

# Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings

Raj Chetty, Harvard and NBER  
John N. Friedman, Harvard and NBER  
Emmanuel Saez, UC Berkeley and NBER

January 2013

# Identifying Policy Impacts

- Two central challenges in identifying the impacts of govt. policies:
  1. Lack of counterfactuals to estimate causal impacts of policies  
[Meyer 1995, Saez et al. 2012]
  2. Difficult to identify long run impacts from short-run responses to tax changes
    - Many people are uninformed about tax and transfer policies  
[Brown 1968, Bises 1990, Chetty and Saez 2009]
    - Workers face switching costs for labor supply  
[Cogan 1981, Altonji and Paxson 1992, Chetty et al. 2011]

# Overview

- We develop a new method of addressing these challenges by exploiting differences across neighborhoods in knowledge about tax policies
  - Individuals with no knowledge of a policy's marginal incentives behave as they would in the absence of a policy
  - Cities with low levels of information about policies yield counterfactuals for behavior in absence of policy
- Apply this approach to characterize the impacts of the Earned Income Tax Credit (EITC) on the earnings distribution in the U.S.
  - EITC provides refunds of up to \$5,000 to approximately 25 million households in the U.S.

## Earned Income Tax Credit Schedule for Single Earners with One Child



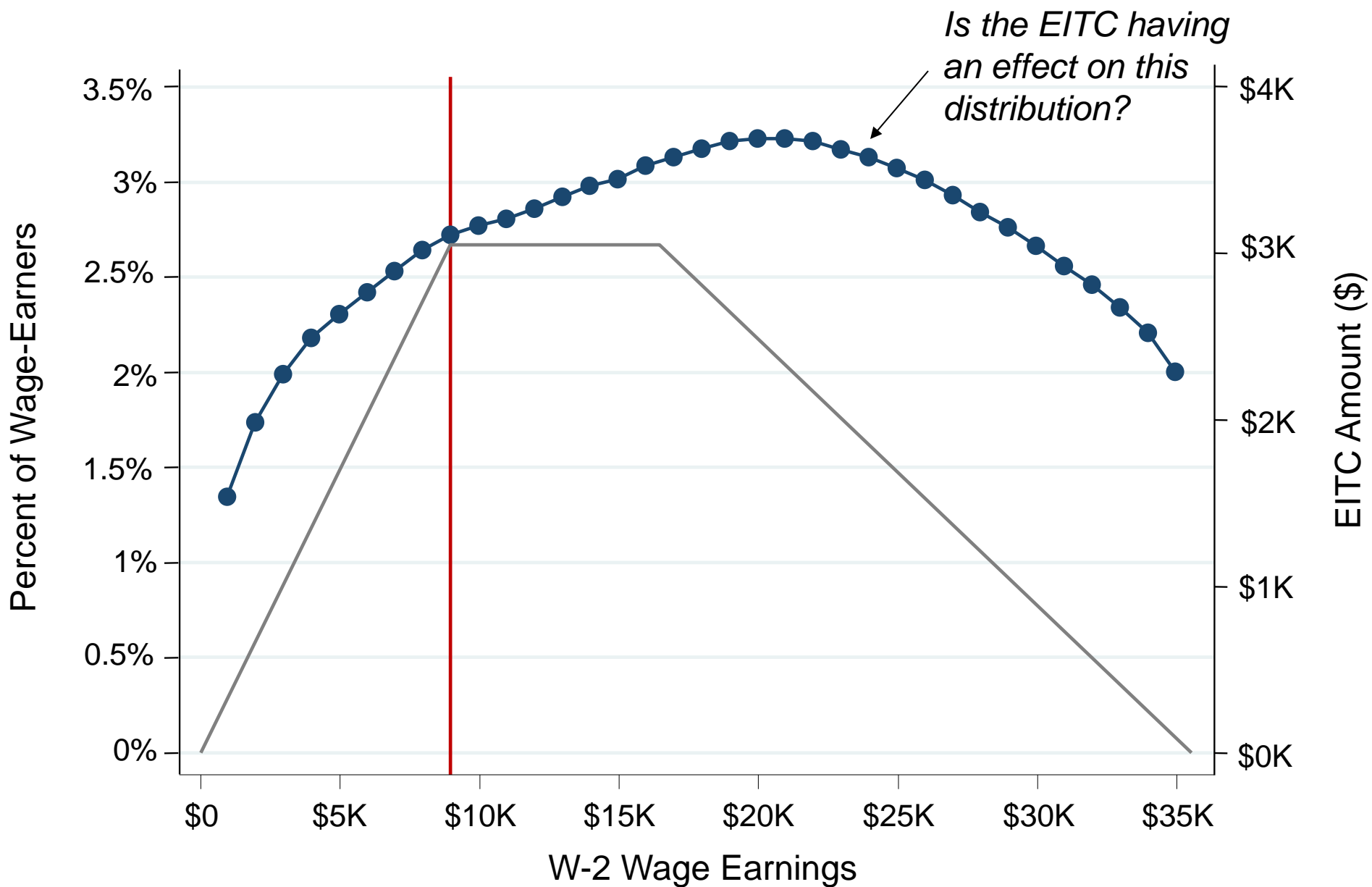
# Relationship to Prior Work

- Large literature has studied the impacts of EITC on labor supply  
[Eissa and Liebman 1996, Meyer and Rosenbaum 2001, Meyer 2002, Grogger 2003, Hoynes 2004, Gelber and Mitchell 2011]
- Clear evidence of impacts on *participation* (extensive margin)
- But no clear, non-parametric evidence on impacts of EITC on *earnings distribution* (intensive margin)
- Same pattern in studies of labor supply elasticities more generally
- Observed extensive responses may be larger because more people know about existence of EITC refund than shape of schedule
- Gains from re-optimization 2<sup>nd</sup>-order on intensive but 1<sup>st</sup> order on ext. margin → frictions attenuate intensive responses [Chetty 2012]

# Income Distribution For Single Wage Earners with One Child



# Income Distribution For Single Wage Earners with One Child



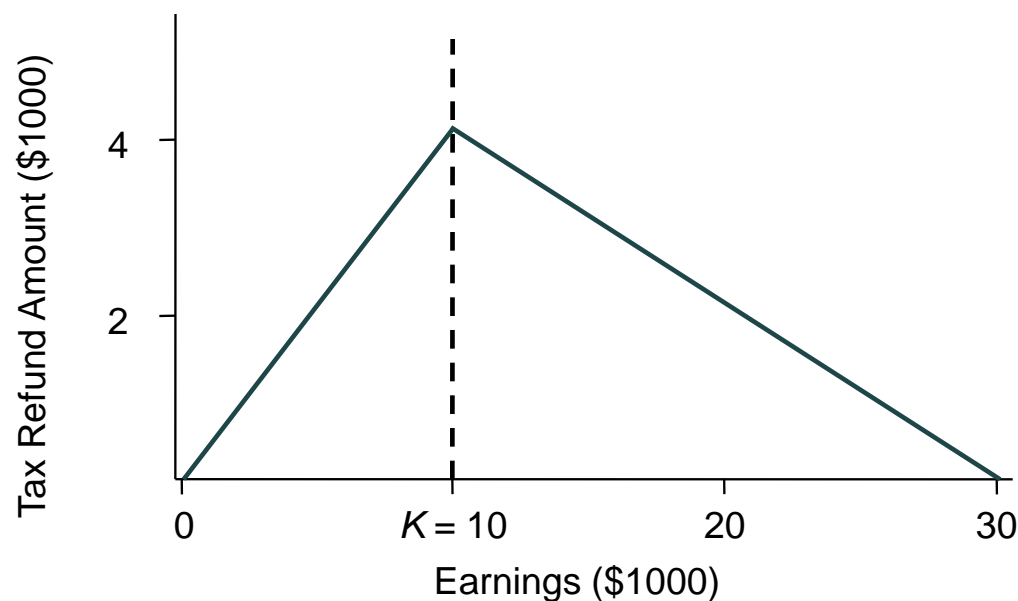
# Outline

1. Conceptual Framework
2. Data and Institutional Background
3. Proxy for Knowledge: Sharp Bunching via Self-Emp Income Manipulation
4. Uncover Wage Earnings Responses
5. Implications for Tax Policy



# Stylized Model: Tax System

- Workers face a two-bracket income tax system  $\tau = (\tau_1, \tau_2)$  and choose earnings  $z=wl$  to maximize quasi-linear utility  $C_i - h(l_i, \alpha_i)$
- Tax rate of  $\tau_1 < 0$  when reported income is below  $K$
- Marginal tax rate of  $\tau_2 > 0$  for reported income above  $K$
- Tax refund maximized when income is  $K \rightarrow$  bunching around  $K$



# Neighborhoods

- Cities indexed by  $c = 1, \dots, N$
- In stylized model, assume that cities differ only in one attribute: knowledge of tax code
  - We relax this assumption in our empirical implementation and instead impose an orthogonality condition for identification
- In city  $c$ , fraction  $\lambda_c$  of workers know about tax subsidy for work
  - Others optimize as if tax rates are 0 (i.e. subsidy is lump-sum)
- Firms pay workers fixed wage rate in all cities

# Identifying Tax Policy Impacts

- Goal: estimate impact of tax system on earnings distribution  $F(z | \tau)$  with average level of knowledge in economy

$$\Delta F(z | \tau) = F(z | \tau \neq 0, \bar{\lambda}_c) - F(z | \tau = 0, \bar{\lambda}_c)$$

- Challenge: potential outcome without taxes  $F(z | \tau = 0, \bar{\lambda}_c)$  unobserved
- Our solution: earnings behavior with no *knowledge* about taxes is equivalent to earnings behavior with no taxes

$$F(z | \tau = 0, \bar{\lambda}_c) = F(z | \tau > 0, \lambda_c = 0)$$

$$\Rightarrow \Delta F(z | \tau) = F(z | \tau > 0, \bar{\lambda}_c) - F(z | \tau > 0, \lambda_c = 0)$$

# Data and Sample Definition

- Selected data from population of U.S. income tax returns, 1996-2009
  - Includes 1040's and all information forms (e.g. W-2's)
- Sample restriction: individuals who at least once between 1996-2009:  
(1) file a tax return, (2) have income < \$50,000, (3) claim a dependent
- Sample size after restrictions:
  - 77.6 million unique taxpayers
  - 1.09 billion taxpayer-year observations on income

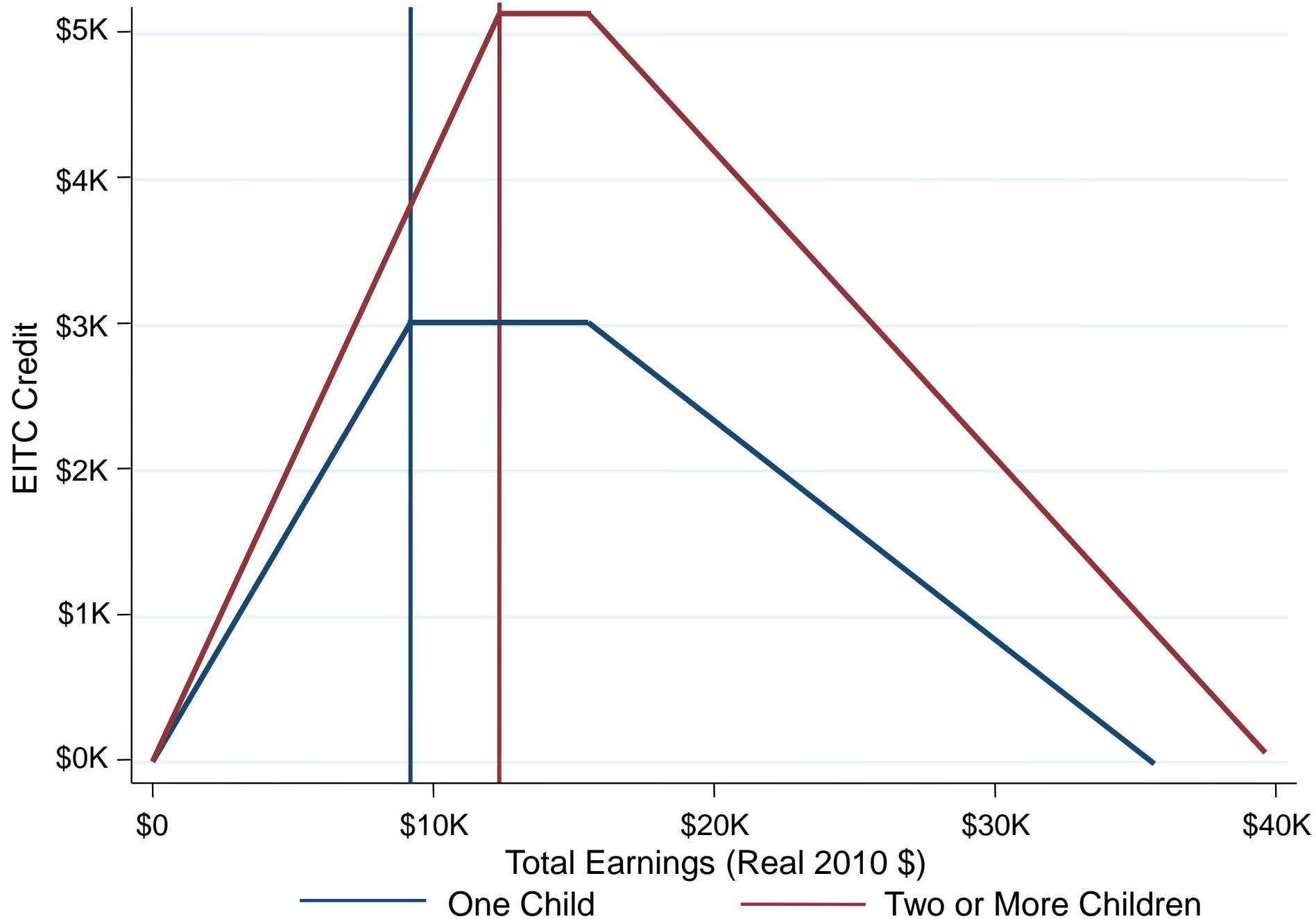
## Summary Statistics for EITC Eligible Individuals

Variable	Mean (1)	Std. Dev. (2)
<u>Income Measures</u>		
Total Earnings	\$20,091	\$10,784
Wage Earnings	\$18,308	\$12,537
Self-Employment Income	\$1,770	\$6,074
Non-Zero Self-Emp. Income	19.6%	39.7%
<u>Tax Credits</u>		
EITC Refund Amount	\$2,543	\$1,454
Claimed EITC	88.9%	31.4%
Professionally Prepared Return	69.6%	46.0%
<u>Demographics</u>		
Age	37	13
Number of Children	1.7	0.8
Married	30.3%	45.9%
Female (for single filers)	73.0%	44.4%
Number of Observations	219,742,011	

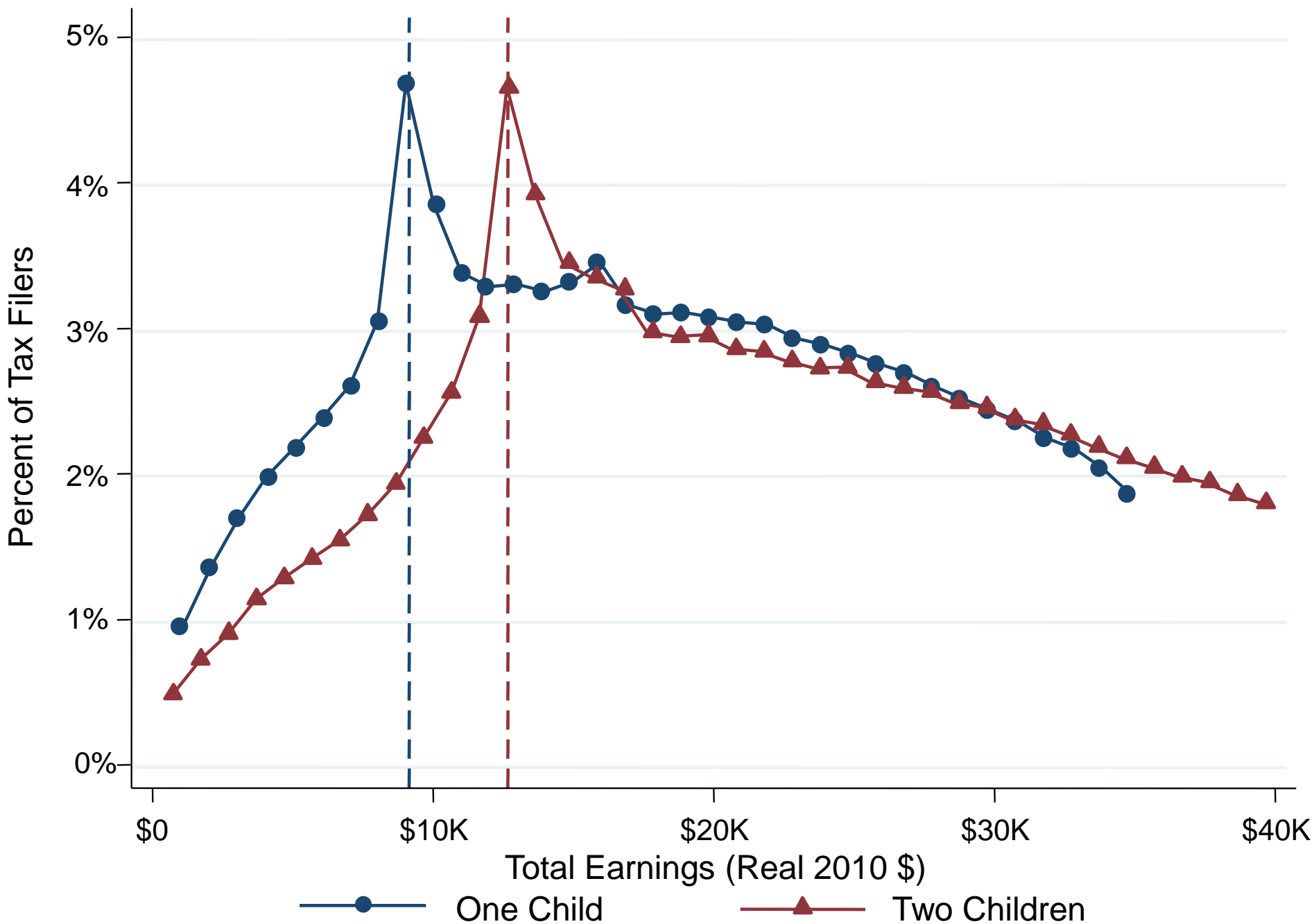
# Self Employment Income vs. Wage Earnings

- To measure local knowledge, we rely on a critical distinction between wage earnings and self-employment income
- Self-employment income is self-reported → easy to manipulate
- Wage earnings are directly reported to IRS by employers
  - Therefore more likely to reflect “real” earnings behavior

## 2008 Federal EITC Schedule for a Single Filer with Children



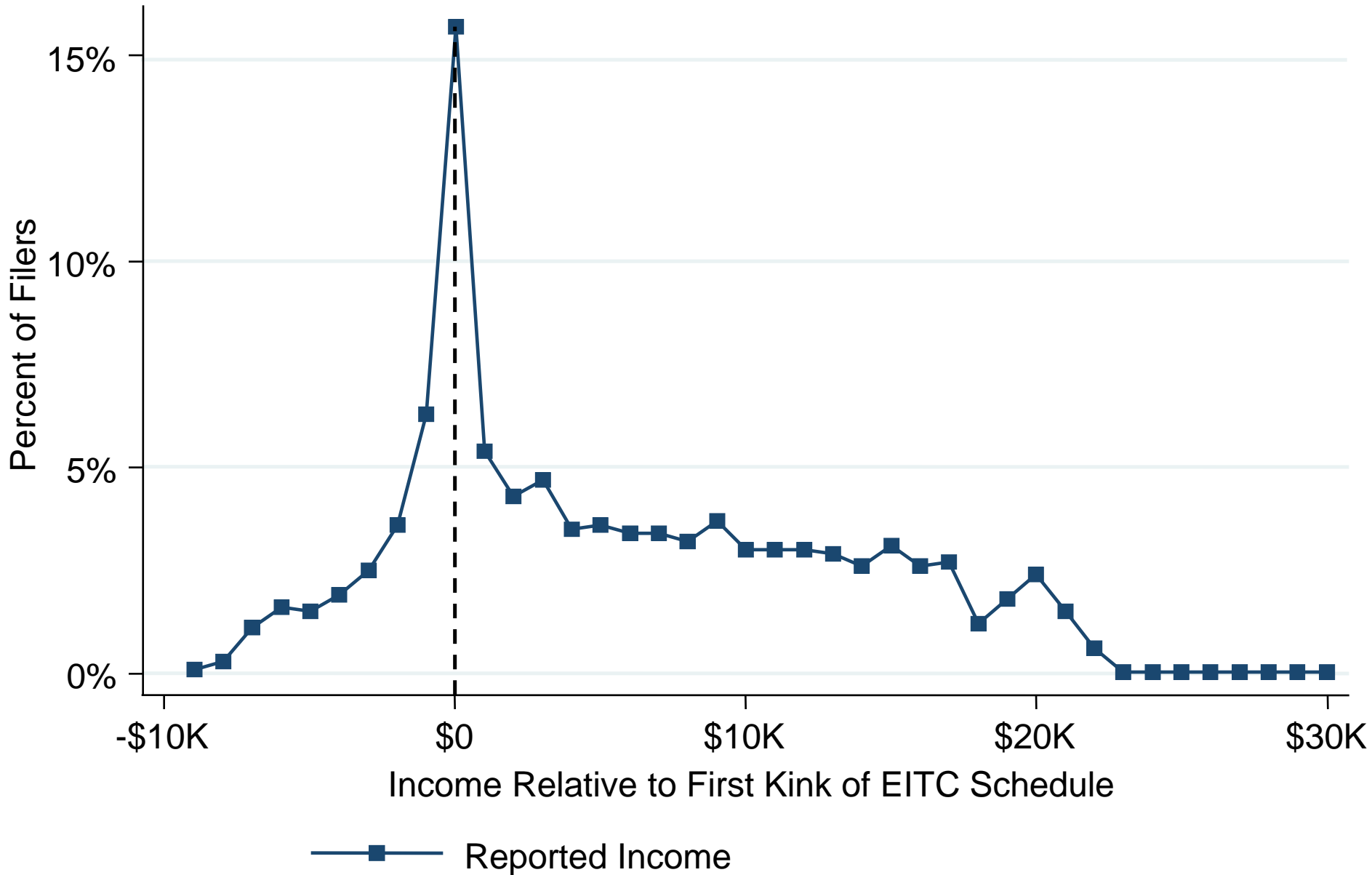
# Income Distributions for Individuals with Children in 2008





# Reported vs. Audited Income Distributions for SE EITC Filers in 2001

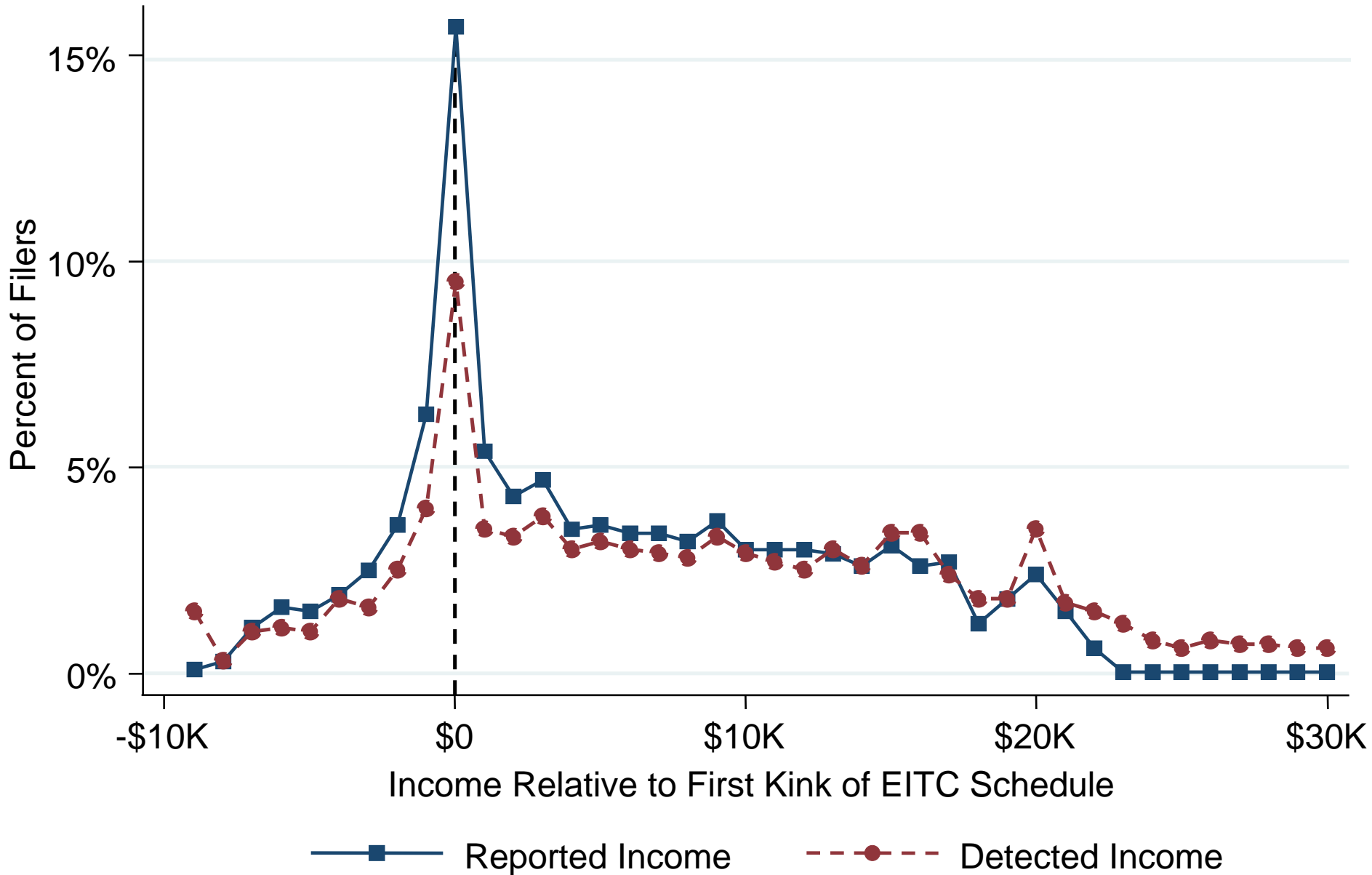
## National Research Program Tax Audit Data



Source: IRS TY01 NRP reporting compliance study of individual income tax returns for those reporting dependent children; amounts reflect only what was detected by the auditors, weighted to population levels.

# Reported vs. Audited Income Distributions for SE EITC Filers in 2001

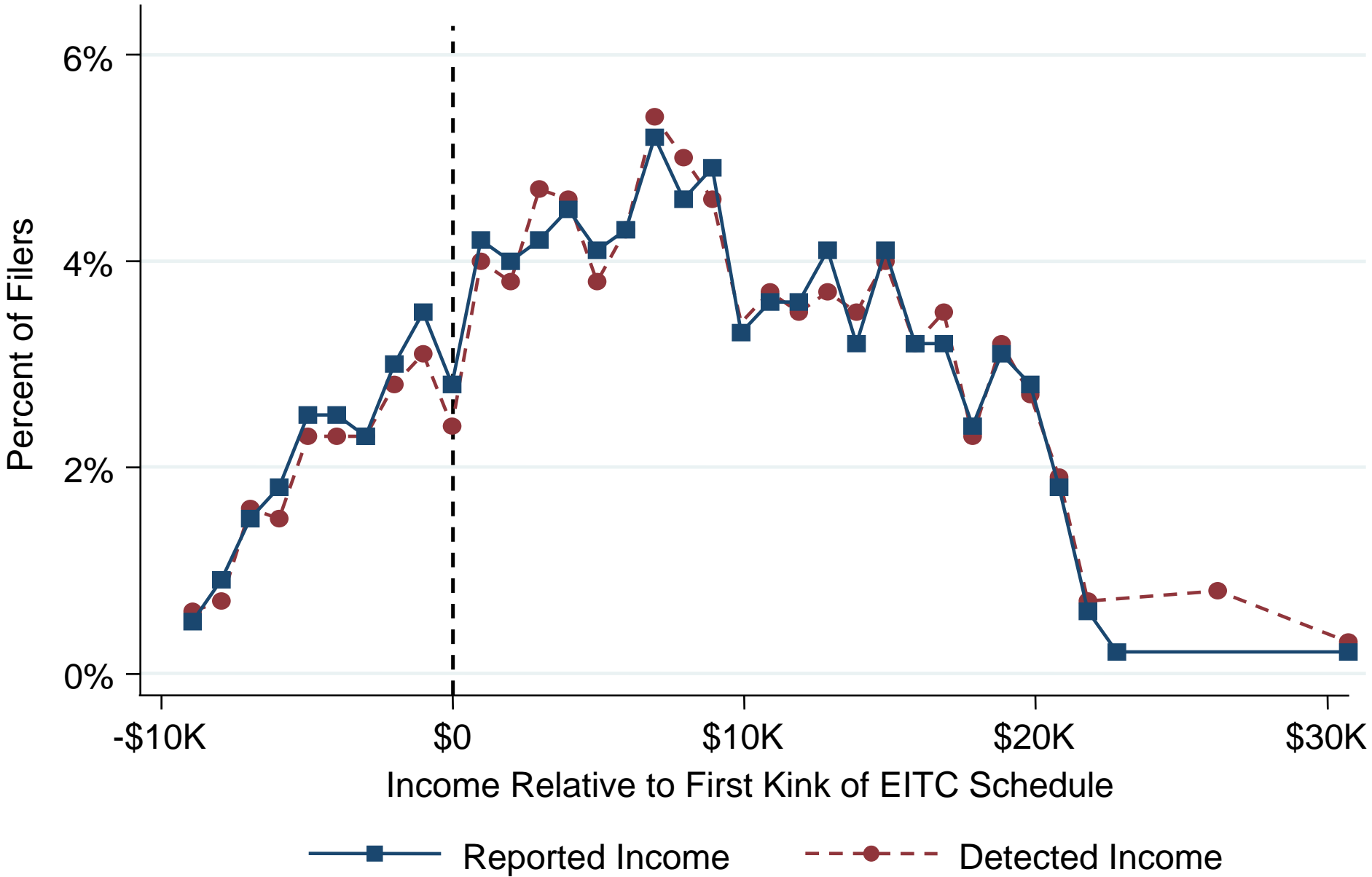
## National Research Program Tax Audit Data



Source: IRS TY01 NRP reporting compliance study of individual income tax returns for those reporting dependent children; amounts reflect only what was detected by the auditors, weighted to population levels.

# Reported vs. Audited Income Distributions for EITC Wage Earners with Children

## National Research Program Tax Audit Data



Source: IRS TY01 NRP reporting compliance study of individual income tax returns for those reporting dependent children; amounts reflect only what was detected by the auditors, weighted to population levels.

# Empirical Implementation: Proxy for Knowledge

- We proxy for knowledge  $\lambda_c$  using sharp bunching at refund-maximizing kink among the self-employed
  - Intuition: use amount of misreporting to measure local tax knowledge
- Workers make two choices: earnings ( $z_i$ ) and reported income ( $\hat{z}_i$ )
  - Fraction  $\theta_c$  of workers face 0 cost of non-compliance  $\rightarrow$  report  $\hat{z}_i = K$
  - Remaining workers face infinite cost of non-compliance  $\rightarrow$  set  $\hat{z}_i = z_i$
- Fraction who report  $\hat{z}_i = K$  is proportional to local knowledge:

$$\phi_c = \theta_c \lambda_c$$

# Empirical Implementation: Proxy for Knowledge

- We use areas with no sharp bunching as counterfactuals for behavior in the absence of the EITC
- Research design rests on two identification assumptions in a model that permits arbitrary differences in distribution of skills  $G_c(\alpha_i)$  across cities

# Identification Assumption 1: Tax Knowledge

**Assumption 1 [Tax Knowledge]** *Individuals in cities with no sharp bunching have no knowledge about EITC schedule and perceive  $\tau = 0$*

$$\phi_c = 0 \rightarrow \lambda_c = 0$$

- Requires that individuals in areas with no sharp bunching behave as if tax policy has no impact on marginal incentives
  - We present evidence supporting this assumption below
  - Violations of this assumption lead us to understate impacts of EITC

# Identification Assumption 2: Counterfactuals

- Cross-sectional estimator: compare aggregate earnings distribution with distribution in neighborhoods with 0 sharp bunching

$$\widehat{\Delta F} = F(z|\tau) - F(z|\tau, \phi_c = 0)$$

**Assumption 2a [Cross-Sectional Identification]** *Individuals' skills  $G_c(\alpha_i)$  do not vary across cities with different levels of knowledge  $\lambda_c$*

# Identification Assumption 2: Counterfactuals

- Cross-sectional estimator: compare aggregate earnings distribution with distribution in neighborhoods with 0 sharp bunching

$$\widehat{\Delta F} = F(z|\tau) - F(z|\tau, \phi_c = 0)$$

**Assumption 2a [Cross-Sectional Identification]** *Individuals' skills  $G_c(\alpha_i)$  do not vary across cities with different levels of knowledge  $\lambda_c$*

- Panel estimator: compare *changes* in aggregate earnings distribution around eligibility due to child birth with changes in  $\phi_c = 0$  nbhds.

$$\widehat{\Delta F}_{DD} = [F_t(z|\tau) - F_t(z|\tau, \phi_c = 0)] - [F_{t-1}(z|\tau) - F_{t-1}(z|\tau, \phi_c = 0)]$$

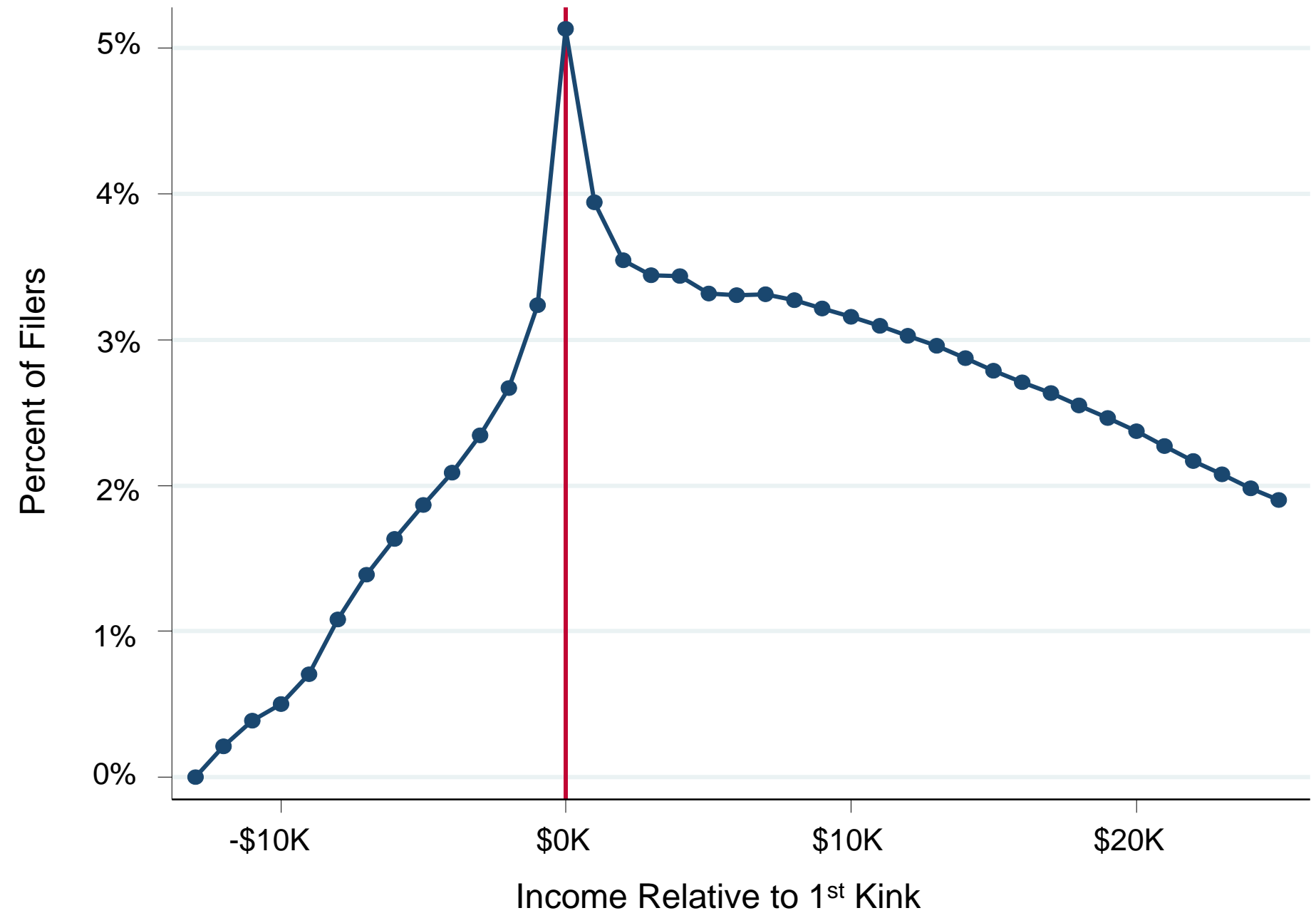
**Assumption 2b [Panel Identification]** *Changes in skills when an individual becomes eligible for credit do not vary across cities with different  $\lambda_c$*



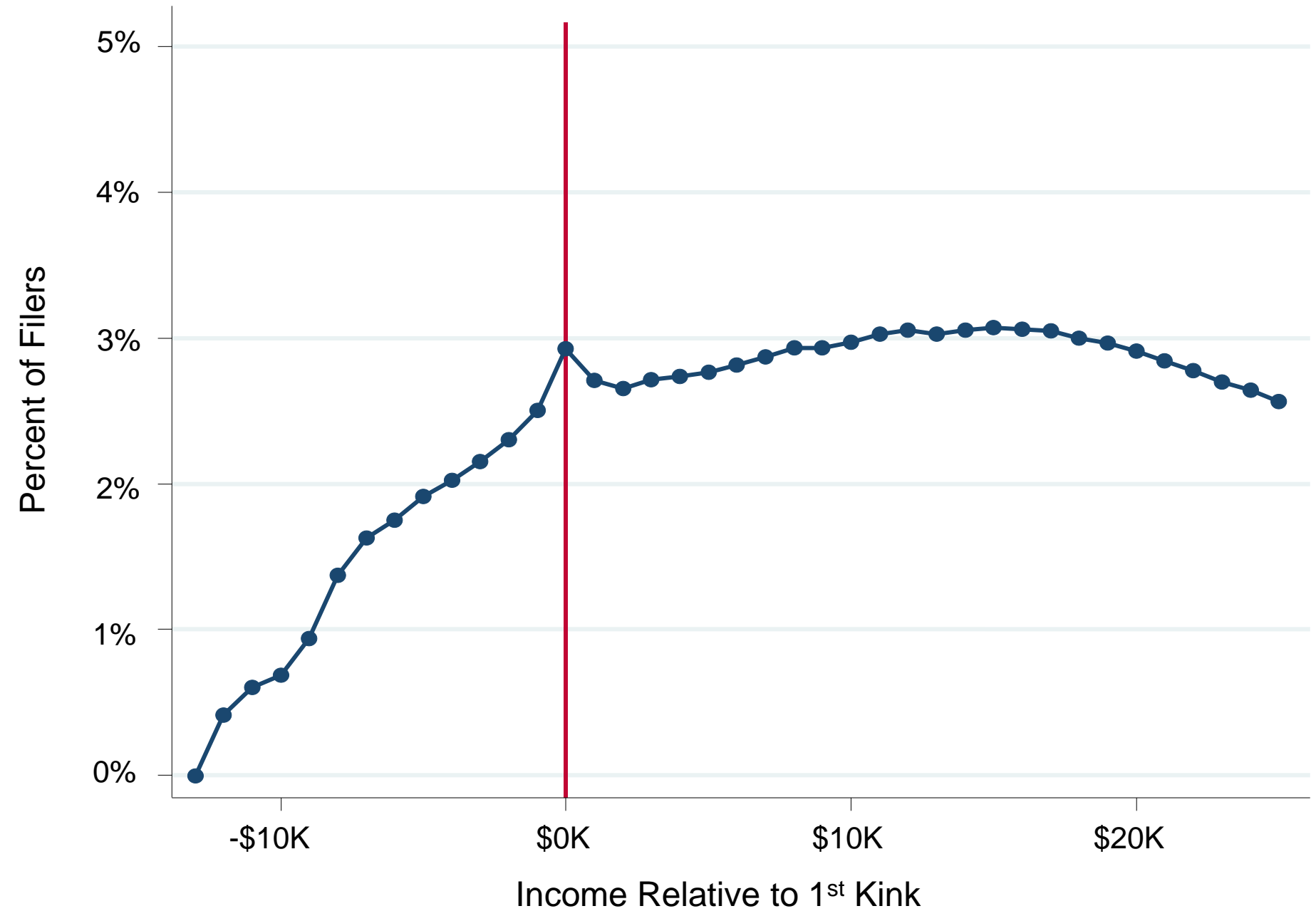
# Outline of Empirical Analysis

- Step 1: Document variation across neighborhoods in sharp bunching among self-employed

## Earnings Distribution in Texas



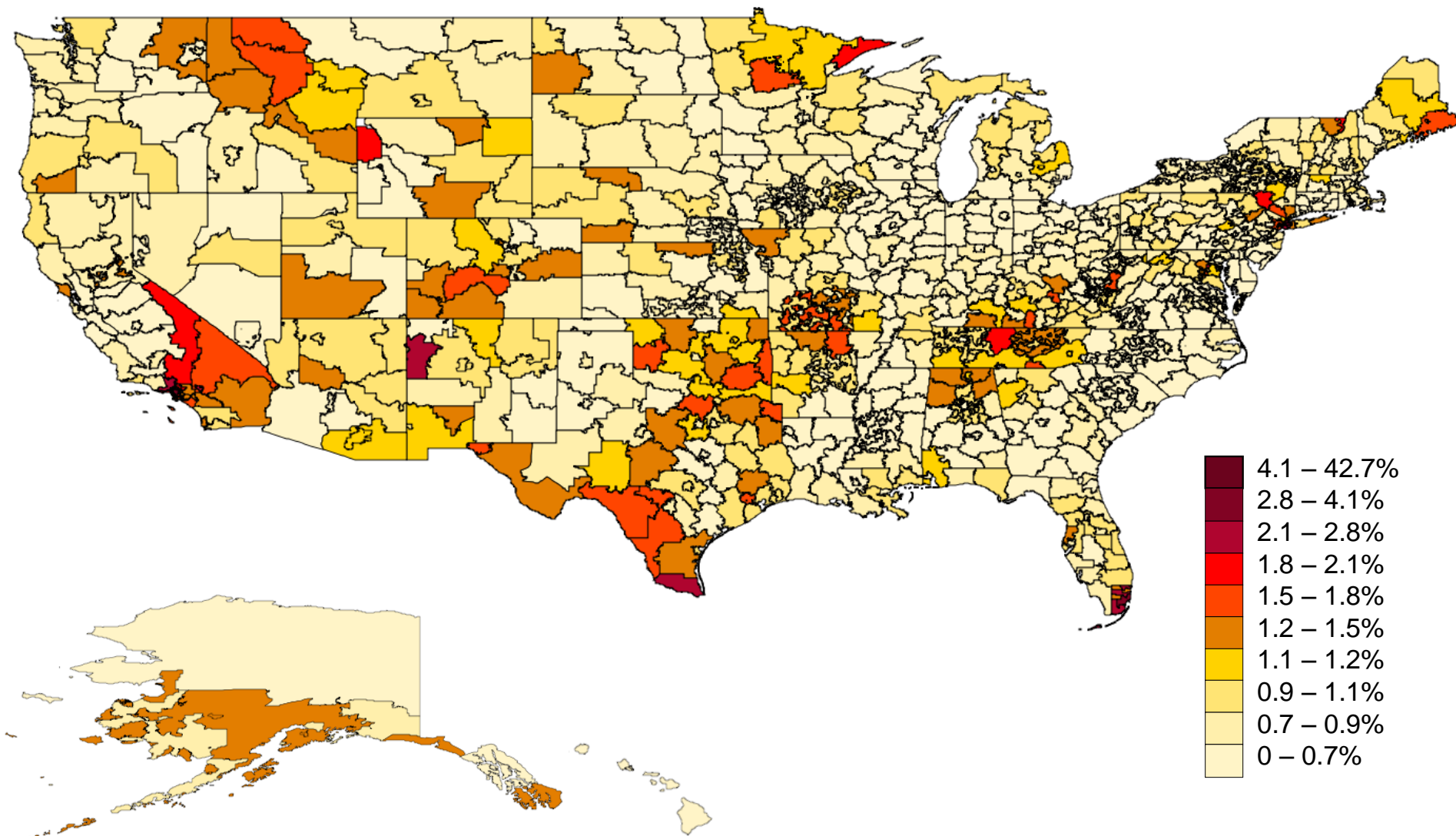
# Earnings Distribution in Kansas



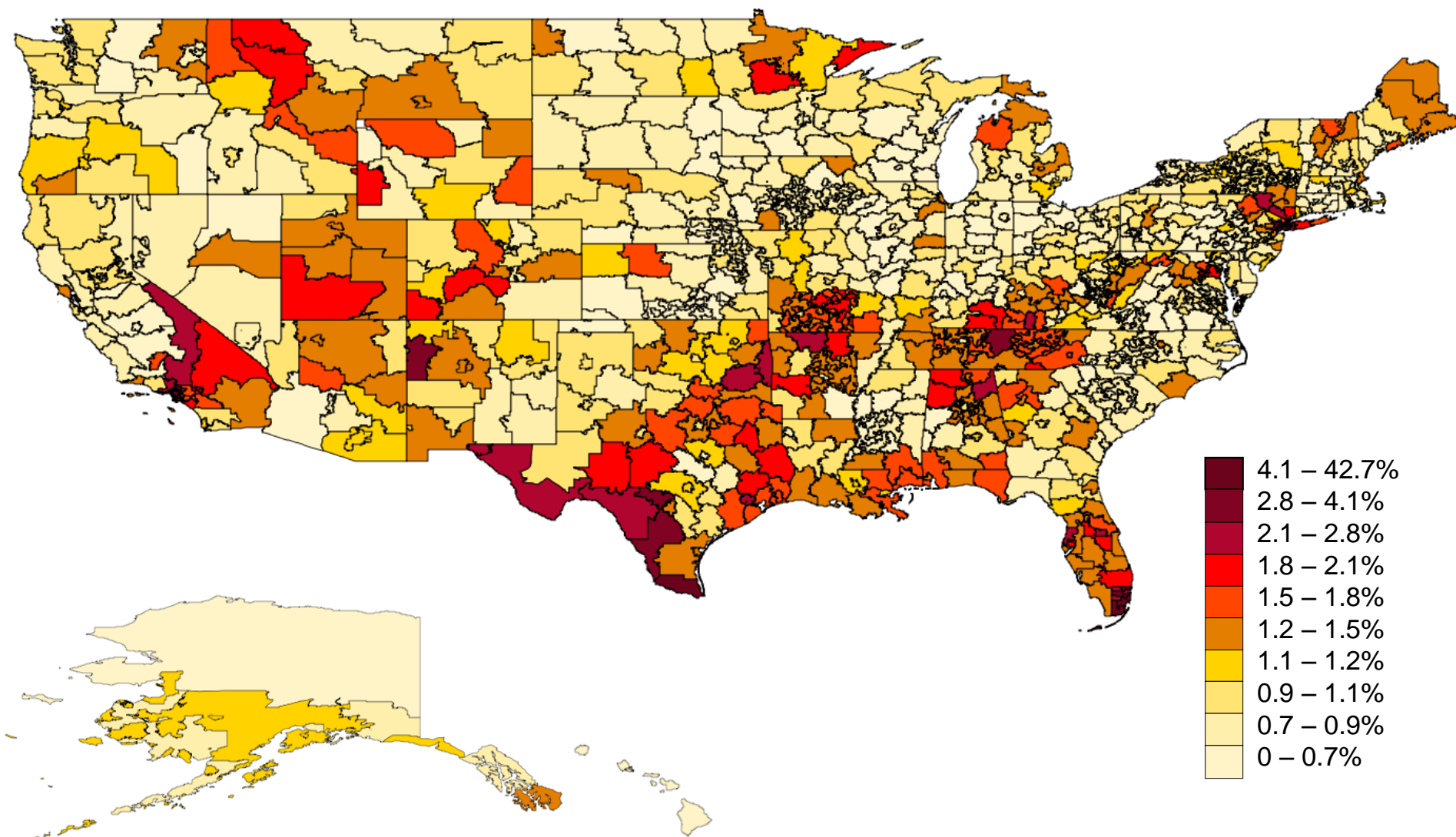
# Neighborhood-Level Measure of Bunching

- Define a measure of “sharp bunching” in each neighborhood
  - Fraction of EITC-eligible tax filers who report income at first kink and have self-employment income
  - Measures fraction of individuals who manipulate reported income to maximize EITC refund in each neighborhood
- Begin by documenting spatial evolution of sharp bunching across U.S.

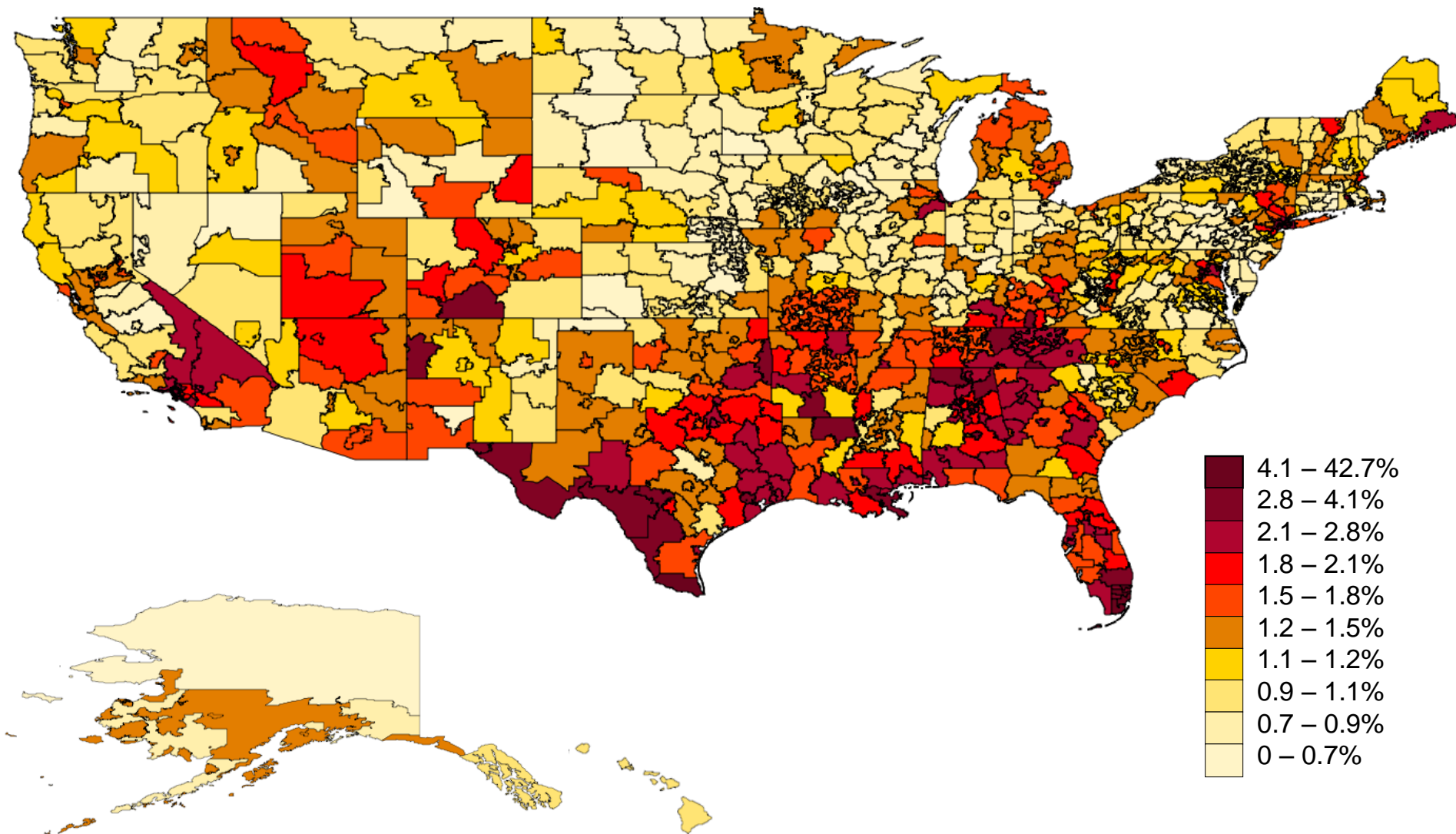
# Fraction of Tax Filers Who Report SE Income that Maximizes EITC Refund in 1996



# Fraction of Tax Filers Who Report SE Income that Maximizes EITC Refund in 1999

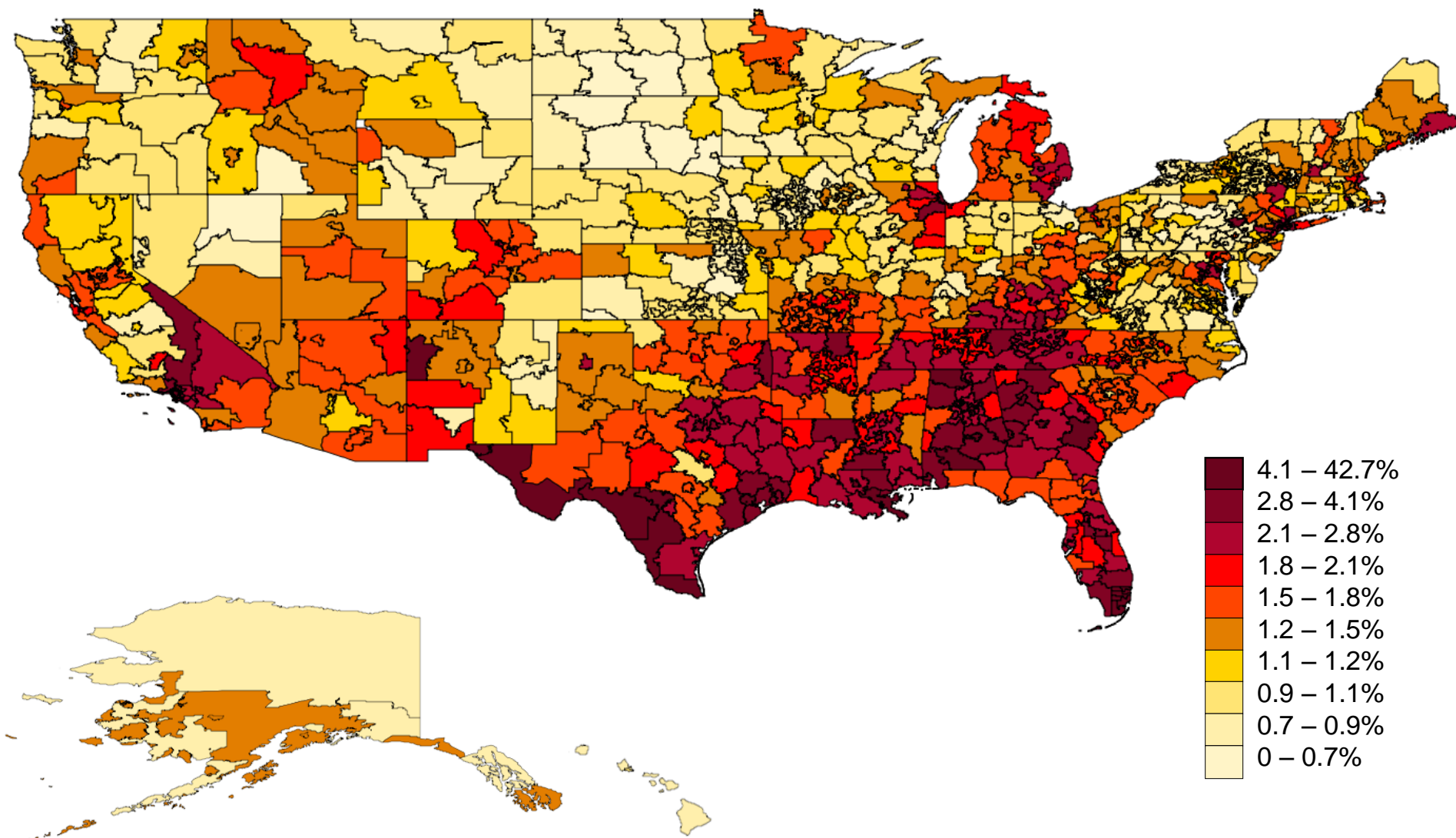


# Fraction of Tax Filers Who Report SE Income that Maximizes EITC Refund in 2002



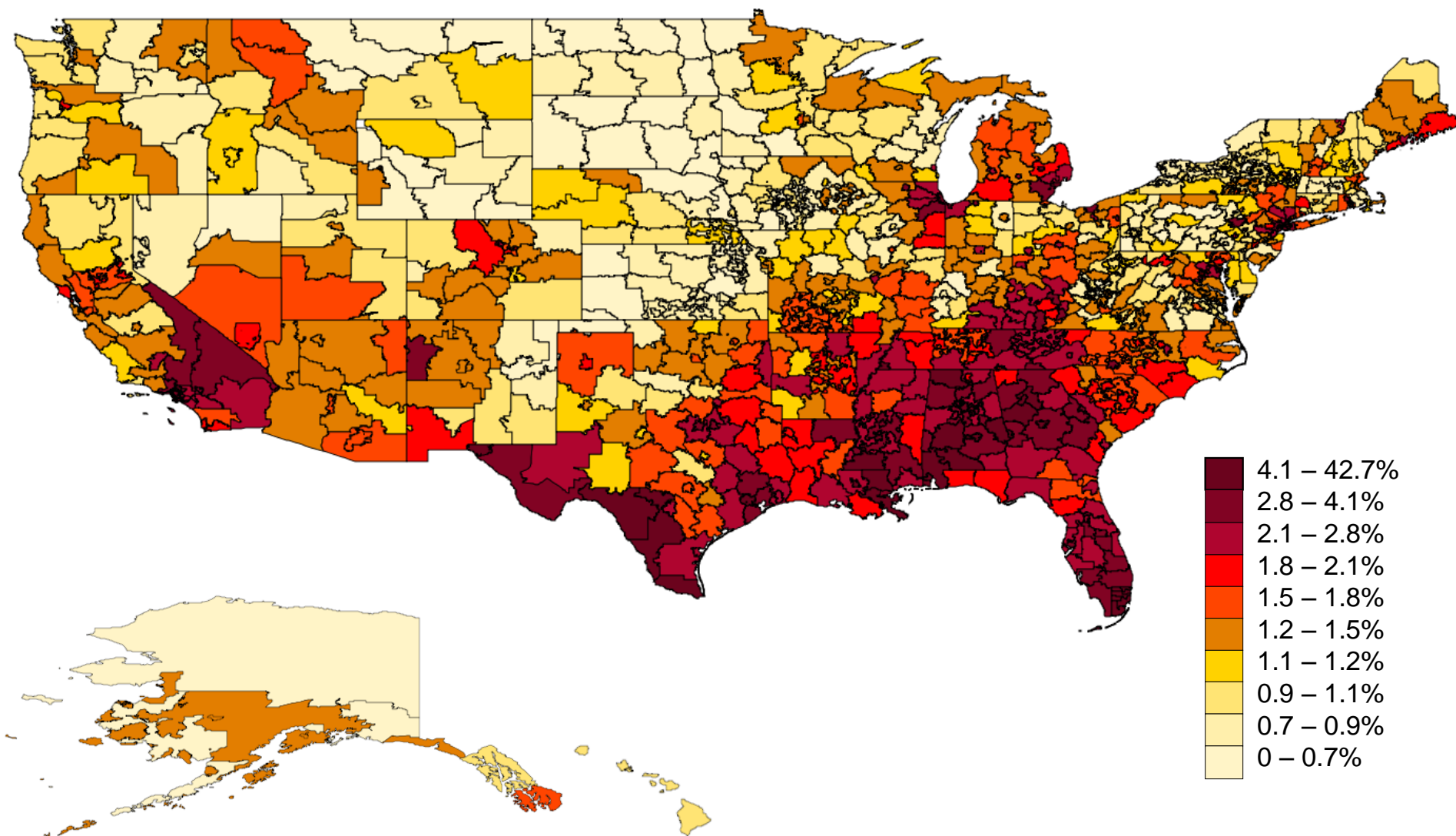


# Fraction of Tax Filers Who Report SE Income that Maximizes EITC Refund in 2005

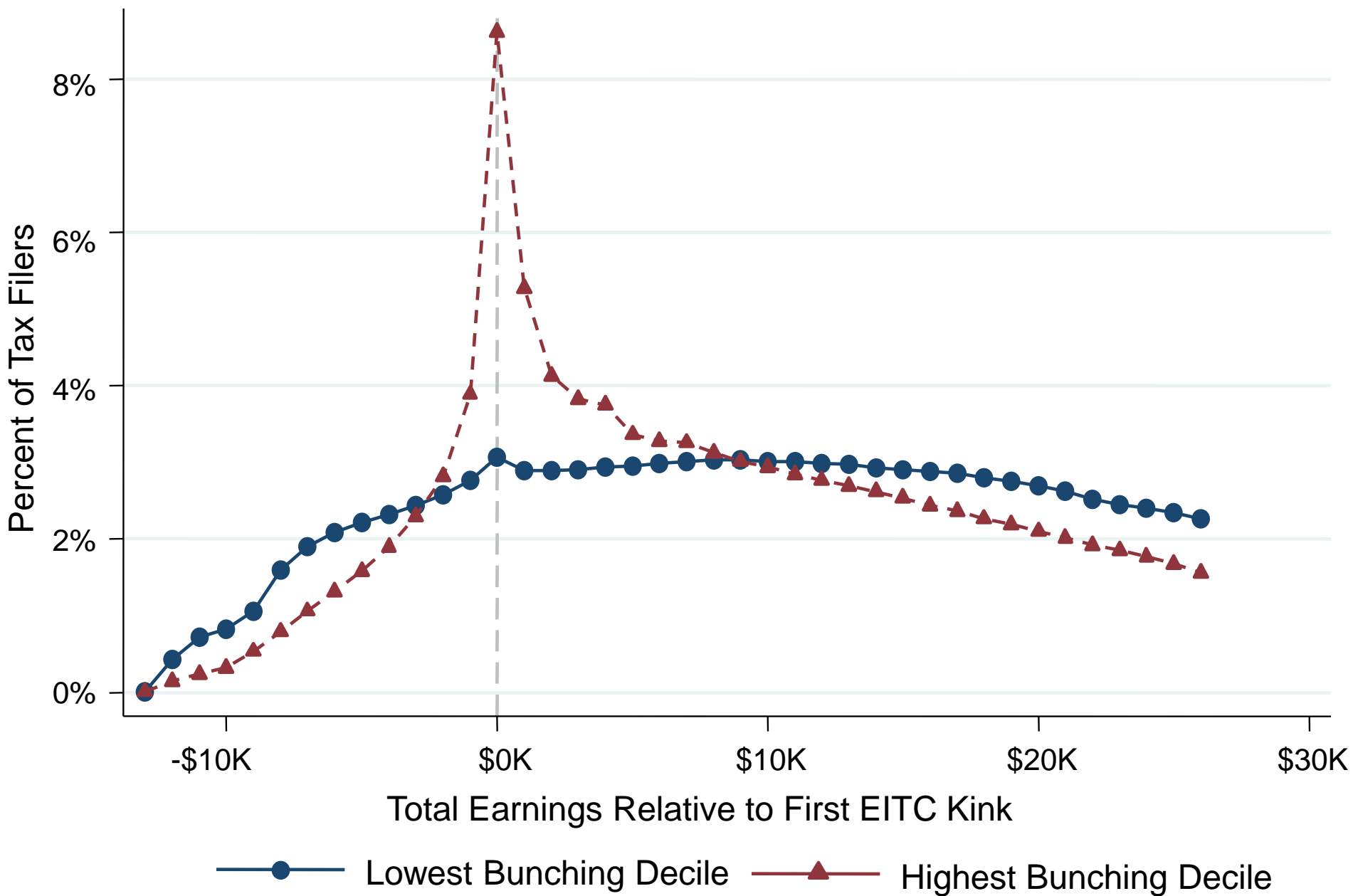




# Fraction of Tax Filers Who Report SE Income that Maximizes EITC Refund in 2008



# Earnings Distributions in Lowest and Highest Bunching Deciles



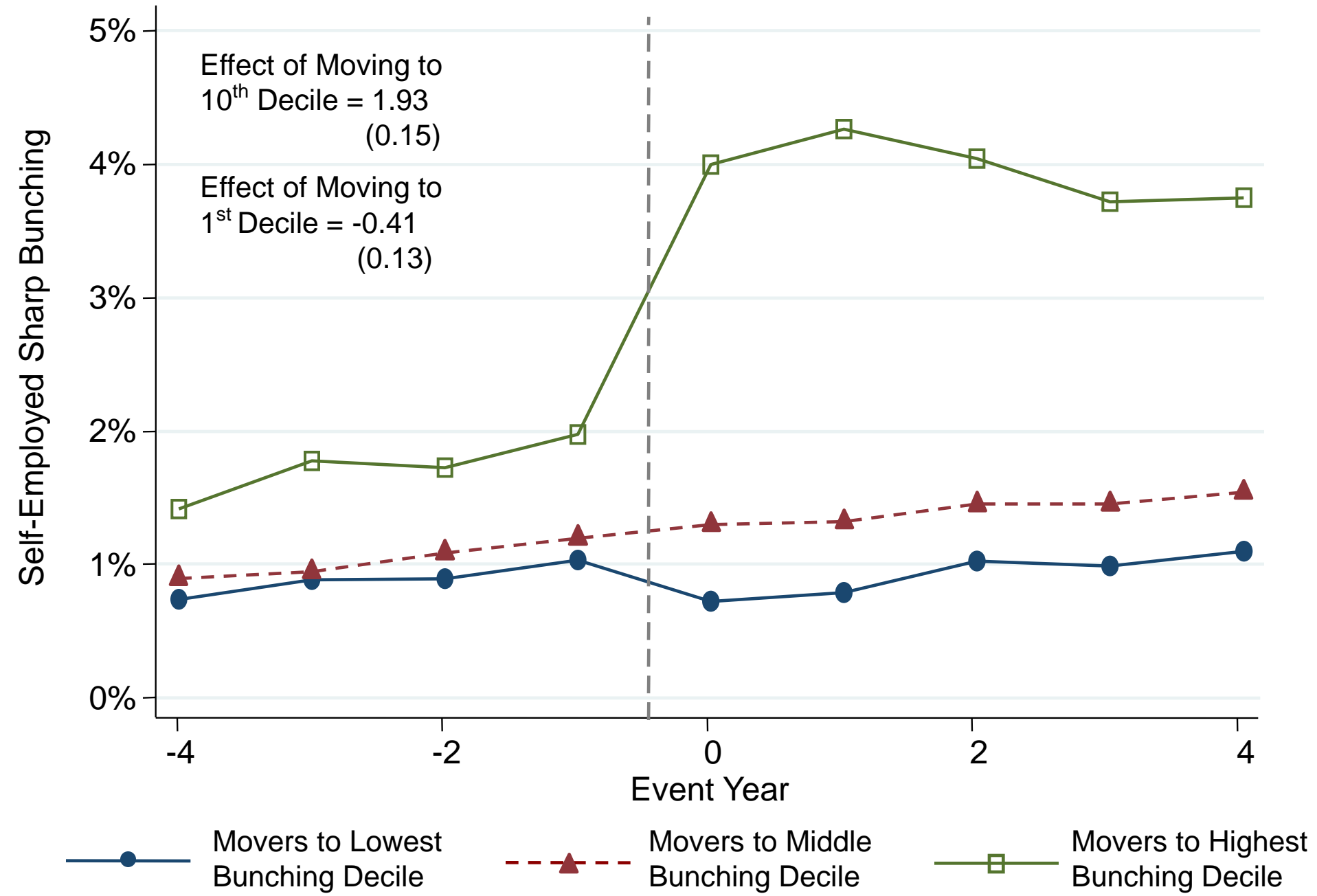
# Outline of Empirical Analysis

- Step 1: Document variation across neighborhoods in sharp bunching among self-employed
- Step 2: Establish that variation in sharp bunching across neighborhoods is driven by differences in knowledge about EITC schedule

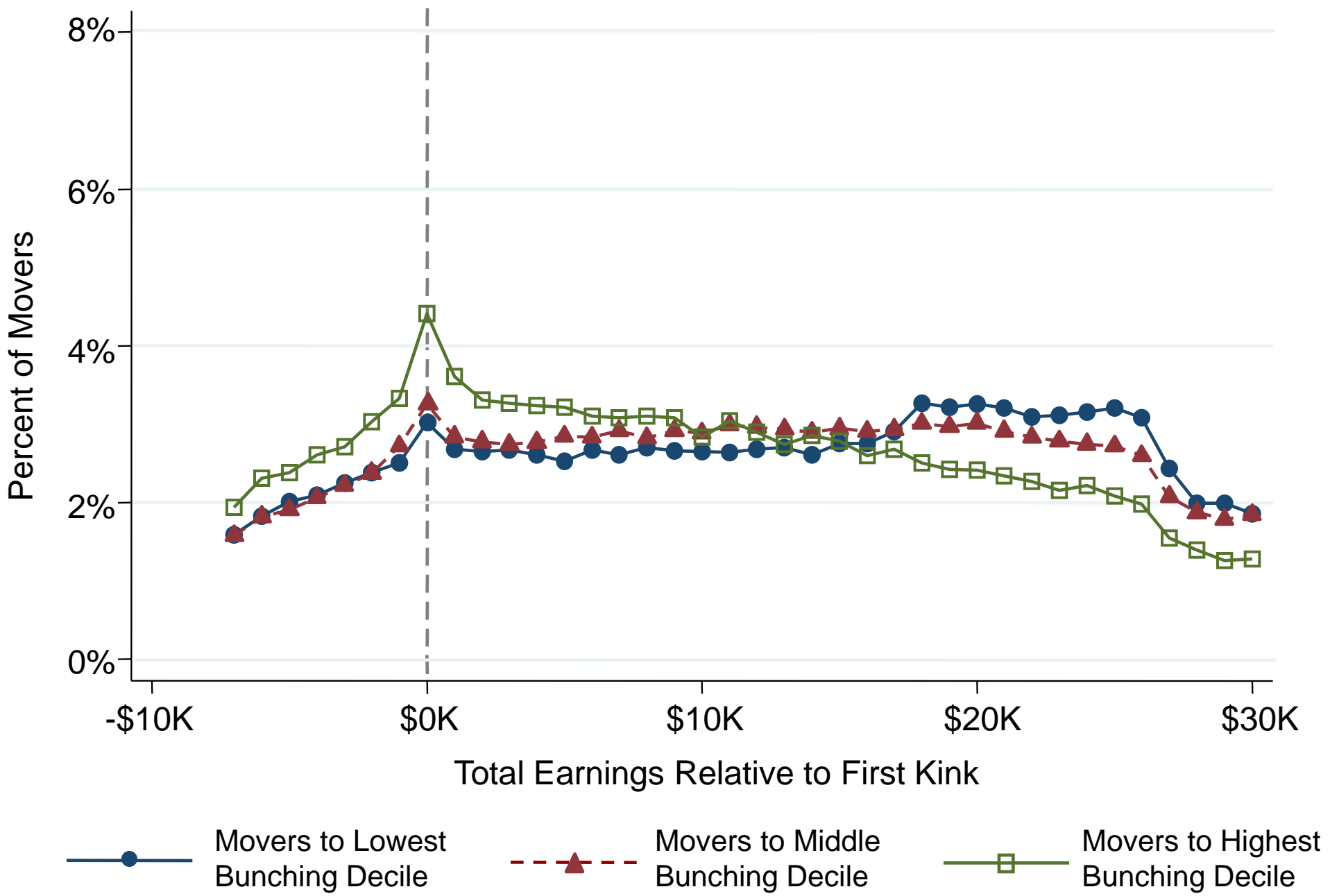
# Movers: Neighborhood Changes

- Consider individuals who move across neighborhoods to isolate causal impacts of neighborhoods on elasticities
  - 54 million observations in panel data on cross-zip movers
- Define “neighborhood sharp bunching” as degree of bunching for *stayers*
- Analyze how changes in neighborhood sharp bunching affect movers’ behavior

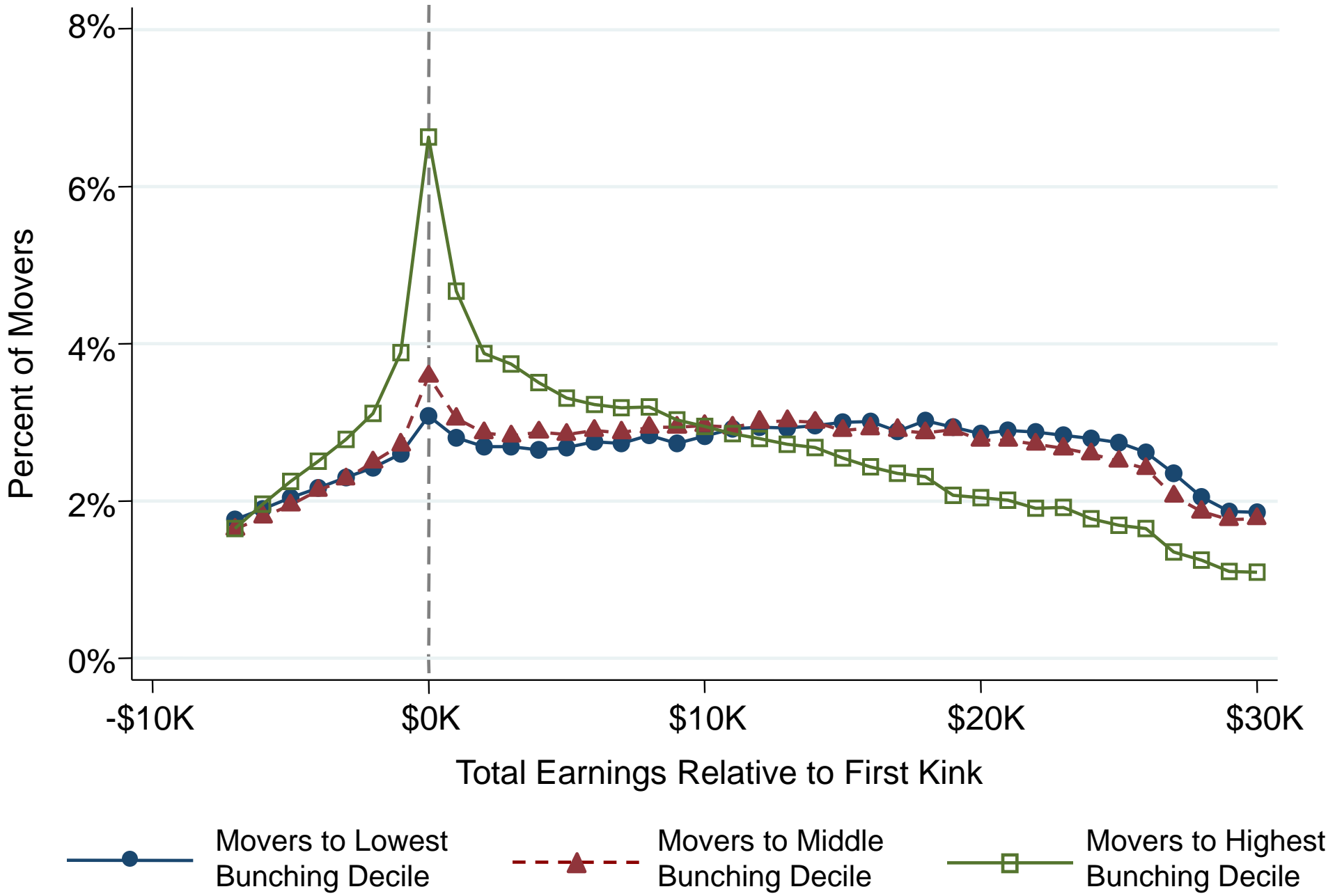
# Event Study of Sharp Bunching Around Moves



# Total Earnings Distribution in Years Before Move



# Total Earnings Distribution in Years After Move

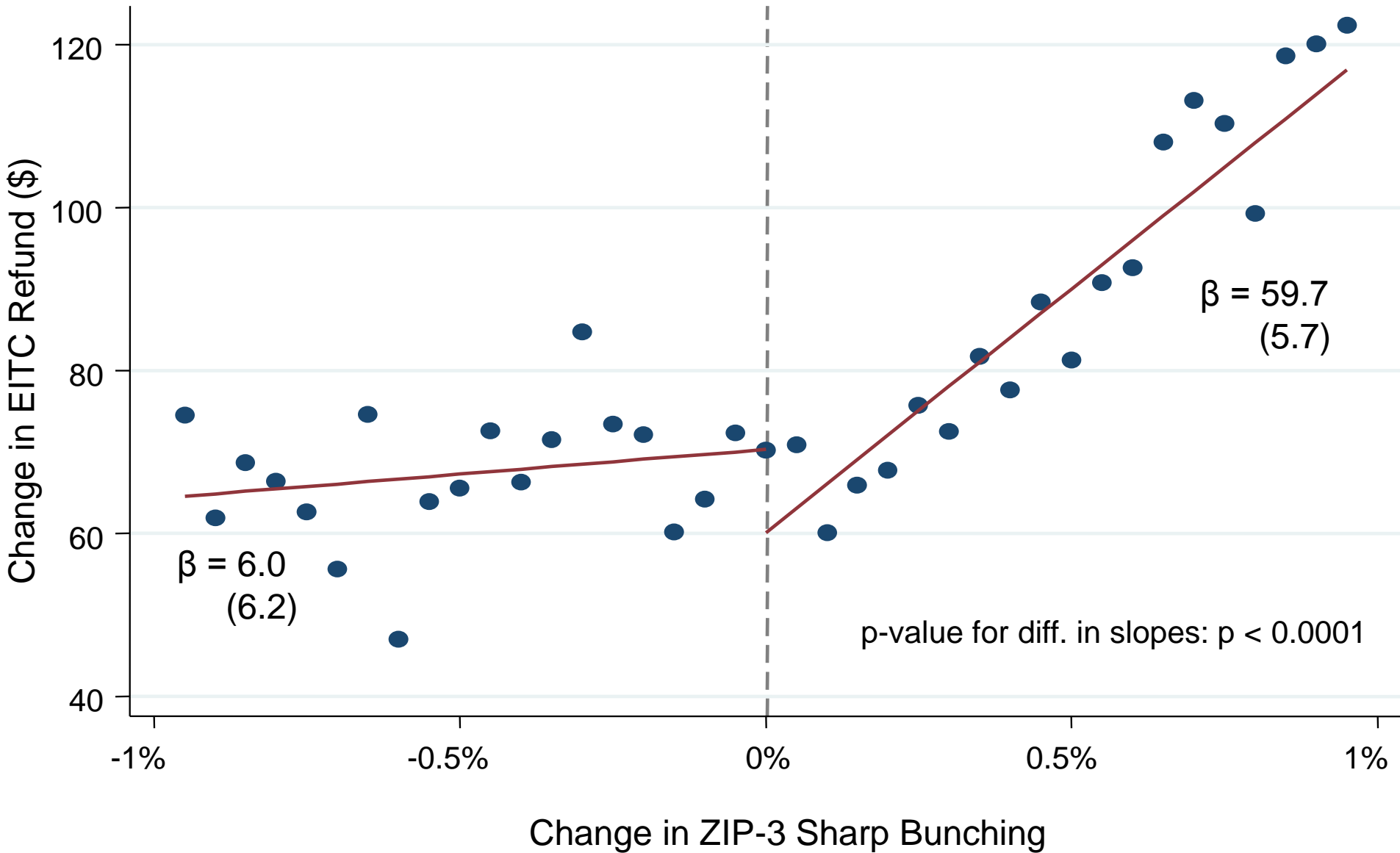


# Learning and Memory

- Knowledge model predicts asymmetric impact of moving:
  - Moving to a higher-bunching neighborhood should raise EITC refund
  - Moving to a lower-bunching should not affect EITC refund



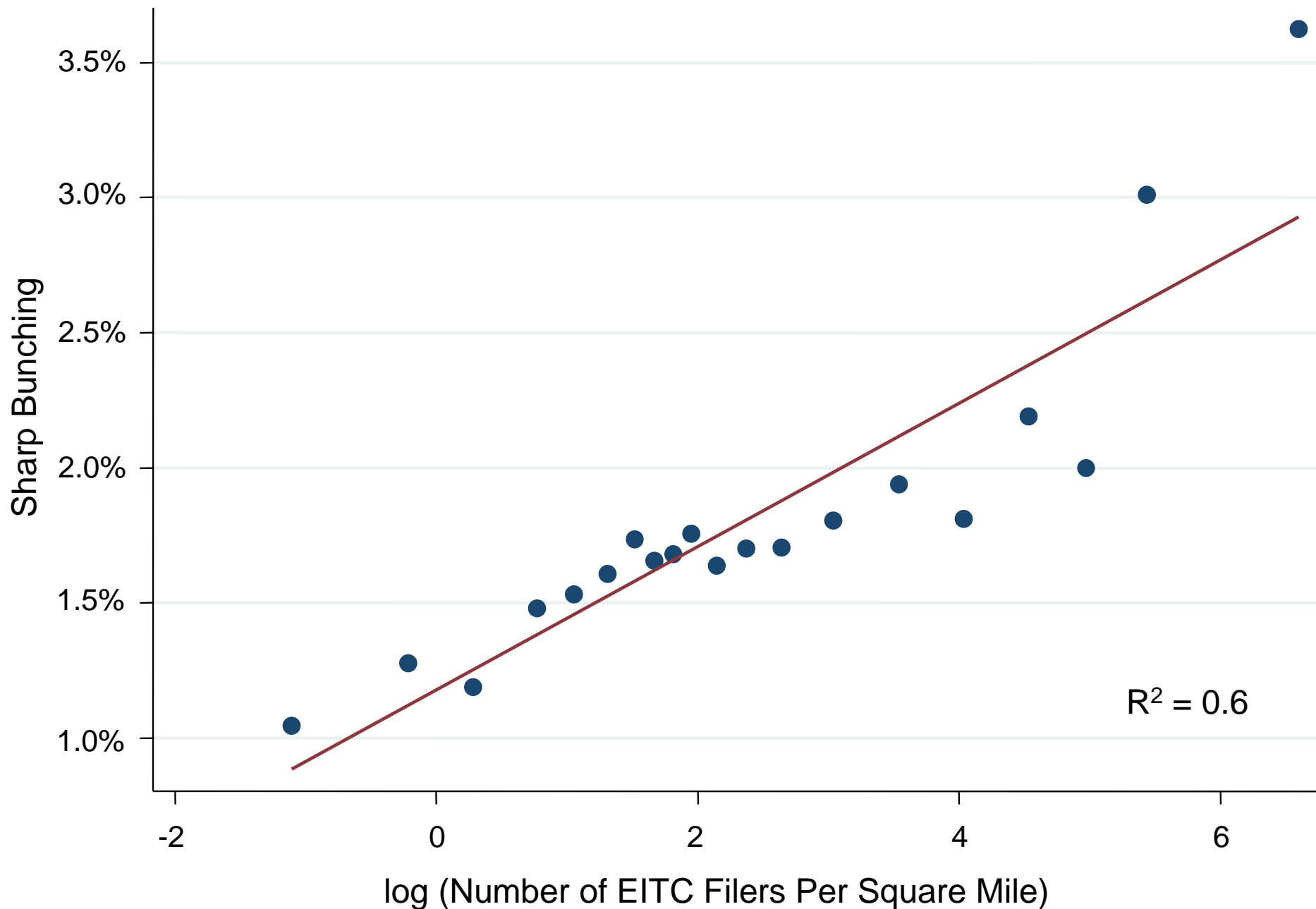
# Change in EITC Refunds vs. Change in Sharp Bunching for Movers



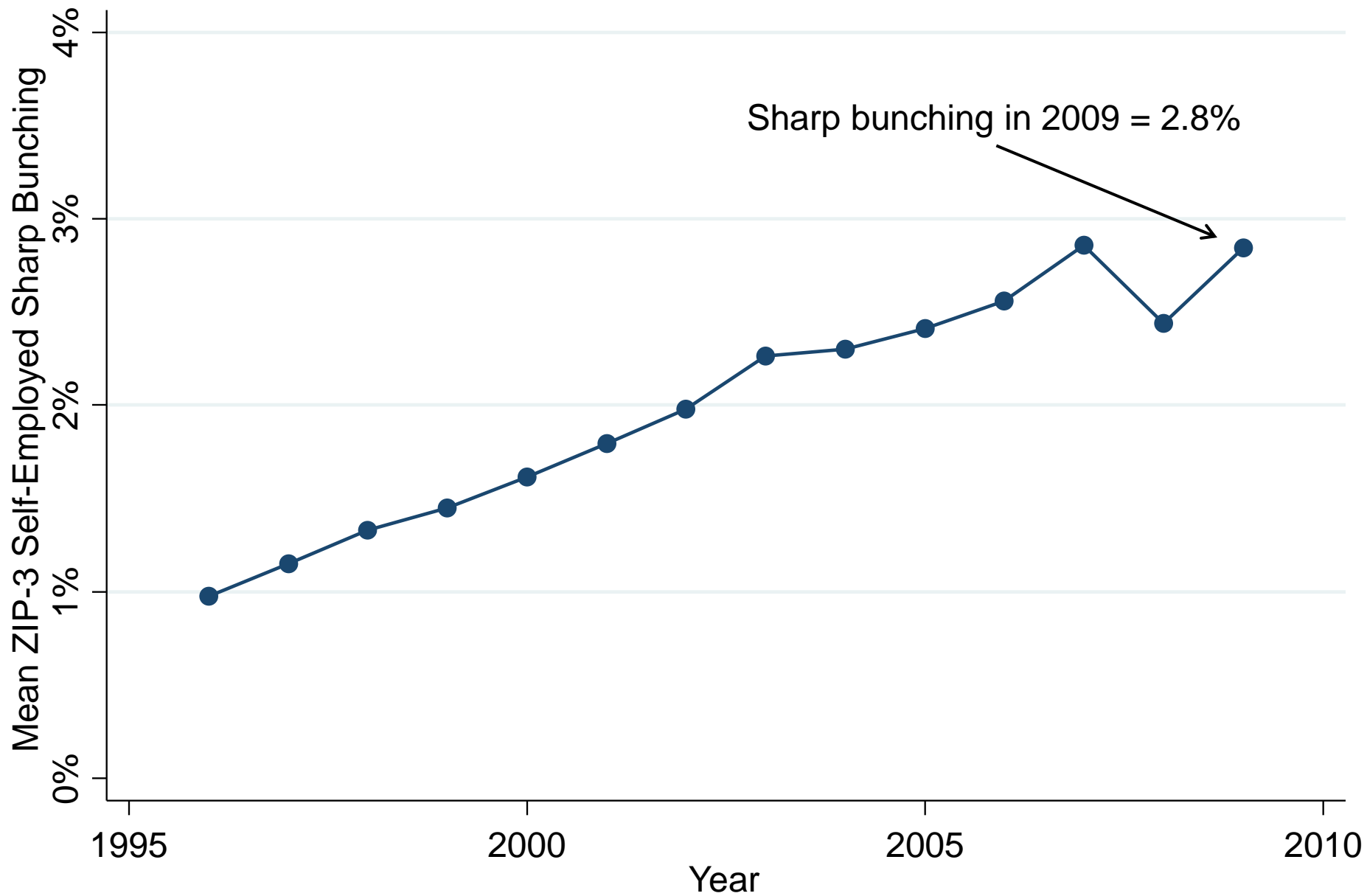
# Cross-Sectional Correlations

- What drives the variation in sharp bunching across neighborhoods?
  - Evaluate predictive power of proxies for information, tax compliance, and other variables

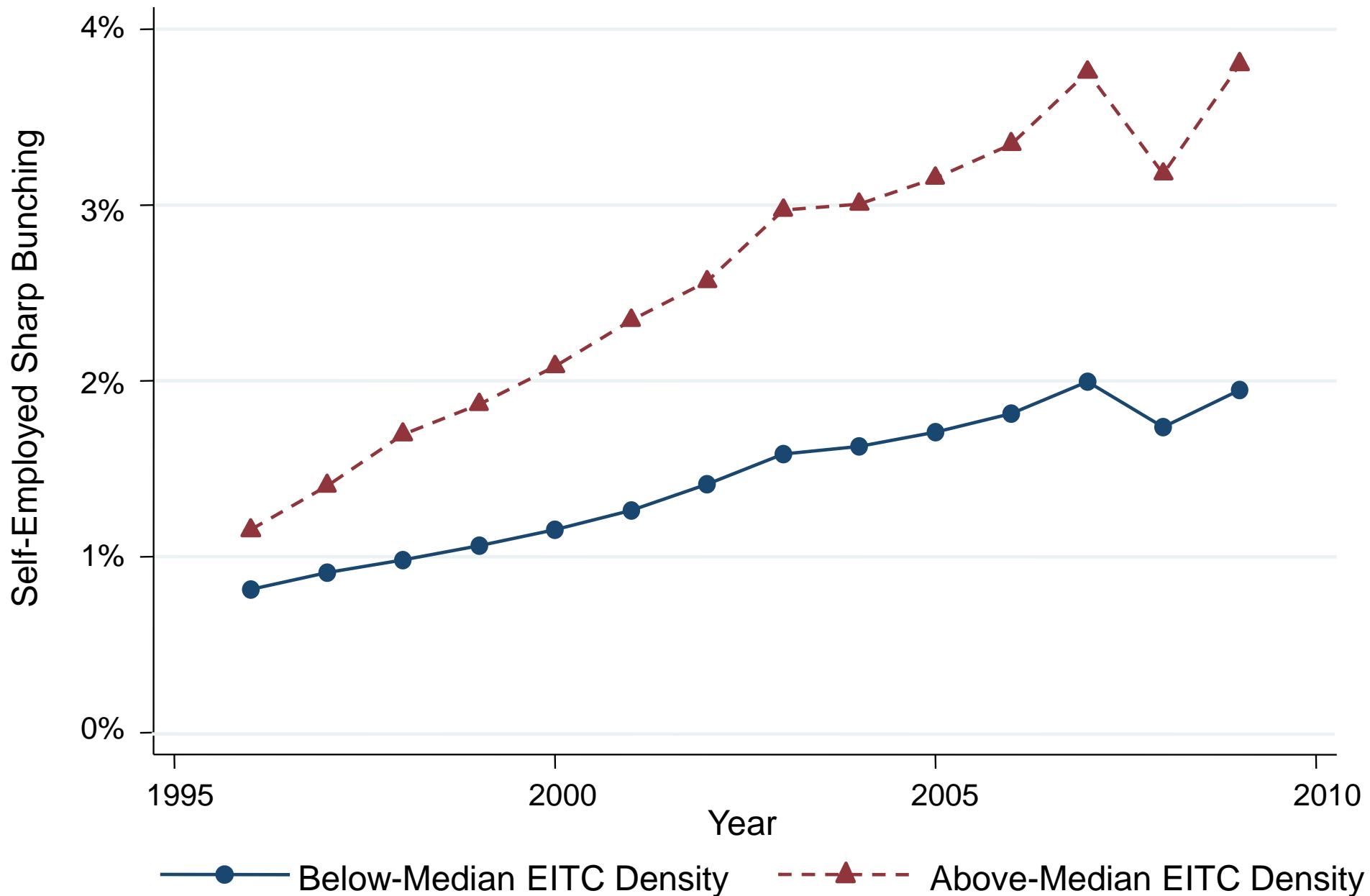
## Agglomeration: Sharp Bunching vs. EITC Filer Density by ZIP Code



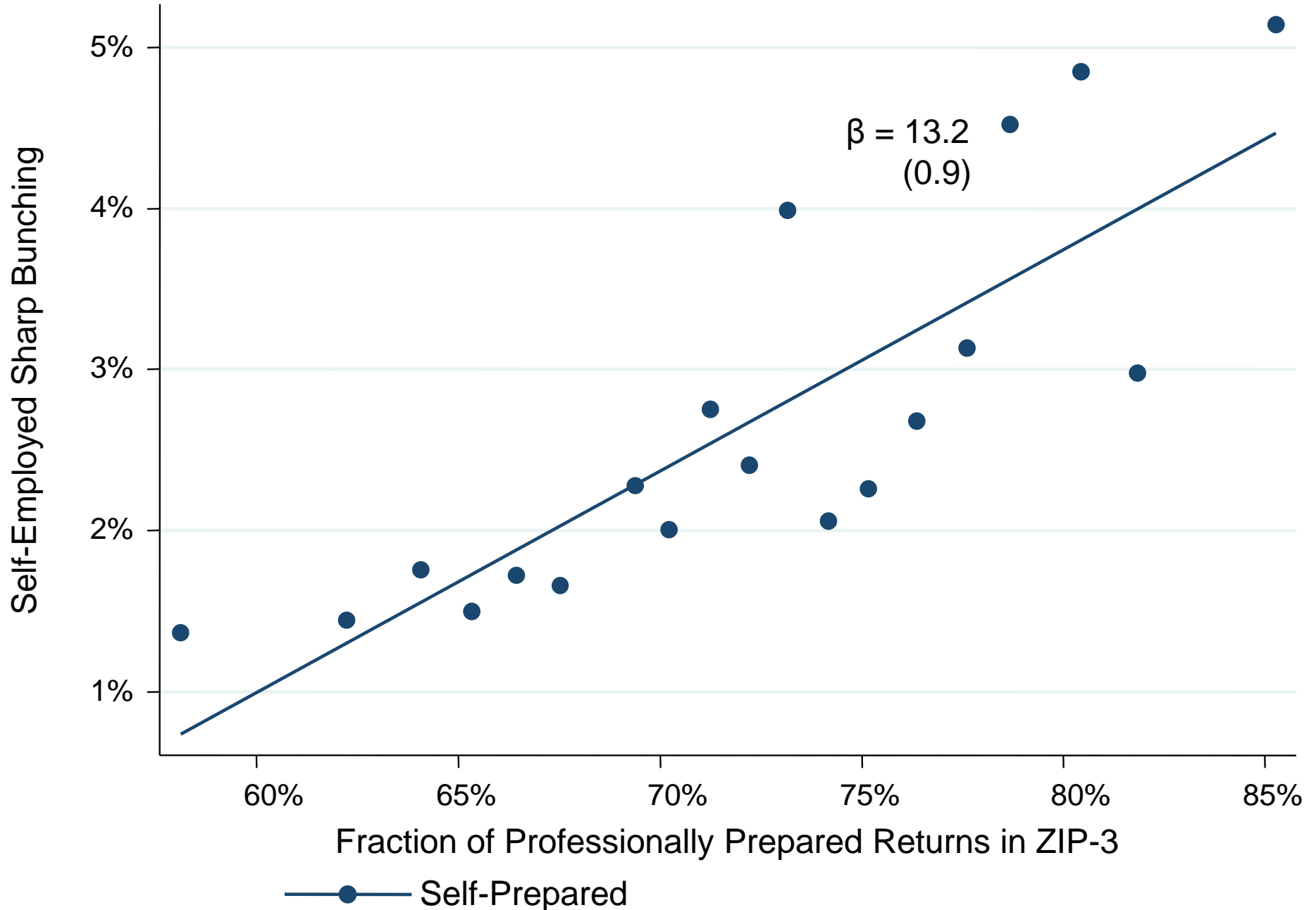
## Self-Employed Sharp Bunching Over Time



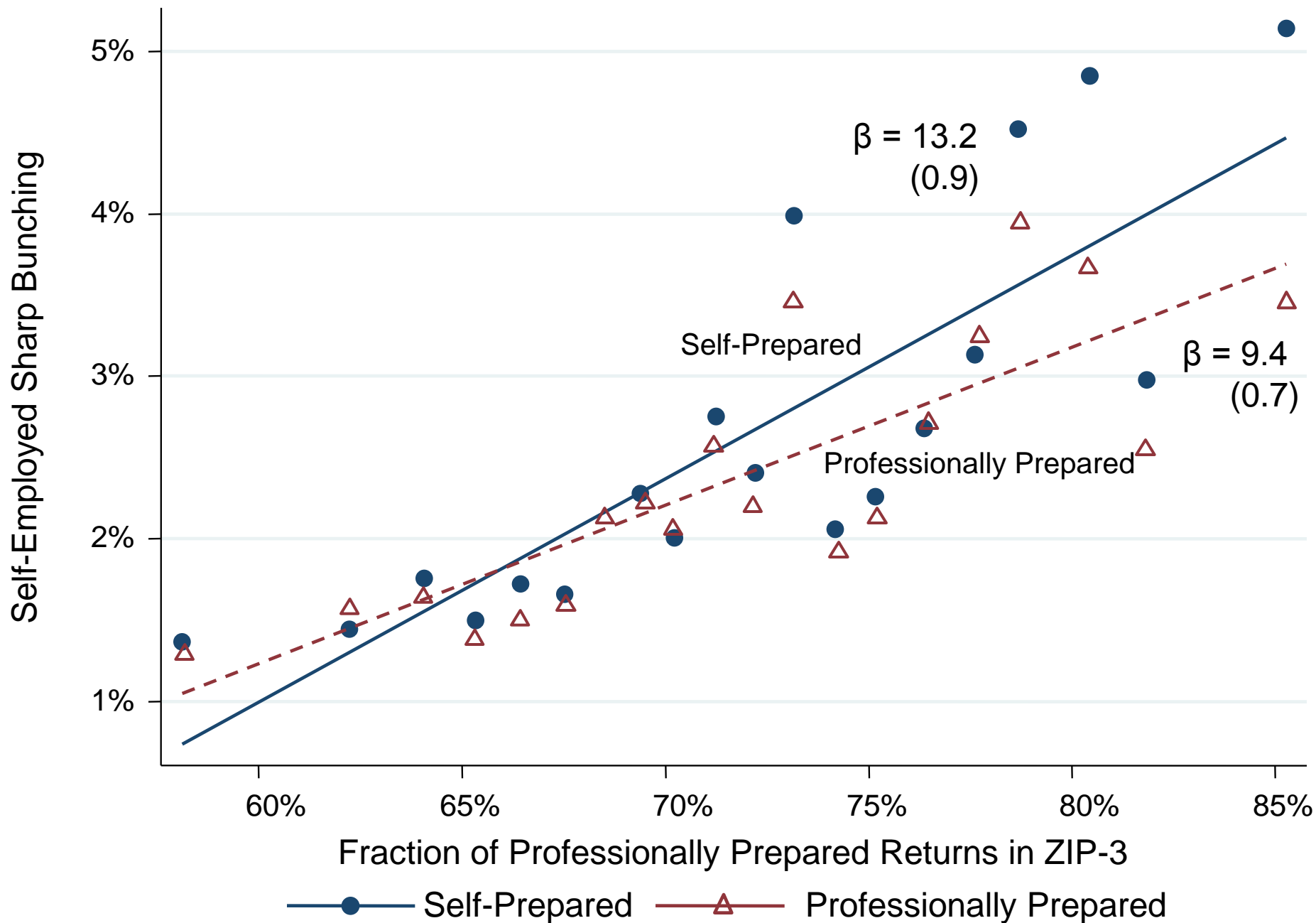
## Evolution of Sharp Bunching in Low vs. High EITC-Density Areas



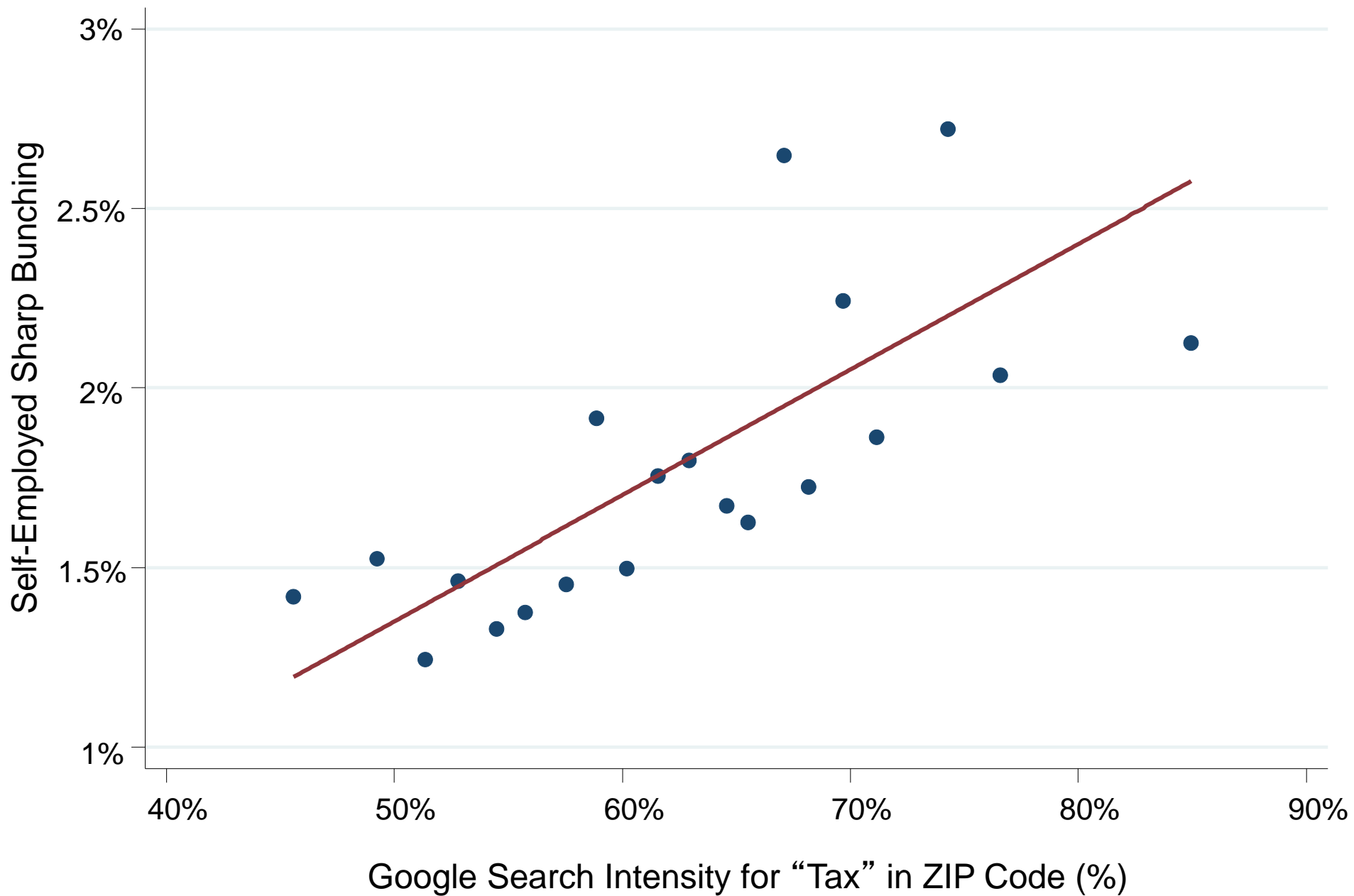
## Sharp Bunching vs. Fraction of Professionally Prepared Returns in ZIP-3



# Sharp Bunching vs. Fraction of Professionally Prepared Returns in ZIP-3



## Correlation Between EITC Bunching and Google Search Patterns





## Cross-Sectional Correlates of Sharp Bunching

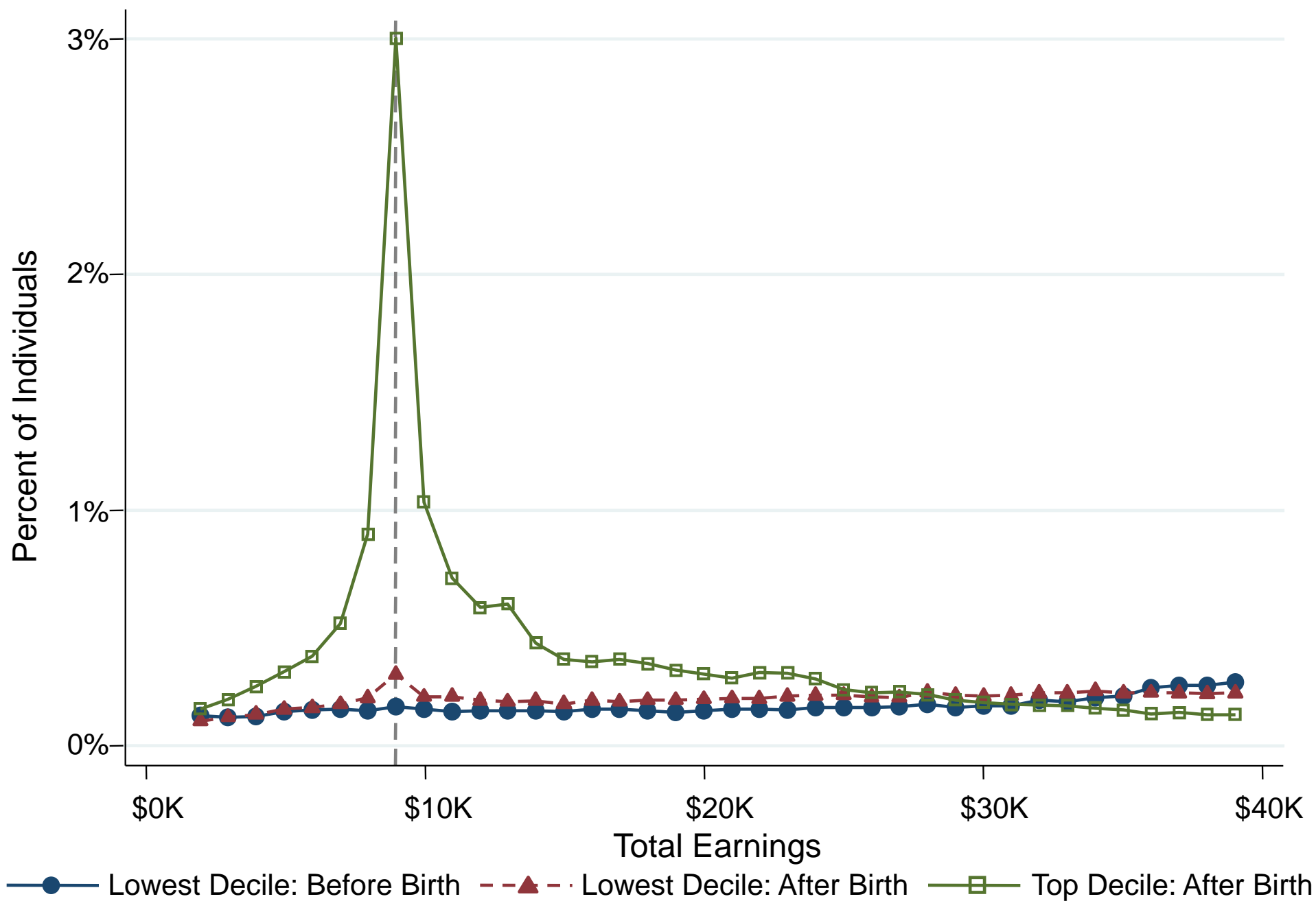
Dep. Var.: Sharp Bunching Rate in ZIP-3 (%)

EITC Filer Density in ZIP-3	1.93 (0.05)	1.82 (0.05)			0.44 (0.06)	0.69 (0.06)		
Fraction of Tax Prepared Returns in ZIP-3			9.86 (1.48)		3.02 (0.51)	3.46 (0.56)		
Google Search Intensity			0.30 (0.05)		0.14 (0.03)	0.19 (0.03)		
State EITC							0.07 (0.05)	
State Non-Compliance Rate								-1.51 (5.32)
Demographic Controls		x			x	x		
State Fixed Effects						x		
Year	2000	2000	2008	2008	2008	2008	2000	2000
R-squared	0.603	0.798	0.169	0.032	0.728	0.848	0.105	0.002
Number of Observations	873	873	883	875	870	870	886	51

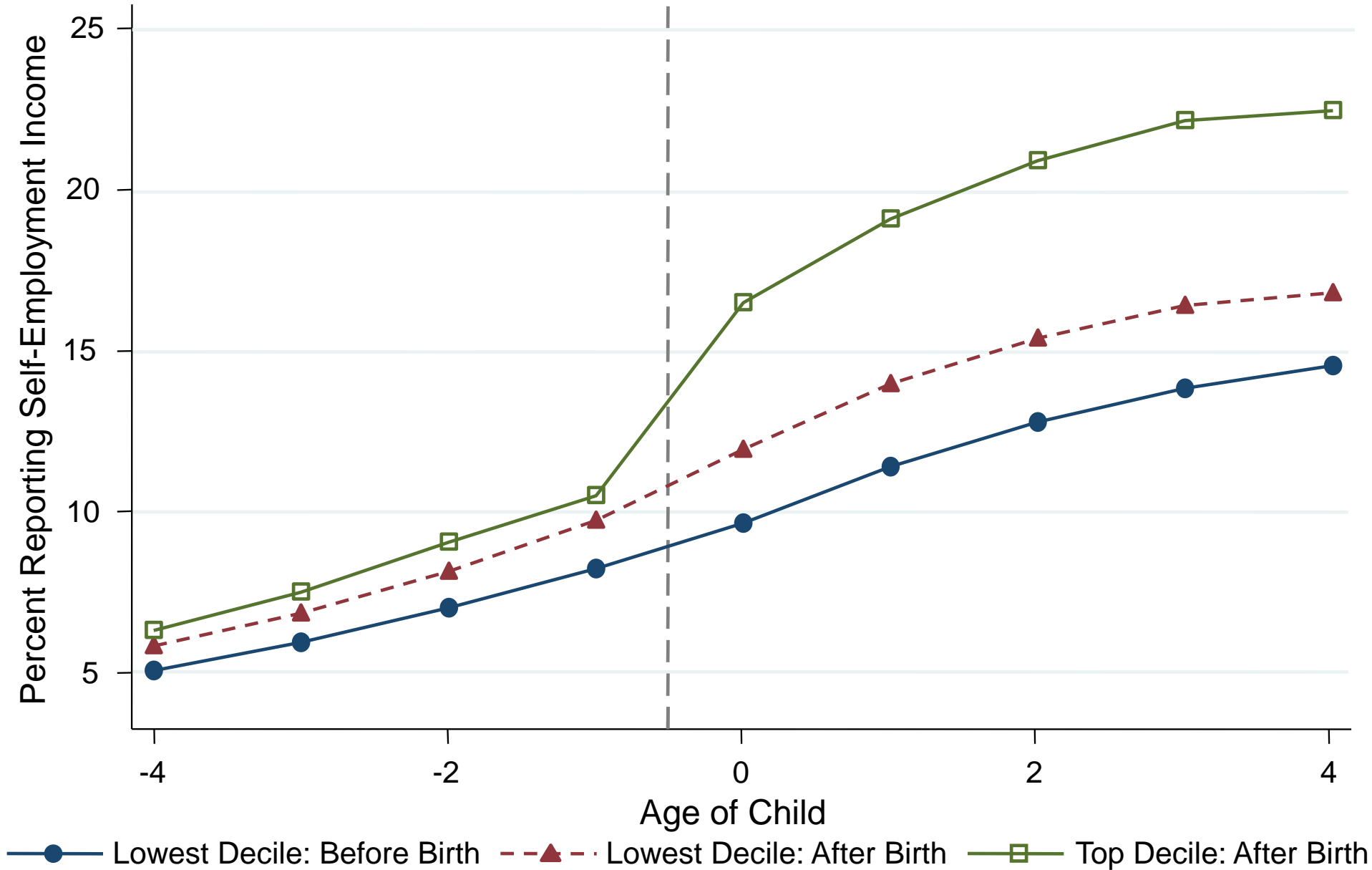
# Perceptions of EITC in Low-Bunching Areas

- Preceding evidence indicates that self-emp. sharp bunching provides a proxy for local knowledge about **first kink** of EITC schedule
- Assumption 1 requires that individuals in low-bunching areas have no knowledge about *entire* EITC schedule and behave as if  $\tau = 0$
- Now assess beliefs about broader EITC schedule in low-bunching areas
  - Analyze reported incomes of self-employed around birth of first child
  - Birth of first child → substantial change in EITC incentives

# Effect of Child Birth on Total Earnings Distribution for the Self-Employed



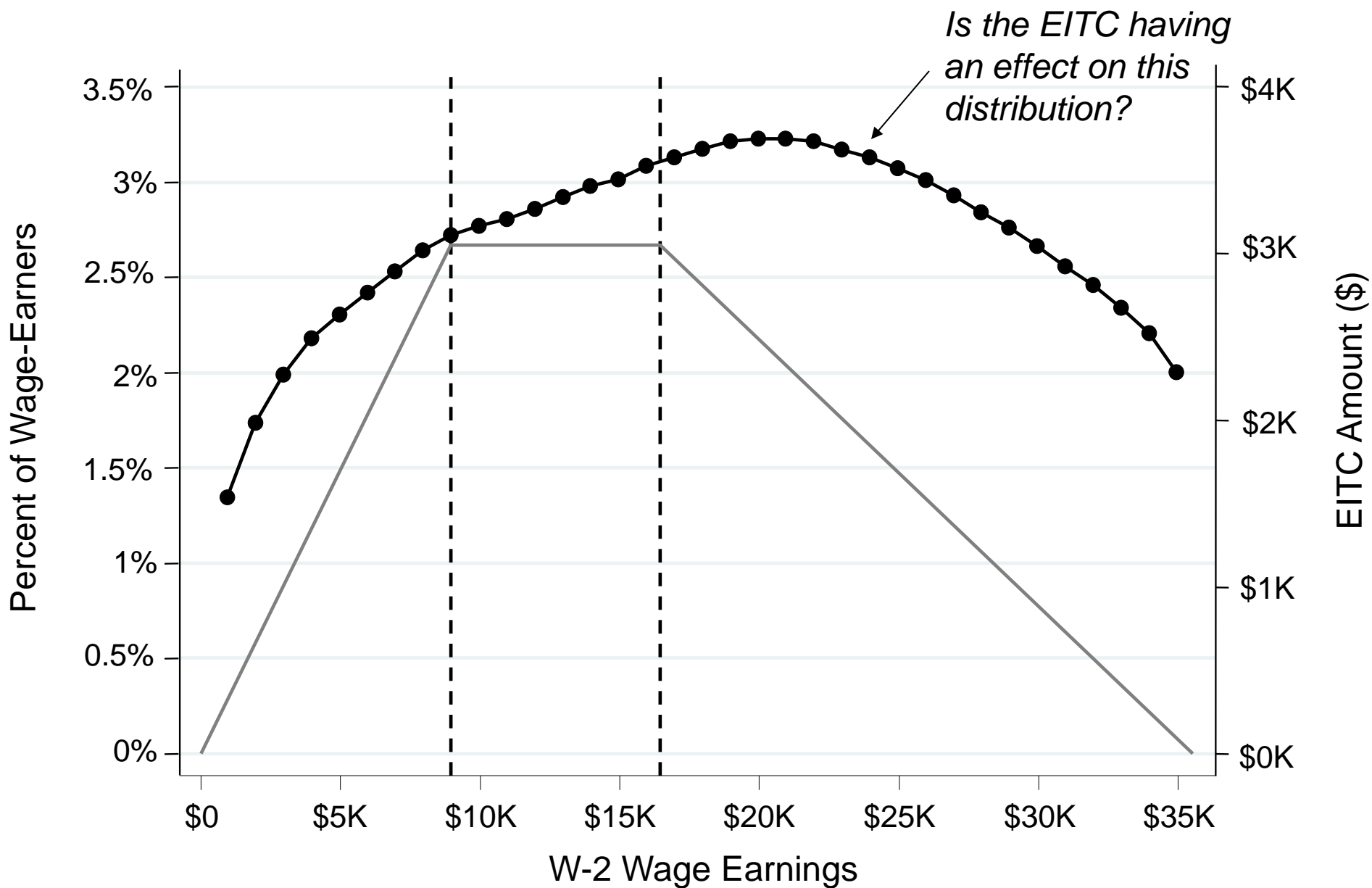
# Fraction of Individuals Reporting Self-Employment Income Around Child Birth



# Outline of Empirical Analysis

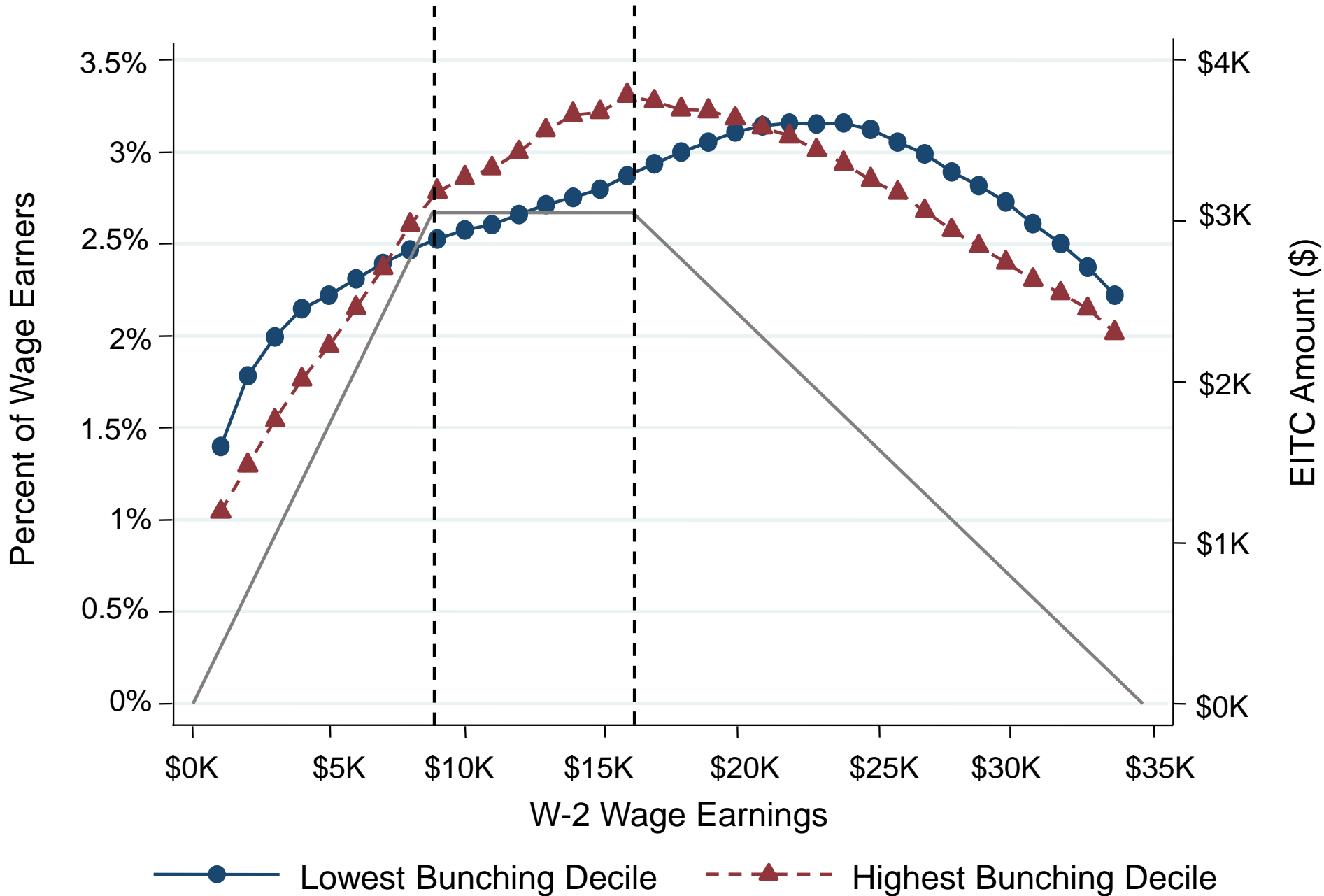
- Step 1: Document variation across neighborhoods in sharp bunching among self-employed
- Step 2: Establish that variation in sharp bunching across neighborhoods is driven by differences in knowledge about EITC schedule
- Step 3: Compare wage earnings distributions across low- and high-knowledge neighborhoods to uncover impacts of EITC on earnings

# Income Distribution For Single Wage Earners with One Child

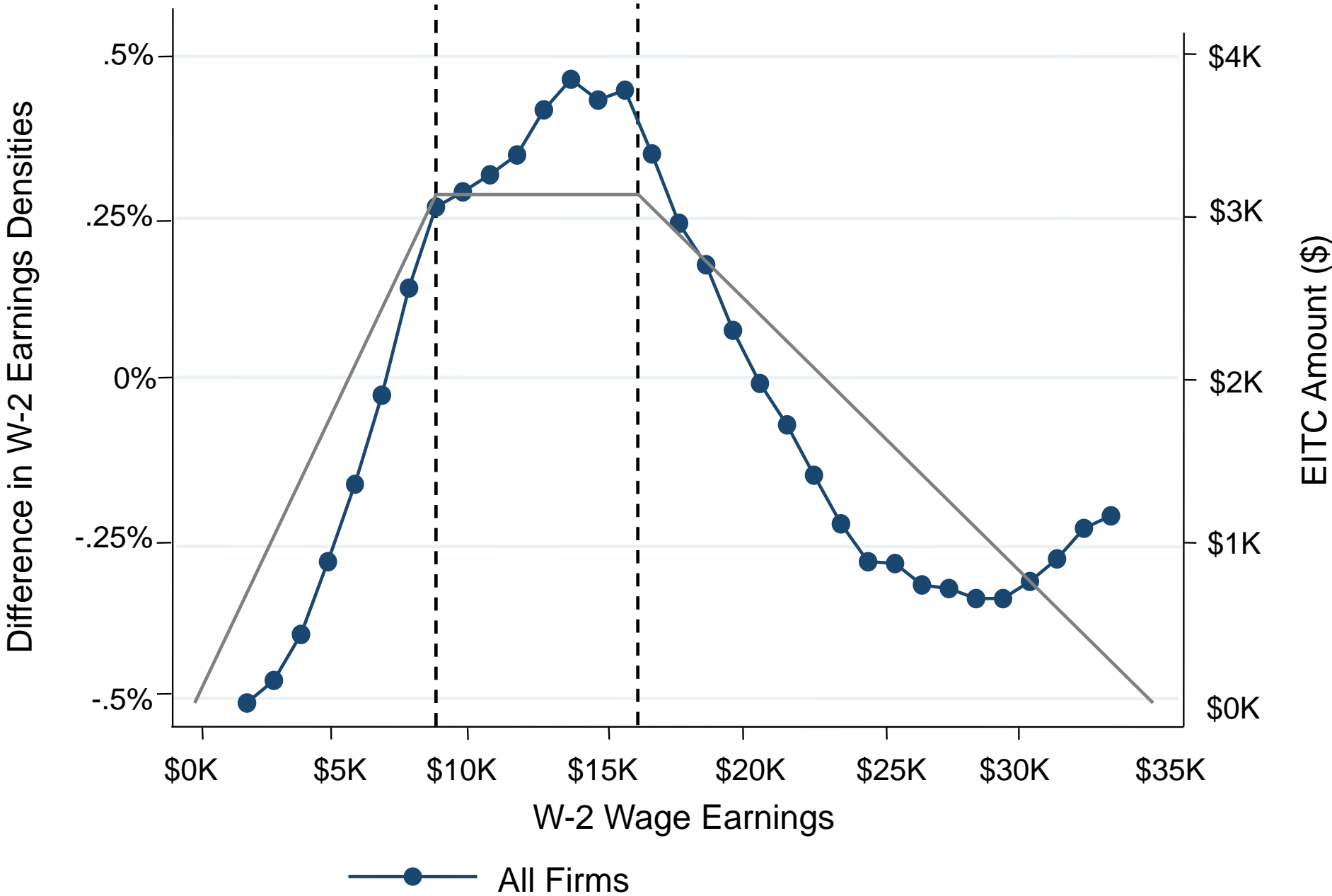


# Income Distribution For Single Wage Earners with One Child

## High vs. Low Bunching Areas



Difference in Wage Earnings Distributions Between Top and Bunching Decile  
Wage Earners with One Child

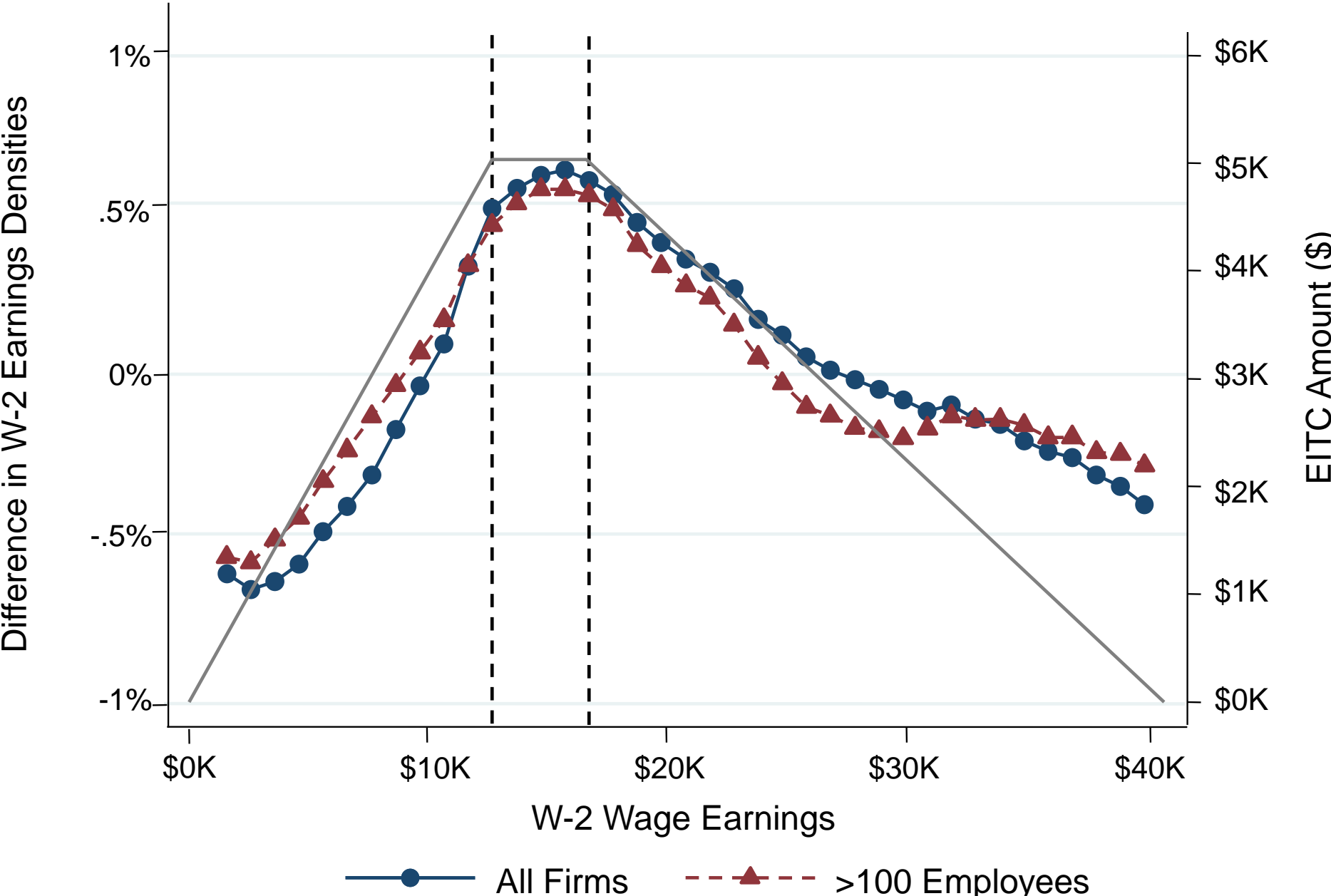




# Difference in Wage Earnings Distributions Between Top and Bunching Decile Wage Earners with One Child



# Difference in Wage Earnings Distribution Between Top and Bunching Decile Wage Earners with Two Children



## EITC Credit Amount for Wage Earners vs. Sharp Bunching



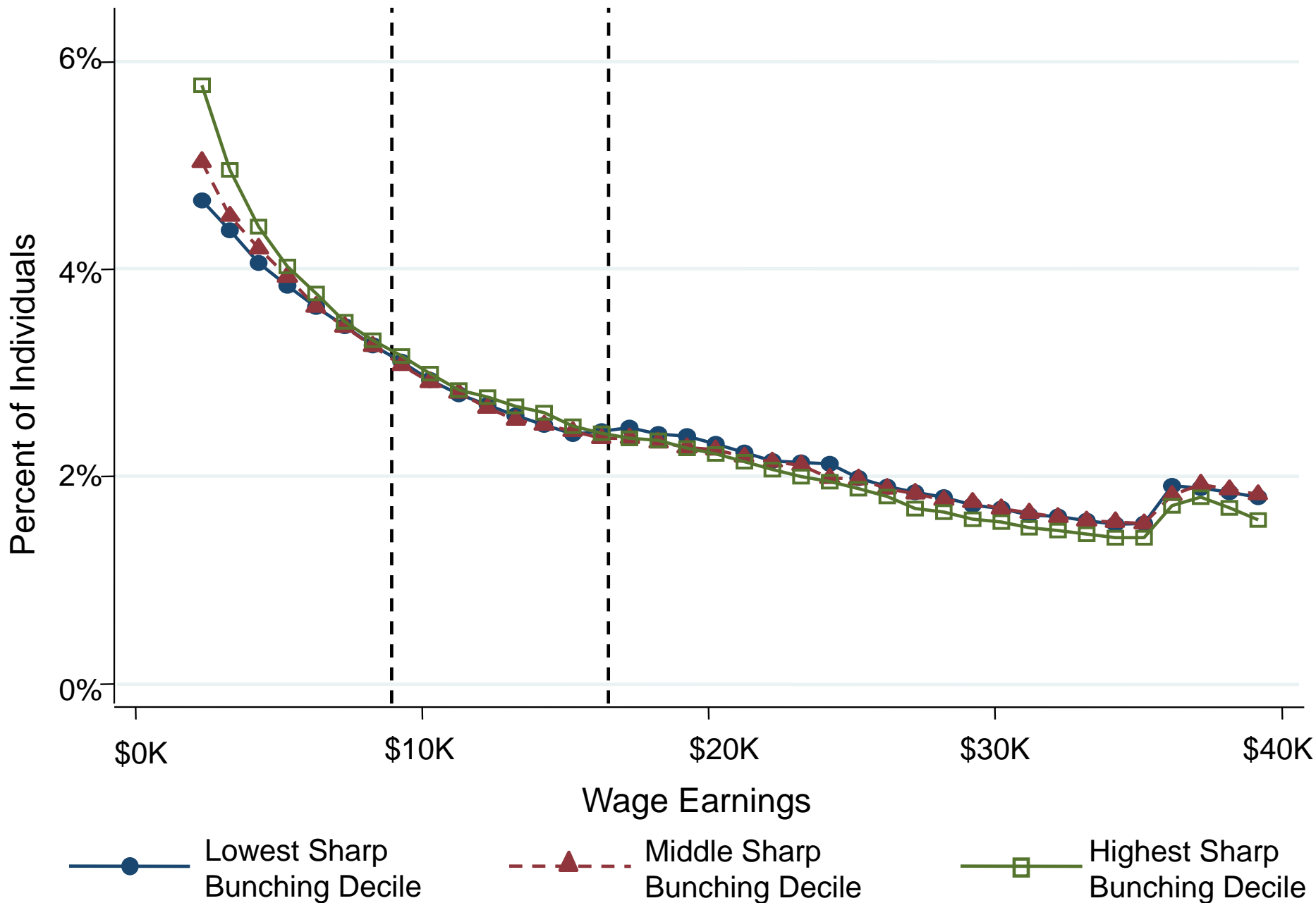
# Outline of Empirical Analysis

- Step 1: Document variation across neighborhoods in sharp bunching among self-employed
- Step 2: Establish that variation in sharp bunching across neighborhoods is driven by differences in knowledge about EITC schedule
- Step 3: Compare wage earnings distributions across low- and high-knowledge neighborhoods to uncover impacts of EITC on earnings
- Step 4: Compare impacts of changes in EITC subsidies on earnings across low vs. high knowledge nbhds. to account for omitted variables

# Child Birth Research Design

- Cross-sectional differences in income distributions could be biased by omitted variables
- To identify causal impacts of EITC, need variation in tax incentives
  - Use child birth as an instrument for EITC eligibility
  - Birth affects labor supply directly, but cross-neighborhood comparisons provide good counterfactuals
- 12 million EITC-eligible individuals give birth within our sample

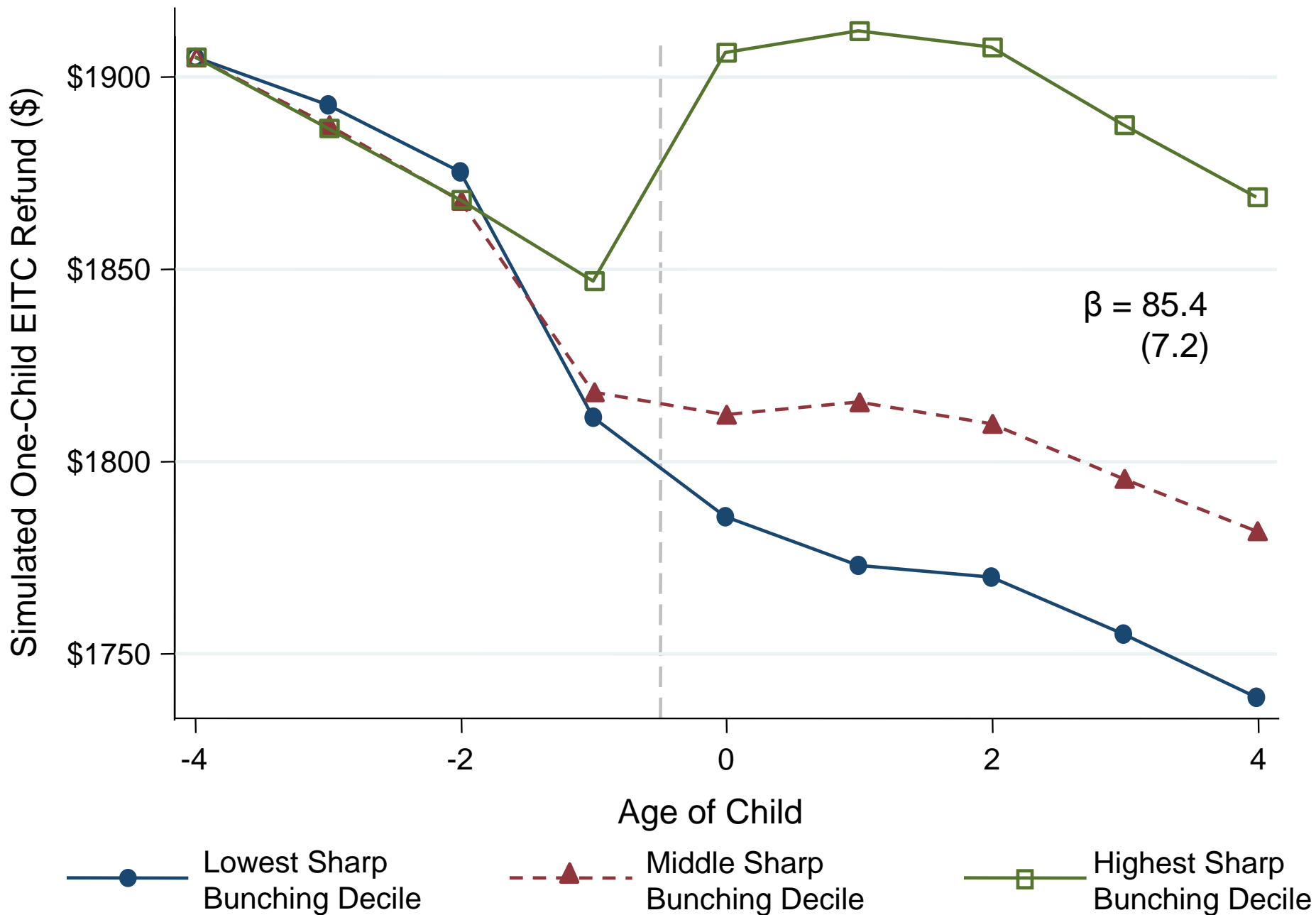
# Earnings Distribution in the Year Before First Child Birth for Wage Earners



# Earnings Distribution in the Year of First Child Birth for Wage Earners



# Simulated EITC Credit Amount for Wage Earners Around First Child Birth

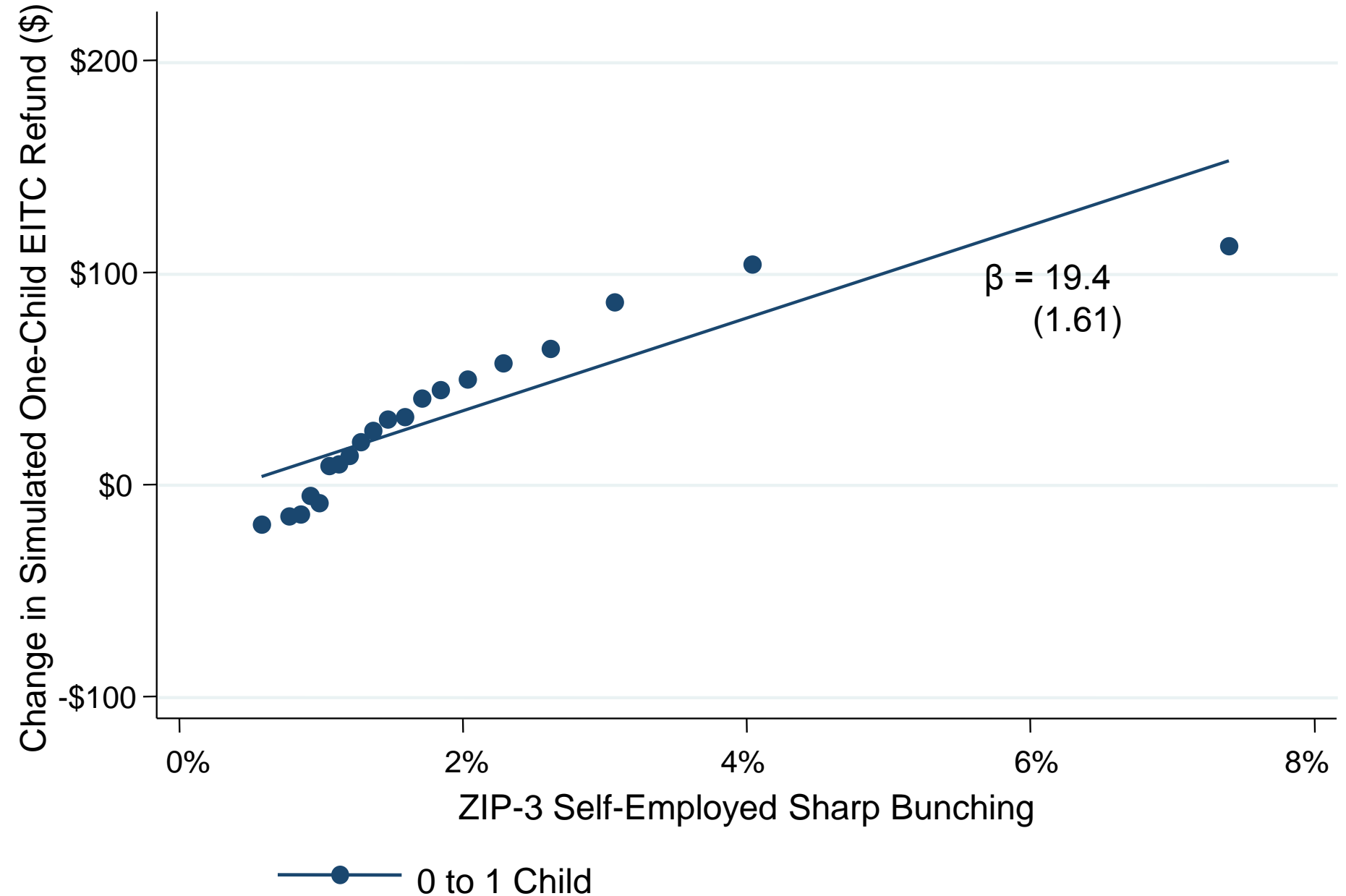




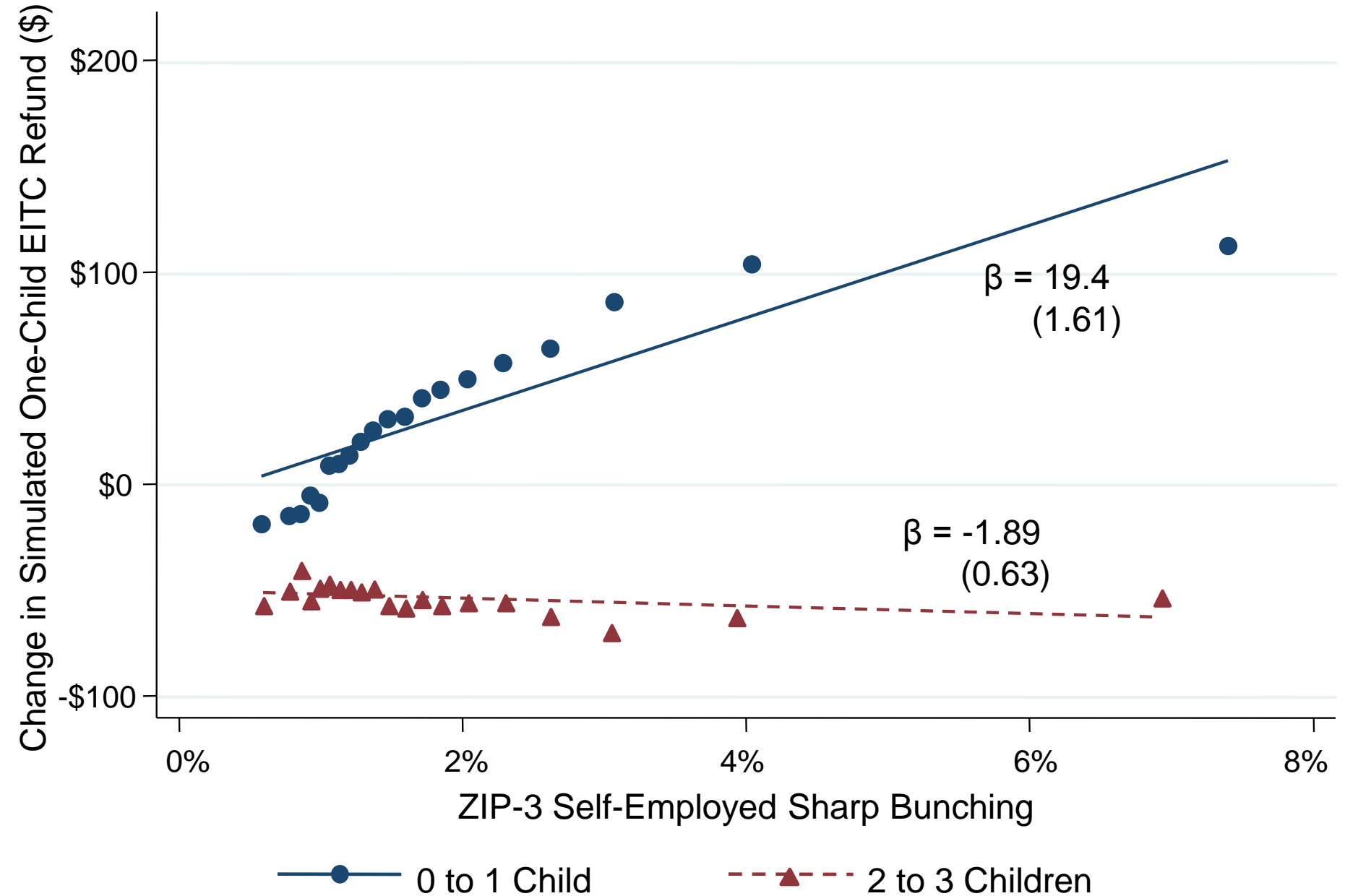
# Composition of Wage Earnings Responses

- Where is the increase in EITC refunds coming from?
  - Phase-in, phase-out, or extensive margin?
  - Important for understanding welfare consequences of EITC
- Compare change in simulated EITC amount (with 1 child) from year -1 to year 0 across low and high information areas

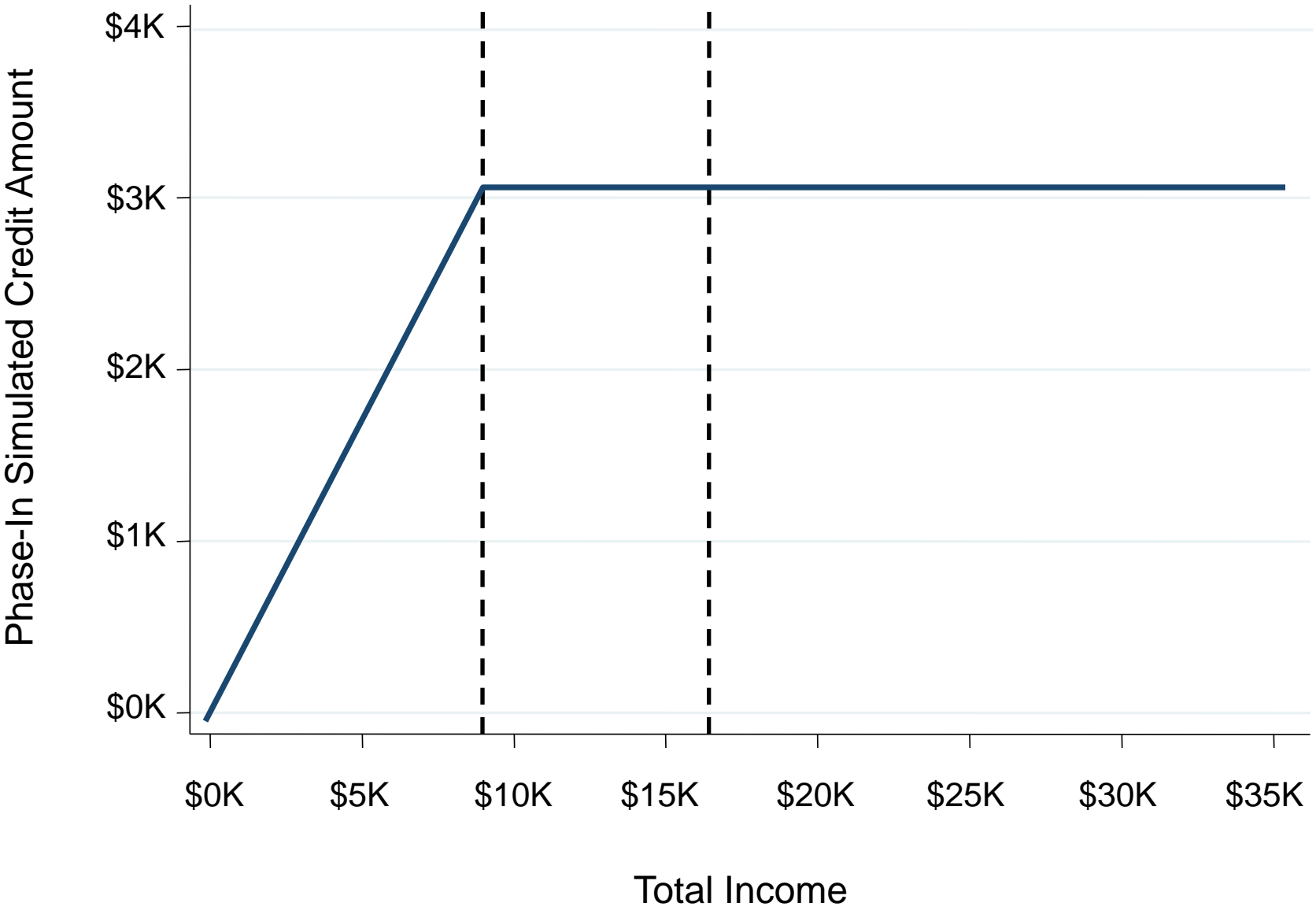
# Changes in W-2 Based Simulated EITC around Child Birth vs. Sharp Bunching



# Changes in W-2 Based Simulated EITC around Child Birth vs. Sharp Bunching



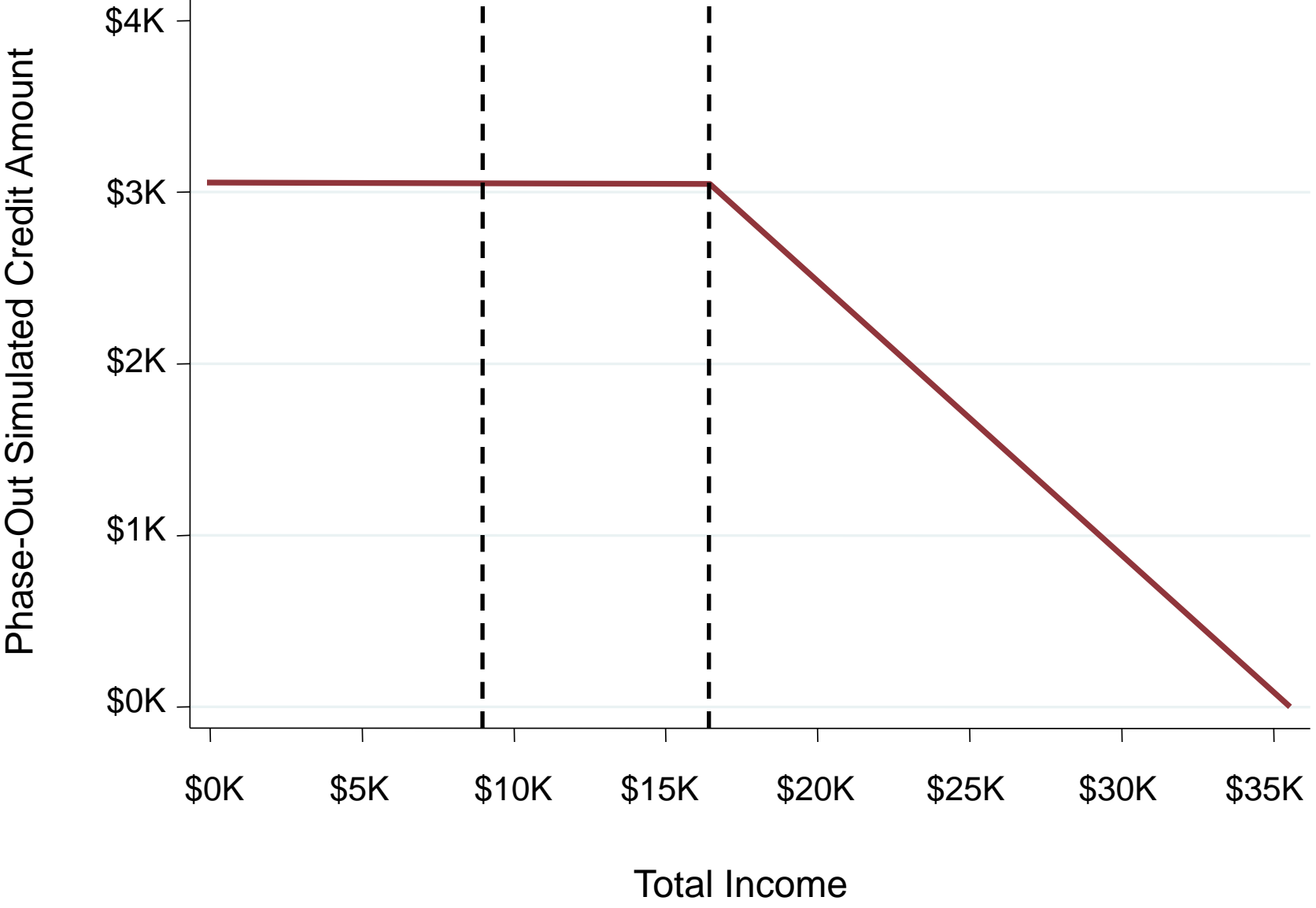
# Simulated Phase-In Credit



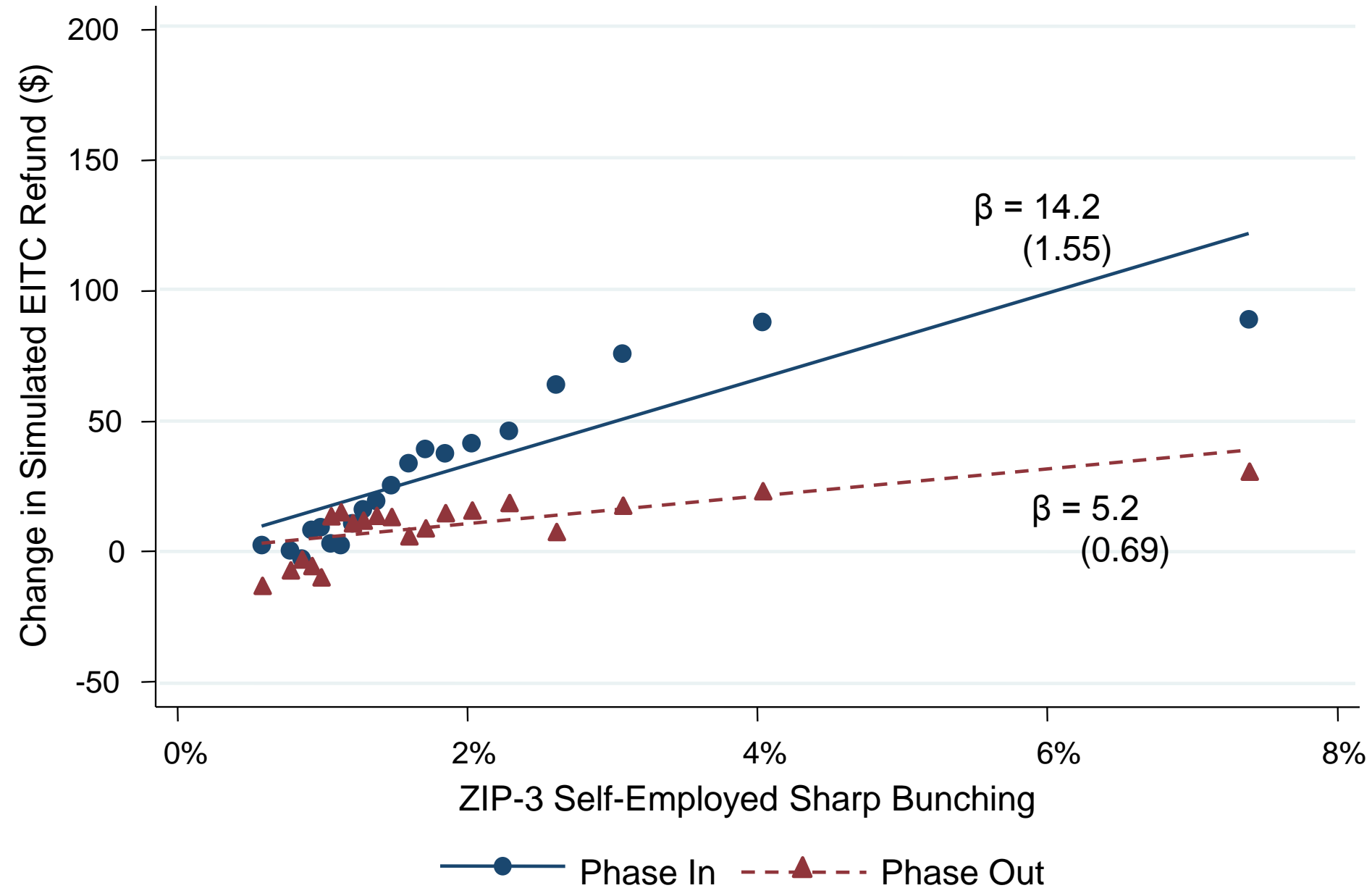
# Changes in W-2 Based Simulated EITC around Child Birth vs. Sharp Bunching



# Simulated Phase-Out Credit

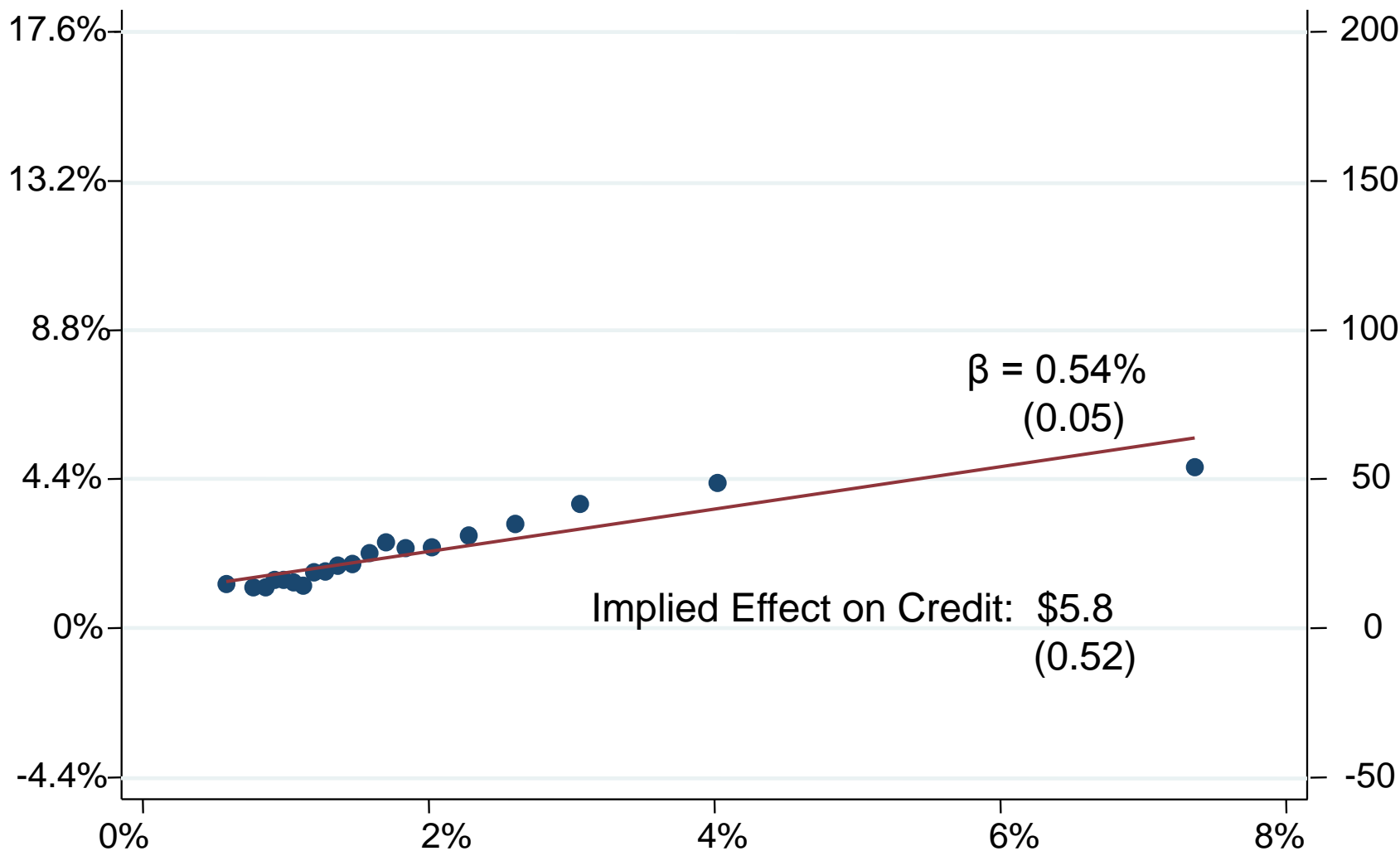


# Changes in W-2 Based Simulated EITC around Child Birth vs. Sharp Bunching



# Extensive Margin: Changes in Fraction Working around First Birth

Change in Percent of Individuals with Positive W-2 Earnings



ZIP-3 Self-Employed Sharp Bunching

Change in Simulated EITC Refund (\$)



## Impact of EITC on Wage Earnings

	Baseline Specification	Large Firms Only	With ZIP-3 Fixed Effects	Placebo Test: 3 <sup>rd</sup> Child
Dependent Variable:	Simulated EITC Refund			
ZIP-3 Sharp Bunching	\$19.4 (1.61)	\$14.4 (1.14)	\$34.7 (3.20)	-\$1.89 (0.63)

## Impact of EITC on Wage Earnings

Dependent Variable:	Phase-in vs. Phase-out		Extensive Margin	
	Sim. Phase-in Credit	Sim. Phase-out Credit	Positive W-2 Earnings	Number of Jobs (W-2's)
ZIP-3 Sharp	\$14.2	\$5.2	0.54%	<b>0.017</b>
Bunching	(1.55)	(0.69)	(0.05)	<b>(0.002)</b>

# Tax Policy Implications

- Our estimates can be used to characterize impact of EITC on income distribution taking into account behavioral responses
- Use neighborhoods in bottom decile of sharp bunching as counterfactual for earnings distribution without EITC

## Impact of EITC on Income Distribution

### Percent of EITC-Eligible Households Below Threshold

	50% of Poverty Line	100% of Poverty Line	150% of Poverty Line	200% of Poverty Line
No EITC Counterfactual	<b>13.2%</b>	31.3%	53.8%	77.1%
EITC, No Behavioral Response	<b>8.9%</b>	21.4%	41.6%	70.8%
EITC, with Avg. Behavioral Response	<b>8.2%</b>	21.0%	42.0%	71.3%
EITC with Top Decile Behavioral Response	<b>6.7%</b>	20.2%	42.6%	72.1%

# Elasticity Estimates Based on Change in EITC Refunds Around Birth of First Child

	Mean Elasticity	Phase-in Elasticity	Phase-out Elasticity	Extensive Elasticity
<i>A. Wage Earnings</i>				
Elasticity in U.S. 2000-2005	<b>0.21</b> <b>(0.012)</b>	<b>0.31</b> <b>(0.018)</b>	<b>0.14</b> <b>(0.015)</b>	<b>0.19</b> <b>(0.019)</b>
Elasticity in top decile ZIP-3's	0.55 (0.020)	0.84 (0.031)	0.29 (0.020)	0.60 (0.034)
<i>B. Total Earnings</i>				
Elasticity in U.S. 2000-2005	0.36 (0.017)	0.65 (0.030)	0.11 (0.006)	0.36 (0.019)
Elasticity in top decile ZIP-3's	1.06 (0.029)	1.70 (0.047)	0.31 (0.010)	1.06 (0.040)

# Conclusion

- EITC has significantly increased incomes of low-income families with children through mechanical effects + behavioral responses
  - Behavioral responses still concentrated in a few areas but continuing to spread across the U.S.
  - Contrary to prior findings, intensive margin responses are substantial and may even be larger than extensive margin responses
- Differences in knowledge can provide useful counterfactuals when traditional approaches are unavailable
  - Characterizing impacts of social security on retirement behavior using social security earnings test
  - Analyzing responses to corporate taxation