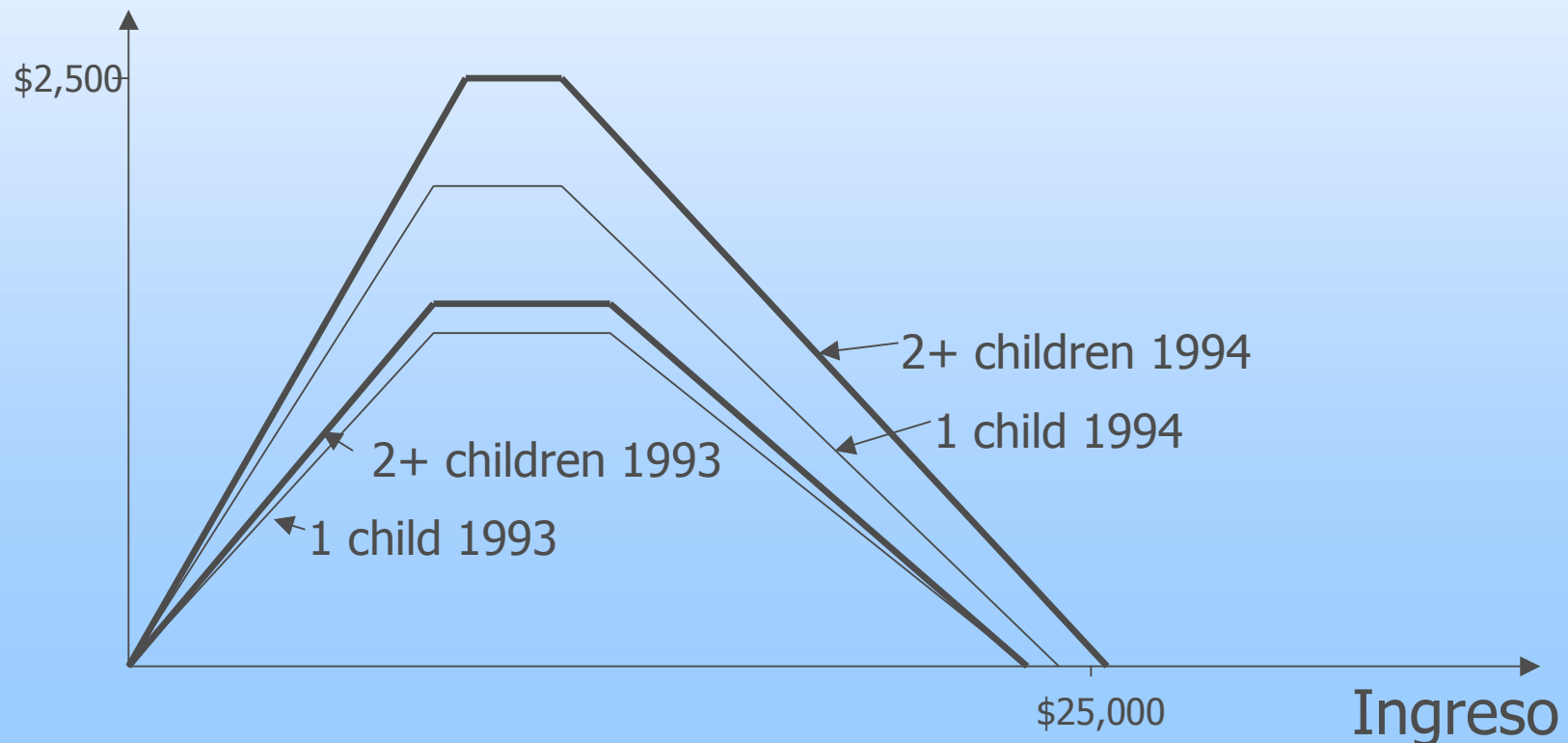
# Introducción

- El Earned Income Tax Credit (EITC), el principal programa de transferencias monetarias en USA
- Cambió los incentivos de los programas tradicionales de asistencia, creando incentivos a trabajar.
- Literatura: Efectos positivos y significativos sobre part. laboral de madres solteras.
- En contextos similares, se ha detectado la presencia de efectos de Interacción Social (decisiones individuales se ven afectadas por comportamiento grupal)
- Si existen efectos externos positivos de mujeres que dejan Welfare y comienzan a trabajar, resultados previos subestimarían impacto social

# La regla de beneficios del EITC

	Credit rate		Maximum Credit	
	1 child	2+ children	1 child	2+ children
1993	18.5%	19.5%	\$ 1,434	\$ 1,511
1994	26.3%	30.0%	\$ 2,038	\$ 2,528

Beneficios



# Pregunta de investigación

- Cual es el efecto externo del EITC en la participación laboral?
- Es decir, el comportamiento del vecino afecta el propio?

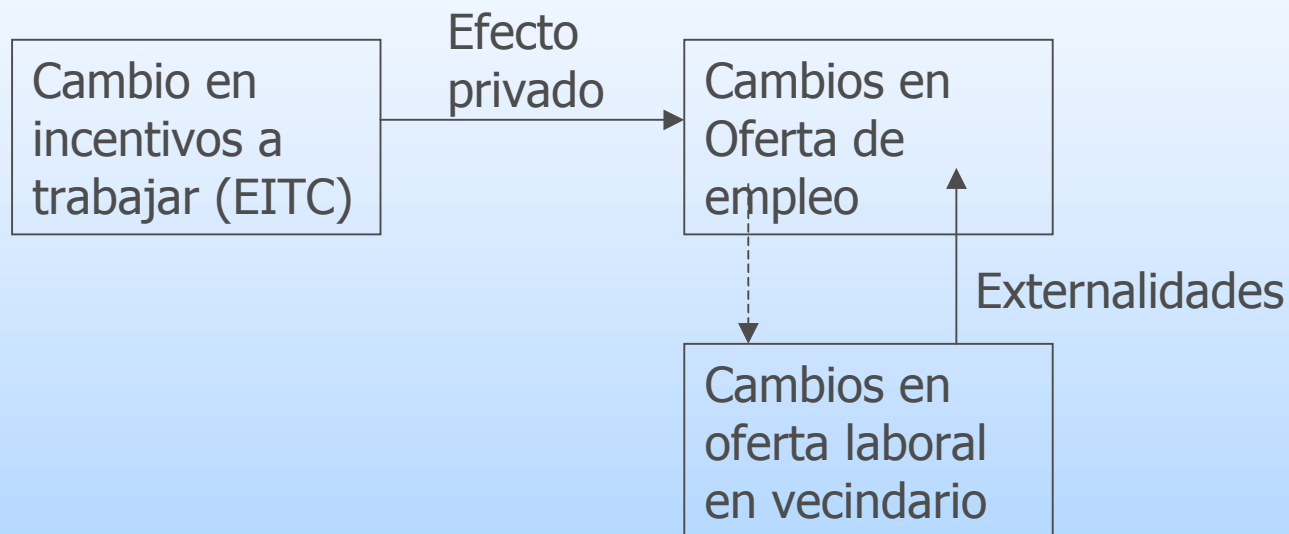
# Motivación

- Si existen efectos externos positivos de mujeres que dejan Welfare y comienzan a trabajar, resultados previos subestimaría el impacto social del EITC

# Efectos directos y externos

- Efecto directo: Como el cambio en incentivos de la expansión del EITC afectó mi decisión de trabajar?
- Externalidad: Cómo el comportamiento de mi vecino (en respuesta a la expansión del EITC) afectó mi propio comportamiento.

# Efectos privados, externalidades e impacto social



$$\text{Impacto Social} = \text{Efecto Privado} + \text{Externalidad}$$

# Objetivo de este artículo

Estimar el impacto del efecto externo del EITC en las decisiones de trabajo de madres solteras.

Metodología: Explotar datos de panel combinado con la expansión diferencial del EITC para madres con 1 o 2+ hijos



# Resumen de resultados

- Efecto privado estimado es similar a literatura previa
- Resultados sugieren un importante efecto externo

# Estructura de la presentación

- Resumen literatura
- Modelo de elección de vecindario y decisión de trabajo
- Identificación
- Modelo empírico y estrategia de estimación
- Descripción de datos
- Resultados y conclusiones

# Revisión literatura

- El impacto del EITC en la oferta laboral de madres solteras
- Identificación en modelos de interacciones sociales
- Literatura empírica en interacciones sociales
  - Experimental
  - No Experimental
- “Sample Selection” con datos de Panel: Kyriazidou (1997)

# Modelo de elección de vecindario /participación laboral

- **Objetivo:** Ilustrar auto-selección en vecindarios
- Decisiones: Vecindario (A or B), Trabaja/No
- Heterogeneidad de preferencias: consumo-ocio ( $\alpha$ ) and aire limpio ( $\rho$ ):

$$U_i = U(Cons, Leis, Qual) = \alpha_i C + (1 - \alpha_i) L + \rho_i Q$$

$$C = \begin{cases} Wage - Rent(n) & \text{if work} \\ Benefits - Rent(n) & \text{if not} \end{cases}$$

$$L = \begin{cases} 0 & \text{if work} \\ 1 & \text{if not} \end{cases}$$

$$Q = \begin{cases} 0 & \text{if } n = A \\ 1 & \text{if } n = B \end{cases}$$

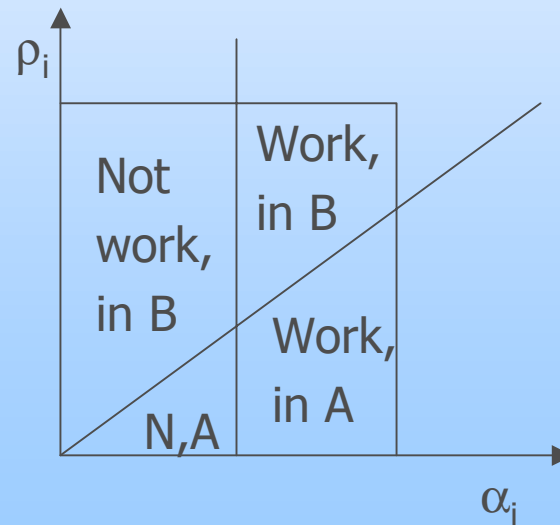
# Modelo de elección de vecindario participación laboral (cont)

- Solucion:

$$Work \Leftrightarrow \alpha_i > \frac{1}{W - B + 1}$$

$$Neighborhood A \Leftrightarrow \alpha_i > \frac{\rho_i}{R_B - R_A}$$

- Equilibrio:



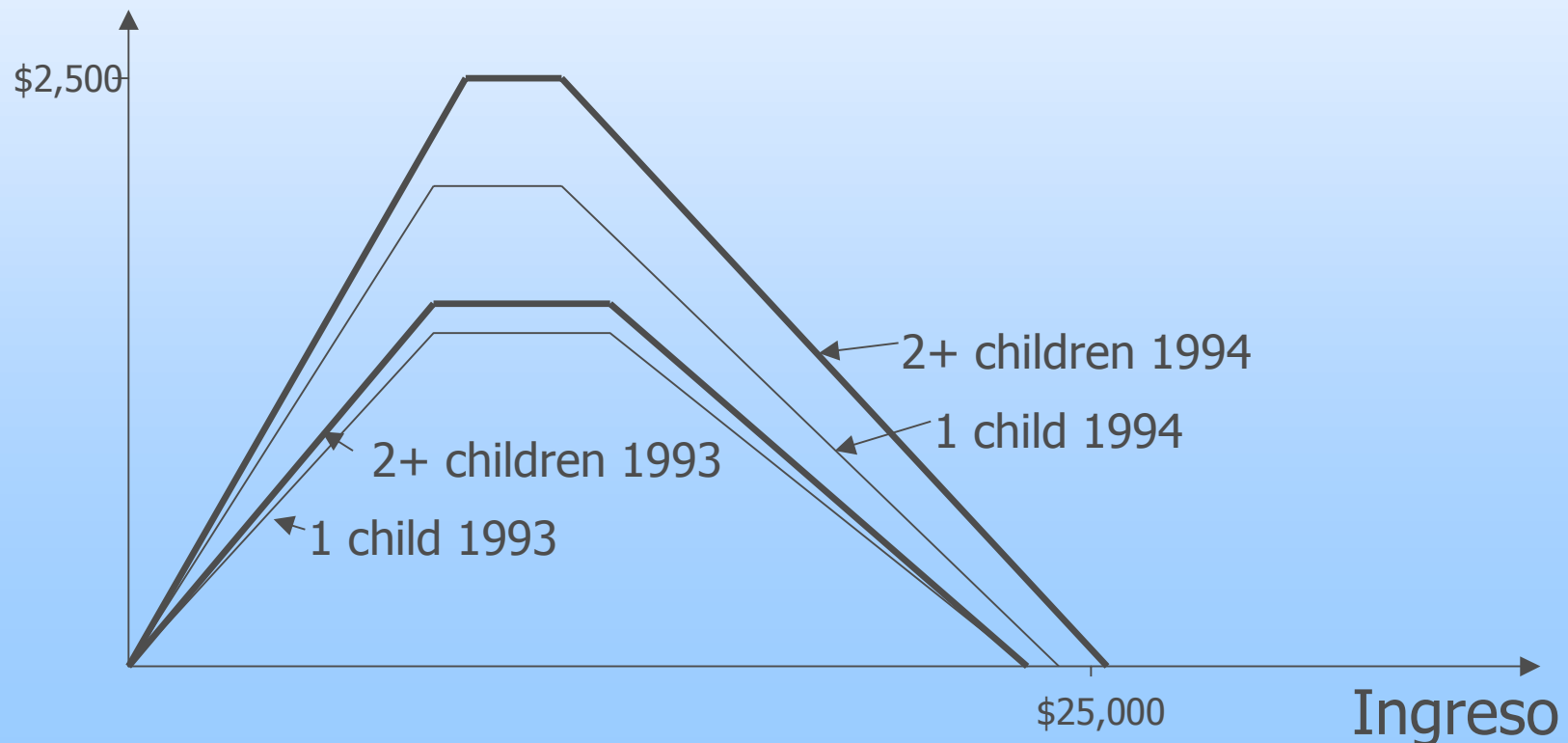
# Problemas de identificación

- Auto-selección en vecindarios:
  - Solución: Datos de panel con Efecto Fijo individual
- Problema de “Reflexión”
- Componente de vecindario no observado
  - Solución: Instrumento en base a expansión del EITC

# La expansión del EITC de 1993

	Credit rate		Maximum Credit	
	1 child	2+ children	1 child	2+ children
1993	18.5%	19.5%	\$ 1,434	\$ 1,511
1994	26.3%	30.0%	\$ 2,038	\$ 2,528

Beneficios



# Modelo empírico

- Modelo de Dif.-en-Dif. + Interacciones sociales

$$LFP_{igt} = \alpha + \beta E[LFP_{igt} | G_{it} = g] + \gamma (KIDS_{it}^{2+} * YEAR_{it}^{1994}) + \delta_1 KIDS_{it}^{2+} + \delta_2 YEAR_{it}^{1994} + \delta_3 AGE_{it} + \delta_4 AGE_{it}^2 + \theta' LEC_{gt} + (\mu_{gt} + v_{igt})$$

- Simplificando notación

$$Y_{igt} = \alpha + \beta \bar{Y}_{gt} + \gamma EITC_{igt} + \delta' X_{igt} + \theta' Z_{gt} + (\mu_{gt} + v_{igt})$$

- Problemas de selección y comp. no observada

$$E[v_{igt} | G_{it} = g] \neq 0 \quad and \quad E[\bar{Y}_{gt} \mu_{gt}] \neq 0$$



# Forma reducida

- Equilibrio con interacciones sociales

$$\bar{Y}_{gt} = \frac{\alpha}{1-\beta} + \frac{\gamma}{1-\beta} \overline{EITC}_{gt} + \frac{\delta}{1-\beta} \bar{X}_{gt} + \frac{\theta}{1-\beta} Z_{gt} + \left( \frac{\mu_{gt}}{1-\beta} + \frac{\bar{v}_{gt}}{1-\beta} \right)$$

- Sustituyendo en ecuación principal, se obtiene una segunda forma reducida

$$Y_{igt} = \frac{\alpha}{1-\beta} + \frac{\gamma\beta}{1-\beta} \overline{EITC}_{gt} + \gamma EITC_{igt} + \frac{\beta\delta}{1-\beta} \bar{X}_{gt} + \delta X_{igt} + \frac{\theta}{1-\beta} Z_{gt} + \left( \frac{\mu_{gt}}{1-\beta} + \frac{\beta\bar{v}_{gt}}{1-\beta} + v_{igt} \right)$$

# Modelo en primeras diferencias (con EF)

- Panel de 2 años, modelo en forma reducida

$$\Delta \bar{Y}_{gt} = \frac{\gamma}{1-\beta} \Delta \bar{EITC}_{gt} + \frac{\theta}{1-\beta} \Delta Z_{gt} + \left( \frac{\Delta \mu_{gt}}{1-\beta} + \frac{\Delta \bar{v}_{gt}}{1-\beta} \right)$$

$$\Delta Y_{igt} = \frac{\gamma\beta}{1-\beta} \Delta \bar{EITC}_{gt} + \gamma \Delta EITC_{igt} + \delta' \Delta X_{igt} + \frac{\theta' \Delta Z_{gt}}{1-\beta} + \left( \frac{\Delta \mu_{gt}}{1-\beta} + \frac{\beta \Delta \bar{v}_{gt}}{1-\beta} + \Delta v_{igt} \right)$$

- Kyriazidou(97): Suponiendo estacionariedad de vecindarios in el corto plazo (1 año), el EF individual debería eliminar el problema de selección para la muestra de los que no cambian ("stayers"):

$$E \left[ \frac{\Delta \mu_{gt}}{1-\beta} + \frac{\beta \Delta \bar{v}_{gt}}{1-\beta} + \Delta v_{igt} \mid \Delta \bar{EITC}_{gt}, \Delta EITC_{igt}, \Delta X_{igt}, \Delta Z_{gt}, G_{it} = G_{it-1} = g \right] = 0$$

# Resumen de estrategias

A) Mínimos cuadrados en 2 etapas:

$$\text{First : } \Delta \bar{Y}_{gt} = \chi_1 \overline{\Delta EITC}_{gt} + \chi_2' \Delta Z_{gt} + \varepsilon_{gt} \Rightarrow \square \bar{Y}_{gt}$$

$$\text{Second : } \Delta Y_{igt} = \beta \square \bar{Y}_{gt} + \gamma \Delta EITC_{igt} + \delta' \Delta X_{igt} + \theta' \Delta Z_{gt} + \eta_{igt}$$

B) Estimador de forma reducida (suponiendo equilibrio):

$$\Delta Y_{igt} = \pi_1 \overline{\Delta EITC}_{gt} + \pi_2 \Delta EITC_{igt} + \delta' \Delta X_{igt} + \tilde{\theta}' \Delta Z_{gt} + \varepsilon_{igt}$$

$$\text{with } \left\{ \begin{array}{l} \pi_1 = \frac{\gamma \beta}{1 - \beta} \\ \pi_2 = \gamma \end{array} \right\} \Rightarrow \beta = \frac{\pi_1}{\pi_1 + \pi_2}$$

# Descripción de base de datos

- Datos cruzados de 3 fuentes en California:
  - MEDS: ubicación (condado, código postal), demográficos (age, race, language), y composición familiar.
  - Ingresos de EDD: empleos formales
  - Condiciones económicas locales: Desempleo y ratio de empleados a población
- Muestra: madres solteras en welfare al menos 1 mes entre 1992 y 1995

# Descripción de base de datos (2)

- PL= ( Ingresos anuales > US\$200 )
- # de individuos = 142,391
- # de zip codes (N>100) = 503
- PL promedio= 33.7% en 1993, 36.7% en 1994
- Fracción promedio de 2+niños. = 38.6% (sd= 6%)
- Raza / etnia: 42.8% Hispanico, 28.8% blanca, 19.5% negra
- Desempleo = 10.3% en 1993, 9.4% en 1994

# Efecto privado del EITC (sin IS) (Tab3)

Dependent Variable = Labor Force Participation (Annual Earnings > \$200)

	Cross Section	Pooled 1993 and 1994	Pooled 1993/94 + FE
<b>(More than 1 child)*(Post 93)</b>		<b>0.016</b>	<b>0.016</b>
<b>(Private effect of EITC)</b>		<b>(7.11)**</b>	<b>(7.07)**</b>
More than 1 child	-0.027 (9.51)**	-0.043 (15.51)**	
Post 93		0.019 (9.85)**	
Mother's age	0.015 (7.64)**	0.017 (9.45)**	
Mother's squared age	-0.000 (7.22)**	-0.000 (8.84)**	-0.000 (2.27)*
Employment / Population - all sectors	-0.072 (5.16)**	-0.095 (7.44)**	-0.101 (1.78)
Observations	142391	284782	142391
R-squared	0.0012	0.0029	0.0004

Absolute value of t statistics in parentheses (standard errors in pooled regression clustered by individual)

\* significant at 5%; \*\* significant at 1%

# Estimadores de IS con diferentes controles (Tab 4)

Dependent Variable = Labor Force Participation (Annual Earnings > \$200)

	FE without IV	IV without FE	Baseline: IV + FE
<b>Average LFP</b>	<b>0.996</b>	<b>0.753</b>	<b>0.651</b>
<b>( Spillover effect )</b>	<b>(472.79)**</b>	<b>(8.47)**</b>	<b>(4.07)**</b>
<b>(More than 1 child)*(Post 93)</b>	<b>0.015</b>	<b>0.015</b>	<b>0.016</b>
<b>( Private effect of EITC )</b>	<b>(6.59)**</b>	<b>(3.83)**</b>	<b>(6.59)**</b>
More than 1 child		-0.041 (14.24)**	
Post 1993		-0.004 (1.09)	
Mother's age		0.015 (9.38)**	
Mother's squared age	-0.000 (2.20)*	-0.000 (8.70)**	-0.019 (2.23)*
Employment / Population - all sectors	-0.104 (1.72)	-0.026 (2.67)**	-0.103 (1.70)
# Observations	142391	284782	142391
# zip codes	503	503	503
R-squared	0.0046	0.0249	0.0041

First Stage: Dependent Variable = Average Labor Force Participation in the zip code

<b>Average ( (More than 1 child)*(Post 93) )</b>	<b>0.046</b>
<b>( Social impact of the EITC )</b>	<b>(2.29)*</b>
Observations	142391
R-squared	0.0015

Absolute value of t statistics in parentheses (standard errors allowing for clustering by zip code)

\* significant at 5%; \*\* significant at 1%

Note: Other covariates in first stage (not reported) include all individual variables in second stage

# Estimadores de IS con diferentes muestras (Tab5)

Second Stage: Dependent Variable = Labor Force Participation (Annual Earnings > \$200)

	zip size > 25	zip size > 50	zip size > 100	zip size > 100	zip size > 100
	Base sample	Base sample	Base sample	Sample B*	Sample C*
<b>Average LFP</b>	<b>0.574</b>	<b>0.544</b>	<b>0.651</b>	<b>0.629</b>	<b>0.603</b>
<b>( Spillover effect )</b>	<b>(3.20)**</b>	<b>(2.32)*</b>	<b>(4.07)**</b>	<b>(1.93)</b>	<b>(1.56)</b>
<b>(More than 1 child)*(Post 93)</b>	<b>0.017</b>	<b>0.016</b>	<b>0.016</b>	<b>0.012</b>	<b>0.012</b>
<b>( Private effect of EITC )</b>	<b>(7.70)**</b>	<b>(7.06)**</b>	<b>(6.59)**</b>	<b>(3.55)**</b>	<b>(3.46)**</b>
Squared age	-0.023	-0.022	-0.019	-0.057	-0.063
	(2.90)**	(2.64)**	(2.23)*	(4.48)**	(4.79)**
Employment / Population - all sectors	-0.111	-0.103	-0.103	-0.091	-0.084
	(2.00)*	(1.80)	(1.70)	(1.33)	(1.21)
# Observations	165139	157762	142391	84853	82146
# zip codes	921	716	503	364	354
R-squared	0.0060	0.0047	0.0041	0.0048	0.0048

First Stage: Dependent Variable = Average Labor Force Participation in the zip code

<b>Average ( (More than 1 child)*(Post 93) )</b>	<b>0.041</b>	<b>0.036</b>	<b>0.046</b>	<b>0.034</b>	<b>0.031</b>
<b>( Social impact of the EITC )</b>	<b>(2.49)*</b>	<b>(2.01)*</b>	<b>(2.29)*</b>	<b>(1.21)</b>	<b>(1.08)</b>
Observations	165139	157762	142391	84853	82146
R-squared	0.0066	0.0057	0.0102	0.0052	0.0045

Absolute value of t statistics in parentheses (standard errors allowing for clustering by zip code)

\* significant at 5%; \*\* significant at 1%

Note: Other covariates in first stage (not reported) include all individual variables in second stage

Sample B = Base sample restricted to individuals on welfare at least twice between 1992 and 1995

Sample C = Base sample restricted to individuals on welfare at least twice between 1992 and 1995, at most one year apart



# Estimadores de IS con EF por condado (tab7)

Second Stage: Dependent Variable = Labor Force Participation (Annual Earnings > \$200)

	zip size > 25	zip size > 50	zip size > 100
	Base sample	Base sample	Base sample
<b>Average LFP</b>	<b>0.518</b>	<b>0.467</b>	<b>0.606</b>
<b>( Spillover effect )</b>	<b>(2.25)*</b>	<b>(1.46)</b>	<b>(3.11)**</b>
<b>(More than 1 child)*(Post 93)</b>	<b>0.017</b>	<b>0.016</b>	<b>0.016</b>
<b>( Private effect of EITC )</b>	<b>(7.69)**</b>	<b>(7.06)**</b>	<b>(6.60)**</b>
Squared age	-0.000	-0.000	-0.000
	(2.88)**	(2.60)**	(2.21)*
# Observations	165139	157762	142391
# zip codes	716	503	364
R-squared	0.0058	0.0045	0.0040

First Stage: Dependent Variable = Average Labor Force Participation in the zip code

<b>Average ( (More than 1 child)*(Post 93) )</b>	<b>0.038</b>	<b>0.032</b>	<b>0.045</b>
<b>( Social impact of the EITC )</b>	<b>(2.19)*</b>	<b>(1.73)</b>	<b>(2.18)*</b>
Observations	165139	157762	142391
R-squared	0.0888	0.0962	0.1154

Absolute value of t statistics in parentheses (standard errors allowing for clustering by zip code)

\* significant at 5%; \*\* significant at 1%

Note: Other covariates in first stage (not reported) include all individual variables in second stage

Sample B = Base sample restricted to individuals on welfare at least twice between 1992 and 1995

Sample C = Base sample restricted to individuals on welfare at least twice between 1992 and 1995, at most one year apart

# Instrumentos débiles?

- Instrumentos débiles implican que la inferencia basada en teoría asintótica tradicional está errada.
- Literatura reciente sugiere utilizar estadísticos pivotaes para hacer inferencia: elegí estadístico de Anderson-Rubin

# Estadístico de Anderson-Rubin

- Idea: Testear hipótesis  $H_0: \beta = \beta_0$  in modelo de ecuaciones simultaneas:

$$\Delta Y_{igt} = \beta \Delta \bar{Y}_{gt} + \gamma \Delta EITC_{igt} + \delta' \Delta X_{igt} + \theta' \Delta Z_{gt} + \eta_{igt}$$

$$\Delta \bar{Y}_{gt} = \chi_1 \Delta \overline{EITC}_{gt} + \chi_2' \Delta Z_{gt} + \varepsilon_{gt}$$

- Se puede estimar un modelo auxiliar:

$$\left( \Delta Y_{igt} - \beta_0 \Delta \bar{Y}_{gt} \right) = \pi_1 \Delta \overline{EITC}_{gt} + \pi_2 \Delta EITC_{igt} + \pi_3' \Delta X_{igt} + \pi_4' \Delta Z_{gt} + \eta_{igt}$$

- Y realizar un test F sobre la hipotesis nula que  $\pi_1 = 0$
- Con un sólo instrumento, este test es equivalente a un test T
- La hipótesis de « No Interacciones Sociales » puede testearse directamente en modelo de forma reducida

# Resultados por zip\*raza

<b>Minimum Distance Results for zip*race groups</b>					
Areas with at least 21 individuals in a zip*race class					
Specifications with County*time fixed effects					
<b>Reduced form parameters</b>	previous	all races	whites	blacks	hispanics
Average Fraction of 2+ child families	0.02705	0.03130	0.01102	0.10381	0.07754
(AR test for $H_0: \beta=0$ )	(1.30)	(2.16)*	(0.32)	(2.04)*	(2.71)**
Dummy=1 if 2 or more children	0.01557	0.01509	0.01691	0.02447	0.01303
	(6.61)**	(6.23)**	(3.99)**	(4.31)**	(3.94)**
Squared Age	-0.00039	-0.00038	0.00040	-0.00122	-0.00055
	(2.24)*	(2.13)*	(1.09)	(2.56)*	(2.58)*
Observations	142391	138477	40392	26750	60657
R-squared	0.0010	0.0010	0.0018	0.0029	0.0017
# zip codes	503	503	439	309	457
Robust t statistics in parentheses					
* significant at 5%; ** significant at 1%					
<b>Structural coefficient (beta)</b>					
beta	0.6346	0.6746	0.3946	0.8092	0.8562
t-stat	<b>3.42</b>	<b>5.98</b>	0.52	<b>9.05</b>	<b>15.06</b>

# Conclusiones

- Efecto privado similar a la literatura
  - Familias con 2 o más niños aumentaron su PL entre 1993 y 1994 en 1.6% p.p. MAS que familias con 1 solo niño
  - Inafectado por inclusión de término de interacciones sociales
- Importantes efectos externos
  - Interacciones sociales entre 0.47(0.32) y 0.61(0.20)
  - Efecto privado agregado = 1.1% (0.2)
  - Efecto externo entre 1.1%(1.3) y 2.1(1.4)

## Conclusiones (2)

- Puede ser más costo-efectivo intervenir directamente en Interacciones Sociales
  - Ej: aumentar difusión de información acerca de empleos in vecindarios de baja PL
- Presencia de IS tiene importantes implicancias para evaluaciones apropiadas de programas sociales