

Local Taxes and Suburbanization: Evidence from Philadelphia's Wage Tax*

Matthew Jacob
Harvard University

René Livas
Harvard University
Job Market Paper

November 9, 2025

[Link to newest version](#)

Abstract

Do city taxes shift economic activity to the suburbs? We study Philadelphia's wage tax, which applies to residents regardless of where they work and to suburban residents who work in the city. Because city residents always pay the tax, it does not distort their workplace choices, whereas suburban residents are penalized only for working in Philadelphia. At the city boundary, rising wage tax rates should sharply reduce commuting to the city in suburban tracts relative to neighboring city tracts, while falling tax rates should increase it. Using a spatial regression discontinuity design, we find that as the wage tax rose from 1.5 to 4.3% between 1960 and 1980, the change in the proportion of residents working in the city fell sharply in suburban tracts just outside the boundary; as the tax fell to 3.4% between 2003 and 2019, the change in that proportion increased sharply in the same tracts. Similar results hold along the boundaries of other cities with wage tax variation, such as Detroit and Cleveland, but not in cities without wage taxes. In our preferred estimate, a 1% increase in the tax rate reduces suburb-to-city commuting by 6.39%, holding wages, rents, and amenities constant. We embed this elasticity in a quantitative spatial model to estimate how the wage tax affects suburbanization once wages and rents adjust. Replacing the wage tax with a non-distortionary land value tax would bring 26,000 jobs from the suburbs into Philadelphia. Such gains triple when we allow for productivity agglomeration forces.

*rlivas@g.harvard.edu. We are indebted to Gordon Hanson, Ed Glaeser, and Gabriel Kreindler for their support and advice. We also thank David Agrawal, Prottoy Akbar, Raj Chetty, Robert French, Claudia Goldin, Anders Jensen, Jeff Liebman, Jeff Lin, Avi Moorthy, Michael Nutter, Allison Shertzer, Stefanie Stancheva, and seminar participants at Harvard for helpful comments and suggestions. Muhammad Sultan Dogar provided excellent research assistance. Livas gratefully acknowledges financial support from the Taubman Center for State and Local Government, the Harvard Mellon Urban Initiative, and the Reimagining the Economy project.

1 Introduction

On January 1st, 1940, Philadelphia became the first U.S. city to introduce a wage tax. By taxing residents and suburban commuters, the wage tax raised a remarkable 20% of the city’s revenue in 1940. Success begat imitation. Of the ten largest U.S. cities in 1940, six had enacted a wage tax by 1967.¹ Yet workers in those cities could avoid the wage tax by moving and taking their jobs with them. By 1980, living and working in the suburbs had become more common in metro areas where the central city had adopted a wage tax (e.g., Philadelphia, Detroit) than in other metro areas (e.g., Chicago, Boston; see Figure 1a). One view attributes this fact to wage taxes. Another is that some cities heavily suburbanized due to manufacturing job losses or the Great Migration, and, in turn, adopted this tax as demand for public services soared. To what extent, then, do wage taxes affect the location of economic activity within local labor markets?

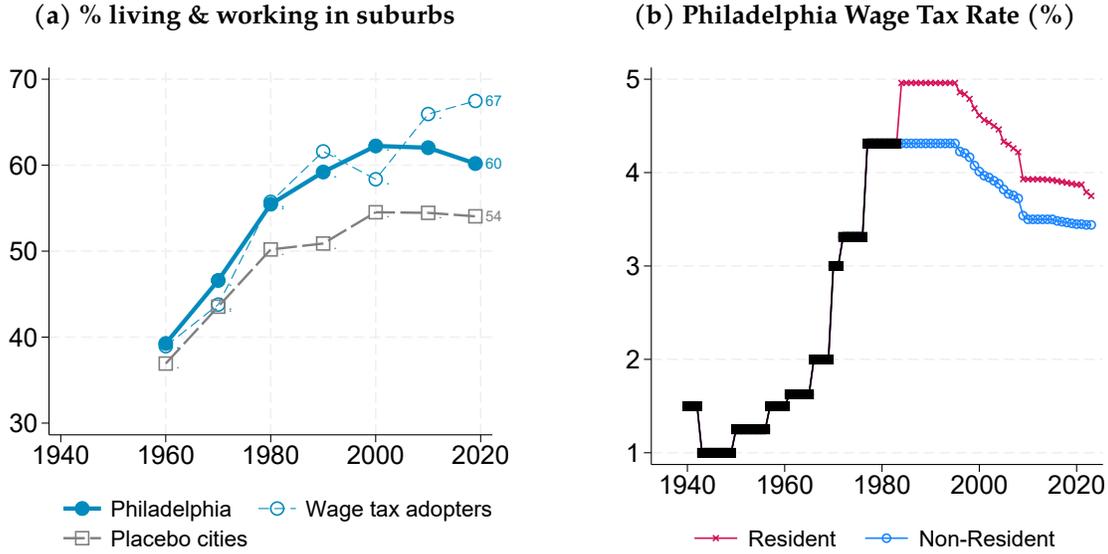
This paper studies Philadelphia’s wage tax and its historical commuting patterns to estimate two quantities: (1) the effect of the tax on location choices holding wages, rents, and amenities constant, and (2) the effect of the tax on the stocks of city residents and employees once wages and rents adjust in equilibrium. We exploit the design of the wage tax to estimate the first quantity. The tax does not distort the workplace choices of city residents because it applies to them regardless of where they work. In contrast, the wage tax applies to suburban residents only if they work in the city. The wage tax thus sharply penalizes residents just outside the city boundary for choosing the city as a workplace. As the wage tax rises and falls, so does this penalty. Philadelphia offers a valuable case study because its wage tax has risen and fallen substantially. The nonresident rate, which penalizes suburban commuting into the city, rose from 1.5 to 4.3% between 1960 and 1976, remained at that level until 1995, and then fell to 3.46% by 2018 (Figure 1b). By 1990, the wage tax raised almost twice as much per resident as New York City’s income tax, while neighboring municipalities had no wage tax or maintained much lower rates.

The analysis is in four parts. We begin by presenting a quantitative spatial model of a metropolitan area where the central city imposes a wage tax. We derive equilibrium expressions for how changes in tax rates differentially affect the workplace choices of city and suburban residents. Guided by these expressions, we then present nonparametric evidence of discontinuous changes in commuting composition at the city boundary as the wage tax rose and fell. In the third part of the paper, we use these discontinuities to estimate how the tax affects location choices holding wages, rents, and amenities constant. Embedding this estimate in our model, we then replace the wage tax with a non-distortionary land value tax to structurally estimate the effect of the wage tax on the stocks of city residents and employees once wages and rents adjust in equilibrium.

Our quantitative spatial model builds on the seminal work of [Heblich, Redding, and Sturm \(2020\)](#). Identically productive workers choose residential and workplace neighborhoods to maximize utility over wages, rents, commuting costs, amenities, and local wage taxes. In addition to bilateral commuting costs, we include a symmetric, time-invariant term that captures frictions

¹The six cities are Philadelphia (1940), St. Louis (1946), Pittsburgh (1949), Detroit (1962), New York City (1966, though its nonresident rate was repealed in 2000), and Cleveland (1967).

Figure 1: Suburbanization and wage taxes



Note: **Panel (a)**: flows that start and end outside of the central city’s limits as a share of commuting within 1960 SMSA. Wage tax adopters are Detroit, Cleveland, St. Louis, and Pittsburgh. Placebo cities include Boston, Brooklyn-Queens, Chicago, and Minneapolis. Shares for wage tax adopters and placebo cities are simple averages across metro areas. **Panel (b)**: residential and nonresidential tax rates from 1940 to 2023. Sources: see Section 4.

to cross-municipality commuting (Loumeau, 2023). We assume that revenue from the wage tax funds an exogenously set level of public expenditure.² The model’s key moment is the commuting elasticity, which governs the decrease in the probability that workers choose a residence-workplace pair as the tax rate associated with that pair increases, holding all else constant.

If access to jobs is continuous along the road network, our model implies that a change in the wage tax discontinuously affects neighborhood commuting patterns at the city boundary. In turn, this discontinuity identifies the commuting elasticity. To derive this result, we first show that city residents’ workplace choice is not distorted by the tax, while suburban residents face the wage tax and time-invariant frictions to commuting only if they work in the city. To disentangle the impact of the tax from these frictions, we study expressions for log changes in the share of neighborhood residents working in the city, and assume access to jobs is continuous along roads that cross the city boundary. As the wage tax rises, working in the city becomes less attractive for suburban residents, so the share of city commuters in suburban neighborhoods falls by more than that in city neighborhoods. When the tax falls, the city becomes relatively more attractive for suburban commuters, so the opposite occurs. In both cases, the discontinuity at the city boundary equals the commuting elasticity multiplied by the change in tax rates.

We bring the model to the data by constructing a panel of consistent census tracts from 1960 to 2016 (Census and American Community Survey) and from 2003 to 2018 (LEHD-LODES). By combining place-of-work tables from the Elizabeth Mullen Bogue file with scanned versions of those

²This assumption follows Wildasin (1985). Wage tax hikes financed the rising cost of existing municipal services, which may have been driven by cost disease (Baumol, 1967) or by higher city employee compensation (Inman, 1995).

tables for 1960 and 1970, we are able to measure the number of census tract residents commuting to Philadelphia and to other suburban counties. We also use tabulations of commuting flows to the central city in 1980 and bilateral commuting flows at the census tract level starting 1990. We cross-walk historical tracts to consistent tract definitions following [Lee and Lin \(2018\)](#).³ We restrict our sample to tracts within a five-minute drive of Philadelphia's boundary with the rest of Pennsylvania using the present-day road network, and limit our outcome to workplaces in Pennsylvania. To test the validity of our design, we digitize and geocode the 1938 *Industrial Directory of Pennsylvania*, and combine city enumeration districts from [Shertzer et al. \(2016\)](#) with our own digitized suburban enumeration districts linked to the 1940 Full Count Census. The area around the city limits had a light manufacturing presence and specialized as a place of residence, but we find no baseline discontinuities in workplace or population density, or across a range of residential characteristics.

We then present nonparametric evidence that our key outcome, the log change in the share of neighborhood residents working in the city, responds sharply to changes in the city's wage tax just outside the city boundary. As the tax increased from 1.5 to 4.3% between 1960 and 1980, the outcome was sharply lower by 0.09 points ($SE = 0.042$) in suburban neighborhoods just beyond Philadelphia's boundary. When the tax fell from 4.0 to 3.4% between 2003 and 2018, the outcome was instead sharply higher by 0.13 points ($SE = 0.021$) in the same neighborhoods. These results are consistent with our model and hold across other base and end years.

We also find discontinuities for our key outcome in other cities with wage tax variation, but not in cities without wage taxes. Suburban neighborhoods just outside Cleveland experienced a sharp decline in city commuting as its wage tax rose from 1 to 2% between 1970 and 2000, but no discontinuity when the rate remained constant from 2000 to 2014. Suburban neighborhoods along Detroit's boundary exhibit sharp declines in city commuting as its wage tax rose from 0.5 to 1.5% between 1970 and 2000, which then reversed as the rate fell to 0.25% by 2014. In contrast, suburban neighborhoods along the boundaries of cities without wage taxes, such as Boston and Chicago, exhibit no discontinuities in our key outcome in any period.

One concern about these results is that they could be confounded by sorting on observables if certain types of workers prefer to live and work in the city. To address this concern, we compute log changes in neighborhood commuting shares for workers of similar race, income level, sector, and uncongested commute time. We find large discontinuities of the expected sign for each of these characteristics. We also find discontinuities along Philadelphia's boundaries but not along placebo city boundaries in other outcomes, including changes in the share of employees commuting from the city and in moving rates to the central city, as well as growth rates in neighborhood establishments and contract rents.⁴ Our nonparametric evidence is consistent with wage tax variation discontinuously affecting workplace choice at the city boundary, and with such discontinuities arising solely from wage tax variation.

Our key theoretical result—that the discontinuity in the log change in the share of residents

³Historical tracts never straddle county lines, so they always align with Philadelphia's border.

⁴Importantly, neighborhood residential rents are not a determinant of workplace choice in our model.

working in the city equals the commuting elasticity multiplied by the change in tax rates—implies that we can estimate the commuting elasticity using a fuzzy spatial regression discontinuity design (RDD). This design identifies the elasticity by dividing the discontinuous change in commuting composition by the discontinuous change in tax differentials at the city boundary. We pool our key outcome across end years to estimate a commuting elasticity for each base year; our preferred estimate of 6.39 is for changes relative to the base year 1960.⁵ Our estimates are robust to using a parametric Instrumental Variables Pseudo-Poisson Maximum Likelihood estimator (IV-PPML), controlling for baseline measures of observable worker heterogeneity, implementing donut-hole specifications, and performing randomization inference with simulated city limits. Using a gravity equation (a fixed-effects estimator on a panel of bilateral commuting flows) and a structural approach (which requires assumptions about production and additional data), we obtain elasticity estimates consistent with those from our fuzzy spatial RDD and from seminal work ([Ahlfeldt et al., 2015](#); [Heblich et al., 2020](#)).

We assess the equilibrium effects of the wage tax on suburbanization through a counterfactual in which Philadelphia replaces it with a revenue-equivalent land value tax. Philadelphia could feasibly fund itself with a non-distortionary land value tax, and has both the legal authority and precedent to implement such a policy.⁶ Holding residential and workplace amenities fixed and calibrating the model with our estimated commuting elasticity, we allow wages, rents, and commuting choices to adjust across the metropolitan area. In an open-city specification with inelastic floorspace supply and no agglomeration forces, replacing the wage tax with a land value tax would increase the city’s resident and nonresident employment by 26,000 jobs, implying a tax base elasticity of 0.86. Pre-tax land values rise by 9% in Philadelphia and by 4.4% across the metropolitan area. Allowing for agglomeration and elastic floorspace supply triples the tax base elasticity and doubles the total land value gain. Agglomeration forces thus *weaken* the case for a wage tax.

We conclude by using the model to estimate sales, property, and wage tax revenues under counterfactual wage tax rates ranging from 0 to 10%. We hold amenities fixed to estimate the endogenous revenue shortfall or surplus as the wage tax varies. We calibrate the sales tax rate as a fixed proportion of residential income after taxes and rental payments, and the property tax rate as a fixed proportion of land values, to match the city’s observed tax revenue composition in 2018. We find that although higher wage tax rates reduce sales and property tax revenues, total tax revenue increases with the wage tax rate over the 0–10% range. This result persists even when we account for elastic floorspace supply and productivity agglomeration. Despite its adverse effects on the location of economic activity, land values, and other tax bases, the wage tax has never approached its revenue-maximizing rate.

Related literature Our research connects two strands of literature. We first emphasize the importance of local labor market effects in the literature on taxation and location choice. To estimate

⁵Because wage tax rates had been stable throughout the 1950s, the 1960 base year is less subject to dynamic adjustments that may overstate the elasticity.

⁶Pittsburgh’s land value tax had a positive impact on building activity during the 1980s ([Oates and Schwab, 1997](#)).

the effect of subnational taxation on location, studies tend to focus on specific types of taxpayers, such as inventors (Moretti and Wilson, 2017), high-income individuals (Bakija and Slemrod, 2004; Young and Varner, 2011; Agrawal and Foremny, 2019; Martínez, 2022), or professional athletes (Agrawal and Tester, 2024). While such focus is useful for identification, it often cannot speak to how taxation changes the distribution of economic activity within local labor markets. Other work on subnational taxation and firm location often does not distinguish between relocation *across* and *within* local labor markets (Duranton et al., 2011; Fajgelbaum et al., 2019; Giroud and Rauh, 2019; Risch, 2024). Indeed, our commuting elasticity estimates are roughly three times as large as the analogous flow elasticities in the cross-local labor market literature.

Second, spatial models used to study transportation (Heblich et al., 2020; Severen, 2023; Tsvanidis, 2024), public works (Franklin et al., 2024), and urban structure (Ahlfeldt et al., 2015; Owens et al., 2020; Kreindler and Miyauchi, 2023) often abstract from decentralized jurisdictions and their taxing powers (recent exceptions include Bordeu, 2023; Loumeau, 2023). We apply a quantitative spatial model to study the relationship between taxes and suburbanization (Gyourko and Tracy, 1989; Mieszkowski and Mills, 1993; Epple and Sieg, 1999) and map our results to flow and stock elasticities, which are estimands of interest in public economics. We find that rising wage tax differentials between cities and suburbs contributed to the suburbanization of residences and workplaces (Redding, 2022), while falling wage tax differentials may have facilitated urban revival (Couture and Handbury, 2020) and reduced fiscal externalities (Agrawal et al., 2024). Our framework could also be applied to location choices across states in the context of state tax reciprocity agreements and remote work (Agrawal and Hoyt, 2018; Agrawal and Brueckner, 2025).

Methodologically, we introduce a transparent and easily implementable approach to estimating an important parameter in quantitative spatial models. While local tax differentials have been shown to lengthen commuting times (Wildasin, 1985; Agrawal and Hoyt, 2018) and reduce a city's share of national employment (Grieson, 1980; Haughwout et al., 2004), we instead leverage them to identify the commuting elasticity using an RDD. This insight can be applied to other policies that discontinuously affect commuting choices, such as congestion charge zones that exempt residents (Herzog, 2024; Cook et al., 2025). Our proposed implementation, which measures the running variable from multiple points on the road network within a tract, is straightforward in contexts where the unit of observation is an area (Baum-Snow and Ferreira, 2015; Cattaneo et al., 2025).

Lastly, we highlight a source of spatial misallocation *within* local labor markets that may be difficult to mitigate because of its importance for tax revenue. Fajgelbaum and Gaubert (2020) and Bilal (2023) instead emphasize spatial misallocation *across* labor markets as a rationale for place-based policy, but some existing place-based policies merely target neighborhoods with mixed results (Freedman, 2012; Busso et al., 2013). Local government taxation may therefore represent a blind spot in U.S. place-based policy (Hanson et al., 2025). Targeting cities such as Philadelphia, Detroit, and St. Louis to reduce wage taxes and shift toward less distortionary taxation may be a productive direction for place-based policy.

2 Policy details and Background

The main sources of revenue for city governments in the United States are property, sales, and income taxes. Local income taxes are more common in older cities and account for a significant share of revenue where adopted. An advantage of local income taxes is their relatively low administrative cost compared to other tax instruments (Gabler, 1970). Since 1970, New York City has raised between 15 and 20% of its general fund revenue from its personal income tax. In 2019, Philadelphia, Detroit, and St. Louis raised 37, 22, and 19% of their revenue, respectively, from their wage taxes (Appendix Figure A.3).

We define the wage tax as any local income tax that applies to both resident and nonresident workers. Wage taxes are distinct from state and federal income taxes. We estimate that 9 million jobs were subject to a wage tax in 2017–2019.⁷

In Philadelphia, the wage tax has provided at least one fifth of general fund revenue since its introduction in 1940. In 1960, this share was one third; by 1980, one half. In dollar terms, Philadelphia raised the equivalent of \$800 per resident in 1990 (in inflation-adjusted dollars), roughly twice the per capita yield of New York City’s personal income tax in the same year (Appendix Figures A.4 and A.5). By 2019, the wage tax generated 37% of Philadelphia’s general fund revenue, with non-residents contributing 38.3% of that amount. The wage tax thus remains important for Philadelphia’s fiscal capacity, owing to specific institutional and political factors described below.

2.1 Policy details

The wage tax applies to wages, salaries, and other compensation earned through employment or services performed within the City of Philadelphia. City residents pay the tax regardless of workplace location, while nonresidents pay the tax if employed in Philadelphia. Contributions to 401(k) plans are not exempt. Incorporated businesses also pay the wage tax on net profits for services performed in Philadelphia on the same resident and nonresident basis. The wage tax does not apply to government transfers such as Social Security benefits.⁸

Pennsylvania grants large cities greater fiscal autonomy, which affects sourcing rules and tax rates. The state legislature authorized Philadelphia to levy the wage tax with the Sterling Act of 1932, and extended similar authority to all municipalities through Act 481 of 1947 and Act 511 of 1965. Wage tax rates for small jurisdictions are capped at 1%. Employers must withhold the higher of the tax rates between a worker’s home and work jurisdictions. Suburban residents employed in

⁷The wage tax is known by different names in different jurisdictions. Our definition is based on sourcing rules applying to both residents and nonresident workers. Some local governments tax earnings only at the place of residence (e.g., New York City, the District of Columbia, counties in Maryland and Indiana) or only at the place of work (e.g., counties in Kentucky, several municipalities in Alabama). To calculate the number of jobs subject to the wage tax, we use LEHD-LODES main job data to count commuting links that start or end in local governments in Pennsylvania, Ohio, Michigan, Missouri, and New York that tax both resident and nonresident income.

⁸Although wage taxes are local ordinances, enforcement is aided by state and federal law. State law requires Philadelphia employers to withhold wage tax, and Pennsylvania employers to report the residence of individual employees. City, state, and federal governments exchange taxpayer information. Misreporting residence or workplace status would generally require misreporting on state and federal returns, which constitutes tax fraud.

Philadelphia are exempt from their home jurisdiction's wage tax and instead pay the nonresident wage tax to Philadelphia.

Two cases allow at least partial refunds. First, New Jersey residents employed in Philadelphia can claim a credit on their state tax return for local wage taxes paid in Pennsylvania. This credit, while distinct from the NJ/PA Reciprocal Agreement covering state income taxes, has also existed since the mid-1970s, when both states began taxing income. If all income for a New Jersey resident is earned in Philadelphia, she pays whichever is higher: Philadelphia's nonresident wage tax rate or her average New Jersey state income tax rate. Second, Philadelphia offers partial tax relief to low-income households. Eligibility depends on state-set thresholds. Households earning less than \$8,750 annually pay a reduced wage tax rate of 1.5%, with a gradual phase-out at higher incomes. But this program is undersubscribed: as of 2024, fewer than 1% of eligible households claimed it.⁹

Given these institutional characteristics, we exclude New Jersey from our spatial RDD. For reduced-form results, we rely on neighborhood-level comparisons between Philadelphia and its neighboring Pennsylvania suburbs. For these comparisons to be valid, Philadelphia's western boundary must be unrelated to wage tax variation, and the variation itself must be exogenous to neighborhoods along the border. We discuss these conditions next.

2.2 Philadelphia County and its suburbs

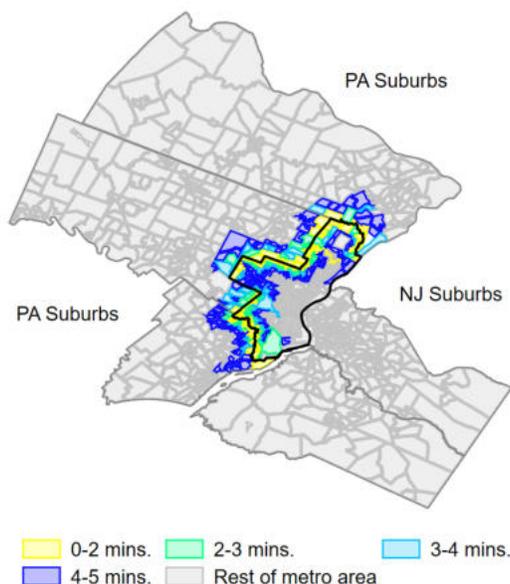
Philadelphia County lies west of the Delaware River, which marks Pennsylvania's eastern border with New Jersey. City and county boundaries, outlined in black in Figure 2, have been coterminous since 1854. Three historical developments shaped them. The northeastern border with Bucks County along Poquessing Creek has remained unchanged since 1682. The western boundary with Delaware County reflects oversubscription to William Penn's land allocation policy, which pushed the line west of the Schuylkill River. The remaining boundaries were established in 1784, when Montgomery County was split from Philadelphia to grant autonomy to rural townships northwest of the city. The last modification occurred in 1915, when 18 blocks were ceded from Montgomery County to Philadelphia.¹⁰ We find no evidence that these boundaries were drawn with the intent of taxing wages at a future date.

Philadelphia's suburbs expanded throughout the 19th and early 20th centuries, first around railroad stops in the 1860s and later with public transit and automobiles (Jackson, 1987). By 1897, Boyd's Philadelphia city directory recorded commuters from adjoining townships in a wide range of occupations (Akbar, Shertzer, and Livas, 2025). By the time the wage tax was enacted, the border region already contained substantial residential and industrial activity. Using the 1940 Census and 1938 Industry Directories of Pennsylvania, we calculate that townships within a five-minute drive of the city limits contained 215,624 residents and 119 industrial plants employing 26,215 workers.

⁹See a complete discussion of the local tax burden in <https://www.pew.org/en/research-and-analysis/reports/2023/03/the-local-tax-burden-on-philadelphia-households>.

¹⁰The boundary between Philadelphia and Montgomery Counties largely follows township borders set as early as 1687 (see Appendix Figure A.7).

Figure 2: Map of census tracts



Note: The map displays the census tracts in the portion of the Philadelphia Commuting Zone used in our analysis. The central region outlined in black is Philadelphia County. Proceeding clockwise around Philadelphia, the remaining counties are Bucks (PA), Camden (NJ), Gloucester (NJ), Delaware (PA), and Montgomery (PA). The tracts in bright colors show ranges for uncongested drive times to the Philadelphia city limits using the present-day road network. The tracts in gray are only used in the gravity equation and structural estimates. See Appendix Figure A.8 for the continuous drive time isochrone.

Philadelphia's boundary, as well as the economic activity of the adjoining neighborhoods, was established well before and independently of the wage tax.

2.3 Variation in the wage tax

Origins The wage tax originated as a temporary measure during the Great Depression. At the time, Philadelphia's property tax was already high, and further increases were seen as regressive (Fogelson, 2001), while local governments still played a dominant role in providing unemployment relief (Glaeser, 2013; Siodla, 2020). In 1932, the city council sought state authorization to enact new taxes. With both Philadelphia and Pennsylvania under Republican control, authority was granted that year, but competing interests delayed implementation until 1937 (Tinkcom, 1982). After proposals for a (resident) income tax collapsed, the wage tax emerged as a compromise and took effect on January 1, 1940, at a flat rate of 1.5%. This rate held until 1960.

Rising rates, 1960–1984 Like other large U.S. cities, Philadelphia faced rising costs after 1960.¹¹ In Philadelphia, collective bargaining rules under the 1951 Home Rule Charter limited council control over payrolls, contributing to rising expenditures (Inman, 1995). Local corruption and patronage

¹¹In a famous paper, Baumol (1967) cited municipal payrolls as an example of cost disease, since they were concentrated in low-productivity-growth sectors such as education and healthcare.

may have further reduced fiscal efficiency (Mandel, 2023). To manage the rising costs, Philadelphia relied on debt and wage tax hikes to finance spending before the 1970s.

Following New York City's 1975 fiscal crisis, interest rates for municipal bonds surged nationwide (Kidwell and Trzcinka, 1983). With the sudden increase in borrowing costs, Mayor Frank Rizzo raised the wage tax by one percentage point in February 1976, breaking a campaign promise. The City Council approved the increase 15–1, as it had done for a 1973 hike. Wage tax increases between 1961 and 1984 frequently appeared on the front page of *The Philadelphia Inquirer* and became a salient political issue, as reflected in reports by the Pennsylvania Economy League (see, e.g., Pennsylvania Economy League, 1999, for a retrospective discussion).

Falling rates, 1995–present In 1990, Philadelphia was unable to meet its short-term debt obligations due to declining state aid, a recession, and creditors unwilling to lend additional funds (Inman, 1995). In response, the state established the Pennsylvania Intergovernmental Cooperation Authority in 1991, with authority to oversee Philadelphia's five-year budgets and the power to withhold wage tax revenues. Oversight, combined with governance and property tax reforms, improved the city's creditworthiness. Since 1995, successive mayors have successfully pursued a program of gradual wage tax reductions (Figure 1). Philadelphia's recovery was further aided by the broader resurgence of U.S. central cities, driven by college-educated workers (Diamond, 2016; Baum-Snow and Hartley, 2020; Couture and Handbury, 2020; Couture et al., 2024), which improved municipal finances (Boustan et al., 2013).

Suburban wage taxes, 1992–present Suburban municipalities had little incentive to enact wage taxes while their tax base was captured by Philadelphia. Opposition to suburban wage taxes was strong (Strumpf, 1999). In our RD sample, the first suburban jurisdiction adopted a wage tax only in 1992. Because a wage tax captures the tax base of other jurisdictions, a township's best response to another municipality enacting a wage tax is to enact a wage tax of its own. Strumpf (1999) documents the chain reaction of adoption across Montgomery County in the 1990s, with Bucks County following in the 2000s. Adoption remains more limited in Delaware County, though still spatially clustered (Appendix Figure A.2).

Exogeneity of variation at city limits We find no evidence that the mayor or council gave special attention to areas near the city limits when setting wage tax rates. From inception, rates were a blunt instrument determined by citywide fiscal conditions. The flat rate structure and Pennsylvania's uniformity clause limited the targeting of particular neighborhoods. We find no references to border neighborhoods or adjacent municipalities in city council minutes, mayoral documents, or prior studies of the tax (Inman, 1995; Haughwout et al., 2004; Gyourko et al., 2005). The evidence supports the view that the wage tax was exogenous to neighborhoods at the boundary.

3 Model

We build on the quantitative spatial model in [Heblich, Redding, and Sturm \(2020\)](#) by augmenting it with wage taxes and border effects. We derive equilibrium expressions for changes in workplace choice of neighborhood residents on either side of the city boundary. Such expressions are the foundations for our nonparametric evidence at the city limits in [Section 5](#) and our approach to estimate the commuting elasticity in [Section 6](#).

3.1 Setup

A metropolitan area is composed of $i = 1, \dots, n$ neighborhoods split between two municipalities: the city (\mathcal{C}) and the suburbs (\mathcal{S}). Either municipality can implement a wage tax. We assume that tax revenue pays for an exogenously set level of public expenditure as in [Wildasin \(1985\)](#).

Workers Workers (ω) choose their residence neighborhood (i) and workplace neighborhood (j). Their indirect utility function is given by

$$U_{ij}(\omega) = b_{ij}(\omega) [B_{ij} w_j (1 - \tau_{ij}) / d_{ij}] [P_i^\alpha Q_i^{1-\alpha}]^{-1}$$

where $b_{ij}(\omega)$ is a worker's idiosyncratic amenity draw, B_{ij} represents residence and workplace amenities common to all workers, w_j is the (gross-of-tax) workplace wage, τ_{ij} is the wage tax rate, $d_{ij} \equiv \exp(t_{ij})$ captures bilateral commuting costs increasing in bilateral driving times (t_{ij}), P_i is the price index for consumption goods (consumed at the neighborhood of residence), and Q_i is the residential floorspace price. Parameter α is the share of income allocated to consumption. Idiosyncratic amenities for worker ω are drawn from a Fréchet distribution for each residence-workplace dyad:

$$G(b) = e^{-b^{-\theta}}, \quad \theta > 1$$

where θ is the Fréchet spread parameter. We describe the composite amenity term (B_{ij}), the wage tax rate (τ_{ij}), and the interpretation of the Fréchet spread parameter (θ) below.

3.2 Commuting probabilities

Standard results for Generalized Extreme Value distributions imply that the probability that workers choose residence neighborhood i and workplace neighborhood j is given by commuting pair ij 's share of commuting in the metropolitan area:

$$\lambda_{ij} = \frac{[B_{ij} w_j (1 - \tau_{ij}) / d_{ij}]^\theta [P_i^\alpha Q_i^{1-\alpha}]^{-\theta}}{\sum_k \sum_l [B_{kl} w_l (1 - \tau_{kl}) / d_{kl}]^\theta [P_k^\alpha Q_k^{1-\alpha}]^{-\theta}}, \quad i, j, k, l \in (\mathcal{C} \cup \mathcal{S}) \quad (1)$$

where the supply of workers to workplace j is increasing in wages and workplace amenities and decreasing in taxes and commuting costs from residence i ; the demand of housing in residence i is increasing in residential amenities and decreasing in consumption prices and rents. A larger value of the Fréchet spread parameter (θ) makes workers more sensitive to each of these forces. The numerator of this expression is the measure of workers commuting from i to j ; the denominator is the measure of all commutes in the metro area.

We define the amenity term B_{ij} as

$$B_{ij} \equiv B_i^R B_j^L (1 - \zeta_{ij}), \quad (2)$$

which captures residential amenities (B_i^R), workplace amenities (B_j^L), and a symmetrical, time-invariant border effect (ζ_{ij}) that is activated whenever a commuter crosses the city limits and shut down otherwise:

Location of workplace j	Location of residence i	
	$i \in \mathcal{C}$	$i \in \mathcal{S}$
$j \in \mathcal{C}$	0	ζ
$j \in \mathcal{S}$	ζ	0

The border effect term captures frictions to commuting across municipalities (Loumeau, 2023).¹²

The statutory wage tax term (τ_{ij}) depends on the municipality of residence and workplace. For ease of exposition, we consider years where suburban municipalities adjoining the city have no wage taxes (see Appendix Figure A.1):

Location of workplace j	Location of residence i	
	$i \in \mathcal{C}$	$i \in \mathcal{S}$
$j \in \mathcal{C}$	τ_C^R	τ_C^L
$j \in \mathcal{S}$	τ_C^R	0

where superscripts R or L denote the tax rate that applies to residents or nonresident workers. The resident wage tax rate (τ_C^R) applies to city residents ($i \in \mathcal{C}$) regardless of the location of their workplace. The nonresident rate (τ_C^L) applies to suburban residents ($i \in \mathcal{S}$) if they commute to any city workplace ($j \in \mathcal{C}$). Only to commutes within the suburbs have a tax rate of 0.

The numerator in (1) says that, all else equal, workers are more likely to commute from places with low amenity-adjusted cost of living, and to places with high amenity-adjusted wages net of commuting costs and wage taxes. To retain their attractiveness, higher-taxed commutes must be compensated by higher wages, lower rents or prices, or higher amenities, as in a Rosen-Roback

¹²Note that we could instead include an isomorphic border effect in the bilateral commuting cost term d_{ij} in the spirit of Allen and Arkolakis (2014). We include ζ in the amenities term since border effects in commuting could originate generally from social or business networks, as in Combes et al. (2005), or, more specifically, from the municipal provision of infrastructure jointly used by workers and firms. This mechanism is consistent with the evidence for firm-level trade in Hillberry and Hummels (2008) and is explicitly tested for in the context of commuting networks across French regions in Loumeau (2023), and within Chilean municipalities in Bordeu (2023).

model with taxes (Gyourko and Tracy, 1989). Unlike Rosen-Roback, however, a wage tax increase not only makes some residence-workplace dyads less attractive, but also affects the attractiveness of every other dyad in the metro area (Allen and Arkolakis, 2023).

3.2.1 Neighborhood commuting shares

We now turn to deriving commuting probabilities conditional on residence, a model object from Hebllich et al. (2020), and compare the resulting expressions for city and suburban neighborhoods. We first condition (1) on residence i , which has the effect of canceling out the i -specific variables,

$$\lambda_{ij|i} = \frac{(1 - \zeta_{ij})^\theta (1 - \tau_{ij})^\theta [B_j w_j / d_{ij}]^\theta}{\sum_l (1 - \zeta_{kl})^\theta (1 - \tau_{kl})^\theta [B_l w_l / d_{kl}]^\theta}. \quad (3)$$

The conditional commuting probability in (3) holds constant local amenities and the local cost of living, since they are captured by the i -specific term $[B_i^\theta (P_i^\alpha Q_i^{1-\alpha})^{-\theta}]$. Policies affecting these terms, such as school districts, or property or sales taxes, are therefore also held constant. In other words, commuting probabilities conditional on residence show the attractiveness of *workplaces* by comparing the different commutes of workers with identical residential amenities and cost of living.

To obtain the share of residents in a tract that commute to the city, we sum (3) over all workplaces in the city and unpack the border effect (ζ_{ij}) and statutory wage tax (τ_{ij}) terms. This yields the share of residents of a city neighborhood commuting to the city,

$$\lambda_{iC|i \in C} = \frac{\Phi_{i \in C}^C}{\Phi_{i \in C}^C + (1 - \zeta)^\theta \Phi_{i \in C}^S} \quad (4)$$

and the share of residents of a suburban neighborhood commuting to the city,

$$\lambda_{iC|i \in S} = \frac{(1 - \zeta)^\theta (1 - \tau_C^L)^\theta \Phi_{i \in S}^C}{(1 - \zeta)^\theta (1 - \tau_C^L)^\theta \Phi_{i \in S}^C + \Phi_{i \in S}^S} \quad (5)$$

where terms such as $\Phi_{i \in \mathcal{K}}^{\mathcal{H}} \equiv \sum_{j \in \mathcal{H}} [B_j^L w_j / d_{ij}]^\theta$ are gross-of-tax resident market access (RMA) variables capturing workplace attractiveness in municipality \mathcal{H} for a tract located in municipality \mathcal{K} . We sum by municipality to draw a distinction between RMA “to the city” (Φ_i^C) and RMA “to the suburbs” (Φ_i^S). The border effect and net-of-tax-rate terms pull out of these sums since they do not vary between neighborhoods within a municipality.

Equation (4) says that the wage tax does not affect workplace choice for residents of a city neighborhood, since they pay the tax regardless of where they work. The expression is only a function of workplace attractiveness in the city ($\Phi_{i \in C}^C$) and in the suburbs ($\Phi_{i \in C}^S$), as well as a border effect (ζ). All else equal, larger wage tax rates do not directly affect the share of city residents working in the city.

In contrast, (5) says that the nonresident wage tax rate directly affects the composition of com-

muting in *suburban* neighborhoods. For suburban residents, workplaces in the city ($\Phi_{i \in \mathcal{C}}^{\mathcal{C}}$) are made less attractive by a border effect (ζ) and the nonresident wage tax ($\tau_{\mathcal{C}}^L$), while workplaces in the suburbs ($\Phi_{i \in \mathcal{S}}^{\mathcal{S}}$) are unaffected. All else equal, the larger the wage tax, the smaller the share of suburban workers working in the city.

Appendix Section C shows that, when the nonresident and resident wage tax rates are equal to each other, similar results hold when conditioning on workplaces to compute the neighborhood share of employees commuting *from* the city. Workplace commuting composition is a function of firm market access (FMA). We show that the share of employees living in the city is only *directly* affected on the suburban side, since employees working in the suburbs finds the city less attractive to live in because of the resident wage tax. Importantly, there is no accounting relationship between residential shares (where neighborhood residents commute to) and workplace shares (where neighborhood employees commute from). Residential and workplace commuting shares are different outcomes similarly affected by the wage tax.

3.2.2 Changes in log commuting shares to the city

Let t_0 and t be subscripts denoting a base period and end period. We apply these subscripts to (4) and (5) to define the log change in commuting shares to the city for neighborhoods on either side of the city limits. The outcome takes the following form for a city neighborhood,

$$\Delta \ln \lambda_{i \in \mathcal{C}, t | i \in \mathcal{C}} = \ln \frac{\lambda_{i \in \mathcal{C}, t | i \in \mathcal{C}}}{\lambda_{i \in \mathcal{C}, t_0 | i \in \mathcal{C}}} = \underbrace{\ln \left(\frac{\Phi_{i, t}^{\mathcal{C}}}{\Phi_{i, t_0}^{\mathcal{C}}} \right)}_{\text{change in RMA to } \mathcal{C}} - \underbrace{\ln \left(\frac{\Phi_{i, t}^{\mathcal{C}} + (1 - \zeta)^\theta \Phi_{i, t}^{\mathcal{S}}}{\Phi_{i, t_0}^{\mathcal{C}} + (1 - \zeta)^\theta \Phi_{i, t_0}^{\mathcal{S}}} \right)}_{\text{change in total RMA} = \Delta \ln \Phi_{i \in \mathcal{C}, t}} \quad (6)$$

and the next form for a suburban neighborhood,

$$\Delta \ln \lambda_{i \in \mathcal{C}, t | i \in \mathcal{S}} = \underbrace{\theta \ln \left(\frac{1 - \tau_{\mathcal{C}, t}^L}{1 - \tau_{\mathcal{C}, t_0}^L} \right)}_{\text{wage tax variation}} + \underbrace{\ln \left(\frac{\Phi_{i, t}^{\mathcal{C}}}{\Phi_{i, t_0}^{\mathcal{C}}} \right)}_{\text{change in RMA to } \mathcal{C}} - \underbrace{\ln \left(\frac{(1 - \tau_{\mathcal{C}, t}^L)^\theta (1 - \zeta)^\theta \Phi_{i, t}^{\mathcal{C}} + \Phi_{i, t}^{\mathcal{S}}}{(1 - \tau_{\mathcal{C}, t_0}^L)^\theta (1 - \zeta)^\theta \Phi_{i, t_0}^{\mathcal{C}} + \Phi_{i, t_0}^{\mathcal{S}}} \right)}_{\text{change in total RMA} = \Delta \ln \Phi_{i \in \mathcal{S}, t}}, \quad (7)$$

which removes time-invariant border effect from the changes in RMA to the city. Equations (6) and (7) allow for wages, rents, or even amenities to respond endogenously to variation in the wage tax (Diamond, 2016; Almagro and Domínguez-Iino, 2025). The equations also say that, if we compare a city tract to a suburban tract where both change-in-RMA terms were comparable, we could isolate the direct effect of the wage tax on workplace choice with the wage tax variation term $[\theta \ln(1 - \tau_{\mathcal{C}, t}^L / 1 - \tau_{\mathcal{C}, t_0}^L)]$.

3.3 Estimating commuting elasticity from model expressions

We next develop a novel approach to use variation in the wage tax to estimate the elasticity of commuting flows to the net-of-tax rate (θ). Introducing two plausible continuity assumptions on

resident market access enables us to exploit the wage tax variation at the city limits to estimate this parameter. Other methods are a panel gravity equation, which absorbs variation in the attractiveness of residences and workplaces through fixed effects, and a structural approach, which imposes Cobb-Douglas technology on production and uses the structure of the model to calibrate it. We develop the latter two approaches following [Heblich, Redding, and Sturm \(2020\)](#) in Appendix Section D.

We use the theoretical expressions in (6) and (7) as the outcome in a spatial RD estimator that compares neighborhoods in the city and suburbs. To isolate the wage tax variation, we treat (6) and (7) as random variables and make spatial continuity assumptions on the RMA terms.¹³ By combining model objects with the empirical estimator, we aim to be precise about what theoretical parameter we estimate and under what assumptions it is identified.

To begin, we briefly formalize the notion of travel time to the city limits (\tilde{X}_i) in terms of bilateral travel times in the model (t_{ij}). Let k be the closest neighborhood across the city limits from neighborhood i . If the city limits lie at a share $\vartheta_i \in (0, 1)$ of the travel time between i and k , then travel time to the city limits is simply $\tilde{X}_i \equiv \vartheta_i t_{ik}$. We normalize travel time to the city limits at 0 by assigning negative values for travel time from suburban neighborhoods, $X_i = \tilde{X}_i [2 \times \mathbf{1}(i \in \mathcal{C}) - 1]$. We use this definition of travel times to the city boundary to state the two following continuity assumptions.

3.3.1 Continuity assumptions for RMA terms

We assume that two objects are continuous: 1) mean log RMA to the city ($\ln \Phi_{i,t}^{\mathcal{C}}$) conditional on travel time to the city limits, and 2) mean change in total RMA inclusive of border effects and taxes ($\Delta \ln \Phi_{i,t}$) conditional on travel time to the city limits. Continuity in mean log RMA to the city conditional on travel time to the city limits can be stated as

$$\lim_{x \uparrow 0} \mathbf{E}[\ln \Phi_{i \in \mathcal{S}, t}^{\mathcal{C}} | X_i = x] = \lim_{x \downarrow 0} \mathbf{E}[\ln \Phi_{i \in \mathcal{C}, t}^{\mathcal{C}} | X_i = x] = \varphi_t^{\mathcal{C}}, \quad (8)$$

which says that, standing at the city limits, average (log) workplace attractiveness across city workplaces varies smoothly if measured in the tract one step to the right, or in the tract one step to the left.¹⁴ Intuitively, this assumption states that access to jobs in Philadelphia is continuous along the roads that cross the city boundary.

Continuity in mean changes in log (total) RMA conditional on distance to the city limits can

¹³In our context, conditional expectation operators should not be misinterpreted as reflecting uncertainty in the model.

¹⁴Equation (8) may appear to hold trivially, but note that RMA to the city ($\Phi_{i,t}^{\mathcal{C}}$) is defined in terms of bilateral commuting costs (d_{ij}) and not in terms of travel times to the city limits (X_i). We do not redevelop the model in terms of a continuous topography as in [Allen and Arkolakis \(2014\)](#), which would allow us to more rigorously state that RMA is distributed continuously in space. Even if we did, however, our empirical implementation would likely be identical. As [Allen and Arkolakis](#) explain, “to calculate an equilibrium [in the data], it is necessary to approximate the continuous space with a discrete number of locations.”

be stated as

$$\lim_{x \uparrow 0} \mathbf{E}[\Delta \ln \Phi_{i \in \mathcal{S}, t} | X_i = x] = \lim_{x \downarrow 0} \mathbf{E}[\Delta \ln \Phi_{i \in \mathcal{C}, t} | X_i = x] = \Delta \varphi_t, \quad (9)$$

which says that the change in RMA inclusive of taxes and border effects must change smoothly at the city limits. Intuitively, this assumption says that the change in access to jobs in the entire metro area is continuous along the roads that cross the city boundary. Given that the tax and border effect terms of log (total) RMA would appear to induce discontinuities on their own, (9) must be tested empirically through various proxies.

With these continuity assumptions in hand, we define our spatial RD estimator (δ_{RD}) with distance to the city limits (X_i) as the running variable and the log change in neighborhood commuting shares to the city ($\Delta \ln \lambda_{i \in \mathcal{C}, t}$) as the outcome. Using (6) and (7) as outcomes in an RD estimator (Lee and Lemieux, 2010), and plugging in our continuity assumptions from (8) and (9), yields the following expression,

$$\begin{aligned} \delta_{RD} &\equiv \lim_{x \uparrow 0} \mathbf{E}[\Delta \ln \lambda_{i \in \mathcal{C}, t} | X_i = x] - \lim_{x \downarrow 0} \mathbf{E}[\Delta \ln \lambda_{i \in \mathcal{S}, t} | X_i = x] \\ &= \left(\varphi_t^{\mathcal{C}} - \varphi_{t_0}^{\mathcal{C}} - \Delta \varphi_t \right) - \left[\theta \ln \left(\frac{1 - \tau_{\mathcal{C}, t}^L}{1 - \tau_{\mathcal{C}, t_0}^L} \right) + \left(\varphi_t^{\mathcal{C}} - \varphi_{t_0}^{\mathcal{C}} - \Delta \varphi_t \right) \right] \\ &= -\theta \ln \left(\frac{1 - \tau_{\mathcal{C}, t}^L}{1 - \tau_{\mathcal{C}, t_0}^L} \right), \end{aligned} \quad (10)$$

which links the discontinuity in the change in log commuting shares to the city estimated at the city limits (δ_{RD}) to the model parameter (θ) and the observed wage tax rates ($\tau_{\mathcal{C}, t}^L$ and $\tau_{\mathcal{C}, t_0}^L$).¹⁵

Equation (10) delivers three intuitive predictions. The first two relate to the sign of the discontinuity. When the wage tax rate increases, suburban neighborhoods (to the left of the city-limits threshold) face a rising penalty to commute into the city, yielding a positive discontinuity. Second, when the wage tax rate is falling, suburban commuters face a falling penalty to commute into the city, yielding a negative discontinuity. The third insight is that dividing the RD estimate ($\hat{\delta}_{RD}$) by the log change in net-of-tax rates ($\ln(1 - \tau_{\mathcal{C}, t}^L / 1 - \tau_{\mathcal{C}, t_0}^L)$) yields the negative of the commuting elasticity (θ).¹⁶ The expression thus suggests a fuzzy RD design (where the first stage is the wage tax variation at the city limits) to estimate the elasticity of commuting to the net-of-tax rate.

Several confounders are held constant in (10). Confounders affecting residential choice (which vary by i and t) are held constant by using commuting probabilities conditional on residence in (3). These confounders include changes in the local cost of living (Handbury, 2021), in school district

¹⁵We use the log change in commuting shares to the city as our principal outcome, but we could instead define a log odds ratio for two periods to remove the change in total RMA. Because the initial shares for 1960 and 1970 are sometimes equal or close to 1, however, the log odds ratio becomes numerically unstable. When using 2003 as the base year, we obtain similar results to our main specification. Observe that, had our outcome been cross-sectional shares (4) and (5) instead of their changes, the RD estimate would be confounded by the border effect (ζ).

¹⁶Implicitly, we are assuming workers find it is not optimal to illegally evade the wage tax. While we have no data to test for tax evasion, the assumption is not unreasonable given the enforcement tools described in Section 2.1.

quality (Bayer et al., 2007), or in amenities (Almagro and Domínguez-Iino, 2025). Confounders affecting workplace choice (j, t -varying) are handled by the continuity assumptions in (8) and (9). Examples of these confounders include deindustrialization (Hanson and Livas, 2025), corporate taxes such as Philadelphia’s Business Income & Receipts Tax (BIRT) that may pass through to wages (Giroud and Rauh, 2019; Risch, 2024), or the general equilibrium effect of the wage tax on wages and workplace amenities. A feature of our framework is that there is a unique wage in each j , so forces like deindustrialization and corporate taxes affect the workplace wages of city and suburban workers equally. Finally, time-invariant confounders affecting commuting between municipalities are handled by taking changes in commuting probabilities in (6) and (7). Examples here include road roughness (Currier et al., 2023) and infrastructure quality (Loumeau, 2023; Bordeu, 2023). Any remaining confounders must differentially affect suburb-to-city commuting and in the same direction as changes in the net-of-tax rate.

There are a few key limitations to (10). Our expressions originate from identical workers whose different preferences for residences and workplaces are drawn from the same Fréchet distribution. But preferences for some workers could be drawn from different distributions: the commuting elasticity may vary by skill (Kreindler and Miyauchi, 2023; Tsivanidis, 2024) or demographic characteristics (Bunten et al., 2024). Second, with no dynamics in this model, workers and firms adjust immediately to changes in the wage tax rate. In reality, adjustments to changes in tax policy take time, particularly when mediated by job switching (Kleven et al., 2023).¹⁷ Third, differences in floor space supply elasticities between cities and suburbs may affect the ease with which workers and firms can switch locations (Glaeser and Gyourko, 2005, 2025). Finally, expectations about future wage tax rates, for which we have no data, may also be a determinant of residential-workplace decisions (Coglianese et al., 2017).

4 Data

We provide an overview of the data sources we use for the analyses in the rest of the paper. Appendix Section B details our sources and data construction. We detail the sources for the key outcome, the change in the share of neighborhood residents that commute to Philadelphia, and our sources for tax rates and other outcomes. We also describe the implementation of our spatial RDD, which uses the road network to measure travel times to the city boundary. We conclude with some balance checks and summary statistics for the city boundary sample and the Philadelphia metro area in 1940, the year the wage tax was implemented.

4.1 Census tract commuting data

Residential commuting shares in the 1960, 1970, and 1980 censuses The 1960 U.S. Census was the first to provide workplace information at the city, county, and state levels. Combining 1960

¹⁷In a dynamic model with adjustment costs and depreciating building capital, Heblich, Redding, and Sturm (2020) (Online Appendix F) obtain commuting probabilities that yield exactly the same predictions as their static model.

and 1970 tract-level place-of-work tabulations from the Elizabeth Mullen Bogue file with scanned versions of Census reports for those two years allows us to identify flows by workplace county for a given census tract of residence (see Appendix Figures B.27 and B.28). For 1980, we combine census tract tabulations that include total commuting, commuting outside the state of residence, and commuting to the SMSA's central city; we deduct out-of-state commuting from total commuting to exclude NJ. Census tabulations before 1990 do not allow us to construct tract-level workplace commuting shares.

Census tract-to-tract commuting (1990 to 2016) We use four publicly available Census-based data products for bilateral commuting flows: the 1990 and 2000 Census Transportation Planning Packages (CTPP), and the 2006-2010 and 2012-2016 bilateral commuting samples from the ACS. To preserve confidentiality, tract-to-tract flows of fewer than 3 records are suppressed in these data sets. The 2000 CTPP is the earliest data set in which tract-to-tract commuting can be further split by race and nominal income bins, enabling us to compute race- or income-bin-specific workplace and residential shares. Data suppression rules further apply to commuting split by characteristics, so more income- or racially-diverse neighborhoods may be subject to a higher degree of suppression. Subsequent ACS data are further subject to sampling error, which we do not take into account. We refer to the ACS samples by their middle year (i.e., 2008 and 2014).

LEHD Origin-Destination Employment Statistics (LODES, 2001 to 2019) We complement our Census and ACS measures with the LODES 7 data sets, which provide yearly job tabulations at the 2010 census-block level for 2001 to 2019. We use data for main jobs to measure commuting flows. LODES features no data suppression; to preserve confidentiality, synthetically generated flows are instead added at random to block-to-block flows each year. We follow [Owens et al. \(2020\)](#) and [Dingel and Tintelnot \(2021\)](#) by aggregating blocks to census tracts and taking three-year rolling averages of raw commuting flows. We refer to these three-year samples by their middle year (e.g., 2003, 2006, ..., 2018).

Imputation is a possible caveat of LODES in our analysis. The data are constructed with probabilistic imputation of establishments to workers, known as the unit-to-worker (U2W) procedure. State UI records identify firms using the State Employer Identification Number, but, with the exception of Minnesota, do not typically identify specific establishments. Using a model estimated with worker-establishment data from Minnesota, the U2W imputes establishments to workers for all other states based on exactly two variables: the distance between the place of work and place of residence, and the distribution of employment across establishments within multiestablishment firms ([Abowd et al., 2009](#)). A potential source of bias in principle, our RD results using LODES data are nonetheless robust to the inclusion of controls for this type of establishments.

To compute our commuting outcomes, we harmonize the data to 2010 census tract definitions using area-based crosswalks following ([Lee and Lin, 2018](#)). Census tracts never straddle a county border. For our spatial RDD, we only use commuting flows starting or ending in Bucks, Delaware,

Montgomery, and Philadelphia counties. We then construct log changes in the neighborhood share of commutes to or from Philadelphia. Since [Green et al. \(2017\)](#) find a lower rate of within-county commuting and longer commutes in LODES relative to the Census/ACS, we separately consider Census and LODES data in our analysis.

4.2 Tax rates for bilateral commutes

We calculate tax rates for all possible tract-to-tract commuting flows in the Philadelphia metro area from 1960 to 2018. We use historical statutory wage tax rates for each municipality from [Strumpf \(1999\)](#) and Pennsylvania’s Municipal Statistics agency. We use the NBER-TAXSIM model to compute state and federal tax rates for the median wage and salary income in the Philadelphia CZ, which we compute using Census/ACS data.¹⁸

We account for the institutional details in Section 2.1 to assign a wage tax rate to any given commute. From 1980 onward, after state income taxes in PA and NJ (and tax reciprocity agreements) were enacted, we include the flat personal income tax rate for Pennsylvania and the average income tax rate paid by the median worker in the Philadelphia CZ. We assume all workers residing in NJ claim a refund for taxes paid to Philadelphia. Because federal income taxes may magnify changes in the log net-of-tax rates due to the concavity of the log function, we also include the average federal income tax rate. To allow for some dynamic adjustment, we average the net-of-tax rates in years t , $t - 1$, and $t - 2$ for each base and end year. We consider no other taxes.

For our fuzzy RD, we construct a tax differential term using tax rates for bilateral commutes between Philadelphia and its suburbs. The term mimics the comparisons in our model. To construct the term for the share of residents commuting to Philadelphia, we compute the log net-of-tax rate paid by commuting to Philadelphia minus the log net-of-tax rate paid by commuting to the suburbs. The changes in this difference are discontinuous at the city limits. On the Philadelphia side, the changes always equal zero. On the suburban side, the changes reflect the changing gap between Philadelphia’s nonresident wage tax rate and other municipalities’ wage taxes.¹⁹ Analogously, to construct the term for the share of residents commuting from Philadelphia, we compute the analogous log net-of-tax rate paid by commuting from Philadelphia minus the log net-of-tax rate paid by commuting from the suburbs.²⁰

4.3 Other data sets for reduced form analyses

To study residential moving patterns, we use Infutor residential address histories (1990–2019), which cover most U.S. adults ([Diamond, McQuade, and Qian, 2019](#)). We do not observe place of

¹⁸We do not include the SALT deduction since it is only claimed by 5.7 to 13.2% of filers in our sample, which is in PA’s first to fifth congressional districts. See <https://www.congress.gov/crs-product/R46246>.

¹⁹To compute the net-of-tax rates for suburb-to-suburb commutes, we consider the average net-of-tax rate that would be paid in all possible tract-to-tract commutes within the Pennsylvania suburbs. We do not weight the net-of-tax rates by observed commuting patterns, but our procedure places more weight on municipalities with more census tracts.

²⁰Note that, on the Philadelphia side, this term is not necessarily equal to 0, as differentials arise between Philadelphia’s resident and nonresident rates after 1984.

work information in the Infutor data. The moving rate to the central city is the share of residential moves between year $t - 1$ and year t whose new address is in the central city. We compute the log change in this moving rate between years 2003 and 2018.

To study outcomes for firms, we use annual Dun & Bradstreet Private Company Listings (1969–2023), which provide establishment-level addresses, industries, ownership, employment, and sales. We geocode establishments and aggregate to census tracts. Our proxy for wages is the tract average of log sales per worker residualized by industry and state-year fixed effects.

Finally, we use census tract tabulations of median rents and house values (1960 onward), harmonized to 2010 tracts using area-based crosswalks and deflated with the CPI. Using our crosswalks, we take spatially-weighted averages to compute values at the 2010 census tract definitions across time, and deflate using the CPI. We then take log changes between census years.

4.4 Data for quantitative analysis

For counterfactual analyses, we impute historical land values and predict bilateral commuting flows. Land values are projected backward from 2018 estimates (Davis et al., 2021) using tract characteristics selected by LASSO.²¹ Our measure for model rents (R_i) is the predicted land value for a 1 acre lot. We combine the 2018 LODES sample with a bilateral travel time matrix for 2017 (McDonald, 2017) and indicators for cross-county and cross-state commutes, and estimate predicted bilateral commuting shares using PPML estimation following Redding (2024).

4.5 Spatial RDD implementation

Three problems motivate our use of drive times along the road network to the city limits. The first and most natural concern is that Euclidean distance may not be robust to natural geography (e.g., rivers) that can act as barriers to commuting. Such barriers affect the supply of infrastructure and therefore travel time (Loumeau, 2023).

Second, Cattaneo et al. (2025) indicate that misspecification bias may arise near kinks of geographic boundaries when using univariate Euclidean distance as the running variable. The kinks truncate Euclidean distance, meaning that the conditional expectation function is not differentiable everywhere. Cattaneo et al. (2025) propose new methodological tools based on bivariate location to address this issue, but do not address concerns on natural geography.

Third, the unit of analysis in our context is an area, but location is usually measured as a point. To see why this matters, first suppose that distances are measured between the area centroid and the nearest point in the boundary, as is conventional (Keele and Titiunik, 2015). This approach could yield the same distance for a large tract that borders the city limits and for a smaller tract that doesn't, in which case a spatial RDD would treat the two tracts as comparable units. Suppose

²¹The selected predictors are log acres, the square of log acres, the product of log acres and log median house value, log rent, the square of log household income, and the square of log median house value. The post-LASSO R^2 equals 0.86. Our imputation approach implies that changes in land value across time come exclusively from changes in values for the underlying predictors.

instead that distances are measured from the closest point to the boundary, as in [Loumeau \(2023\)](#). This results in a distance measure of 0 for all tracts bordering the city limits, yielding a mass point at the threshold but almost no observations near it. Note that this measurement issue is not addressed by *point-wise* bivariate location. Our argument is that allocating a single point to an area may potentially bias the estimated conditional expectation function in a spatial RDD.

To address these three issues, we use an isochrone from points on the road network located at the boundary. The road network captures actual travel routes, neutralizing barriers from natural geography. While travel time may be truncated by individual roads (e.g., dead ends), travel times are never truncated along a city’s road network, addressing the key issue identified by [Cattaneo et al. \(2025\)](#). We use ArcGIS to generate drive-time isochrones around the Philadelphia boundary, which approximate the (log of) commuting costs in our model. We overlay the isochrones on census tracts, which creates multiple tract-level observations. This approach addresses point-wise location but introduces some inference issues. We thus weight each drive-time-surface-by-tract observation by its share of tract area and cluster standard errors at the grid-cell level. We use the same present-day road network for all years in our sample (1960–2018) because the suburbs tended to get earlier access to interstate highways than more urban tracts in Philadelphia. In addition, location decisions likely incorporated expectations about future road network expansions.²² Figure 2 shows our RD sample along the Philadelphia city limits with discrete intervals for drive times, while Appendix Figure A.8 presents the continuous 5-minute drive time isochrone.

4.6 Balance checks and summary statistics

We next show evidence that workplace and residential characteristics were continuous at Philadelphia’s city limits the year the wage tax came into effect. Several factors could be a potential source of bias by inducing discontinuities in resident or firm market access at the city limits. Two examples are natural amenities ([Lee and Lin, 2018](#)) or persistent historical shocks ([Ambrus et al., 2020](#)). Pre-policy differences in the business environment in Philadelphia relative to its suburbs could have also affected firm density at baseline ([Holmes, 1998](#)). And differences in land use regulation could also affect the price of floorspace, as well as residential and commercial development ([Turner et al., 2014](#)). For our spatial RDD to be valid, observable determinants of commuting around the Philadelphia county border must be continuous prior to the implementation of the wage tax.

To test for continuity of baseline characteristics prior to the wage tax, we digitize two historical sources. First, we digitize the 1938 Industrial Directory of Pennsylvania, which provides plant-level information on location, employment, and sector. We geocode plants for Philadelphia and surrounding counties. Second, we use full-count 1940 Census microdata linked to digitized enumeration district maps in the suburbs and for enumeration districts within Philadelphia from [Shertzer, Walsh, and Logan \(2016\)](#), yielding 773 districts within five minutes of the city boundary.

²²The sections of the I-95 connecting Bucks County to Northeast Philadelphia were completed as early as 1964, but protests and redesigns delayed completion of the highway through the city proper until 1979. Sections of I-76 and I-676 running along the Schuylkill River were completed by 1960, but more central segments (e.g., the Vine Street Expressway) would not be completed until 1991.

Table 1 tests for continuity of observable characteristics in a sample of 1940 enumeration districts within a five-minute drive of the Philadelphia County border. The first set of columns displays the relevant parameters in a spatial RDD: the estimate from the left (Suburbs), the estimate from the right (Philadelphia), and the discontinuity estimate (Δ). The second set of columns helps interpret the magnitude of the discontinuity by showing county means for the suburbs and for Philadelphia as a whole.

Table 1: Continuity of workplace and residential characteristics at city limits at baseline

	RD estimates at city limits			County means	
	Suburbs	Philadelphia	Δ	Suburbs	Philadelphia
<i>Panel A: Workplace characteristics, 1938</i>					
Manufacturing plants per km ²	0.024	0.720	0.696	0.304	15.875
Mfg employees per km ²	2	19	17	32	870
Mfg employees per km ² (\perp industry FEs)	8	51	44	31	874
<i>Panel B: Main residential characteristics, 1940</i>					
Log population per km ²	6.997	6.858	-0.139	5.359	8.621
Log residential rents	3.577	3.751	0.174	3.338	3.328
Log house values	8.576	8.480	-0.096	8.393	8.034
Log male wage (\perp Mincer ctrls)	7.169	7.181	0.013	6.945	6.886
<i>Panel C: Other residential characteristics, 1940</i>					
Share population nonwhite	0.027	0.081	0.053	0.050	0.127
Share population foreign born	0.095	0.129	0.033***	0.087	0.152
Household size	4.299	4.456	0.157	4.827	4.746
Homeownership rate	0.552	0.495	-0.057	0.530	0.394
Employment/population (males 16-64)	0.799	0.770	-0.029	0.786	0.722
Mfg emp/pop (males 16-64)	0.222	0.236	0.014	0.358	0.345
Share males 25-54 w/ some college	0.232	0.186	-0.046	0.095	0.051

Note: The table presents estimates for baseline workplace (Panel A) and residential characteristics (Panels B and C). The first set of columns shows the estimate for each variable at the suburbs' limits with Philadelphia (i.e., the estimate from the left), Philadelphia's limits with the suburbs (estimate from the right) and the discontinuity at the city limits. The final two columns show average values for the entirety of Delaware, Montgomery, and Bucks counties under Suburbs, and for all of Philadelphia County in the last column. The RD sample consists of the 773 enumeration districts within a 5 minute drive of the city limits. Standard errors to obtain statistical significance are clustered at the grid cell level. Omnibus test for discontinuities in Panels A, B, and C has a p-value of 0.17. See Appendix Figure A.25 for RD plots for selected baseline characteristics. Stars in the Δ column denote statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

There is balance across different workplace characteristics (Panel A). We detect no statistically significant discontinuities in the density of manufacturing plants or in the number of manufacturing employees in 1938. We obtain similar results for plant size adjusted for industry fixed effects, suggesting our result for raw manufacturing employee density are unconfounded by industry sorting. The magnitudes of the discontinuities are statistically insignificant and are only about 2-4% of the average densities observed in Philadelphia. Altogether, we find no evidence of statistically or economically significant differences in workplace densities at the city limits in 1938.

There is also balance across the chief residential determinants of commuting in the 1940 Cen-

sus (Panels B and C).²³ In Panel B, we estimate no statistically significant discontinuities for log population density, log rents, log house values, or log residualized wages for males ages 16-64.²⁴ Because the 1940 Census collected no data on housing characteristics (reflected in the somewhat imprecisely estimated and opposite-signed discontinuities for the two raw housing cost variables), we examine some demographic correlates in Panel C. Household size is virtually identical on either side of the city limits, and we detect no statistically significant discontinuities for the nonwhite population share or working-age male employment rates. We only detect a statistically significant discontinuity of 3 *p.p.* in the foreign born share of the population. With that exception, we find no evidence of discontinuities at the city limits for other variables in Panels B and C.

We pool all discontinuities in an omnibus test and obtain a p-value of 0.17, which supports the assumption that baseline workplace and residential determinants of commuting were continuous at the city limits the year the wage tax was introduced. Since spatial RD designs may mask geographic heterogeneity (Keele and Titiunik, 2015), we split the sample into 5 clusters of enumeration districts at the city limits and iterate our balance tests in Appendix Table A.3. While we find some statistically significant discontinuities across a few secondary residential characteristics in some subsamples, this exercise continues to support balance for key dimensions of sorting. Across all subsamples in Appendix Table A.3, we show evidence supporting continuity in manufacturing plant and population density, residential rents, house values, residualized male wages, and the share of the population that is nonwhite. The region around the city limits was more similar to the collar counties of Philadelphia than to the city. The area had a light manufacturing presence in 1938, and specialized as a place of residence for richer and more educated individuals relative to the rest of Philadelphia County.

5 Nonparametric evidence at the city limits

We show nonparametric evidence of changes in neighborhood commuting shares at the city boundary consistent with the expressions derived in Section 3. We show evidence of discontinuities at the boundary of Philadelphia as the wage tax rose and fell, at the boundaries of other cities with wage tax variation, and in placebo cities, as well as heterogeneity in these discontinuities by worker characteristics. We end the section by providing evidence in support of the exclusion restriction to estimate the elasticity of commuting to the net-of-tax rate.

²³The year 1940 is not a strict pre-period since the wage tax went into effect on January and the Census was lifted in April. But the 1940 Census was the first to collect information on wages and education, which is important to assess sorting at the city limits. To the extent that households were imperfectly mobile within that four-month time-span, we can use the 1940 Census to test our identification assumption.

²⁴We adjust wages with indicators for education and race, and an age quartic using the universe of working-age male workers in Philadelphia and its collar counties in 1940.

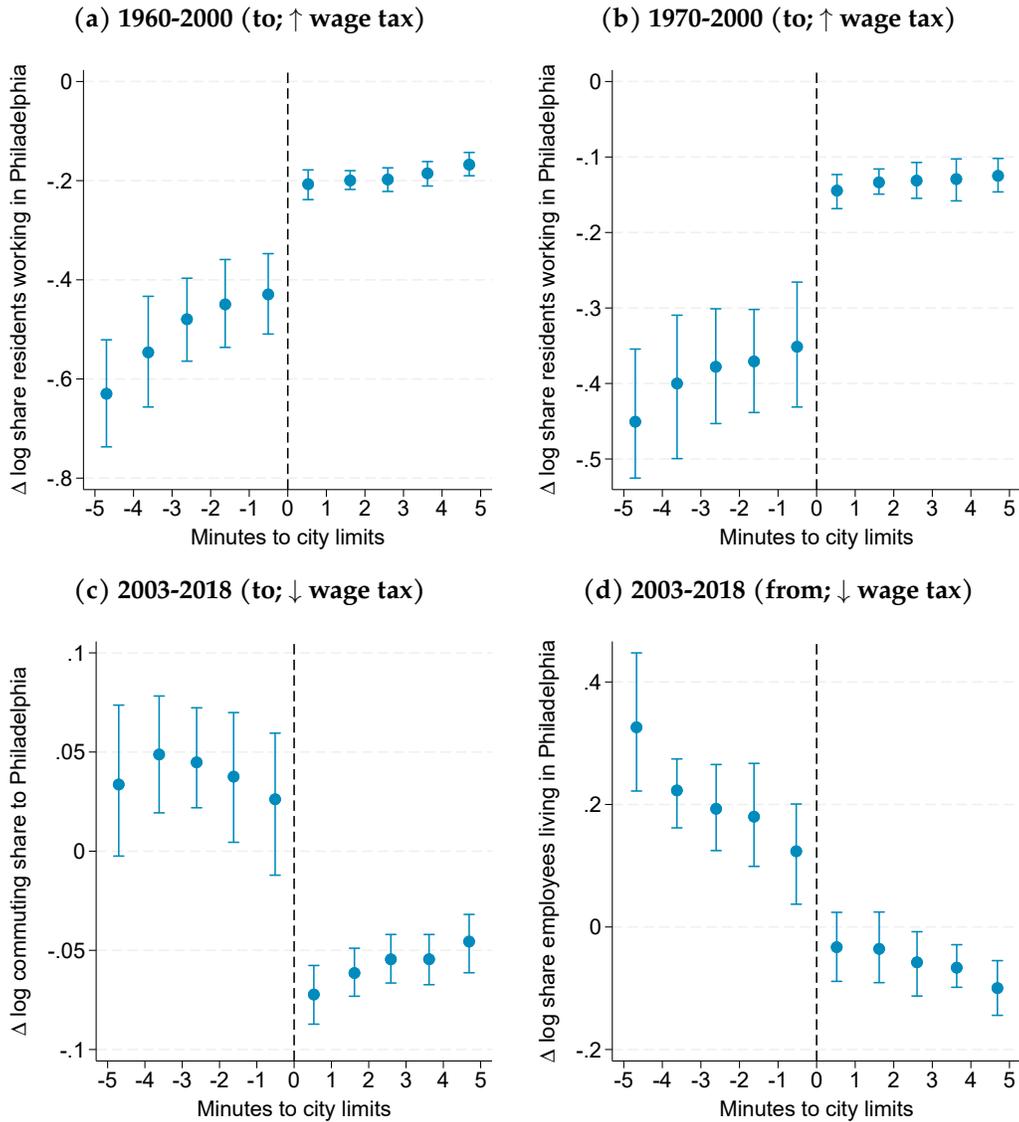
5.1 Changes in the share of residents working in the city

Rising wage tax rates Figure 3a shows our RD plot for the change in log share of residents commuting to Philadelphia between 1960 and 2000. Philadelphia neighborhoods are to the right of x-axis value of zero, marked by the dashed line. In Philadelphia neighborhoods, the log commuting share to Philadelphia on average fell by 20 log points, consistent with declining workplace attractiveness in the city during this period. As the nonresident wage tax increased from 1.5 to 4% in this period, however, so did the penalty of commuting into Philadelphia from the suburbs. In suburban neighborhoods, the change in the share of commutes to Philadelphia falls discontinuously relative to Philadelphia neighborhoods. Consistent with our theoretical prediction in (10), we estimate that the rising wage tax induces a positive discontinuity at the city limits of 10 log points ($SE = 3.7$). A similar result holds with 1970 as the base period (Figure 3b).

Falling wage tax rates Figure 3c repeats the exercise over 2003-2018. Philadelphia's nonresident wage tax rate fell while several of its neighbors introduced new wage taxes, reducing the average tax gap between Philadelphia and its suburbs by more than a full percentage point. Per (10), these narrowing tax differentials would discontinuously increase commuting into the city from the suburbs. Consistent with that theoretical prediction, the sign of the discontinuity is now negative. We estimate a discontinuity of -11 log points ($SE = 1.67$).

We obtain similar results when iterating this exercise for the log change in the share of workplace employment commuting *from* Philadelphia for 2003-2018 (Figure 3d). Conditional on working in the suburbs, Philadelphia became a discontinuously more attractive residence due to the falling resident wage tax rate. Figures 3c and 3d support the notion that the two different outcomes are similarly affected by the wage tax.

Figure 3: RD plots for change in log commuting share to or from Philadelphia



Note: Spatial RDD plots display changes in log commuting shares for tracts within a 5-minute drive of Philadelphia’s city limits. Positive x-values are inside the city. Panel (a): change in log share of residents commuting to Philadelphia, 1960–2000. Panel (b): change in log share of residents commuting to Philadelphia, 1970–2000. Panel (c): change in log share of residents commuting to Philadelphia, 2003–2018. Panel (d): change in log share of employees commuting from Philadelphia, 2003–2018. Standard errors clustered at the grid-cell level.

Appendix Figures A.9 to A.12 show robustness to the choice of end period for each of the three base periods, a total of 21 RD plots. With one exception, we find discontinuities that are positive as the wage tax rose and negative as the tax fell. The exception is Appendix Figure A.11a, where we find a positive discontinuity for 2003–2006 despite a falling wage tax. We attribute this discontinuity to the fact that there is a positive discontinuity in the change in log total commuting flows, which, as we discuss below, violates our continuity assumption in (9). We discuss and test this assumption in Section 6.1.1 below.

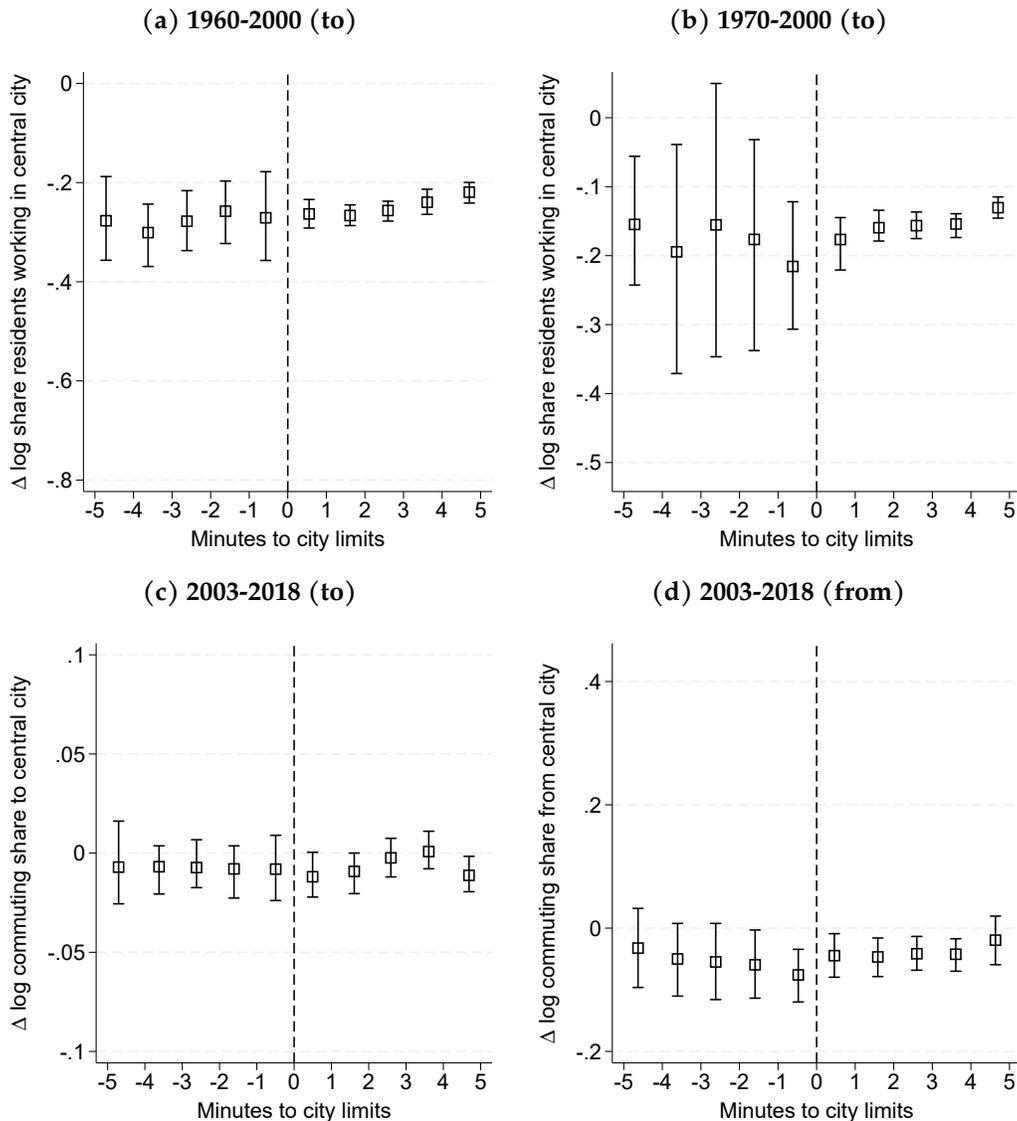
Changes in commuting shares in other cities with wage taxes To adjudicate the discontinuities seen at the Philadelphia city limits to the wage tax, we examine other cities that also adopted a wage tax. In Appendix Section E, we perform separate case studies for Cleveland and Detroit, which implemented wage tax ordinances in 1966 and 1962. One strength of this analysis is that, by taking 1960 as the base year, we can study the effect of enacting a wage tax. One caveat is that, while we have historical wage tax variation for these two cities, we lack good data on historical wage tax rates in neighboring jurisdictions. We detect positive and statistically significant discontinuities as the wage tax rose in Cleveland and Detroit. We also detect a large, but noisy, negative discontinuity as the wage tax fell modestly in Detroit, and no significant discontinuity for Cleveland in periods in which the wage tax did not vary. The results for Cleveland and Detroit are consistent with both our theoretical framework and our empirical estimates for Philadelphia, and indicate that Philadelphia is not the only city with a wage tax that experienced discontinuous changes in commuting composition at the city limits.

Changes in commuting shares in placebo cities As a validation test for the same base and end periods in Figure 3, we ask whether the discontinuities we document above are present in jurisdictions where labor earnings are not taxed differentially across the city limits. Selecting jurisdictions comparable to Philadelphia in terms of temperature, population size, and industrial and demographic composition, our set of placebo cities includes Boston, Brooklyn-Queens, Chicago, and Minneapolis.²⁵ We construct analogous samples of census tracts within a 5-minute drive of the city limits. To account for level differences in the composition of commuting between cities, we remove city-year fixed effects from the outcomes prior to the spatial RDD analysis (Lee and Lemieux, 2010).

In Figure 4, we iterate the RD plots for the outcomes in Figure 3 for our (pooled) sample of placebo cities. We find no evidence of a discontinuity for the change in log commuting shares to the central jurisdiction over 1960 to 2000 (4a), 1970 to 2000 (4b), 2003-2018 (4c), or the change in log commuting shares from the central jurisdiction over 2003-2018 (4d). There is little evidence that the discontinuities we find for Philadelphia in Figure 3 were present in this sample of cities that did not adopt a wage tax. The results for placebo cities validate that the discontinuities in Figure 3 are particular to Philadelphia.

²⁵Boston, Chicago, and Minneapolis do not have a wage tax. New York City enacted a wage tax in 1966 and abolished the nonresident tax rate in 1999, but labor earnings have never been taxed differentially across the boroughs of Brooklyn and Queens, whose borders we use for this analysis. With the exception of New York City, we exclude historically comparable cities that enacted their own wage tax ordinances, such as Detroit, St. Louis, Newark, Pittsburgh, Cleveland, or Columbus. Outcomes with base 2003 exclude Boston due to LODS data availability. We also restrict to commuting flows starting or ending in the counties adjacent to the central city in each of these placebo jurisdictions.

Figure 4: Continuity check: RD plots for change in log commuting share to or from central city in placebo jurisdictions



Note: For a sample of placebo cities (Boston, Brooklyn-Queens, Chicago, and Minneapolis), the spatial RDD plots display changes in log commuting shares for tracts within a 5-minute drive of the city limits. Positive x-values are inside the central city. Panel (a): change in log share of residents commuting to the central city, 1960–2000. Panel (b): change in log share of residents commuting to the central city, 1970–2000. Panel (c): change in log share of residents commuting to the central city, 2003–2018. Panel (d): change in log share of employees commuting from the central city, 2003–2018 (excluding Boston). Displayed data are orthogonal to city fixed effects. Standard errors clustered at the grid-cell level.

Robustness: outcomes by race, income, commute time, and sector One possible concern is that the discontinuities in the changes in commuting composition are merely the result of changes in sorting on observables at the city limits, which in turn correlate with commuting decisions. Examples could include race, income, or preferences for commute time. Moreover, discontinuities could potentially arise because of (perhaps unobserved) changes to residency requirements for city gov-

ernment employees.²⁶ To test for these possibilities, we compute the log change in the commuting share to Philadelphia for different classes of workers and iterate our spatial RDD.²⁷

In Appendix Figure A.18, we split tract-to-tract commuting flows in 2000 and 2014 by two nominal earnings bins: one for workers earning up to 50,000 dollars, and another for workers earning more than that quantity. Displaying the RD plot for the change in the share of residents earning up to \$50,000 that commute to Philadelphia, Appendix Figure A.18a shows a negative discontinuity as the wage tax fell between 2000 and 2014. Appendix Figure A.18b also shows a negative discontinuity of similar size when considering commuters earning more than \$50,000. Using the same nominal earnings bins, but instead considering the change in the log share of employees commuting *from* Philadelphia, we also find negative discontinuities for workers in either group (A.18c and A.18d).

In Appendix Figure A.19, we split tract-to-tract commuting flows in 2000 and 2014 by race and recompute our outcomes. Negative discontinuities continue to appear when considering white residents (A.19a), or white or nonwhite employees (A.19c and A.19d). While we obtain an insignificant estimate for the change in the log share of nonwhite residents commuting to Philadelphia (A.19b), the other panels support the notion that the discontinuities we document in Figure 3 are not the indirect result of race- or income-based sorting.

In Appendix Figure A.20, we split changes in commuting shares in the LODES data by 1) commutes to/from tracts up to 15 minutes away, and 2) commutes to/from tracts more than 15 minutes. We use bilateral commute times on the uncongested road network. The results indicate that the negative discontinuities persist when comparing residents or employees with similar commuting times. However, the discontinuities are somewhat larger when we restrict to short commutes. While this is outside of our model, it is intuitive that workers who commute for longer may be less sensitive to changes in economic fundamentals like the wage tax.

In Appendix Figure A.21, we restrict commuting shares in the LODES data to the private sector. Reassuringly, we observe similar discontinuities to Figure 3 when we study changes in private sector commuting shares to (A.21a) or from (A.21b) Philadelphia.

5.2 Evidence in support of exclusion restriction

We examine several margins of adjustment coinciding with variation in the wage tax. We analyze changes in workplace density as the wage tax increased between 1970 and 1976. We also examine changes in residualized log sales per worker and in raw median house values to provide some suggestive results on the incidence of the wage tax.

Unlike changes in the composition of commuting above, the outcomes in this section may be affected by other policy variation in Philadelphia. For instance, workplace density may be affected

²⁶Residency requirements for city employees were relaxed by the Supreme Court in *McCarthy v. Philadelphia Civil Svc. Comm'n*, 424 U.S. 645 (1976). To our knowledge, no other large modifications have been effectuated since.

²⁷Note that this test may be sensitive to suppression rules in Census/ACS data, as we explain in Section 4, because more racial or income-diverse tracts may feature more suppression. Our data sources unfortunately have no information on commuting flows by college attainment.

by business taxes; rents may be affected by variation in property taxes. The interpretation of our results for mechanisms is therefore suggestive. We show continuity of the outcomes in placebo cities, however, to validate that what we document in this section is particular to Philadelphia. We are also deliberate in choosing base and end periods to ensure we the variation captured along the city limits can be reasonably attributed to changes in the wage tax rather than to other policy changes, as recommended in [Keele and Titiunik \(2015\)](#).

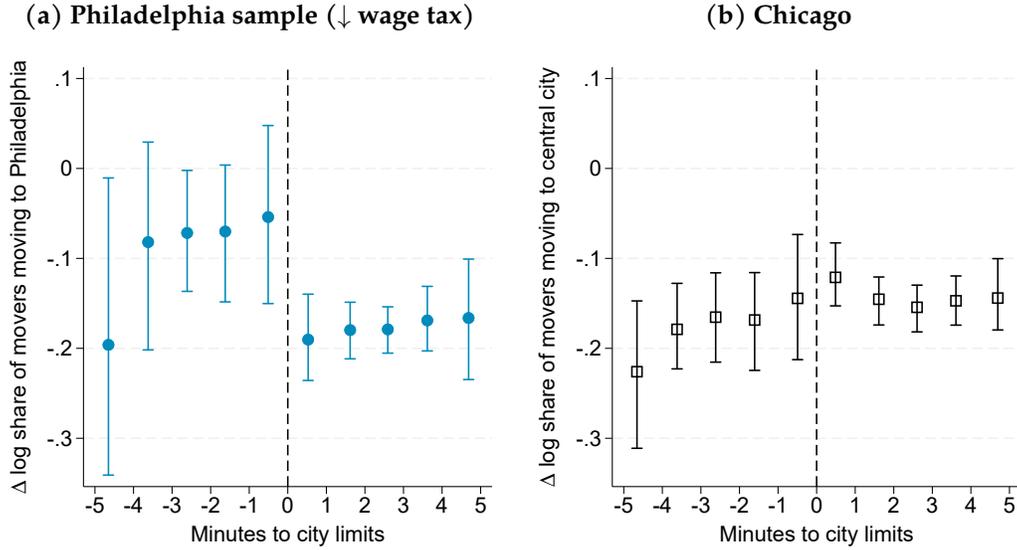
Changes in Infutor moving rates As Philadelphia decreased its residential wage tax, did outcomes other than the composition of commuting move in the same direction? We use geocoded Infutor data to study the change in a census tract’s movers that move to Philadelphia. By conditioning on moving as in [Agrawal and Foremny \(2019\)](#), we capture changes in the attractiveness of Philadelphia as a residence.

A related outcome to our changes in commuting shares, the change in moving rates to Philadelphia captures the sensitivity of workers’ residential choices to the wage tax, and holds constant time-invariant differences in neighborhood size or turnover. There are two caveats in interpreting this outcome. First, there is no expression for moving rates in the model, since the moving rate is equal to zero in equilibrium. Second, unlike the composition of residential commuting, changes in moving rates could potentially respond to changes in neighborhood rents and amenities, meaning spatial RDD estimates for this variable could include variation other than that induced by the wage tax.

Figure 5a displays an RD plot for the change in moving rates to Philadelphia between 2003 and 2018. Figure 5b does the same for Chicago, a placebo jurisdiction. We see a negative discontinuity in the log change in moving rates at Philadelphia’s city limits, but no such discontinuity for Chicago. We estimate a discontinuity of -12.5 log points per decade ($SE \approx 6.00$). While this impact is less precisely estimated and more suggestive than our main commuting-based outcome, the magnitude of the effect is consistent with the discontinuities in the change in log commuting shares for base year 2003.²⁸

²⁸Infutor data at the tract level is sparse prior to 2000.

Figure 5: RD plots for change in log moving rates to central city, 2003-2018



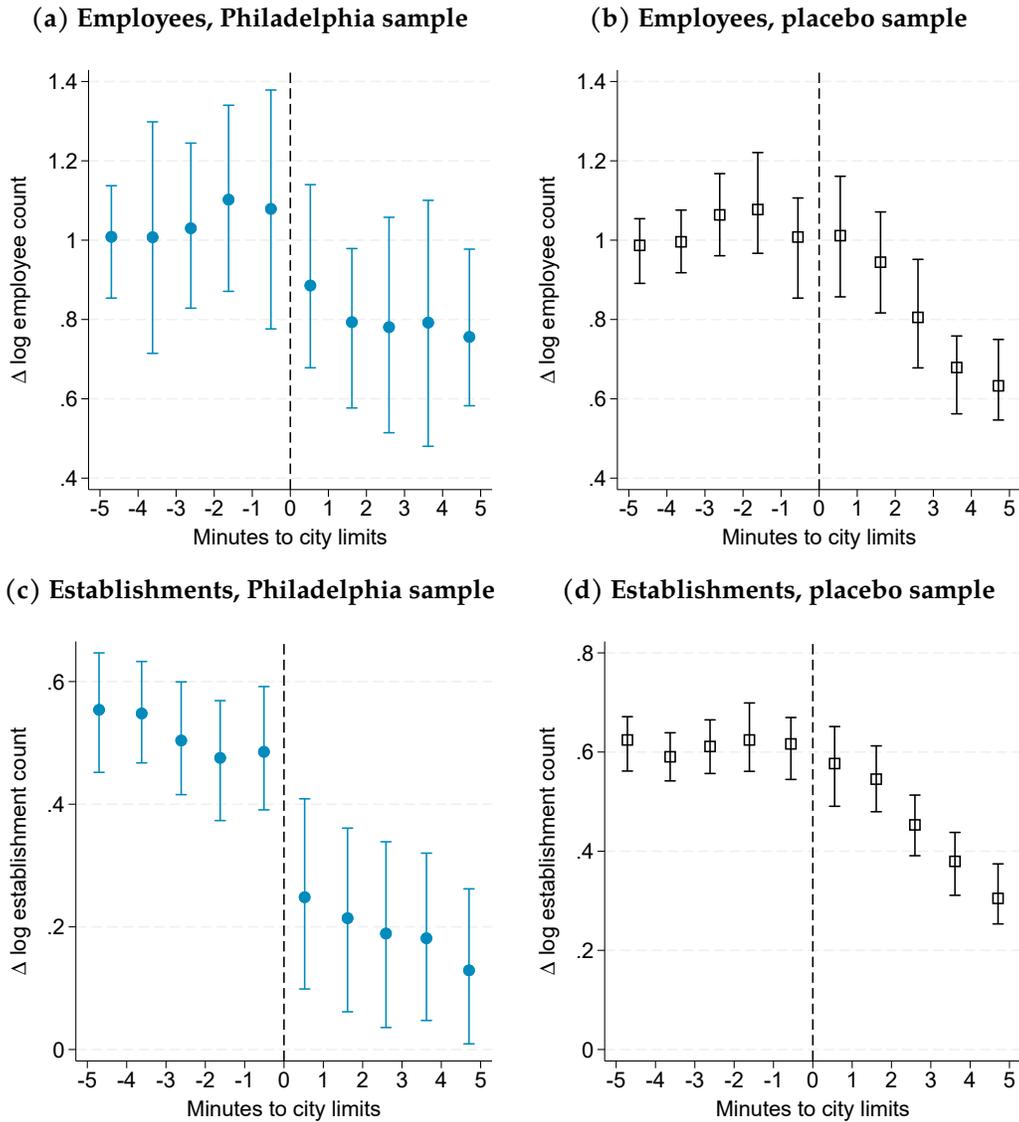
Note: Spatial RDD plots display changes in log share of movers moving to the central city for tracts within a 5-minute drive of each city’s limits. Positive x-values are inside each city. Panel (a): changes in log share of movers moving to Philadelphia, 2003–2018. Panel (b): changes in log share of movers moving to Chicago, 2003–2018. Standard errors clustered at the grid-cell level.

Changes in employment and establishment density The wage tax applies to both worker compensation and firm profits, as we discussed in Section 2.1. Labor demand is thus potentially reduced through two channels: by reducing employment within firms, and by reducing the number of firms. To isolate each channel, we study changes in the density of employees and establishments separately (Holmes, 1998; Suárez Serrato and Zidar, 2016), and restrict the analysis to changes between 1970 and 1976. The period features stark variation in the wage tax rate, no policy variation in property tax rates, and predates the implementation of Philadelphia’s Business Income & Receipts Tax (BIRT) in 1984.²⁹ We use our geocoded Dun & Bradstreet data, which begins in 1969, to compute changes in census tract employee and establishment counts between the two years.

Figure 6 shows our RD plots for the tract-level changes in log employee and log establishment counts. We find no statistically significant discontinuity in the change in employee count at the city limits (6a). Indeed, the log change in employee counts for Philadelphia is similar to that in the placebo sample (6d). In contrast, establishment growth was discontinuously weaker in Philadelphia by a magnitude of 40 log points (6c). At the same time, in our placebo sample, we see a similar rate of suburban establishment growth to that seen in the Philadelphia suburbs, but a smooth decline in establishment growth towards more central neighborhoods moving right along the x-axis (6d). These findings suggest the wage tax reduces labor demand because it reduces the number of small establishments rather than the number of employees.

²⁹Philadelphia’s property tax remained at 4.775% from 1970 to 1976. By 1980, it was at 6.175%.

Figure 6: RD plots for change in workplace density, 1970-1976 (\uparrow wage tax)



Note: For our Philadelphia and placebo samples, the spatial RDD plots display the 1970–1976 changes in log employee and in log establishment counts for tracts within a 5-minute drive of the city limits. Positive x-values are inside the central city. Panel (a): change in log employees (Philadelphia sample). Panel (b): change in log employees (placebo sample). Panel (c): change in log establishments (Philadelphia sample). Panel (d): change in log establishments (placebo sample). Placebo sample is composed of tracts along the city limits of Boston, Brooklyn-Queens, Chicago, and Minneapolis. Displayed data for placebo sample is orthogonal to city fixed effects. Standard errors clustered at the grid-cell level.

Changes in rents and wages If establishment growth was discontinuously weaker in Philadelphia, what happened to wages? Because we do not directly observe wages at the workplace census tract, we use our median residual log sales per worker measure as a wage proxy and show its 1970-1976 change along Philadelphia’s city limits in Figure 7. We see no statistically significant discontinuity for our wage proxy in Philadelphia (7a) or in our placebo sample (7b), although the sign of the discontinuity is small and negative in both cases. That we find no statistically signif-

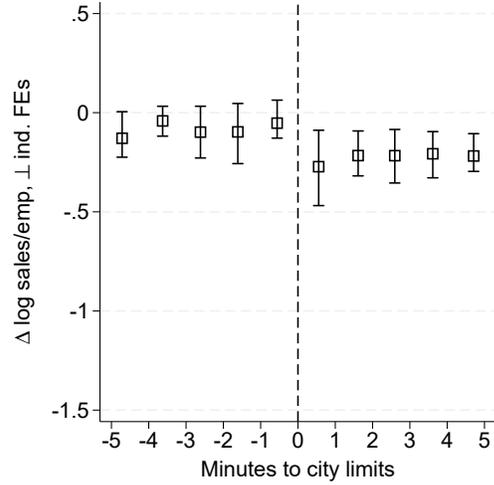
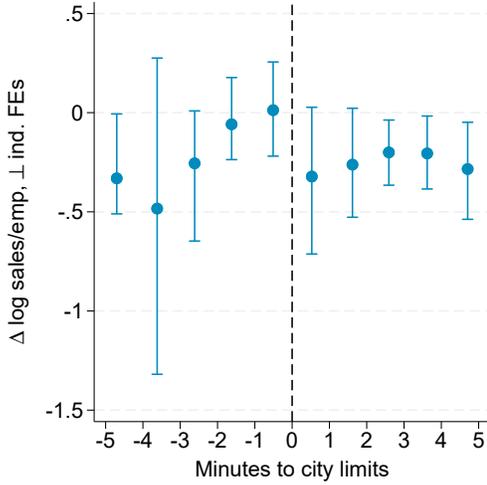
icant discontinuity in wage changes is also consistent with the lack of a significant discontinuity in workplace employment in Figure 6: workplaces along the border were not becoming discontinuously unattractive at the border. The graphical evidence suggests that wages did not respond significantly to variation in the wage tax.

In Figure 7c, we show the RD plot for the real change in log median contract rent in our Philadelphia sample over 1960-1970. In the intervening years, the increase in Philadelphia's property tax was small relative to future periods (3.8% to 4.775%), while the wage tax rate doubled (from 1.5% to 3%).³⁰ The change in residential rents over 1960-1970 discontinuously dropped at the city limits by 20 log points, suggesting that the incidence of the wage tax fell on landlords. While some of this variation may reflect changes in housing characteristics and to differential selection of properties into the rental market, the sign of the discontinuity suggests stark capitalization effects of the wage tax. Although much larger than what Brülhart et al. (2025)'s estimates for the tax-real estate elasticity would imply, the negative discontinuity we find in Figure 7c is consistent with their findings on the capitalization effects of local income taxes. Our findings suggest that the incidence of the wage tax is borne by immobile land, not mobile labor.

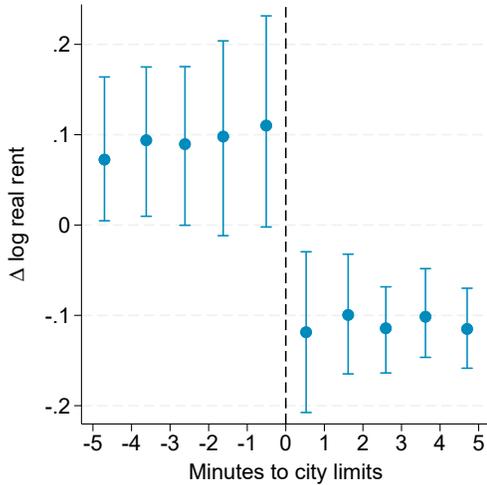
³⁰We unfortunately have no data on assessed property values in Philadelphia, which were likely significantly lower than the market values. Indeed, Philadelphia had increasingly higher property tax rates while its property tax revenues were declining sharply during this period (Figure A.5). That assessed values and property tax rates were moving in opposite directions complicates measuring the property tax burden in Philadelphia. See <https://whyy.org/articles/nutter-says-old-property-tax-system-is-dead/>.

Figure 7: Mechanisms: RD plots for change in factor prices

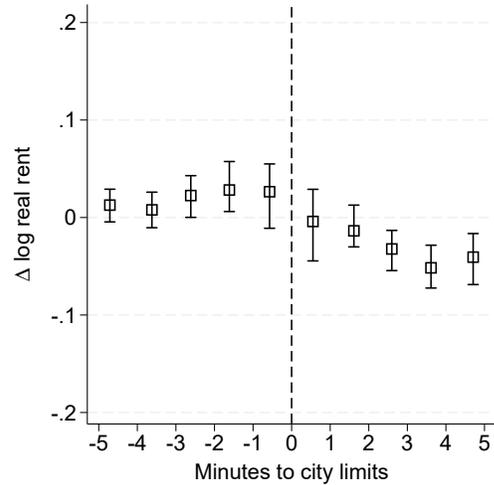
(a) Wage proxy, 1970-1976, Phila. sample **(b) Wage proxy, 1970-1976, placebo sample**



(c) Rent, 1960-1970, Philadelphia sample



(d) Rent, 1960-1970, placebo sample



Note: For our Philadelphia and placebo samples, the spatial RDD plots display changes in factor prices for tracts within a 5-minute drive of the city limits. Positive x-values are inside the central city. Panel (a): change in log sales per worker residualized with state-industry-year FEs, 1970-1976 (Philadelphia sample). Panel (b): change in log sales per worker residualized with state-industry-year FEs, 1970-1976 (placebo sample). Panel (c): change in log real median contract rent, 1960-1970 (Philadelphia sample). Panel (d): change in log real median contract rent, 1960-1970 (placebo sample). Placebo sample is composed of tracts along the city limits of Boston, Brooklyn-Queens, Chicago, and Minneapolis. Displayed data for placebo sample is orthogonal to city fixed effects. Standard errors clustered at the grid-cell level.

In summary, our nonparametric evidence at the city limits is consistent with our model. First, we find that the change in the share of residents commuting to Philadelphia exhibits a positive discontinuity at the city limits as the wage tax rises, and a negative discontinuity as the wage tax falls. We also find similar discontinuities in cities with wage tax variation, but no such discontinuities in placebo cities, and that the discontinuities we document in Philadelphia persist when computing both types of commuting shares for workers of comparable race, income, commute time, and

sector. We then provide evidence in support of the exclusion restriction to estimate the commuting elasticity with a fuzzy spatial RD design. Together with the institutional details discussed in Section 2, the evidence supports the relevance and exclusion restrictions needed to estimate the elasticity of commuting flows to the wage tax (Baum-Snow and Ferreira, 2015).

6 Elasticity of commuting flows to the wage tax

We estimate the elasticity of commuting flows to the wage tax by pooling the discontinuities at the Philadelphia boundary we document in Section 5.1. We do so using a fuzzy spatial RDD, as suggested by the model expressions derived in Section 3 and by the nonparametric evidence supporting the relevance and exclusion restriction in Section 5. We perform a variety of robustness checks and validate our estimate using a panel gravity equation and a structural approach.

6.1 Fuzzy spatial RDD estimates

We can obtain base- and end-year-specific commuting elasticities by taking each of the discontinuities in Section 5.1 and dividing them by the change in the net-of-tax rate. By stacking different end periods, we can instead estimate a pooled commuting elasticity for each base period in a fuzzy spatial RD design. To fix ideas, let $D_i \equiv \mathbf{1}(i \in \mathcal{C})$ be an indicator for census tracts in Philadelphia, and consider a simple linear specification,

$$\begin{aligned}\Delta \ln(1 - \tau_{i,t}^{\mathcal{P}}) - \ln(1 - \tau_{i,t}^{\mathcal{S}}) &= \alpha_t + \varrho_t X_i + \pi D_i + \rho_t D_i X_i + \varepsilon_{i,t} \\ \Delta \ln \lambda_{i\mathcal{C},t} &= \gamma_t + \varsigma_t X_i + \delta D_i + \nu_t D_i X_i + \xi_{i,t},\end{aligned}$$

where $\Delta \ln(1 - \tau_{i,t}^{\mathcal{P}}) - \ln(1 - \tau_{i,t}^{\mathcal{S}})$ is the change in the net-of-tax differential that would be paid by commuting to Philadelphia relative to the suburbs; $\Delta \ln \lambda_{i\mathcal{C},t}$ is the log change in commuting shares to Philadelphia; α_t and γ_t are end-period fixed effects; and ϱ_t , ρ_t , ς_t , and ν_t are end-period-specific slopes. Our estimate of the pooled commuting elasticity ($\hat{\theta}$) is the ratio of the reduced form ($\hat{\delta}$) to the first stage ($\hat{\pi}$).

We consider a PPML estimator to allow for zero-valued commuting shares in an end period and a Local Polynomial Fuzzy RD estimator (hereafter, Fuzzy RD).³¹ The PPML specification includes two quadratics in travel times to the city limits, one for either side of the threshold, fully interacted with end-year fixed effects.³² The Fuzzy RD specification partials out end-year fixed effects from the change in log shares, but does not allow for end-year-specific slopes. To compare census tracts with similar exposure to agglomeration forces, we include a control for log distance to Philadelphia City Hall across all specifications (Calonico et al., 2019). Lastly, to allow for spatial correlation

³¹Zero-valued commuting shares are the result of Philadelphia airport being expanded over a residential area over the course of the 1970s. The 8 affected census tracts only appear in the variation between 1960 and 1980, but our results are similar if we remove them to obtain a balanced panel.

³²We know of no tools that combine Poisson and local linear regression, so parametric assumptions are necessary. But a parametric RD design is not overly restrictive in light of Figure 3.

between neighboring tracts, we cluster standard errors at the grid cell level.

Main results Table 2 shows our main empirical results. The columns show results for different base years, changes in commuting shares, and specifications. Our estimates of the commuting elasticities on the first row are the ratio of the reduced form to the first stage, which we display in the rows immediately below.

Taking 1960 as a base year and using a fuzzy RD specification, we estimate that a 1% increase in the net-of-tax rate reduced commuting to Philadelphia by approximately 6.39% (SE = 1.39). We obtain higher estimates taking 1970 as the base year for tax hikes (10.7; SE = 1.95), when taking 2003 as the base year for tax reductions (12.07; SE = 2.53), or when using changes in the share of employees commuting from Philadelphia as the outcome (10.59; SE = 5.15). These results indicate that the magnitude of the estimate is somewhat sensitive to the choice of base year (and, hence, to the magnitude of the first stage), but much less sensitive to specification choice. Because of the stability of wage tax rates prior to 1960, our preferred estimate of 6.39 is the Fuzzy RD estimate for base year 1960 in Table 2 Column (2).

Table 2: Commuting elasticity estimates from pooled spatial RD designs

Base year Δ log commuting share Specification	1960		1970		2003			
	To Phila.		To Phila.		To Phila.		From Phila.	
	IVPPML (1)	Fuzzy RD (2)	IVPPML (3)	Fuzzy RD (4)	IVPPML (5)	Fuzzy RD (6)	IVPPML (7)	Fuzzy RD (8)
Commuting elasticity (θ)	5.11 (1.93)	6.39 (1.39)	11.17 (2.50)	10.69 (1.95)	11.30 (2.98)	12.07 (2.53)	10.60 (5.12)	10.59 (5.15)
Reduced form ($\times 100$) Δ log commuting share	10.68 (4.02)	13.47 (2.94)	15.39 (3.40)	14.99 (2.74)	-7.15 (1.80)	-7.73 (1.51)	-9.83 (4.94)	-9.83 (5.01)
First stage ($\times 100$) Δ log tax differential	2.11 (0.01)	2.10 (0.01)	1.40 (0.02)	1.39 (0.01)	-0.63 (0.03)	-0.64 (0.03)	-0.92 (0.07)	-0.92 (0.07)
Census Tracts	300	300	300	300	292	292	302	267
Grid cells	100	100	100	100	99	99	100	97

Note: This table displays pooled commuting elasticity estimates from RD designs pooling various end years. The outcome is the change in the log share of commuting **to** (or **from**) Philadelphia between a base period and end period. For the 1960 base, the end years are 1970, 1980, 2000, 2008, and 2014. For 1970, they are the same excluding 1970. For 2003, they are 2006, 2009, 2012, 2015, and 2018. Except for Columns (7) and (8), the outcome is the change in the log share of residents commuting to Philadelphia. The outcome in Columns (7) and (8) is the change in the log share of employees commuting from Philadelphia. Fuzzy RD specifications use the optimal bandwidth in [Calonico et al. \(2022\)](#). All specifications include a control for log distance to City Hall. Standard errors clustered at the grid cell level are shown in parentheses.

For context, we benchmark Fréchet spread parameter and tax elasticity estimates from several studies in Appendix Table A.1. The Fréchet spread parameter estimates from recent studies of *within-city* variation in Panel A range from 2.08 to 8.34 and average 4.7, implying our estimates lie

in the upper range. Analogizing our commuting elasticity with flow elasticities in the tax literature allows us to contextualize our results. Our estimates are much larger than those from studies of taxation and migration in Panel B, whose flow elasticities average 2.1. As we discussed previously, the variation studied in Panel B is typically *across* local labor markets, and, with the possible exception of Italian municipalities, the jurisdictions are typically more expansive than Philadelphia. Our results are most consistent with [Martínez \(2022\)](#)'s flow elasticity of 7.2 in Swiss cantons. Our findings suggest that location elasticities may be higher within local labor markets, and that Fréchet parameters from within-city studies may offer a conservative benchmark.

The strictest interpretation for the estimates in Table 2 is a local average treatment effect: they retrieve the elasticity of commuting to the wage tax at the city limits. In principle, our model in Section 3 permits a more general interpretation. If the location preferences of workers near the city limits are drawn from the same distribution as that for workers elsewhere in the city, then the shape parameter is a constant, and, therefore, the local average treatment effect would be equivalent to the average treatment effect.

Observable worker heterogeneity and sorting qualify the average treatment effect interpretation. In a version of this model with worker heterogeneity, [Tsivanidis \(2024\)](#) shows that the shape parameters within each class of worker are policy-invariant constants. Intuitively, our spatial RD estimate would then recover some weighted average of the shape parameters for the worker mix that predominates at the city limits. Although we find little evidence of pre-policy sorting at the city limits in Table 1, we also find some evidence that this region was, as a whole, whiter and more educated than Philadelphia County. The estimated elasticities would therefore correspond to whiter, more educated suburban compliers, which, as we discuss in Section 3.3, could be potentially lower than the elasticities for other groups. Since our data does not allow us to rule out that workers are drawn from different distributions, our commuting elasticity should be interpreted as a local average treatment effect.

6.1.1 Assumption and robustness checks

We test the continuity assumption in the total market access terms from Section 3.3. We also show robustness of our Fuzzy Spatial RD results to controls for initial worker and establishment heterogeneity, to different donut hole specifications, and to randomization inference with simulated city limits.

Assumption: continuity in the expected change in log RMA We check the continuity of (unobserved) RMA through several measures, most straightforward of which is the log change in total commuting flows originating from a census tract. For a given residence tract, total commuting flows depend not just on RMA but on neighborhood-level determinants such as rents and amenities. Equation (9) would be unlikely to hold, however, if there was *on average* a discontinuity in the change in commuting flow volumes at the city limits.

Appendix Figures A.13 to A.16 display, for each of the base years in Figure 3, the RD plots for the

associated changes in *total* commuting flows for various end years. We detect no discontinuities between 1960 and 2000 (A.13d), 1970 and 2000 (A.14c), and 2003 to 2018 (A.15e), supporting the assumption that RMA is continuous. Commuting measured at the place of work also varied smoothly at the city limits between 2003 to 2018 (A.16e), which also supports the assumption that FMA is continuous. With the exception of the log change in residential commuting between 2003 and 2006 (A.15a), the empirical results generally support our continuity assumption.

In Appendix Section A.17, we construct additional proxy measures of RMA and FMA. We use a travel time matrix for census tracts (McDonald, 2017) and tract-level workplace and residential employment counts for Bucks, Delaware, Montgomery, and Philadelphia counties. We consider different hypothesized values for the commuting elasticity ($\theta = \{1.25, 1.5, \dots, 12\}$), and for the border effect parameter ($\zeta = \{0.1, 0.2\}$), and the corresponding net-of-tax rates for each year. In the simulations, we only detect statistically significant discontinuities for log changes in RMA with small hypothesized commuting elasticity values (< 2.75).³³ Aside from those simulated values, there are no other statistically significant discontinuities in RMA or FMA for the time periods considered in Figure 3. The point estimates for the discontinuities are, at most, half a standard deviation of the change in log market access, but are generally less than a quarter standard deviation and statistically insignificant when the hypothetical value of the commuting elasticity is close to our empirical estimates. In summary, the assumption that the expected change in log RMA is continuous at the city limits is supported in the data.

Controls for worker and firm heterogeneity We test robustness to the possibly confounding effects of baseline sorting on observables. In Appendix Table A.4, we display our baseline estimate in the first column for base years 1960, 1970, and 2003, and then display the robustness of our results to the inclusion of the share of white population in the census tract, the college share of tract working age population, and the log median household income by tract (all measured at the base year or the closest census year before it) as covariates in our fuzzy RD design (Calonico et al., 2019). If changes in sorting were driving changes in commuting decisions, then conditioning on initial demographic, skill, and income composition would reduce the value of our estimated elasticities. As we add covariates, however, the resulting estimates vary little from our pooled specification in Table 2. Because imputation is used in the LODS data (base 2003), we include additional robustness checks for including the share of tract establishments that are corporate owned, or the share of tract employees that work in corporate-owned establishments. Our results for changes in workplace commuting from Philadelphia are robust to the inclusion of these variables.

Donut hole specification Appendix Figure A.22 shows that our fuzzy RD estimates from Table 2 are robust to donut hole specifications (Barreca et al., 2011). We sequentially exclude portions of census tracts within 0.2, 0.4, ..., 2 minutes of drive time to the city boundary. The results remain stable across these thresholds, with the exception of workplace shares in base year 2003, where

³³These small magnitudes, however, are outside the 95% CI of our estimates.

estimates become noisier once we exclude areas more than 1 minute from the boundary. Even in our strictest specification, where we exclude up to 40% of tracts, the elasticity estimates remain generally stable and consistent with our main specification.

Randomization inference with simulated city limits Following [Lehner \(2024\)](#), we simulate city limits for Philadelphia, compute Euclidean distance from a census tract to the simulated city limits, and estimate a spatial RD design for the base years in [Table 2](#). We repeat this procedure 2,500 times using straight-line boundary simulations that are tangent to the actual county border in at least one point. We then estimate the probability that we would have obtained reduced form estimates at least as extreme as those in [Table 2](#) had we randomly drawn the Philadelphia County border some other way.³⁴ In [Appendix Figure A.23](#), we show the CDF plots of RD estimates for all simulations of the Philadelphia city limits, and display the reduced form RD estimates from [Table 2](#) in dashed vertical lines. The randomization inference p-values for base years 1960, 1970, 2003 (residence), and 2003 (workplace) are 0.015, 0.01, 0.004, and 0.03. The small p-values allow us to reject the sharp null of no effect on the change in the composition of commuting.

6.2 Validation with gravity and structural estimates

Based on the model in [Section 3](#), we present two alternative methods to estimate the elasticity of commuting flows to the net-of-tax rate: a panel gravity equation that leverages variation in bilateral commuting flows across time, and a structural approach that requires additional assumptions on production. We derive these methods in [Appendix Section D](#).

6.2.1 Panel gravity equation

A panel gravity equation estimates the parameter of interest by controlling for bilateral commuting costs and absorbing variation in border effects time-varying residential and workplace determinants through fixed effects. It imposes no additional assumptions on the continuity of resident market access (as our fuzzy spatial RDD does) or on the structure of production, but at the cost of estimating thousands of additional parameters.

We use a panel of bilateral commuting flows in the Philadelphia metro area. We stack ACS/Census bilateral commuting flows for 1990, 2000, 2008, and 2014. Separately, we also stack LODES bilateral commuting data for every three-year period between 2003 and 2018, which may be preferable to the ACS/Census data since it features noise but no suppression. We impute zeroes to every tract-to-tract dyad with no recorded commuting flows. We stack log net-of-tax rates for every possible tract-to-tract dyad in the sample.

With these data in hand, we estimate the following extended gravity equation,

$$\lambda_{i,j,t} = \exp \left[\alpha_{j,t} + \gamma_{i,t} + \theta \ln(1 - \tau_{i,j,t}) + \theta \phi_t \ln d_{i,j} + \kappa_t \text{County}_{i,j} + \mu_t \text{State}_{i,j} \right] \quad (11)$$

³⁴“Extreme” in this context means larger than the reduced form coefficients for 1960 and 1970, and smaller than the reduced form coefficient for 2003. We only use the reduced form for this exercise.

where workplace-by-year fixed effects ($\alpha_{j,t}$) capture time-varying wages and workplace amenities, and residence-by-year fixed effects ($\gamma_{i,t}$) capture time-varying rents and residential amenities. Because we do not observe the opportunity cost of commuting at each point in time, we can only identify the parameter cluster $\theta\phi_t$, where $\phi_t > 0$ is the elasticity of commuting costs at time t to bilateral travel times in 2017. To flexibly estimate the border effect term, we allow for time-varying indicators for commutes that cross a state or county line (excluding Philadelphia, which we treat as a municipality). Equation (11) says that if tax rates are exogenous to census tracts, we can identify the parameter of interest (θ) by using a PPML estimator. We cluster standard errors at the municipality-by-municipality level, which is consistent with the level at which the tax rates are assigned as treatment (Abadie et al., 2023).

Table 3 shows our results for the extended gravity equation estimates of the commuting elasticity to log net-of-tax rates. We consider unconditional commuting probabilities from the Census/ACS data in Columns 1 to 3, and from LODES in Columns 4 to 6. All specifications include workplace tract-by-year and residence tract-by-year fixed effects, as well as some time-varying control for bilateral commuting costs. The estimated tax elasticities can be interpreted as comparing commutes with similar changes in amenity-adjusted cost of living and similar changes in amenity-adjusted wages net of some measure of commuting costs.

Table 3: Commuting elasticity estimates from extended gravity equation

	Log unconditional commuting probabilities					
	$\ln \lambda_{i,j,t}$ (1)	$\ln \lambda_{i,j,t}$ (2)	$\ln \lambda_{i,j,t}$ (3)	$\ln \lambda_{i,j,t}$ (4)	$\ln \lambda_{i,j,t}$ (5)	$\ln \lambda_{i,j,t}$ (6)
Log net of tax rate ($\ln(1 - \tau_{i,j,t})$)	9.50 (1.60)	12.92 (1.70)	20.47 (3.52)	8.57 (1.59)	12.39 (1.71)	12.10 (1.69)
Controls						
Travel time-by-year slopes	✓		✓	✓		✓
Road distance-by-year slopes		✓			✓	
Workplace-by-year fixed effects	✓	✓	✓	✓	✓	✓
Residence-by-year fixed effects	✓	✓	✓	✓	✓	✓
Workplace-by-residence fixed effects			✓			✓
County border-by-year fixed effects	✓	✓		✓	✓	
State border-by-year fixed effects	✓	✓		✓	✓	
Sample	Census/ACS	Census/ACS	Census/ACS	LODES	LODES	LODES
Estimation	PPML	PPML	PPML	PPML	PPML	PPML
Census Tracts	1,072	1,072	1,072	1,072	1,072	1,072
Municipalities	212	212	212	212	212	212

Note: This table shows PPML estimates with (log) unconditional commuting probabilities between home tract i and workplace tract j in year t as the outcome. Missing-valued bilateral commutes coded as 0; total observation count per year is 1, 072². Years are 1990, 2000, 2008, and 2014 for the Census/ACS sample. Years are 2003, 2006, 2009, 2012, 2015, and 2018 for the LODES sample. We use the yearly statutory wage tax rate for every possible tract-to-tract commute in Bucks, Camden, Delaware, Gloucester, Montgomery, and Philadelphia counties, and include state and federal taxes for the median wage and salary income in the Philadelphia CZ. Travel time and road distance data measured in 2017. Standard errors clustered at the level of residence municipality-by-workplace municipality shown in parentheses.

In Column (1), we obtain a commuting elasticity of 9.50 (SE = 1.60) when using our Cen-

sus/ACS sample and allowing for time-varying slopes for (uncongested) travel times between all census tracts. We obtain a higher and less precisely estimated elasticity when we instead control for road network distances in columns (2) and (5), which are a noisier control for commuting costs. The elasticities in columns (2) and (5) may thus be biased upwards, as both commuting shares and net-of-tax rates correlate negatively with travel times. Columns (3) and (6) partial out border effects by using workplace-by-residence fixed effects. This control overstates the magnitude of the elasticity given the suppression in Census/ACS data, but results in similar estimates when using LODES data. In Column (4), we obtain a lower elasticity of 8.57 (SE = 1.59) when we use our LODES sample with the travel time slopes and state and border effects. This is our preferred specification because the added noise in the LODES data does not bias the coefficient, whereas suppression of small commuting flows in the Census/ACS data may.

Because the results capture commuting between a state with a flat income tax (PA) and a state with a progressive income tax (NJ), a possible concern is that they may be sensitive to the choice of income percentile at which the tax rates are calculated. We test the robustness of our extended gravity equation estimates to tax rates at different percentiles for wage and salary income in the Philadelphia CZ. We calculate net-of-tax rates for each income decile and for the 99th income percentile in each year, and then iterate our estimation of (11).

Our estimates in Appendix Figure A.24 are remarkably stable if net-of-tax rates are calculated at any point between the 10th and 90th income percentiles. We only obtain smaller elasticities at the 99th income percentile, which is reasonable given that the average federal tax rates at high income levels magnify the log net-of-tax rates inclusive of the wage tax.

6.2.2 Structural estimates

We now turn to using the model structure to estimate the commuting elasticity. This method requires us to assume Cobb-Douglas technology for production and to assume values for several parameters. While it is the least transparent approach to estimate the elasticity, it can also be viewed as a way to validate the structure of the model, which we will rely on in Section 7.

The estimation procedure requires us to assume values for parameters on preferences and production, as we detail in Appendix Section D. In our model, a final good is produced in each neighborhood with Cobb-Douglas technology and three factors: commercial floor space (H_j^L), physical capital (M_j), and labor (L_j). Cost minimization implies that payments to each of these factors are constant shares of revenue (X_j):

$$w_j L_j = \beta^L X_j, \quad q_j H_j^L = \beta^H X_j, \quad r M_j = \beta^M X_j, \quad \beta^L + \beta^H + \beta^M = 1$$

We assume values for each of these parameters as of 2018 or the closest available year. We assume that the share of expenditure on consumption is $\alpha = 0.7$, which is consistent with the observed rent-to-income ratio in the Philadelphia CZ in 2018.³⁵ Although the labor share declined during

³⁵The median rent-to-household income ratio in the Philadelphia CZ in the 2018 ACS is 0.265. To compute this value,

our study period (see Grossman and Oberfield, 2022), we use the estimate in van Vlokhoven (2024) for 2019 and set its value to $\beta^L = 0.59$. We assume $\beta^H = 0.1$ for the floor space share of revenue, which is on the high end of lease-to-revenue ratio benchmarks for industries such as restaurants and retail.³⁶ The latter two values imply a capital share of $\beta^M = 1 - \beta^L - \beta^H = 0.31$.

We then proceed as follows. First, we take 2018 as our estimation year and use data on land values and predicted commuting shares for that year to estimate model-implied wages. To do so, we use the model’s land market clearing condition, which relates land values to wages using the assumed parameters above. Given the change in tax rates for a particular commute, the model yields a unique vector of wage changes for each location. We use this vector to solve for workplace employment going back to 1960 for a range of hypothesized values for the Fréchet shape parameter (θ). Finally, we calibrate the parameter by minimizing the sum of squared deviations between the observed and predicted shares of workers living and working in the suburbs, as in Figure 1a.

Appendix Figure A.26 compares the observed and model-predicted suburb-to-suburb commuting shares for the portion of the Philadelphia CZ we use for this quantitative exercise. With the exception of 2008, our calibrated model generally predicts the suburb-to-suburb commuting share with small errors. Across the range of hypothetical parameter values, the smallest error corresponds to a value of 6.5, which is consistent with our commuting elasticity estimates for 1960 in Table 2.

In summary, this section estimates the parameter that tells us how the wage tax directly affects location decisions holding wages, rents, and amenities constant. Our main fuzzy spatial RDD estimates are robust to including baseline demographic, income, and workplace controls (Appendix Table A.4), to different donut hole specifications (Appendix Figure A.22), and to randomization inference tests (Appendix Figure A.23). We estimate slightly higher elasticities in our extended gravity regression framework relative to our fuzzy spatial RDD. Our structural estimate of 6.5 is more consistent with our preferred estimate from our fuzzy spatial RDD. In the context of the literature, our pooled commuting elasticities lie in the upper range of Fréchet parameter estimates, but our 95% CIs admit estimates from seminal work (see Appendix Table A.1). We turn to using our parameter estimate to study the effects of the wage tax on the stocks of residents, employees, nonresident workers, wages, and land values below.

7 Stock elasticities and counterfactuals

In this section, we estimate how the *stocks* of residents, workers, and nonresident workers of Philadelphia respond to the wage once wages and rents adjust. Understanding the magnitude of the stock elasticity of the tax base is essential to assess the wage tax’s efficiency (Oates, 1972), job relocation (Fajgelbaum et al., 2019), and revenue effects (Kleven et al., 2020). We also use the model to assess how property, sales, and wage tax revenue changes with a range of counterfactual wage tax rates.

we only consider households with positive household income and rent values.

³⁶See <https://www.occupier.com/blog/lease-to-rent-revenue-ratio/>.

7.1 Stock elasticities from removing the wage tax

Stock elasticities cannot be estimated using changes in *unconditional* commuting probabilities [i.e., changes in (1)] along the city boundary. Using changes in unconditional commuting probabilities as the outcome of the spatial RDD would reintroduce neighborhood-specific characteristics that can be affected by policy variation other than the wage tax, weakening the causal interpretation of the estimates. Instead, we treat stock elasticities as a general equilibrium object that can be more rigorously estimated with the structure of the model.

The policy experiment used to estimate the stock elasticities is to set Philadelphia’s wage tax rates to 0. Held constant in this counterfactual are wage taxes in the suburbs, state and federal income taxes, and residential, workplace, and bilateral amenities. The outcomes that endogenously adjust are wages, land values, residents, and employees across neighborhoods in the entire commuting zone. Importantly, our counterfactual assumes that expected worker utility is the same across all locations, so welfare gains from abolishing the wage tax are entirely captured by absentee landlords.

While removing the wage tax may in practice affect the funding of public amenities, the City of Philadelphia could potentially tax land if counterfactual land values are larger than the foregone wage tax revenue. If this condition is met, Philadelphia’s transition to a land value tax would have no additional effects: since the gains from eliminating the wage tax are capitalized into land values, they can be taxed at zero cost to efficiency (Arnott and Stiglitz, 1979). The counterfactual change in land values we obtain is *before* the city steps in to tax land to replace wage tax revenue, so it is overstated and can be interpreted as an upper bound on welfare gains.

Using a commuting elasticity estimate of 6.5 and the assumed values for the share of expenditure on consumption and the shares of labor, land, and capital in production in Section 6.2.2, we estimate the effects of removing Philadelphia’s wage tax. We use 2018 as our base year and an open-city specification that assumes population mobility with the rest of the United States (Haughwout and Inman, 2001).³⁷ We consider three specifications. First, we assume a perfectly inelastic supply of floorspace and no agglomeration externalities, which guarantees a unique equilibrium (Heblich et al., 2020). In our second specification, we allow for a census tract’s supply of land to be elastic to land values. Third, we allow for elastic floorspace supply and productivity agglomeration forces, which allow neighborhood productivity and wages to increase with its own employment density. We borrow estimates of floorspace supply and productivity agglomeration elasticities from the literature.

Table 4 shows the results for removing the wage tax. Panel A displays the stock elasticities of several employment measures, as well as the revenue-maximizing rate of $\tau^* = (1 + \eta)^{-1}$, where η is the stock elasticity of the tax base to the net-of-tax rate (Kleven et al., 2020). Panel B displays the effect on several measures of land values; changes in land values across the metro area capture wel-

³⁷This population mobility condition is a function of the commuting elasticity, as we show in Appendix Section D. Intuitively, this condition tells us that expected utility for workers is the same regardless of where they live and work (including the rest of the U.S.). Welfare gains from the wage tax are thus entirely captured by landlords.

fare effects. Column 1 displays our base specification. Column 2 allows for a tract-level floorspace supply elasticity of 2.5 following [Baum-Snow and Han \(2024\)](#). Column 3 uses this floorspace supply elasticity and a productivity agglomeration elasticity of 0.04, following high-income-country estimates in [Ahlfeldt and Pietrostefani \(2019\)](#).

Table 4: Results from removing Philadelphia’s wage tax

	No agglomeration Inelastic floorspace (1)	No agglomeration Elastic floorspace (2)	Agglomeration Elastic floorspace (3)
<i>Panel A: Stock elasticities</i>			
Residential employment	0.64	1.59	2.19
Workplace employment	1.16	1.66	2.75
Tax base (resident and non-resident workers)	0.86	2.09	3.04
Revenue-maximizing tax rate (%)	53.88	32.32	24.77
<i>Panel B: Counterfactual land values</i>			
Philadelphia: multiple of wage tax revenue	12.28	12.69	13.66
Philadelphia: percent change	9.41	13.12	21.75
Suburbs: percent change	2.37	-1.20	2.79
Metro area: percent change (welfare)	4.44	3.01	8.37

Note: Stock elasticities are obtained by dividing the log change in employment by the log change in the average net-of-tax rate for Philadelphia residences, workplaces, and residences or workplaces. Tax base is the mass of employment living or working in Philadelphia. Revenue-maximizing tax rate equals $\tau^* = (1 + \eta)^{-1}$, where η is the stock elasticity of the tax base to the net-of-tax rate ([Kleven et al., 2020](#)). In Panel B, Philadelphia’s counterfactual land value is normalized by model-implied wage tax revenue in 2018. Percent changes are relative to observed values in 2018. Floorspace supply elasticity is 2.5 ([Baum-Snow and Han, 2024](#)). Production agglomeration elasticity is 0.04 ([Ahlfeldt and Pietrostefani, 2019](#)).

In the specification with no agglomeration and inelastic floorspace in column (1), the stock of workers living in Philadelphia increases by 0.64 log points for every log point increase in their net-of-tax rate. Philadelphia workplace employment has an elasticity that is nearly twice as large at 1.16. The stock elasticity of the tax base, which captures Philadelphia residents and its nonresident workers, is 0.86. This elasticity is the relevant input to obtain the revenue-maximizing tax rate, which is $54\% = (1 + 0.86)^{-1}$. Counterfactual land values in Philadelphia are at more than 12 times the model-implied wage tax revenue, validating the counterfactual exercise. Relative to their observed levels, land values would increase by 9.4% in Philadelphia, 2.4% in the suburbs, and 4.4% in the entire metro area. The open-city specification places all incidence on immobile landlords, whose welfare would increase even in *suburban* census tracts in this specification. The changes in census tract residential employment, workplace employment, and land values, displayed in [Figure 8](#), suggest that economic activity would significantly re-centralize even in absence of agglomeration forces or responses in the real estate market.

The specification in Column (2) allows for an elastic supply of floorspace that significantly magnifies the effect on the stocks of residents, employees, the tax base, and land values. Under this specification, the more elastic tax base implies a lower revenue-maximizing tax rate of 32%. Philadelphia’s counterfactual land values are more than 12 times the revenue raised by the wage tax and increase more than in the base scenario in Column (1). We obtain a net decline of 1.2% in suburban land values, however, implying that the total supply of floorspace in the suburbs would decline as economic activity re-centralizes, and a smaller welfare increase of 3%. This result highlights that the interaction between the wage tax and Philadelphia’s real estate market intensifies the behavioral responses of worker location.

In Column (3), we allow for production agglomeration in addition to an elastic floorspace supply. We obtain even starker responses across both panels. Here the estimated elasticity of the tax base is 3.09, which implies a revenue-maximizing tax rate of 25%. The increase in Philadelphia land values is the starkest across the three specifications at 21.75%. Combined with a more modest change in suburban land values of 2.79%, this implies that removing the wage tax would increase welfare by 8.37%. Put simply, removing the wage tax has the largest effects on economic activity when accounting for gains in productivity and in the built environment.

Across all three specifications, the endogenous responses of wages and rents compensate workers for higher taxes, and therefore reduce the magnitude of the stock elasticity of the tax base relative to the *ceteris paribus* commuting elasticity. That the magnitude of the stock elasticities is lower than the magnitude of the commuting elasticities is also consistent with studies of taxation and location decisions in Appendix Table A.1 Panel B. Nonetheless, the stock elasticities are higher when we allow for elastic floorspace supply and agglomeration forces in production.

How large are the magnitudes of the welfare effects and the revenue-maximizing wage tax rate? In a closely related exercise of tax harmonization across states, [Fajgelbaum et al. \(2019\)](#) obtain welfare effects of at most 1.2%. Depending on the specification, the magnitudes in Table 4 are 2.5 to almost 7 times larger than those estimates. Qualitatively, our welfare results are nevertheless consistent with previous analyses of Philadelphia’s wage tax and its national share of employment ([Grieson, 1980](#); [Haughwout and Inman, 2001](#)). Welfare gains from removing the wage tax are large but in line with previous findings.

Our estimates of revenue-maximizing wage tax rates in excess of 24% are larger than those in previous work. Interpreting the revenue-maximizing tax rate as the “top” of the revenue hill, our estimates imply that Philadelphia has never come close to it, whereas [Haughwout and Inman \(2001\)](#) find that Philadelphia would maximize wage tax revenue at rates slightly larger than 4%.³⁸ This suggests that, in equilibrium, Philadelphia would have been able to obtain more revenue by raising wage tax rates far beyond their observed values *despite* the significant costs to welfare and economic activity. To more rigorously explore this finding, we next consider a series of counterfactual wage tax rates and compute revenues from the wage tax, the property tax, and the sales tax.

³⁸A similar exercise in [Haughwout et al. \(2004\)](#) shows that wage tax revenue is increasing between 0 and 10%.

7.2 Total tax revenue for different wage tax rates

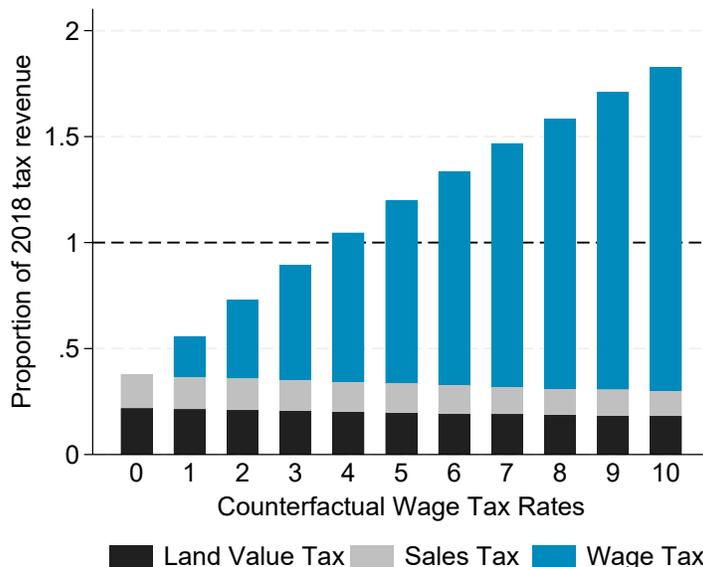
We now estimate how much *total* tax revenue Philadelphia would collect under different wage tax rates. Our findings that the wage tax reduces land values and incentivizes residents to move suggest that the property and sales tax bases are negatively impacted by the wage tax. We use the structure of the model to study these forces instead of estimating the revenue-maximizing tax rate with a simple formula that abstracts away from how the wage tax interacts with other tax bases.

We iterate the procedure in the previous section for counterfactual wage tax rates ranging from 0 to 10%. For simplicity, we assume a unique wage tax rate for residents and nonresident workers. We compute the wage tax revenue for each scenario allowing for endogenous adjustment of wages, rents, and commuting patterns. To account for land value tax revenues, we calibrate the land value tax rates such that their revenue equals observed property tax revenue as a proportion of wage tax revenue in 2018. To compute sales tax revenues, we assume that Philadelphia collects a constant share of residential income net of taxes and rental payments, and calibrate this share such that the sales tax revenue equals its observed proportion of wage tax revenue in 2018. In the counterfactuals, revenues for the land value tax and the sales tax are fixed proportions of land values and after-tax residential income net of rents; we do not model the effect that sales taxes have on the location of economic activity. We consider no other taxes such as BIRT or the component of the wage tax that applies to firm profits. These simplifying assumptions imply that the three taxes studied in this counterfactual empirically amount to 57% of Philadelphia's general fund revenue in 2018. The wage tax accounts for $\frac{2}{3}$ of that revenue.

Figure 9 shows revenue collected under the different counterfactual wage tax rates using a specification with no agglomeration and inelastic floorspace supply. The y-axis displays counterfactual revenue as a proportion of the sum of the wage tax, the property tax, and the sales tax revenue in 2018. The x-axis displays a range of counterfactual wage tax rates. We decompose total tax revenue in each scenario by property, sales, and wage taxes.

In the scenario in which wage taxes are abolished, revenues for land value and sales taxes amount to 38 cents for each dollar raised in 2018. As counterfactual wage tax rates increase from 0 to 10%, land value and sales tax revenues decline from 38 to 30 cents on the dollar, implying the wage tax acts as a fiscal externality by reducing other revenue sources. In contrast, as wage tax rates increase, wage tax revenues also increase, albeit at a diminishing rate. With a wage tax of 10%, Philadelphia would raise 83% more than it raised in 2018. The results for Figure 9 imply that Philadelphia's revenue-maximizing wage tax rate is substantially larger than 10%, even as it diminishes other tax bases.

Figure 9: Model-implied Philadelphia tax revenues under different counterfactual wage tax rates



Note: The figure shows results for a range of counterfactual wage tax rates using a specification with no agglomeration and inelastic floorspace supply. The y-axis displays Philadelphia’s tax revenues divided by the sum of the model-implied wage tax, sales tax, and land value tax revenue. We calibrate sales tax and land value tax rates to match the observed composition of the wage, sales, and property tax revenues in 2018. Calibrated sales and land value tax rates are held constant in each counterfactual scenario and equal fixed proportions of after-tax earnings and land values. Wage tax revenue is endogenous to model-implied commuting patterns for each counterfactual wage tax rate.

We assess the robustness of these results to alternative specifications in which we allow for elastic floorspace supply and agglomeration forces. We use the same parameters for the counterfactual above and display results in Appendix Table A.5. In either of these alternative specifications, we find that abolishing the wage tax would expand the property and sales tax bases by larger amounts than in our base specification. At the same time, higher counterfactual wage tax rates would result in larger declines in these two tax bases. The declines are starkest in the specification with elastic floorspace supply and production agglomeration forces. Since wage tax revenue more than makes up for these declines, we find that total tax revenue is nonetheless increasing in the wage tax rate between 0 and 10% under across the three specifications.³⁹ Floorspace supply and agglomeration forces each *magnify* the fiscal externalities of the wage tax.

A limitation of this exercise is that it does not consider the effect of the wage tax on other sources of tax revenue, such as local corporate taxes. Our estimates nevertheless highlight why it was possible for the City of Philadelphia to become heavily dependent on the wage tax in the first place: it raised more revenue despite its significant impacts on welfare and on other tax bases.

³⁹A wage tax of 10% is well above the wage tax rates historically used in Philadelphia or in other cities with local income taxes. The top marginal income tax rates for New York City and Washington D.C. are 10.9 and 10.75%, for instance.

8 Conclusion

Government in the United States has been slowly decentralizing since 1950 (Agrawal et al., 2024). As state and local governments have taken on more (and perhaps more expensive) responsibilities, so too has their need for tax revenue. Because local labor markets frequently straddle different tax jurisdictions, the effect of state and local taxation on location decisions has come to the forefront. In this paper, we used Philadelphia's wage tax as a case study of a particularly stark tax differential that affected commuting decisions within a local labor market. We argued that a quantitative spatial model of commuting provided an appropriate lens through which to study the wage tax.

We found that the elasticity of commuting flows to the net-of-tax rate was 6.39, which is similar to estimates in urban economics but twice as large as cross-local labor market estimates in recent papers in public economics. Our estimate for the stock elasticity of the tax base to the wage tax, which accounts for endogenous changes in wages, rents, and commuting patterns, was 0.86, similar in magnitude to other estimates in the public economics literature. The implications for urban structure are stark. Even in the absence of agglomeration economies, we found that Philadelphia would gain 26,000 jobs from its suburbs if it abolished its wage tax. Our other results are remarkably consistent with Adam Smith's critique of the (predominantly local) income taxes of his time:

If direct taxes upon the wages of labour have not always occasioned a proportionable rise in those wages, it is because they have generally occasioned a considerable fall in the demand for labour. The declension of industry, the decrease of employment for the poor, the diminution of the annual produce of the land and labour of the country, have generally been the effects of such taxes. In consequence of them, however, the price of labour must always be higher than it otherwise would have been in the actual state of demand, [...] and must always be finally paid by the landlords and consumers.

Future applications of quantitative spatial models to questions of local taxation could aim to understand three margins. First, unlike Philadelphia's flat wage tax, several cities in the United States tax income progressively. This fact complicates assigning tax rates to commutes, since the tax rate depends on the (endogenous) workplace wage. Second, local earnings taxes may affect the decision to work itself. Taxes may cause individuals to substitute away from taxed income sources (such as wages, salaries, and profits) and toward untaxed sources (such as government benefits). Jobless individuals could then potentially sort into the taxing jurisdiction if it provides lower rents compared to suburban locations (Notowidigdo, 2020). Third, rewriting the framework to allow for positive firm profits would allow us to investigate the effect of local corporate taxes (such as the BIRT) on firm location (Suárez Serrato and Zidar, 2016).

References

- ABADIE, A., S. ATHEY, G. W. IMBENS, AND J. M. WOOLDRIDGE (2023): "When Should You Adjust Standard Errors for Clustering?*", *The Quarterly Journal of Economics*, 138, 1–35.
- ABOWD, J. M., B. E. STEPHENS, L. VILHUBER, F. ANDERSSON, K. L. MCKINNEY, M. ROEMER, AND S. WOODCOCK (2009): "The LEHD infrastructure files and the creation of the quarterly workforce indicators," *Producer dynamics: New evidence from micro data*, 68, 149–230, publisher: University of Chicago Press Chicago.
- AGRAWAL, D. R. AND J. K. BRUECKNER (2025): "Taxes and telework: The impacts of state income taxes in a work-from-home economy," *Journal of Urban Economics*, 145, publisher: Elsevier.
- AGRAWAL, D. R., J. K. BRUECKNER, AND M. BRÜLHART (2024): "Fiscal Federalism in the Twenty-First Century," *Annual Review of Economics*, 16, 429–454, publisher: Annual Reviews.
- AGRAWAL, D. R. AND D. FOREMNY (2019): "Relocation of the Rich: Migration in Response to Top Tax Rate Changes from Spanish Reforms," *The Review of Economics and Statistics*, 101, 214–232.
- AGRAWAL, D. R. AND W. H. HOYT (2018): "Commuting and Taxes: Theory, Empirics and Welfare Implications," *The Economic Journal*, 128, 2969–3007.
- AGRAWAL, D. R. AND K. TESTER (2024): "State Taxation of Nonresident Income and the Location of Work," *American Economic Journal: Economic Policy*, 16, 447–481.
- AHLFELDT, G. M. AND E. PIETROSTEFANI (2019): "The economic effects of density: A synthesis," *Journal of Urban Economics*, 111, 93–107.
- AHLFELDT, G. M., S. J. REDDING, D. M. STURM, AND N. WOLF (2015): "The Economics of Density: Evidence from the Berlin Wall," *Econometrica*, 83, 2127–2189, [tex.eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA10876](https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA10876).
- AKBAR, P. A., A. SHERTZER, AND R. LIVAS (2025): "The Emergence of Public Transit and the Transformation of the American Downtown," *Working Paper*.
- ALLEN, T. AND C. ARKOLAKIS (2014): "Trade and the topography of the spatial economy," *The Quarterly Journal of Economics*, 129, 1085–1140, publisher: MIT Press.
- (2022): "The welfare effects of transportation infrastructure improvements," *The Review of Economic Studies*, 89, 2911–2957.
- (2023): "Economic activity across space," *Journal of Economic Perspectives*, 37, 3–28.
- ALMAGRO, M. AND T. DOMÍNGUEZ-IINO (2025): "Location Sorting and Endogenous Amenities: Evidence From Amsterdam," *Econometrica*, 93, 1031–1071, [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA21394](https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA21394).
- AMBRUS, A., E. FIELD, AND R. GONZALEZ (2020): "Loss in the Time of Cholera: Long-Run Impact of a Disease Epidemic on the Urban Landscape," *American Economic Review*, 110, 475–525.
- ARNOTT, R. J. AND J. E. STIGLITZ (1979): "Aggregate Land Rents, Expenditure on Public Goods, and Optimal City Size," *The Quarterly Journal of Economics*, 93, 471–500, publisher: Oxford University Press.

- AUTOR, D. H. AND D. DORN (2013): "The growth of low-skill service jobs and the polarization of the US labor market," *American Economic Review*, 103, 1553–97.
- BAKIJA, J. AND J. SLEMRD (2004): "Do the Rich Flee from High State Taxes? Evidence from Federal Estate Tax Returns," w10645.
- BARRECA, A. I., M. GULDI, J. M. LINDO, AND G. R. WADDELL (2011): "Saving Babies? Revisiting the effect of very low birth weight classification*," *The Quarterly Journal of Economics*, 126, 2117–2123.
- BAUM-SNOW, N. AND F. FERREIRA (2015): "Causal Inference in Urban and Regional Economics," in *Handbook of Regional and Urban Economics*, ed. by G. Duranton, J. V. Henderson, and W. C. Strange, Elsevier, vol. 5 of *Handbook of Regional and Urban Economics*, 3–68.
- BAUM-SNOW, N. AND L. HAN (2024): "The Microgeography of Housing Supply," *Journal of Political Economy*, 132, 1897–1946, publisher: The University of Chicago Press.
- BAUM-SNOW, N. AND D. HARTLEY (2020): "Accounting for central neighborhood change, 1980–2010," *Journal of Urban Economics*, 117, 103228.
- BAUMOL, W. J. (1967): "Macroeconomics of unbalanced growth: The anatomy of urban crisis," *American Economic Review*, 57, 415–426, publisher: American Economic Association.
- BAYER, P., F. FERREIRA, AND R. McMILLAN (2007): "A Unified Framework for Measuring Preferences for Schools and Neighborhoods," *Journal of Political Economy*, 115, 588–638, publisher: University of Chicago Press.
- BERKES, E., E. KARGER, AND P. NENCKA (2023): "The census place project: A method for geolocating unstructured place names," *Explorations in Economic History*, 87, 101477.
- BILAL, A. (2023): "The Geography of Unemployment*," *The Quarterly Journal of Economics*, 138, 1507–1576.
- BORDEU, O. (2023): "Commuting Infrastructure in Fragmented Cities," *Working Paper*.
- BOUSTAN, L., F. FERREIRA, H. WINKLER, AND E. M. ZOLT (2013): "The Effect of Rising Income Inequality on Taxation and Public Expenditures: Evidence from U.S. Municipalities and School Districts, 1970–2000," *The Review of Economics and Statistics*, 95, 1291–1302.
- BRYAN, G. AND M. MORTEN (2019): "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia," *Journal of Political Economy*, 127, 2229–2268, publisher: The University of Chicago Press.
- BRÜLHART, M., J. DANTON, R. PARCHET, AND J. SCHLÄPFER (2025): "Who Bears the Burden of Local Taxes?" *American Economic Journal: Economic Policy*, 17, 464–505, publisher: American Economic Association.
- BUNTEN, D. M., E. FU, L. ROLHEISER, AND C. SEVEREN (2024): "The Problem Has Existed over Endless Years: Racialized Difference in Commuting, 1980–2019," *Journal of Urban Economics*, 141, 103542.
- BUSSO, M., J. GREGORY, AND P. KLINE (2013): "Assessing the Incidence and Efficiency of a Prominent Place Based Policy," *American Economic Review*, 103, 897–947.
- CALONICO, S., M. CATTANEO, M. FARRELL, AND R. TITIUNIK (2019): "Regression Discontinuity Designs Using Covariates," *The Review of Economics and Statistics*, 101, 442–451, publisher: MIT Press.

- (2022): “RDROBUST: Stata module to provide robust data-driven inference in the regression-discontinuity design,” Publisher: Boston College Department of Economics.
- CATTANEO, M. D., R. TITUNIK, AND R. R. YU (2025): “Estimation and Inference in Boundary Discontinuity Designs,” *Working Paper*, arXiv:2505.05670 [econ].
- COGLIANESE, J., L. W. DAVIS, L. KILIAN, AND J. H. STOCK (2017): “Anticipation, Tax Avoidance, and the Price Elasticity of Gasoline Demand,” *Journal of Applied Econometrics*, 32, 1–15, _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/jae.2500>.
- COMBES, P.-P., M. LAFOURCADE, AND T. MAYER (2005): “The trade-creating effects of business and social networks: evidence from France,” *Journal of International Economics*, 66, 1–29.
- COOK, C., A. KREIDIEH, S. VASSERMAN, H. ALLCOTT, N. ARORA, F. VAN SAMBEEK, A. TOMKINS, AND E. TURKEL (2025): “The Short-Run Effects of Congestion Pricing in New York City,” .
- COUTURE, V., C. GAUBERT, J. HANDBURY, AND E. HURST (2024): “Income Growth and the Distributional Effects of Urban Spatial Sorting,” *The Review of Economic Studies*, 91, 858–898.
- COUTURE, V. AND J. HANDBURY (2020): “Urban revival in America,” *Journal of Urban Economics*, 119, 103267.
- CURRIER, L., E. L. GLAESER, AND G. E. KREINDLER (2023): “Infrastructure Inequality: Who Pays the Cost of Road Roughness?” .
- DAVIS, M. A., W. D. LARSON, S. D. OLINER, AND J. SHUI (2021): “The price of residential land for counties, ZIP codes, and census tracts in the United States,” *Journal of Monetary Economics*, 118, 413–431.
- DEKLE, R., J. EATON, AND S. KORTUM (2008): “Global rebalancing with gravity: Measuring the burden of adjustment,” *IMF Staff Papers*, 55, 511–540.
- DIAMOND, R. (2016): “The determinants and welfare implications of US workers’ diverging location choices by skill: 1980 – 2000,” *American Economic Review*, 106, 479–524.
- DIAMOND, R., T. MCQUADE, AND F. QIAN (2019): “The Effects of Rent Control Expansion on Tenants, Landlords, and Inequality: Evidence from San Francisco,” *American Economic Review*, 109, 3365–3394.
- DINGEL, J. I. AND F. TINTELNOT (2021): “Spatial economics for granular settings,” Tech. rep., National Bureau of Economic Research Working Paper 27287.
- DURANTON, G., L. GOBILLON, AND H. G. OVERMAN (2011): “Assessing the Effects of Local Taxation using Microgeographic Data,” *The Economic Journal*, 121, 1017–1046, _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-0297.2011.02439.x>.
- EPPLE, D. AND H. SIEG (1999): “Estimating Equilibrium Models of Local Jurisdictions,” *Journal of Political Economy*, 107, 645–681, publisher: The University of Chicago Press.
- FAJGELBAUM, P. D. AND C. GAUBERT (2020): “Optimal spatial policies, geography, and sorting,” *The Quarterly Journal of Economics*, 135, 959–1036.
- FAJGELBAUM, P. D., E. MORALES, J. C. S. SERRATO, AND O. ZIDAR (2019): “State Taxes and Spatial Misallocation,” *The Review of Economic Studies*, 86, 333–376, publisher: Review of Economic Studies Ltd.

- FISZBEIN, M., J. LAFORTUNE, E. G. LEWIS, AND J. TESSADA (2020): "New Technologies, Productivity, and Jobs: The (Heterogeneous) Effects of Electrification on Us Manufacturing," *SSRN Electronic Journal*.
- FOGELSON, R. M. (2001): *Downtown: Its Rise and Fall, 1880–1950*, Yale University Press.
- FRANKLIN, S., C. IMBERT, G. ABEBE, AND C. MEJIA-MANTILLA (2024): "Urban Public Works in Spatial Equilibrium: Experimental Evidence from Ethiopia," *American Economic Review*, 114, 1382–1414.
- FREEDMAN, M. (2012): "Teaching new markets old tricks: The effects of subsidized investment on low-income neighborhoods," *Journal of Public Economics*, 96, 1000–1014.
- GABLER, R. (1970): "The Commuter and the Municipal Income Tax," *Advisory Commission on Intergovernmental Relations*.
- GIROUD, X. AND J. RAUH (2019): "State Taxation and the Reallocation of Business Activity: Evidence from Establishment-Level Data," *Journal of Political Economy*, 127, 1262–1316, publisher: The University of Chicago Press.
- GLAESER, E. L. (2013): "Urban Public Finance," in *Handbook of Public Economics*, ed. by A. J. Auerbach, R. Chetty, M. Feldstein, and E. Saez, Elsevier, vol. 5 of *Handbook of Public Economics*, 195–256.
- GLAESER, E. L. AND J. GYOURKO (2005): "Urban decline and durable housing," *Journal of Political Economy*, 113, 345–375.
- (2025): "America's Housing Supply Problem: The Closing of the Suburban Frontier?" .
- GREEN, A., M. J. KUTZBACH, AND L. VILHUBER (2017): "Two perspectives on commuting: A comparison of home to work flows across job-linked survey and administrative files," Publisher: Center for Economic Studies.
- GRIESON, R. E. (1980): "Theoretical analysis and empirical measurements of the effects of the Philadelphia income tax," *Journal of Urban Economics*, 8, 123–137.
- GROSSMAN, G. M. AND E. OBERFIELD (2022): "The Elusive Explanation for the Declining Labor Share," *Annual Review of Economics*, 14, 93–124, publisher: Annual Reviews.
- GYOURKO, J., R. A. MARGO, AND A. F. HAUGHWOUT (2005): "Looking back to look forward: Learning from Philadelphia's 350 years of urban development," *Brookings-Wharton Papers on Urban Affairs*, 1–58, publisher: Brookings Institution Press.
- GYOURKO, J. AND J. TRACY (1989): "The Importance of Local Fiscal Conditions in Analyzing Local Labor Markets," *Journal of Political Economy*, 97, 1208–1231, publisher: University of Chicago Press.
- HANDBURY, J. (2021): "Are Poor Cities Cheap for Everyone? Non-Homotheticity and the Cost of Living Across U.S. Cities," *Econometrica*, 89, 2679–2715, _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA11738>.
- HANSON, G. H. AND R. LIVAS (2025): "The Rise of the South and the Deindustrialization of the North," *Working Paper*.
- HANSON, G. H., D. RODRIK, AND R. SANDHU (2025): "The U.S. Place-Based Policy Supply Chain," *NBER Working Paper Series 33511*.

- HAUGHWOUT, A., R. INMAN, S. CRAIG, AND T. LUCE (2004): "Local revenue hills: evidence from four US cities," *Review of Economics and Statistics*, 86, 570–585.
- HAUGHWOUT, A. F. AND R. P. INMAN (2001): "Fiscal policies in open cities with firms and households," *Regional Science and Urban Economics*, 31, 147–180.
- HEBLICH, S., S. J. REDDING, AND D. M. STURM (2020): "The Making of the Modern Metropolis: Evidence from London," *The Quarterly Journal of Economics*, 135, 2059–2133, tex.eprint: <https://academic.oup.com/qje/article-pdf/135/4/2059/33668557/qjaa014.pdf>.
- HERZOG, I. (2024): "The city-wide effects of tolling downtown drivers: Evidence from London's congestion charge," *Journal of Urban Economics*, 144, 103714.
- HILLBERRY, R. AND D. HUMMELS (2008): "Trade responses to geographic frictions: A decomposition using micro-data," *European Economic Review*, 52, 527–550.
- HOLMES, T. (1998): "The Effect of State Policies on the Location of Manufacturing: Evidence from State Borders," *Journal of Political Economy*, 106, 667–705, publisher: The University of Chicago Press.
- INMAN, R. P. (1995): "How to have a fiscal crisis: lessons from Philadelphia," *American Economic Association Papers & Proceedings*, 85, 378–383, publisher: JSTOR.
- JACKSON, K. T. (1987): *Crabgrass Frontier: The Suburbanization of the United States*, Oxford, New York: Oxford University Press.
- KEELE, L. J. AND R. TITIUNIK (2015): "Geographic Boundaries as Regression Discontinuities," *Political Analysis*, 23, 127–155.
- KIDWELL, D. S. AND C. A. TRZCINKA (1983): "The Impact of the New York City Fiscal Crisis on the Interest Cost of New Issue Municipal Bonds," *Journal of Financial and Quantitative Analysis*, 18, 381–399.
- KLEVEN, H., C. T. KREINER, K. LARSEN, AND J. EGHOLT SØGAARD (2023): "Micro vs Macro Labor Supply Elasticities: The Role of Dynamic Returns to Effort," *NBER Working Paper Series 31549*.
- KLEVEN, H., C. LANDAIS, M. MUÑOZ, AND S. STANTCHEVA (2020): "Taxation and Migration: Evidence and Policy Implications," *Journal of Economic Perspectives*, 34, 119–142.
- KREINDLER, G. E. AND Y. MIYAUCHI (2023): "Measuring Commuting and Economic Activity inside Cities with Cell Phone Records," *The Review of Economics and Statistics*, 105, 899–909.
- LEE, D. S. AND T. LEMIEUX (2010): "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 48, 281–355.
- LEE, S. AND J. LIN (2018): "Natural Amenities, Neighbourhood Dynamics, and Persistence in the Spatial Distribution of Income," *The Review of Economic Studies*, 85, 663–694, publisher: Review of Economic Studies Ltd.
- LEHNER, A. (2024): "A Note on Spatial Regression Discontinuity Designs," *Working Paper*.
- LOUPEAU, G. (2023): "Regional Borders, Commuting and Transport Network Integration," *The Review of Economics and Statistics*, 1–45.

- MANDEL, B. H. (2023): *Philadelphia, Corrupt and Consenting: A City's Struggle Against an Epithet*, Philadelphia, UNITED STATES: Temple University Press.
- MARTÍNEZ, I. Z. (2022): "Mobility responses to the establishment of a residential tax haven: Evidence from Switzerland," *Journal of Urban Economics*, 129, 103441.
- McDONALD, G. (2017): "Urban Institute - Census Tract Origin Destination Driving Times [dataset]," .
- MIESZKOWSKI, P. AND E. S. MILLS (1993): "The Causes of Metropolitan Suburbanization," *Journal of Economic Perspectives*, 7, 135–147.
- MONTE, F., S. J. REDDING, AND E. ROSSI-HANSBERG (2018): "Commuting, migration, and local employment elasticities," *American Economic Review*, 108, 3855–90.
- MORETTI, E. AND D. J. WILSON (2017): "The Effect of State Taxes on the Geographical Location of Top Earners: Evidence from Star Scientists," *American Economic Review*, 107, 1858–1903.
- NOTOWIDIGDO, M. J. (2020): "The incidence of local labor demand shocks," *Journal of Labor Economics*, 38, publisher: The University of Chicago Press Chicago, IL tex.date-modified: 2021-01-08 00:10:52 -0500.
- OATES, W. E. (1972): "Taxation and debt finance in a federal system," in *Fiscal Federalism*, New York: Harcourt Brace Jovanovich, Inc., The Harbrace Series in Business and Economics, 119–161.
- OATES, W. E. AND R. M. SCHWAB (1997): "The impact of urban land taxation: the pittsburgh experience," *National Tax Journal*, 50, 1–21, publisher: The University of Chicago Press.
- OWENS, RAYMOND, I., E. ROSSI-HANSBERG, AND P.-D. SARTE (2020): "Rethinking detroit," *American Economic Journal: Economic Policy*, 12, 258–305.
- PENNSYLVANIA ECONOMY LEAGUE (1999): "The Sterling Act: A Brief History," Tech. rep., Economy League of Greater Philadelphia.
- PIERSON, K., M. L. HAND, AND F. THOMPSON (2015): "The Government Finance Database: A Common Resource for Quantitative Research in Public Financial Analysis," *PLOS ONE*, 10, e0130119, publisher: Public Library of Science.
- REDDING, S. J. (2022): "Suburbanization in the USA, 1970–2010," *Economica*, 89, S110–S136, publisher: Wiley Online Library.
- (2024): "Quantitative Urban Economics," *NBER Working Paper Series* 33130.
- RISCH, M. (2024): "Does Taxing Business Owners Affect Employees? Evidence From A Change in the Top Marginal Tax Rate*," *The Quarterly Journal of Economics*, 139, 637–692.
- ROSE, E. K. (2018): "The rise and fall of female labor force participation during world war II in the united states," *The Journal of Economic History*, 78, 673–711, publisher: Cambridge University Press.
- RUBOLINO, E. AND T. GIOMMONI (2023): "Taxation and Mobility: Evidence from Tax Decentralization in Italy," .
- RUGGLES, S., S. FLOOD, R. GOEKEN, M. SCHOUWEILER, AND M. SOBEK (2022): "IPUMS USA: Version 12.0," *Minnesota Population Center, Minneapolis, MN*.

- RUGGLES, S., S. FLOOD, M. SOBEK, D. BROCKMAN, G. COOPER, S. RICHARDS, AND M. SCHOUWEILER (2021): "IPUMS ancestry full count data: Version 3.0," *Minnesota Population Center, Minneapolis, MN*.
- SEVEREN, C. (2023): "Commuting, Labor, and Housing Market Effects of Mass Transportation: Welfare and Identification," *The Review of Economics and Statistics*, 105, 1073–1091.
- SHERTZER, A., R. P. WALSH, AND J. R. LOGAN (2016): "Segregation and neighborhood change in northern cities: New historical GIS data from 1900–1930," *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 49, 187–197, publisher: Taylor & Francis.
- SIODLA, J. (2020): "Debt and taxes: Fiscal strain and US city budgets during the great depression," *Explorations in Economic History*, 76, 101328.
- STRUMPF, K. S. (1999): "Government credibility and policy choice: evidence from the Pennsylvania earned income tax," *Journal of Public Economics*, 80, 141–167.
- SUÁREZ SERRATO, J. C. AND O. ZIDAR (2016): "Who Benefits from State Corporate Tax Cuts? A Local Labor Markets Approach with Heterogeneous Firms," *American Economic Review*, 106, 2582–2624.
- TINKCOM, M. B. (1982): "Depression and war, 1929-1946," in *Philadelphia: a 300-year history*, W.W. Norton & Company.
- TSIVANIDIS, N. (2024): "Evaluating the impact of urban transit infrastructure: Evidence from Bogotá's Trans-Milenio," *Working Paper*.
- TURNER, M. A., A. HAUGHWOUT, AND W. VAN DER KLAUW (2014): "Land Use Regulation and Welfare," *Econometrica*, 82, 1341–1403, _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA9823>.
- VAN VLOKHOVEN, H. (2024): "Estimating the Cost of Capital and the Profit Share," *The Economic Journal*, 134, 2175–2206.
- WILDASIN, D. E. (1985): "Income taxes and urban spatial structure," *Journal of Urban Economics*, 18, 313–333.
- YOUNG, C. AND C. VARNER (2011): "Millionaire migration and state taxation of top incomes: evidence from a natural experiment," *National Tax Journal*, 64, 255–283, publisher: The University of Chicago Press.
- ZÁRATE, R. D. (2022): "Spatial misallocation, informality, and transit improvements: Evidence from Mexico City," *World Bank, Washington, DC*.

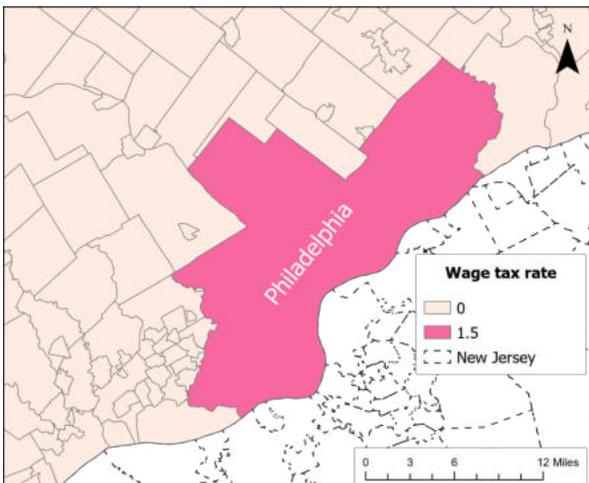
A Appendix Figures and Tables

Figure A.1: Maps of wage tax rates by year (%)

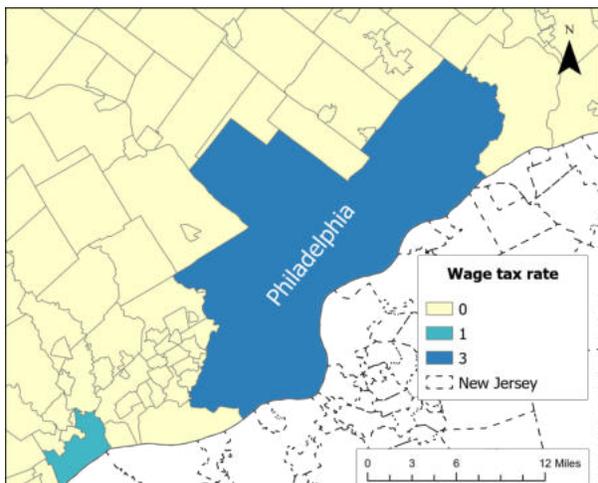
(a) Resident wage tax rate, 1960



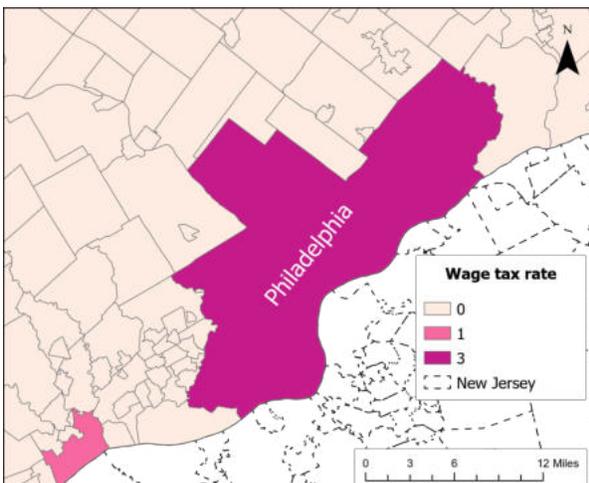
(b) Nonresident wage tax rate, 1960



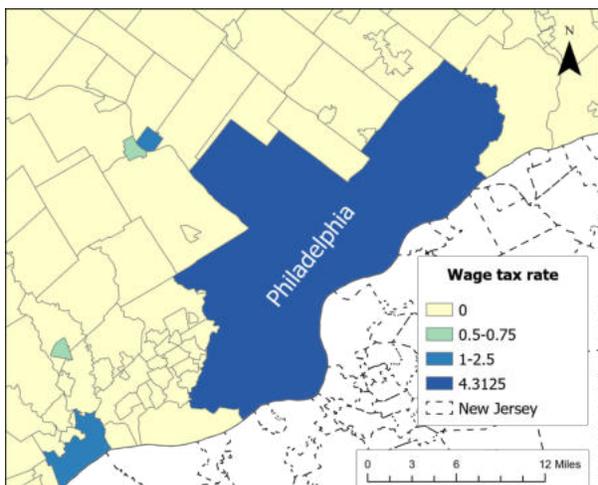
(c) Resident wage tax rate, 1970



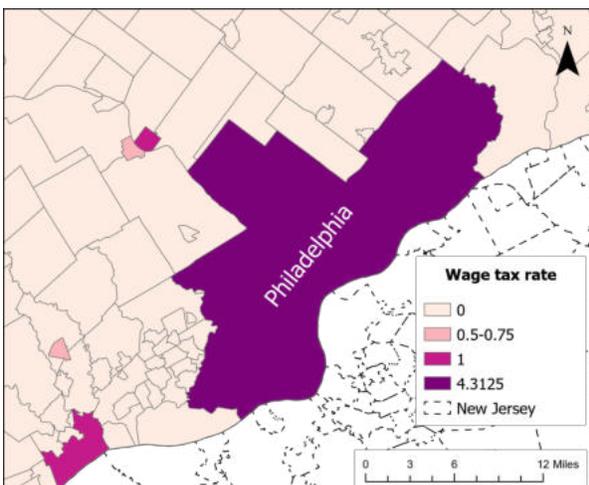
(d) Nonresident wage tax rate, 1970



(e) Resident wage tax rate, 1980



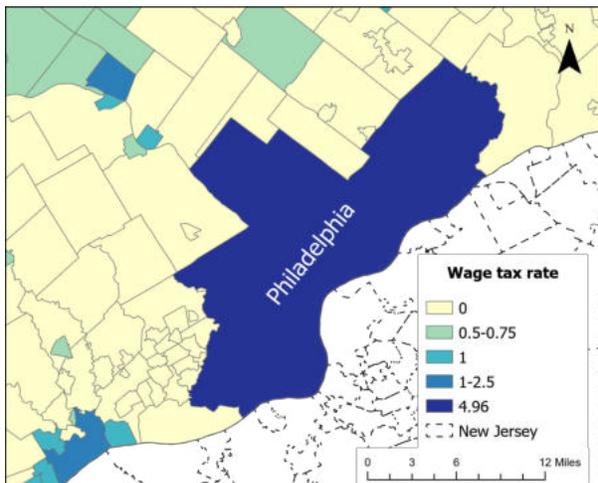
(f) Nonresident wage tax rate, 1980



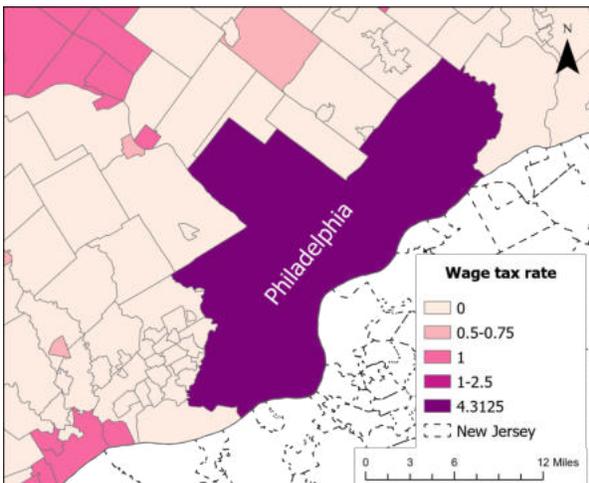
Note: The maps display the values for the resident (Panel (a)) and nonresident (Panel (b)) wage tax rates for municipalities near the City of Philadelphia in various years. State income taxes in PA and NJ, enacted after 1970, were 2.24% and 1.97% in 1980 (not displayed). See Section 2.3 for institutional details. See Figure A.2 for maps of tax rates in other years.

Figure A.2: Maps of wage tax rates by year (%)

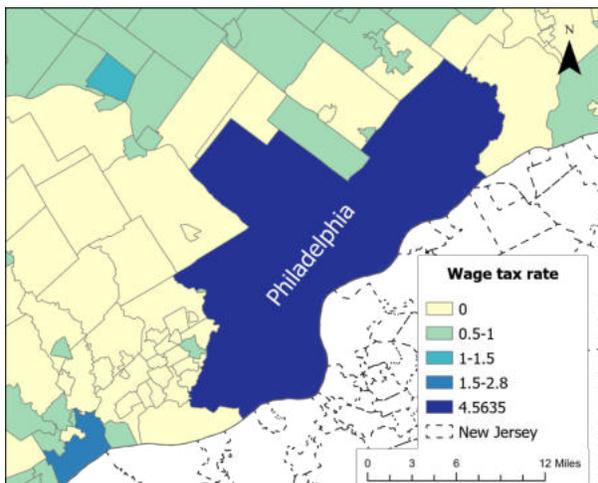
(a) Resident wage tax rate, 1990



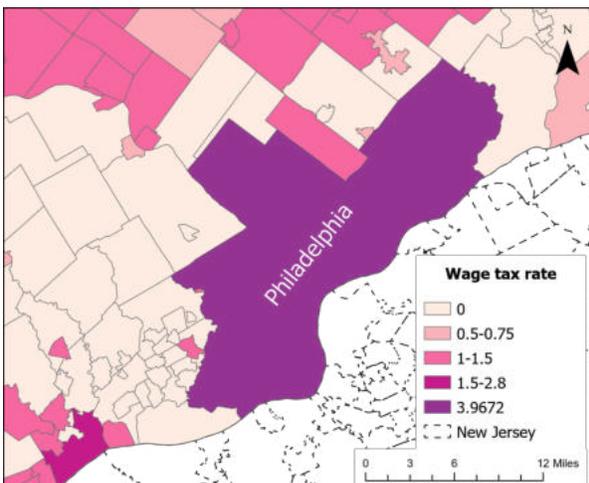
(b) Nonresident wage tax rate, 1990



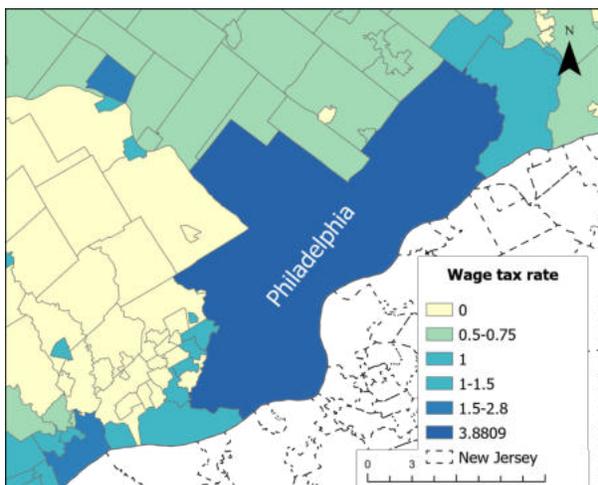
(c) Resident wage tax rate, 2000



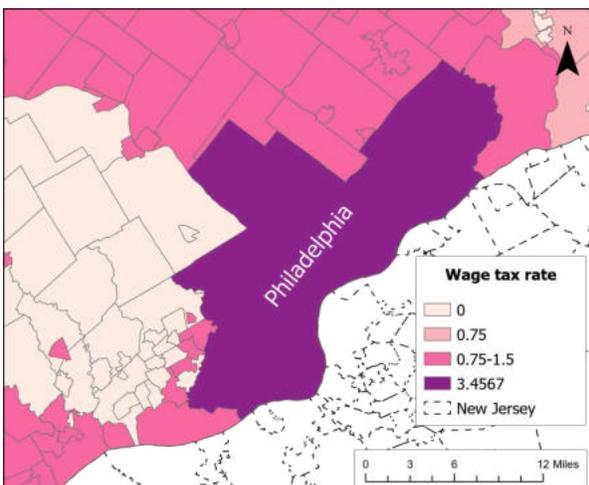
(d) Nonresident wage tax rate, 2000



(e) Resident wage tax rate, 2018

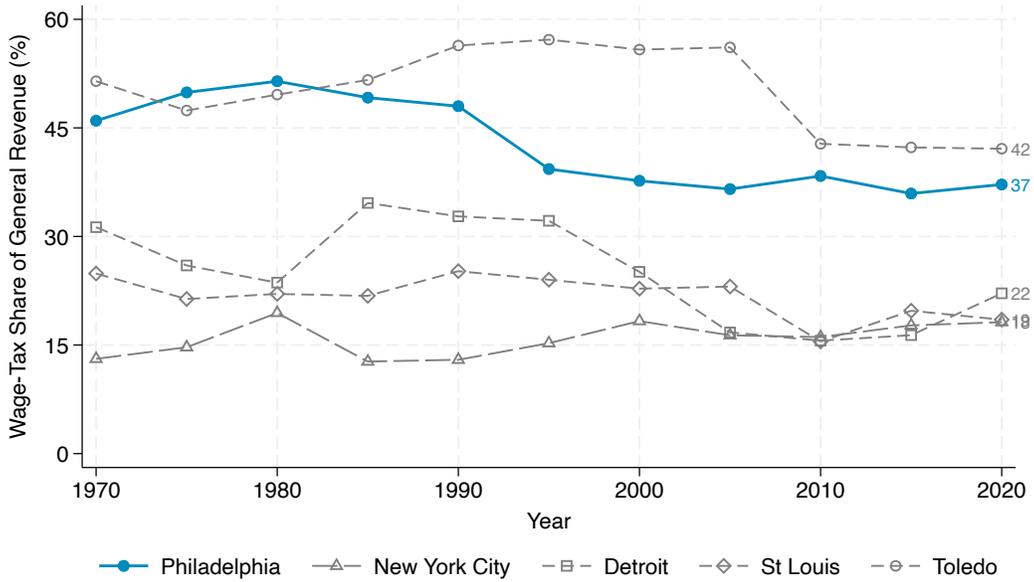


(f) Nonresident wage tax rate, 2018



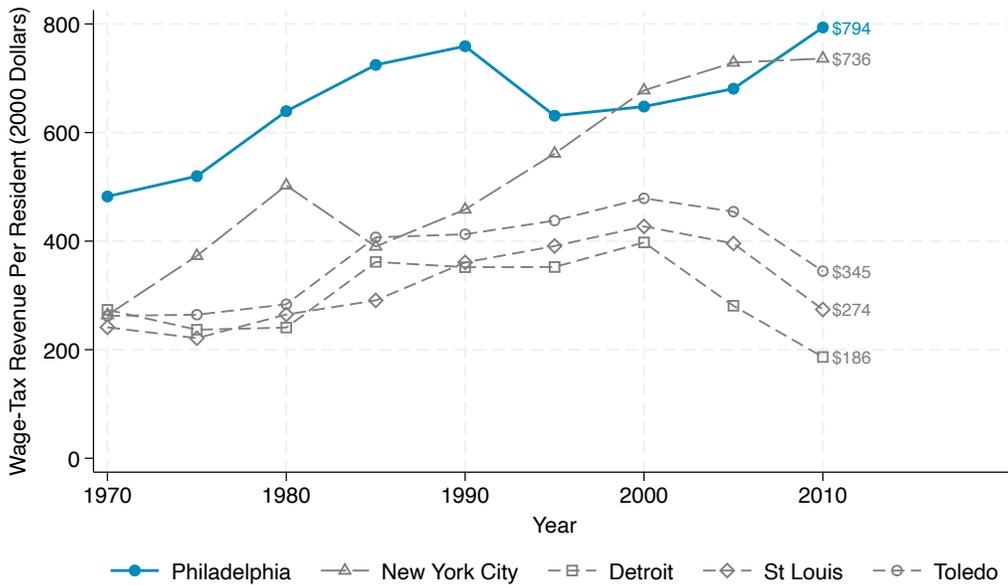
Note: The maps display the values for the resident (Panel (a)) and nonresident (Panel (b)) wage tax rates for municipalities near the City of Philadelphia in various years. State income taxes in PA were 2.1, 2.8, and 3.07% in 1990, 2000, and 2018. Average state income taxes for those same years in NJ were 2.72, 2.04, and 2.51%. See Section 2.3 for institutional details.

Figure A.3: Revenue shares for wage taxes in select jurisdictions



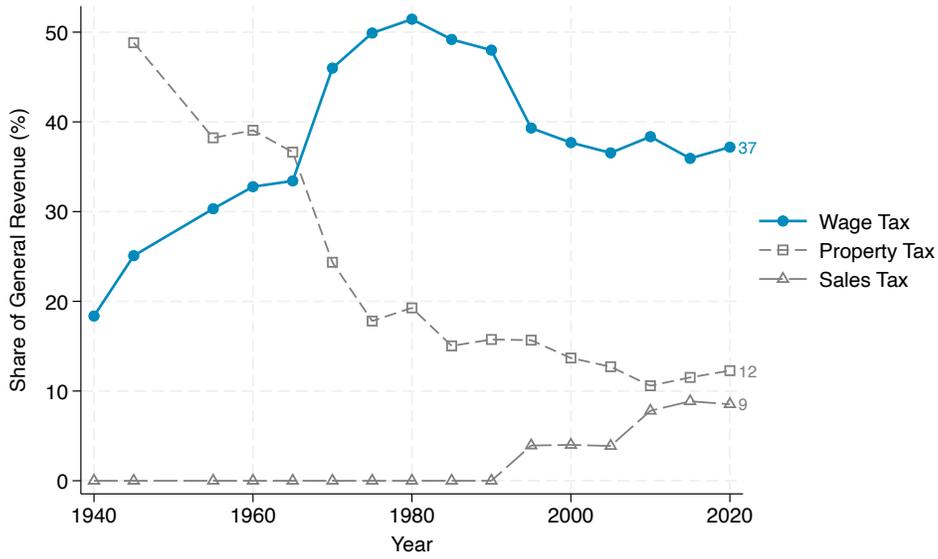
Note: This figure shows wage tax revenues as a share of general revenue in Philadelphia, New York City (which abolished the nonresident rates in 1999), Detroit, St Louis and Toledo. Data from the Government Finance Database (Pierson et al., 2015).

Figure A.4: Wage tax revenue per capita in select jurisdictions



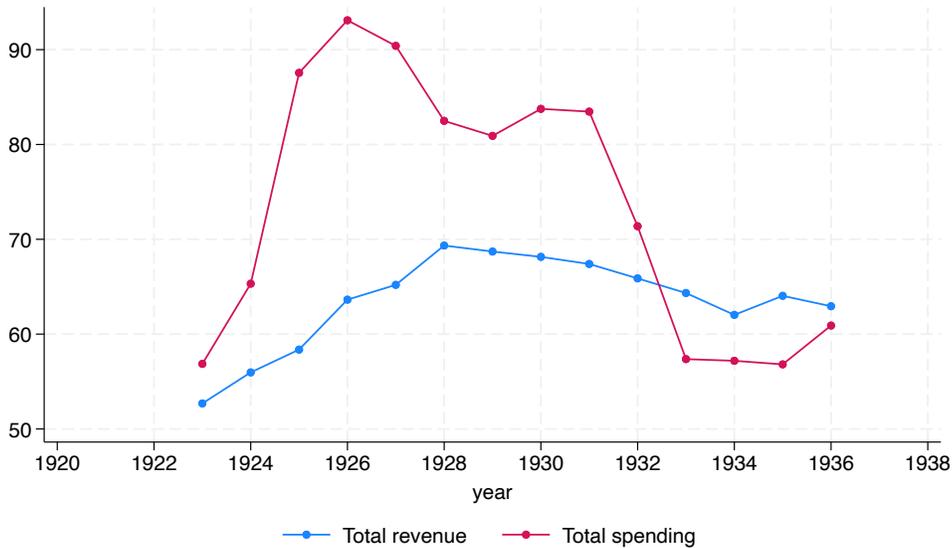
Note: This figure uses the same data as above to show wage tax revenues per resident for Philadelphia, New York City, Detroit, St Louis and Toledo in constant (2020) dollars.

Figure A.5: Importance of wage tax in Philadelphia General Fund revenue, 1940 to 2019



Note: The figure displays time series for the share of General Fund revenue for the wage tax, the property tax, and the sales tax. General Fund revenue includes tax and non-tax revenues for the City of Philadelphia. Sources: [Pennsylvania Economy League \(1999\)](#), City of Philadelphia Income Tax Regulations, and [Pierson et al. \(2015\)](#).

Figure A.6: Revenue and Expenditure in the City of Philadelphia, 1923 to 1936



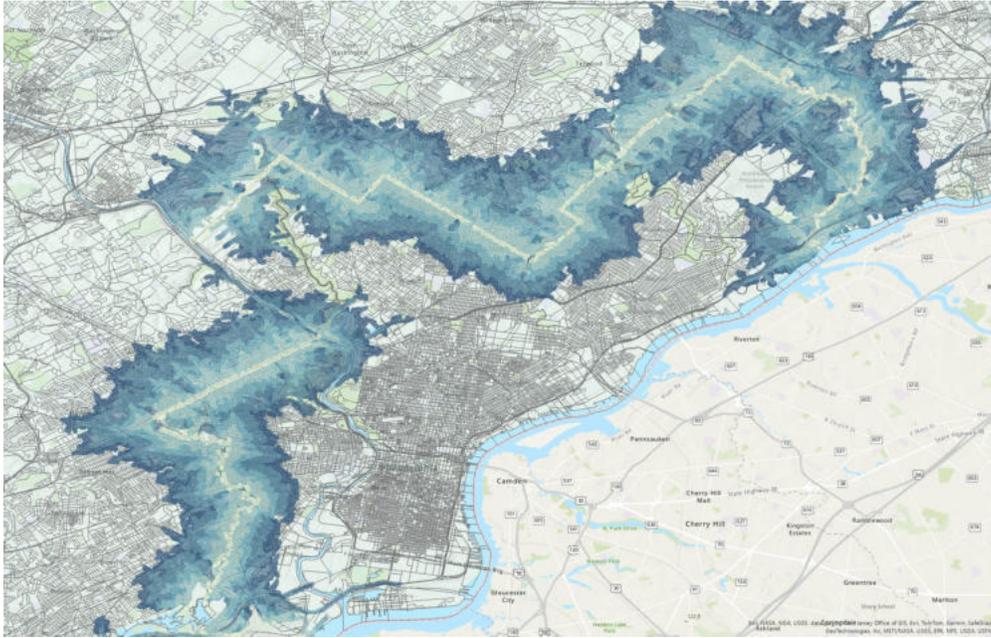
Note: The figure shows the City of Philadelphia’s total revenue and total expenditure from 1923 to 1936. Values are in 1967 dollars. Data from [Siodla \(2020\)](#).

Figure A.7: Holme Map of 1687



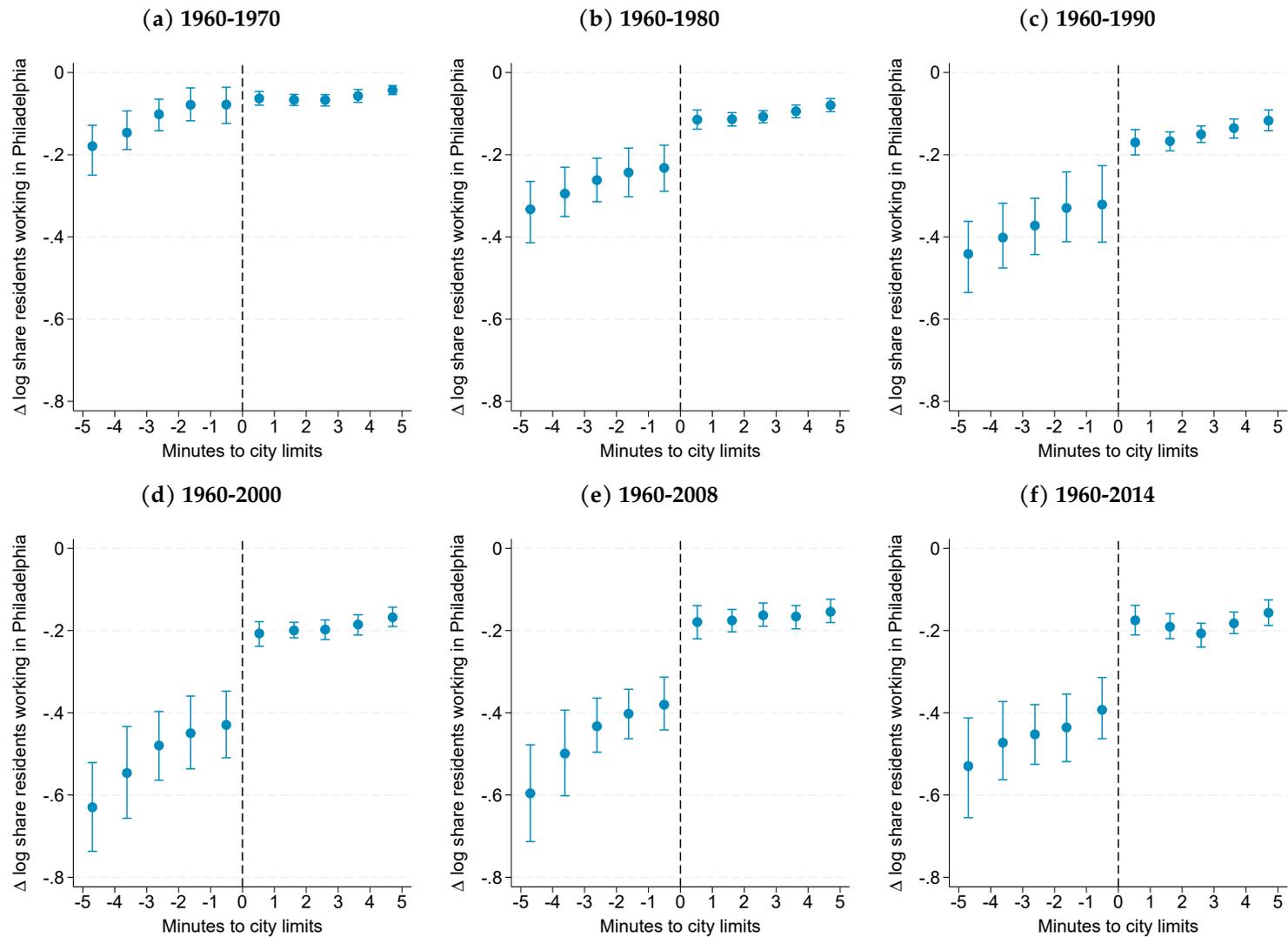
Note: We overlay approximate present-day borders for Philadelphia County (marked in bright green) over township borders (marked in dull red in the original source) from Thomas Home's Map of 1687. See Holme, Thomas, John Harris, and Philip Lea. *A mapp of ye improved part of Pensilvania in America, divided into countyes, townships, and lotts.* [London: Sold by P. Lea at ye Atlas and Hercules in Cheapside, ?, 1687] Map. <https://www.loc.gov/item/81692881/>.

Figure A.8: Five-minute drive time isochrone to the Philadelphia County border



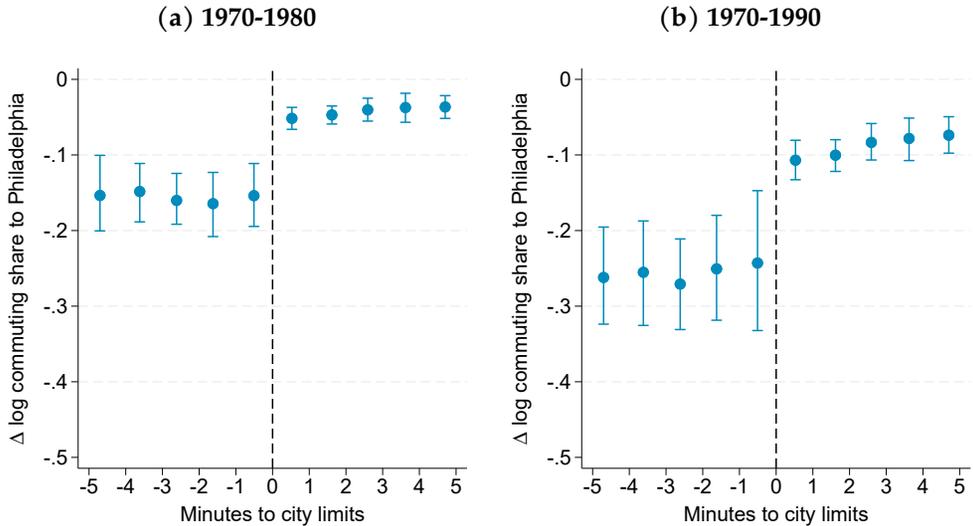
Note: The map displays the uncongested drive time surface to the Philadelphia County border using the present-day road network. The darkest shade of blue corresponds to areas that are within 5 minutes of the border. The lightest shade of blue corresponds to areas within 0.1 minutes of the border. We overlay 2010 census block definitions and a standard topographic map for reference.

Figure A.9: RD plots for change in log commuting share to Philadelphia: end years for 1960 base

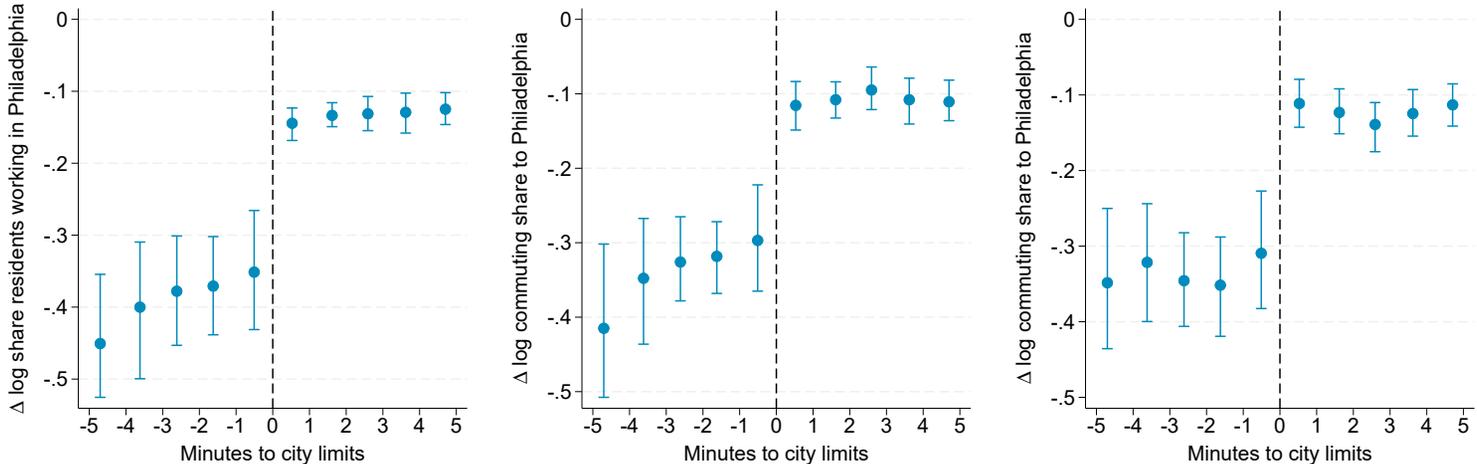


Note: Spatial RD plots display changes in log commuting shares for tracts within a 5-minute drive of Philadelphia's city limits. Positive x-values are inside the city. Panels (a)-(e) show all end years for base year 1960. Standard errors clustered at the grid-cell level.

Figure A.10: RD plots for change in log commuting share to Philadelphia: end years for 1970 base

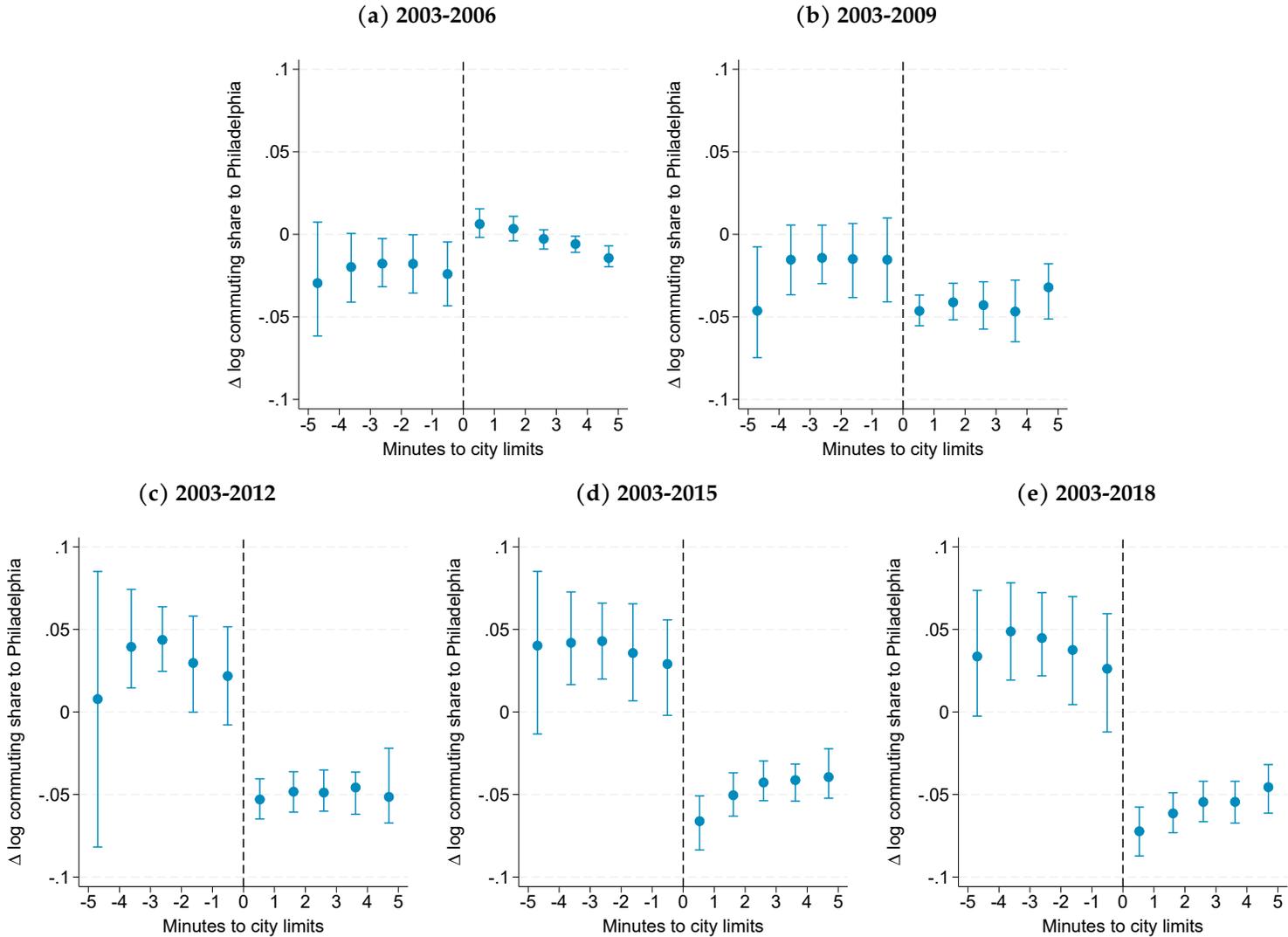


(c) 1970-2000 (d) 1970-2008 (e) 1970-2014



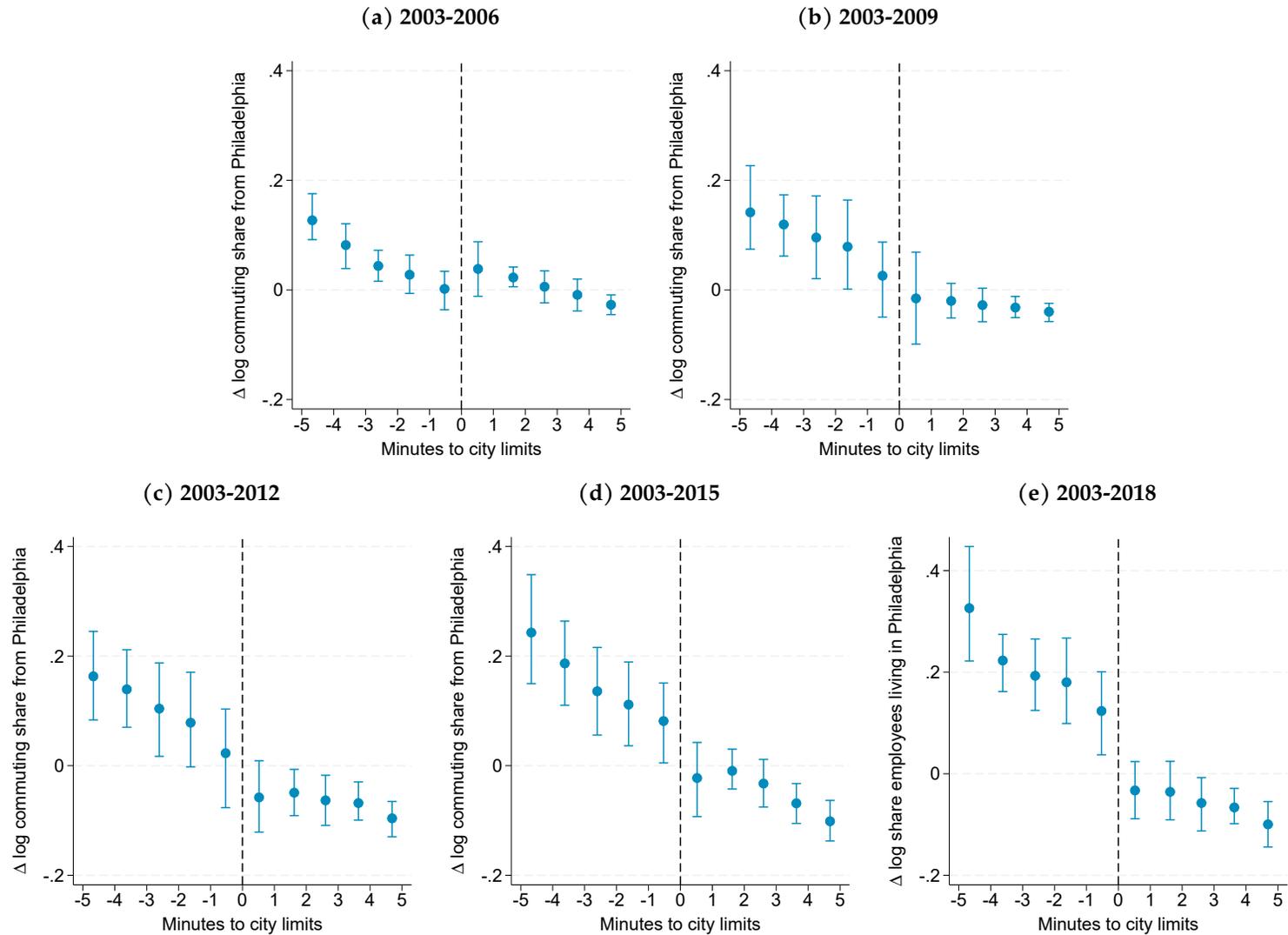
Note: Spatial RD plots display changes in log commuting shares for tracts within a 5-minute drive of Philadelphia’s city limits. Positive x-values are inside the city. Panels (a)-(d) show all end years for base year 1970. Standard errors clustered at the grid-cell level.

Figure A.11: RD plots for change in log commuting share to Philadelphia: end years for 2003 base



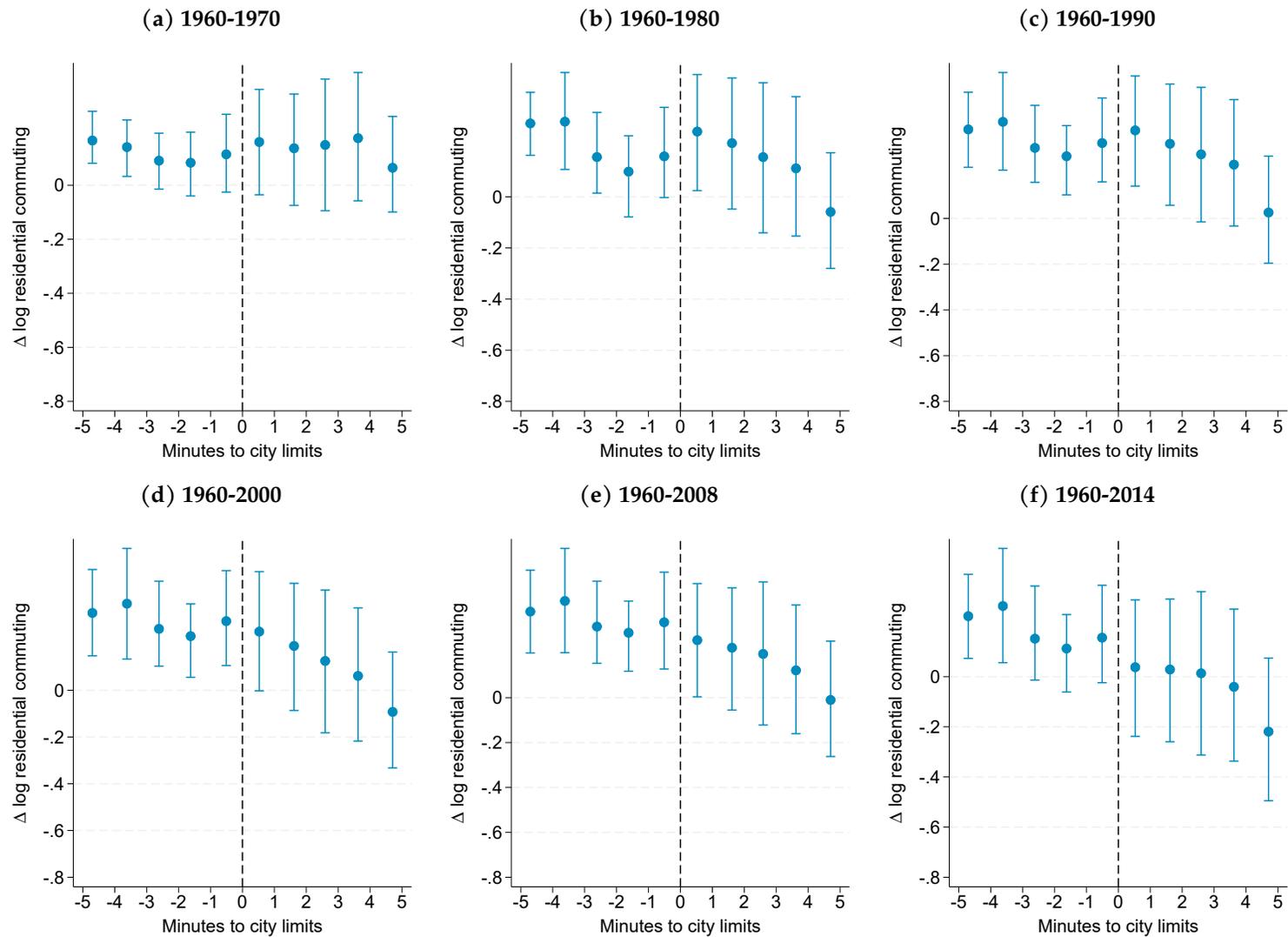
Note: Spatial RD plots display changes in log commuting shares for tracts within a 5-minute drive of Philadelphia’s city limits. Positive x-values are inside the city. Panels (a)-(d) show all end years for base year 2003. Standard errors clustered at the grid-cell level.

Figure A.12: RD plots for change in log commuting share from Philadelphia: end years for 2003 base



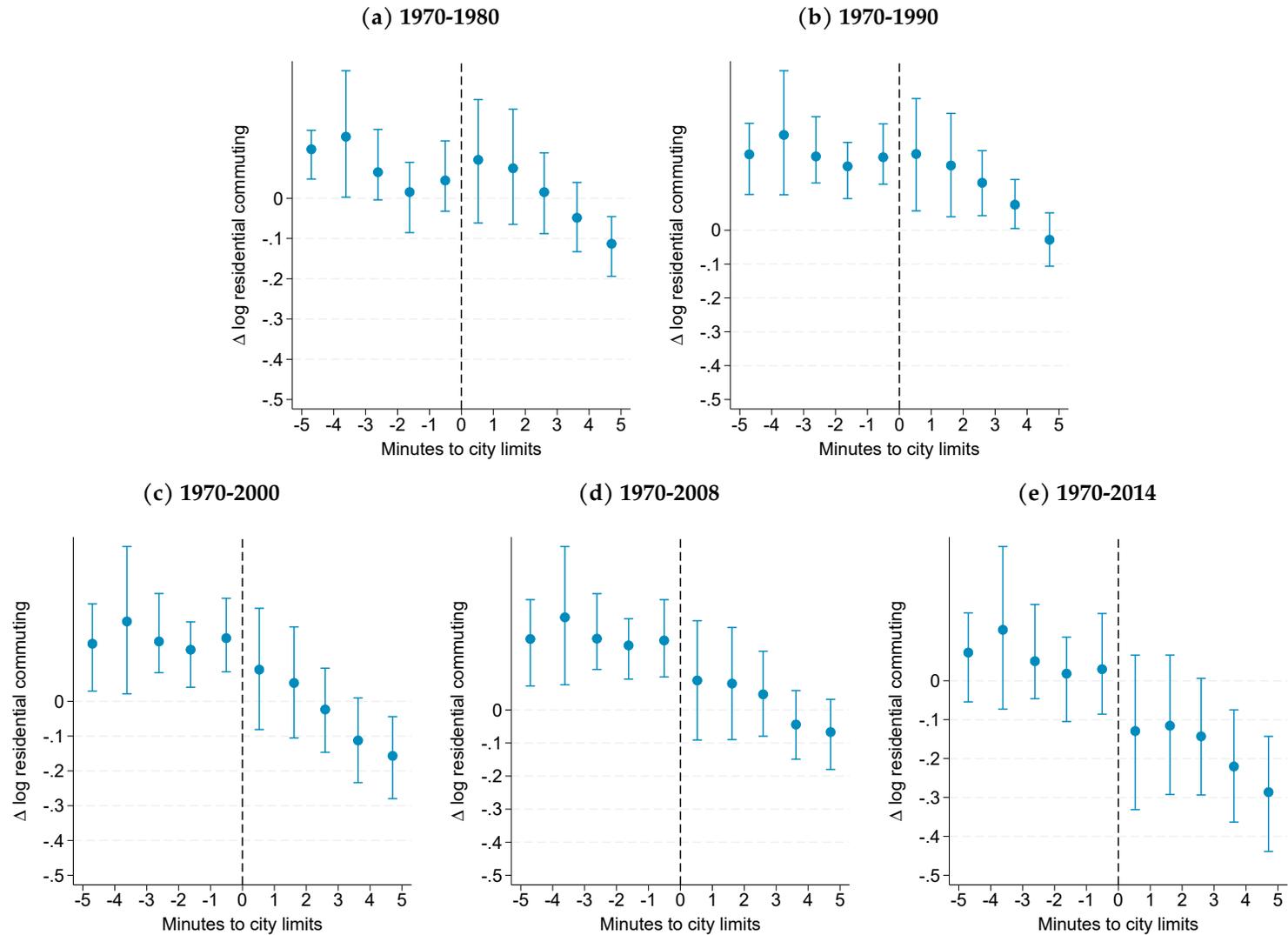
Note: Spatial RD plots display changes in log commuting shares for tracts within a 5-minute drive of Philadelphia's city limits. Positive x-values are inside the city. Panels (a)-(d) show all end years for base year 2003. Standard errors clustered at the grid-cell level.

Figure A.13: Continuity check: RD plots for change in log commuting: end years for 1960 base



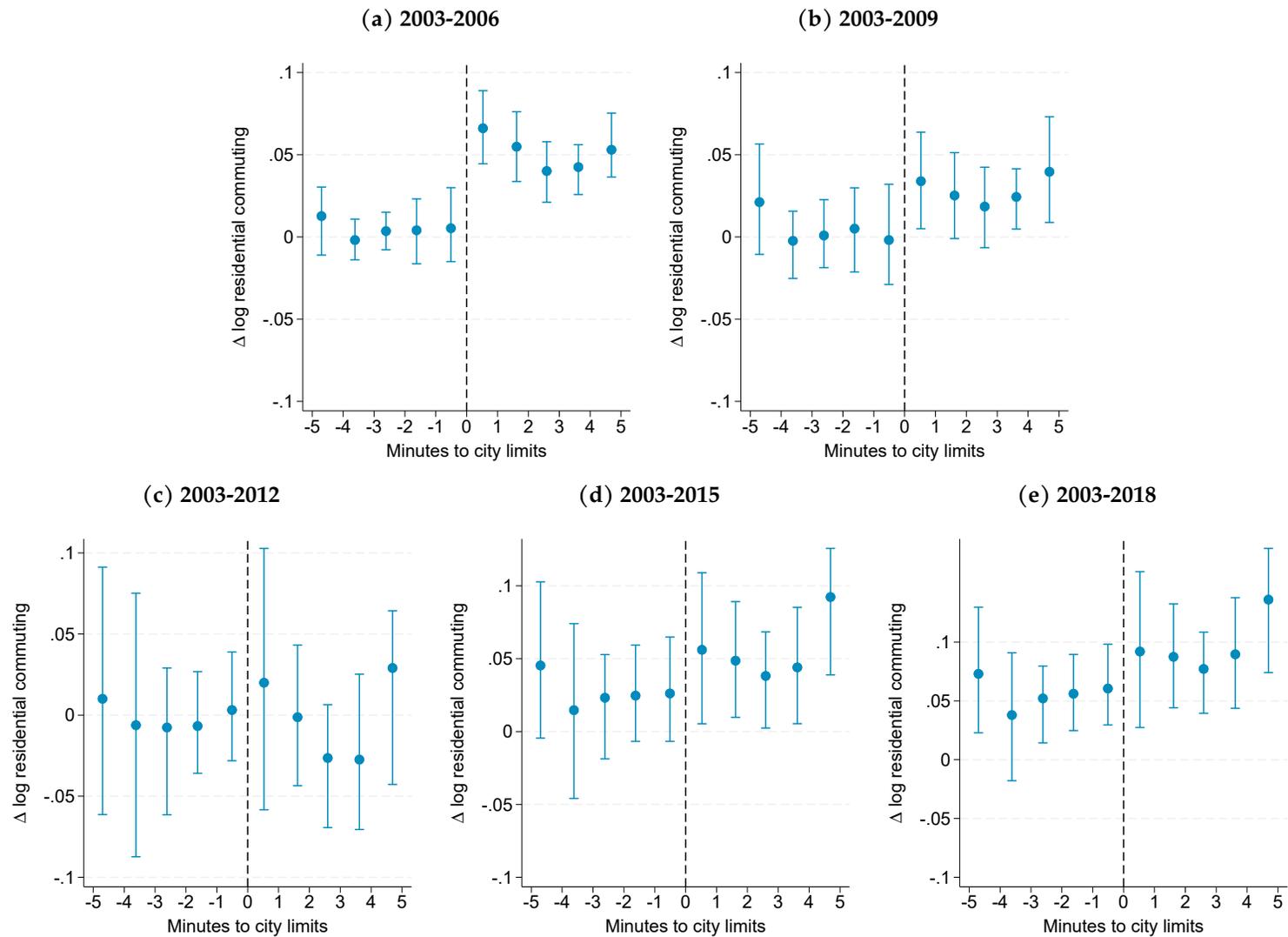
Note: Spatial RD plots display changes in log residential commuting for tracts within a 5-minute drive of Philadelphia’s city limits. Positive x-values are inside the city. Panels (a)-(e) show all end years for base year 1960. Standard errors clustered at the grid-cell level.

Figure A.14: Continuity check: RD plots for change in log commuting: end years for 1970 base



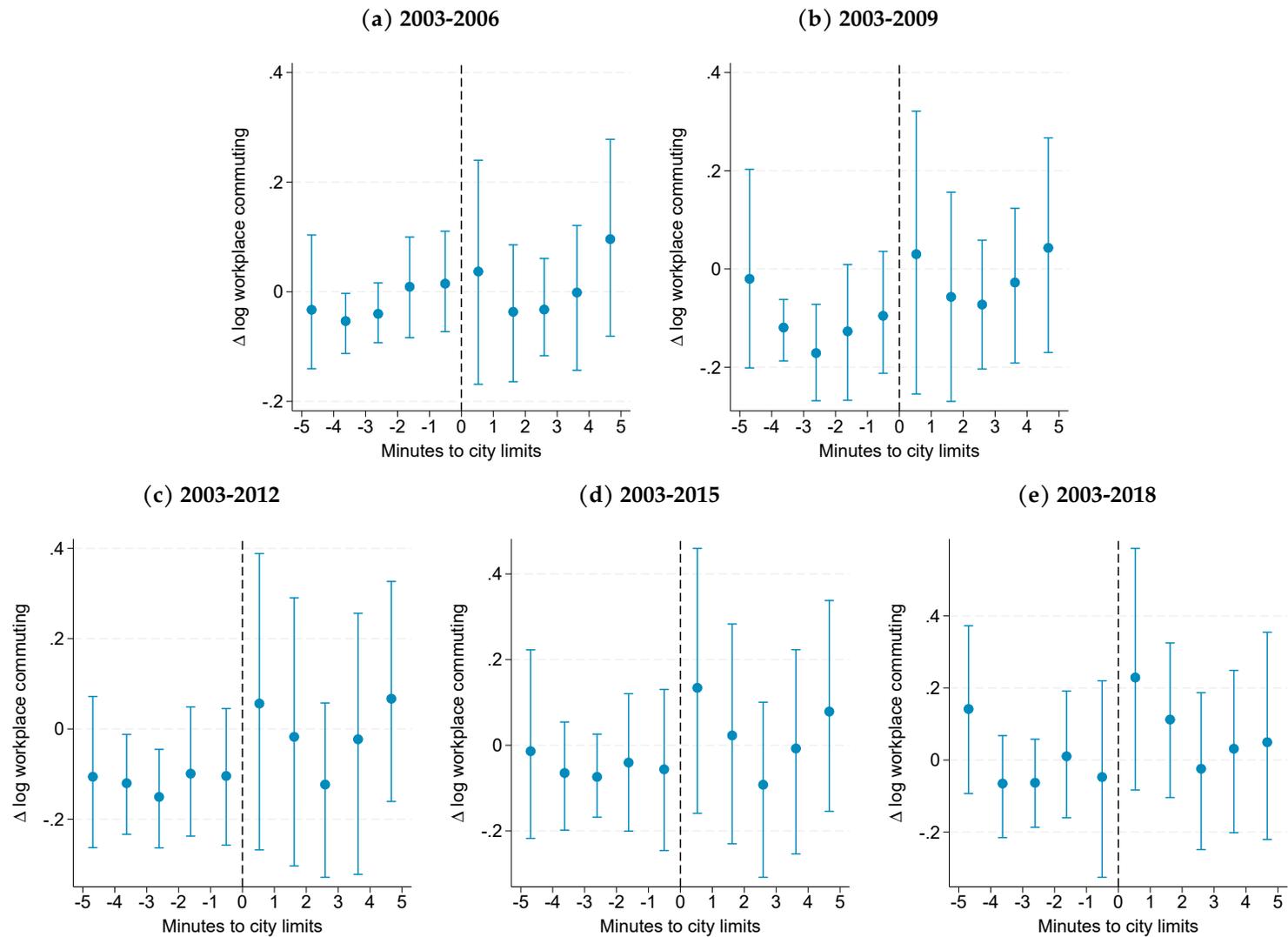
Note: Spatial RD plots display changes in log residential commuting for tracts within a 5-minute drive of Philadelphia’s city limits. Positive x-values are inside the city. Panels (a)-(d) show all end years for base year 1970. Standard errors clustered at the grid-cell level.

Figure A.15: Continuity check: RD plots for change in log commuting: end years for 2003 base



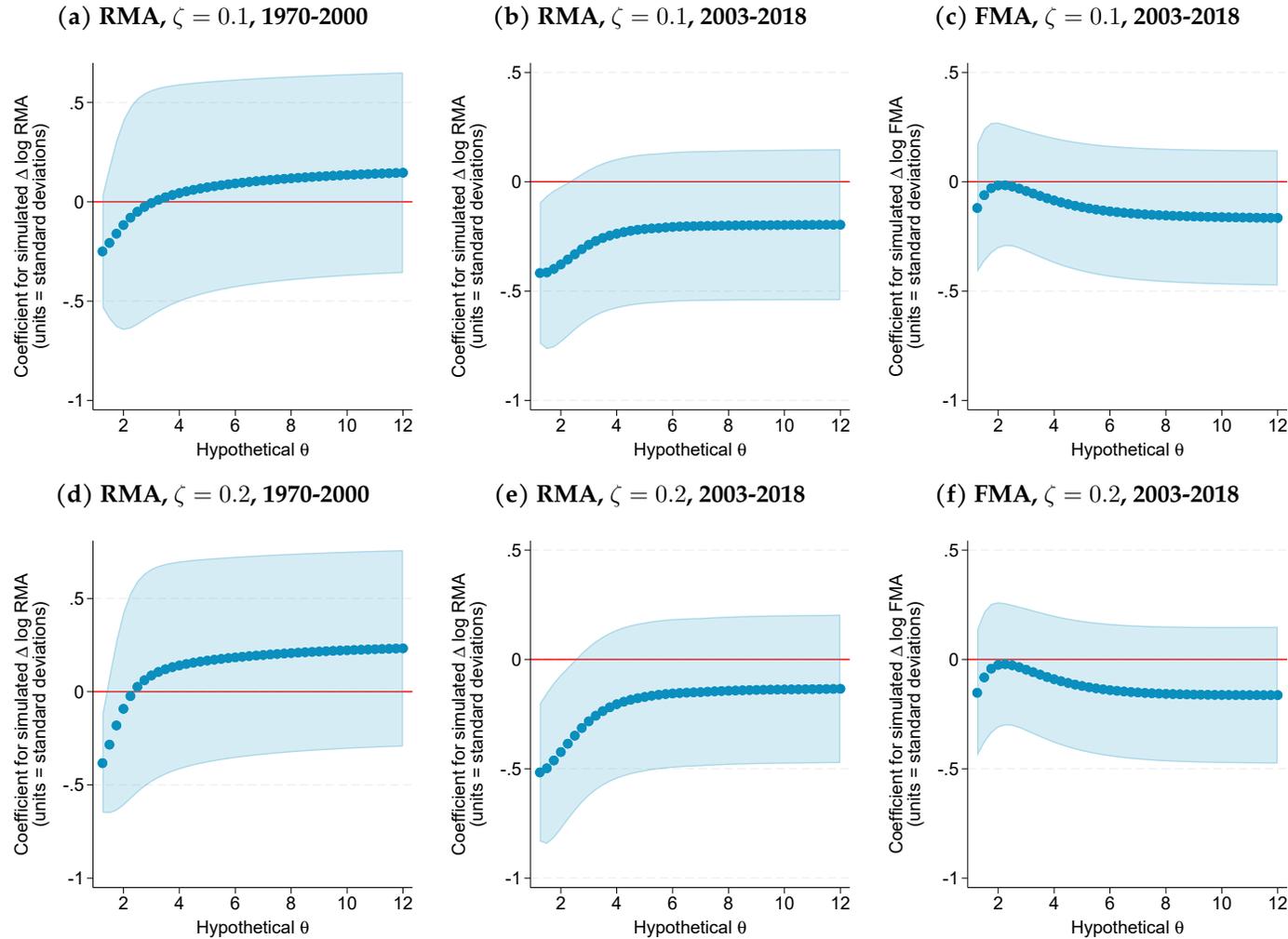
Note: Spatial RD plots display changes in log residential commuting for tracts within a 5-minute drive of Philadelphia's city limits. Positive x-values are inside the city. Panels (a)-(d) show all end years for base year 2003. Standard errors clustered at the grid-cell level.

Figure A.16: Continuity check: RD plots for change in log workplace commuting: end years for 2003 base



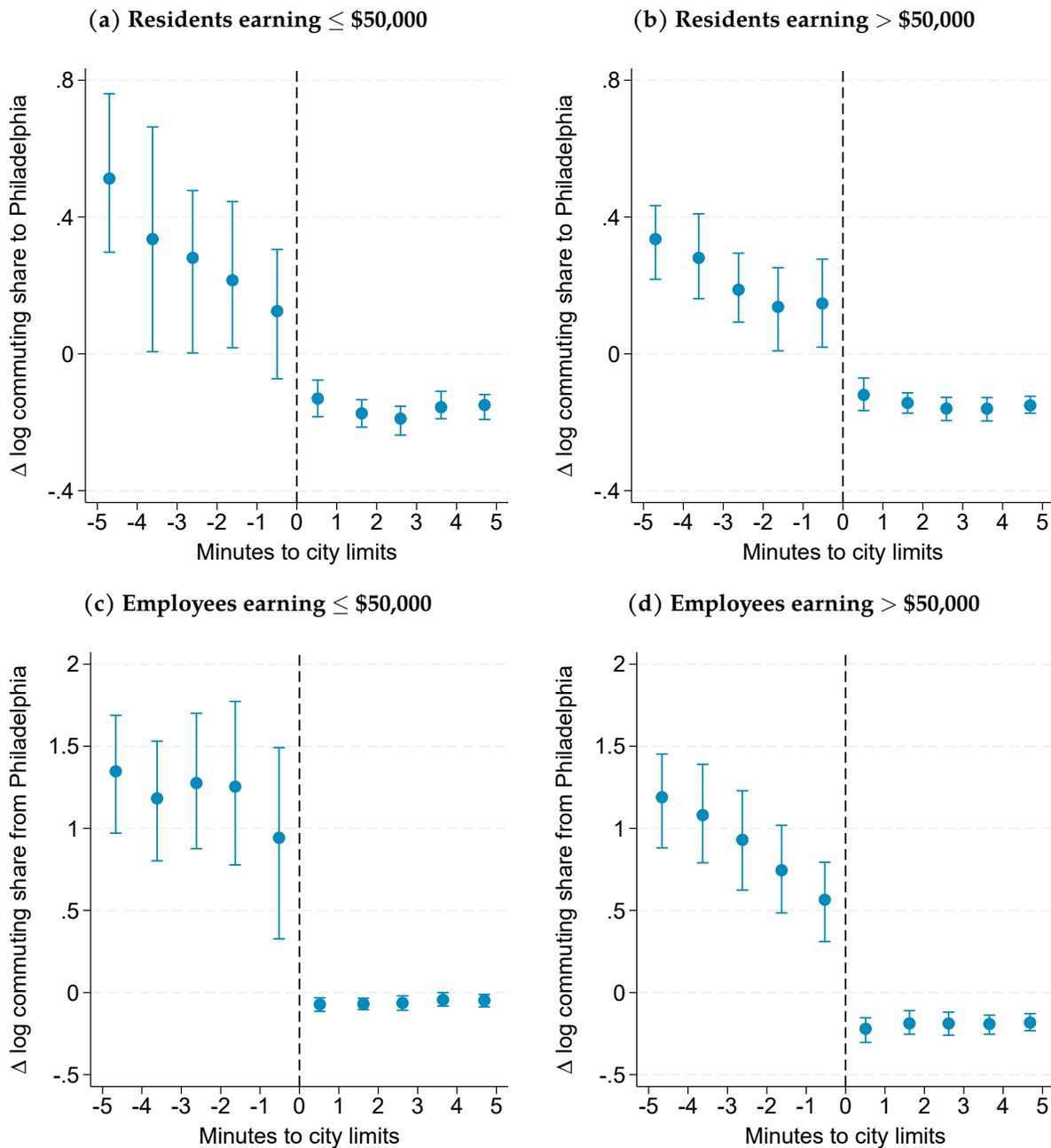
Note: Spatial RD plots display changes in log workplace commuting for tracts within a 5-minute drive of Philadelphia's city limits. Positive x-values are inside the city. Panels (a)-(d) show all end years for base year 2003. Standard errors clustered at the grid-cell level.

Figure A.17: Robustness: Continuity in Simulated Resident Market Access or Firm Market Access



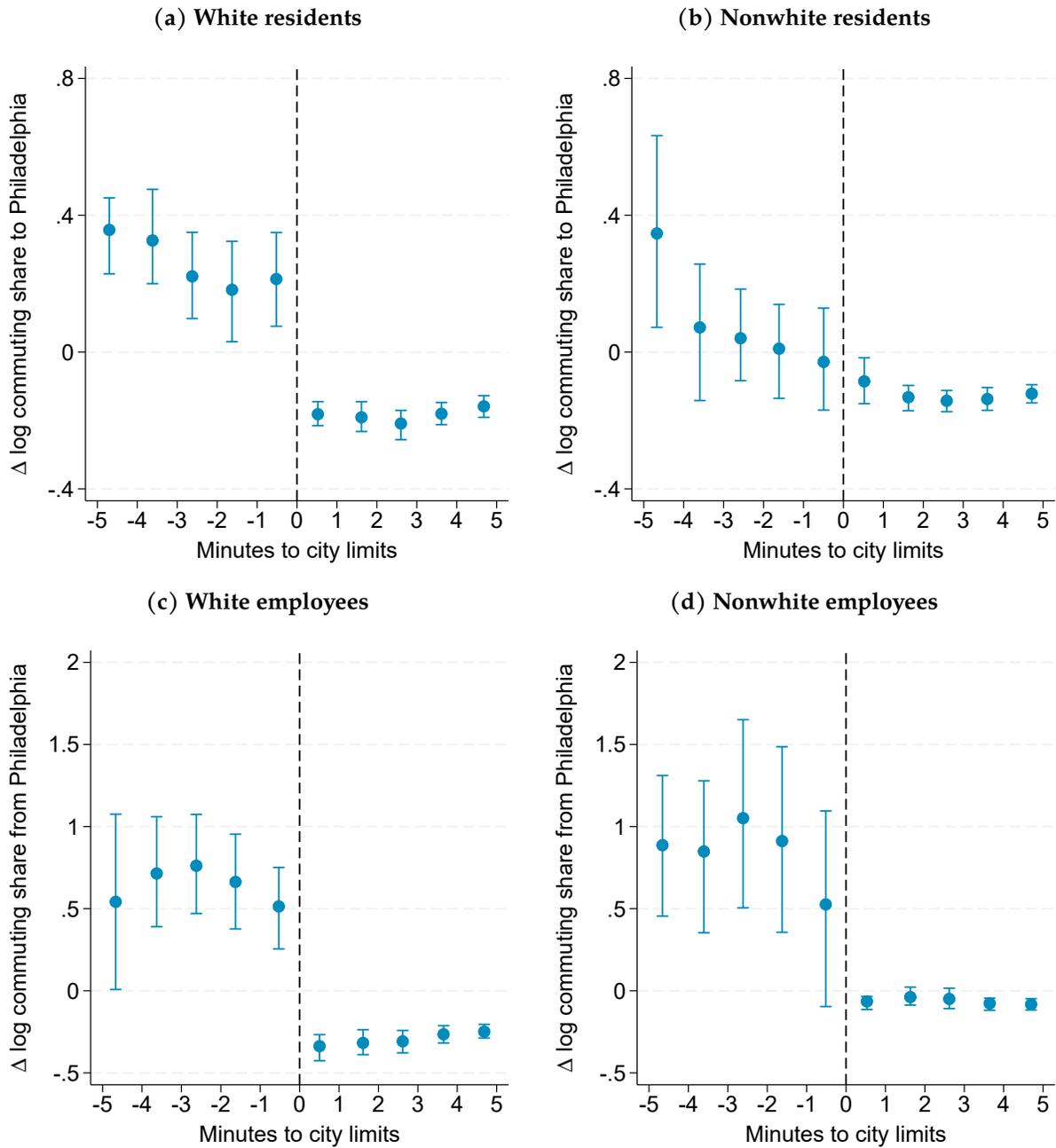
Note: The coefficient plots show the estimated RD coefficients for (simulated) changes in RMA and in FMA terms for different hypothesized values of the border effect parameter (ζ) and the Fréchet parameter (θ). To construct RMA, we use a travel time matrix between census tracts in our RD sample and every other census tract in Philadelphia and its collar counties. We then merge employee counts for every census tract in Philadelphia and the other counties, divide by travel times, raise to the power of the hypothesized value of the Fréchet spread parameter, and then aggregate across workplace tracts to obtain gross-of-tax RMA. We then compute the change in log total RMA terms from Equations 6 and 7 for different base-end year pairs (1970-2000, 2003-2018), and different hypothetical values for the border effect parameter (ζ) and the commuting elasticity (θ). To construct the change in log total FMA, we perform a similar procedure using working age population counts for census tracts instead of employee counts. Employee counts from Dun & Bradstreet not available before 1969. Standard errors clustered at the grid-cell level.

Figure A.18: Robustness: RD plots for changes in log commuting shares to or from Philadelphia by nominal income bin, 2000-2014 (\downarrow wage tax)



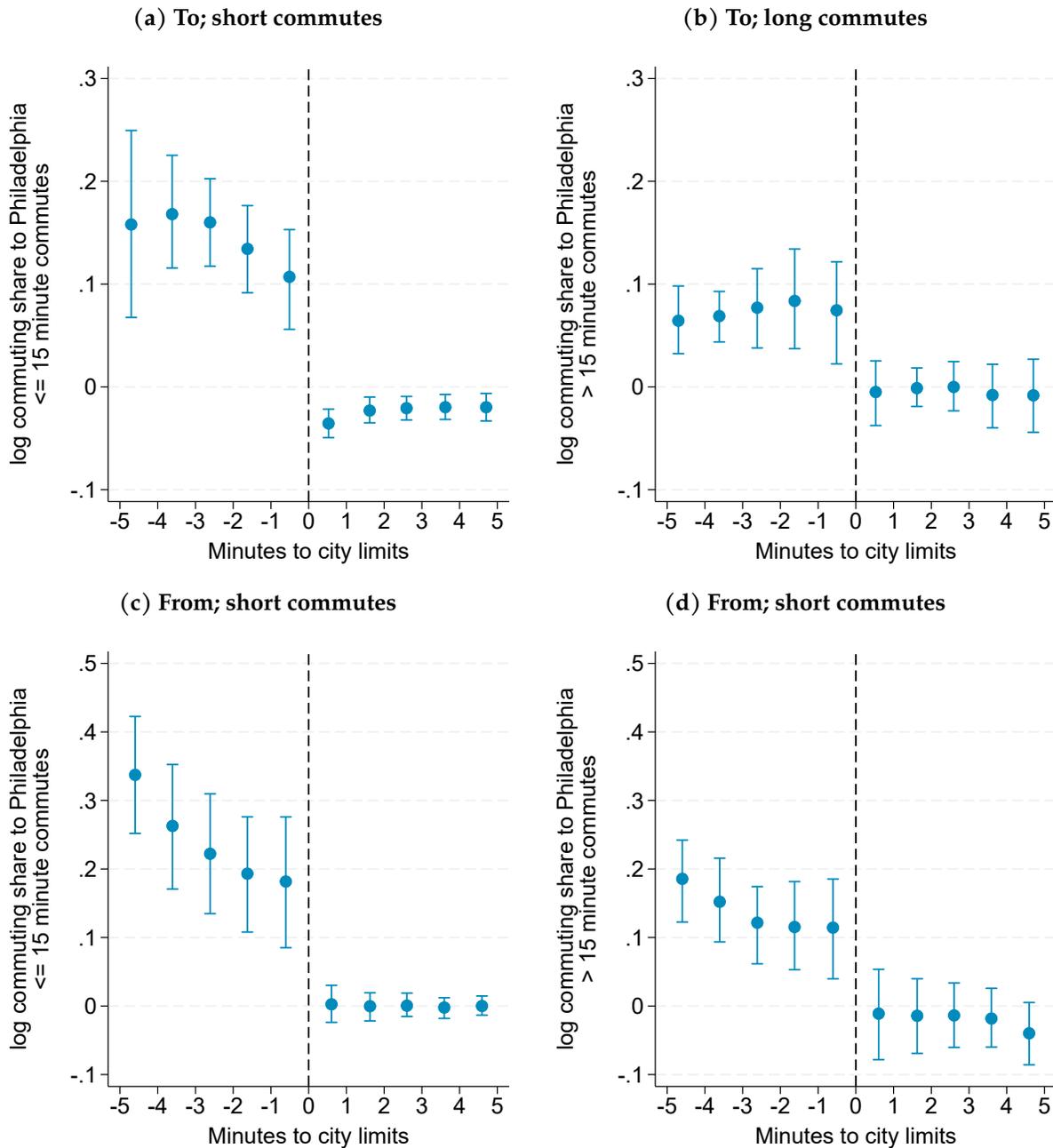
Note: Spatial RD plots display 2000–2014 changes in log commuting shares for tracts within a 5-minute drive of Philadelphia’s city limits. Positive x-values are inside the city. Earnings bins are in nominal dollars. Panel (a): change in log share of residents earning \leq \$50,000 that commute to Philadelphia, 2000–2014. Panel (b): change in log share of residents earning more than \$50,000 that commute to Philadelphia, 2000–2014. Panel (c): change in log share of employees earning \leq \$50,000 that commute from Philadelphia, 2000–2014. Panel (d): change in log share of employees earning more than \$50,000 that commute from Philadelphia, 2000–2014. Standard errors clustered at the grid-cell level.

Figure A.19: Robustness: RD plots for changes in log commuting shares to or from Philadelphia by race, 2000-2014 (↓ wage tax)



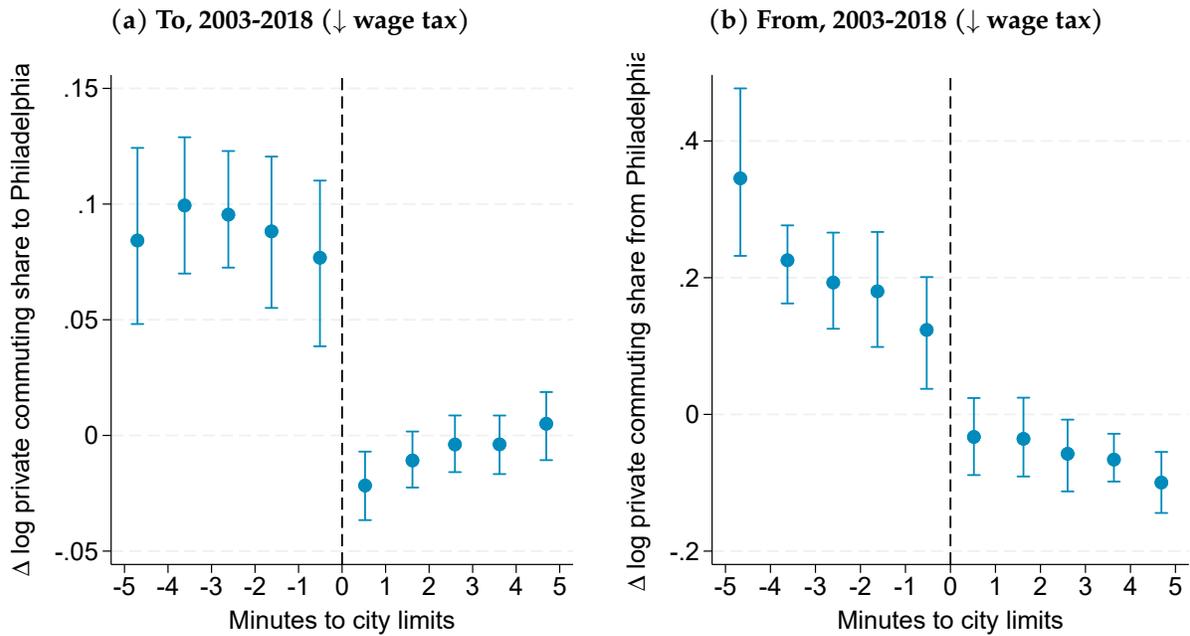
Note: Spatial RD plots display 2000–2014 changes in log commuting shares for tracts within a 5-minute drive of Philadelphia’s city limits. Positive x-values are inside the city. Panel (a): change in log share of white residents that commute to Philadelphia, 2000–2014. Panel (b): change in log share of nonwhite residents that commute to Philadelphia, 2000–2014. Panel (c): change in log share of white employees that commute from Philadelphia, 2000–2014. Panel (d): change in log share of nonwhite employees that commute from Philadelphia, 2000–2014. Standard errors clustered at the grid-cell level.

Figure A.20: Robustness: RD plots for changes in log commuting shares to or from Philadelphia by commute time, 2003-2018 (\downarrow wage tax)



Note: Spatial RD plots display 2003–2018 changes in log commuting shares for tracts within a 5-minute drive of Philadelphia’s city limits. Positive x-values are inside the city. Panel (a): change in log share of residents with short commutes (<15 minutes) that commute to Philadelphia, 2003–2018. Panel (b): change in log share of residents with long commutes (\geq 15 minutes) that commute to Philadelphia, 2003–2018. Panel (c): change in log share of employees with short commutes (<15 minutes) that commute from Philadelphia, 2003–2018. Panel (d): change in log share of employees with long commutes (\geq 15 minutes) that commute from Philadelphia, 2003–2018. Commute times between census tracts calculated using uncongested road network (McDonald, 2017). Standard errors clustered at the grid-cell level.

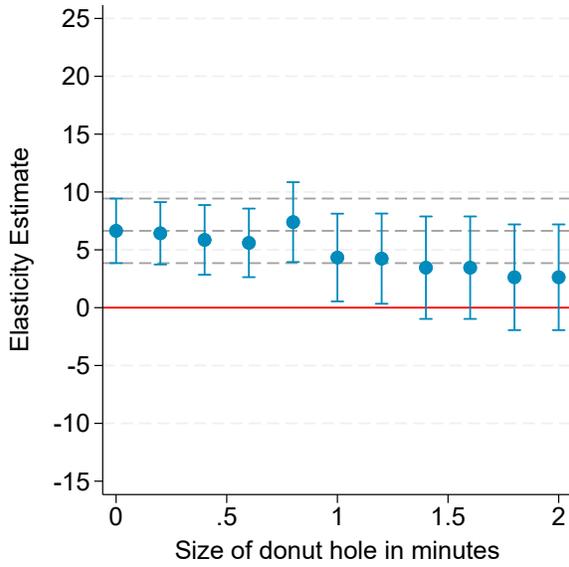
Figure A.21: Robustness: RD plots for change in log private sector commuting share to or from Philadelphia



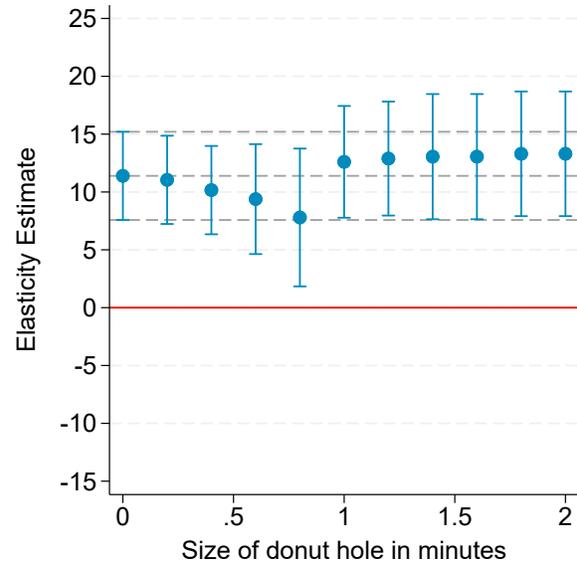
Note: Spatial RD plots display changes in log private sector commuting shares for tracts within a 5-minute drive of Philadelphia’s city limits. We exclude government jobs from the LODES data. Positive x-values are inside the city. Panel (a): change in log share of residents commuting to Philadelphia, 2003–2018. Panel (b): change in log share of employees commuting from Philadelphia, 2003–2018. Standard errors clustered at the grid-cell level.

Figure A.22: Robustness: Elasticities from donut hole specifications

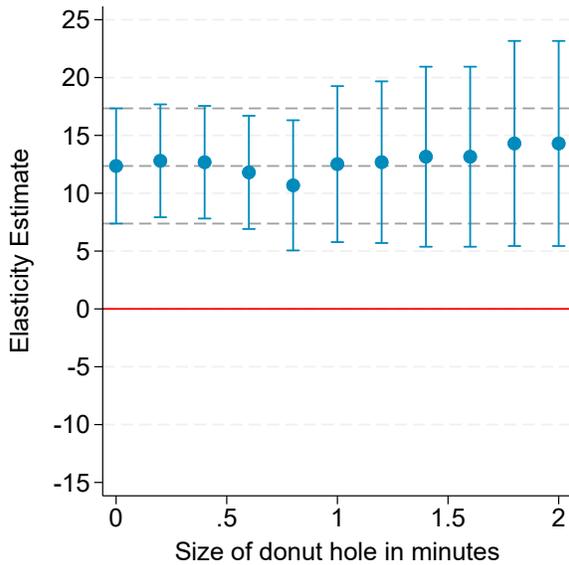
(a) To \mathcal{C} , 1960 base year



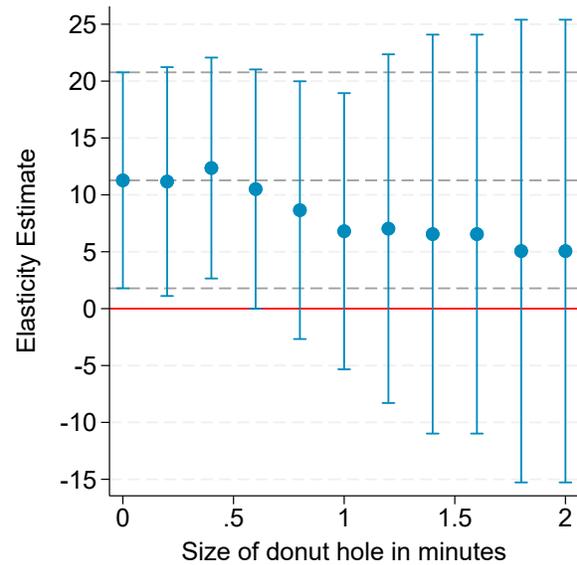
(b) To \mathcal{C} , 1970 base year



(c) To \mathcal{C} , 2003 base year

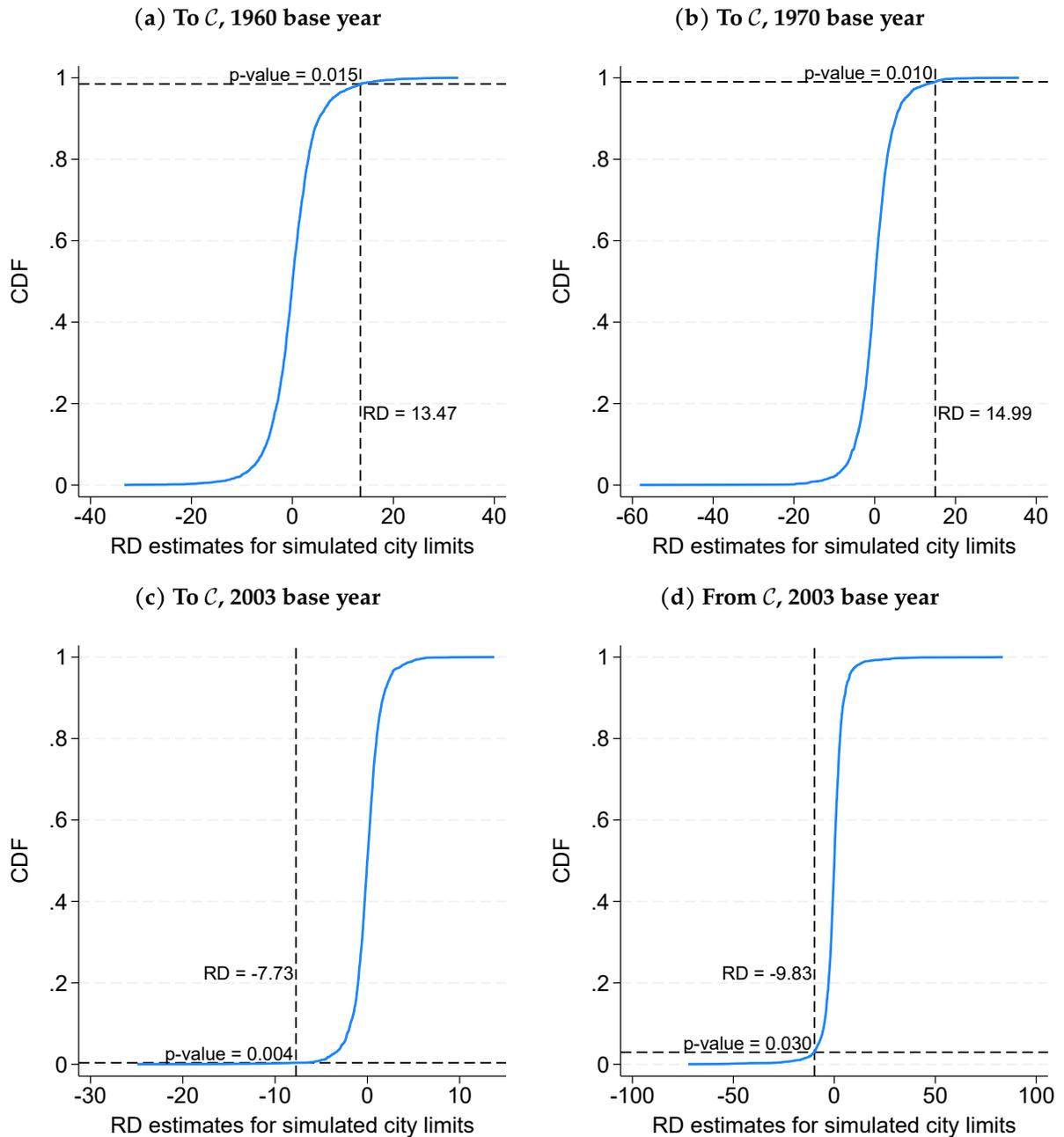


(d) From \mathcal{C} , 2003 base year



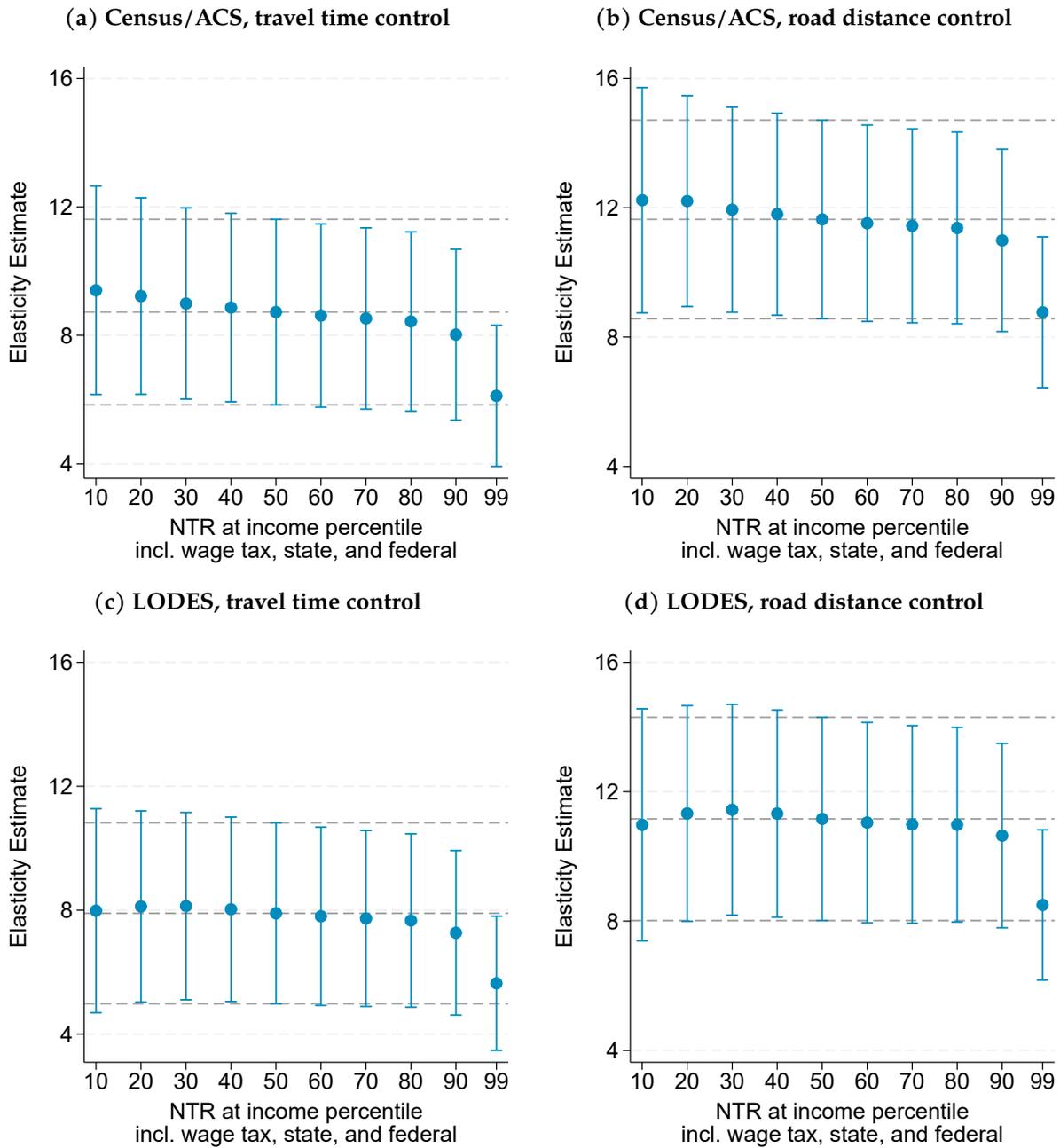
Note: The coefficient plots display the point estimate and 95% CI for donut hole RD specifications. The x-axis is the size of the hole, in minutes to the Philadelphia city limits. Panel (a): pooled commuting elasticity from Δ log commuting share to Philadelphia, base 1960. Panel (b): pooled commuting elasticity from Δ log commuting share to Philadelphia, base 1970. Panel (c): pooled commuting elasticity from Δ log commuting share to Philadelphia, base 2003. Panel (d): pooled commuting elasticity from Δ log commuting share from Philadelphia, base 2003. The first coefficient in each panel (where the size of the donut hole equals 0) is equivalent to that shown in the Fuzzy RD column in Table 2. Standard errors clustered at the grid-cell level.

Figure A.23: Robustness: Randomization inference with simulated city limits



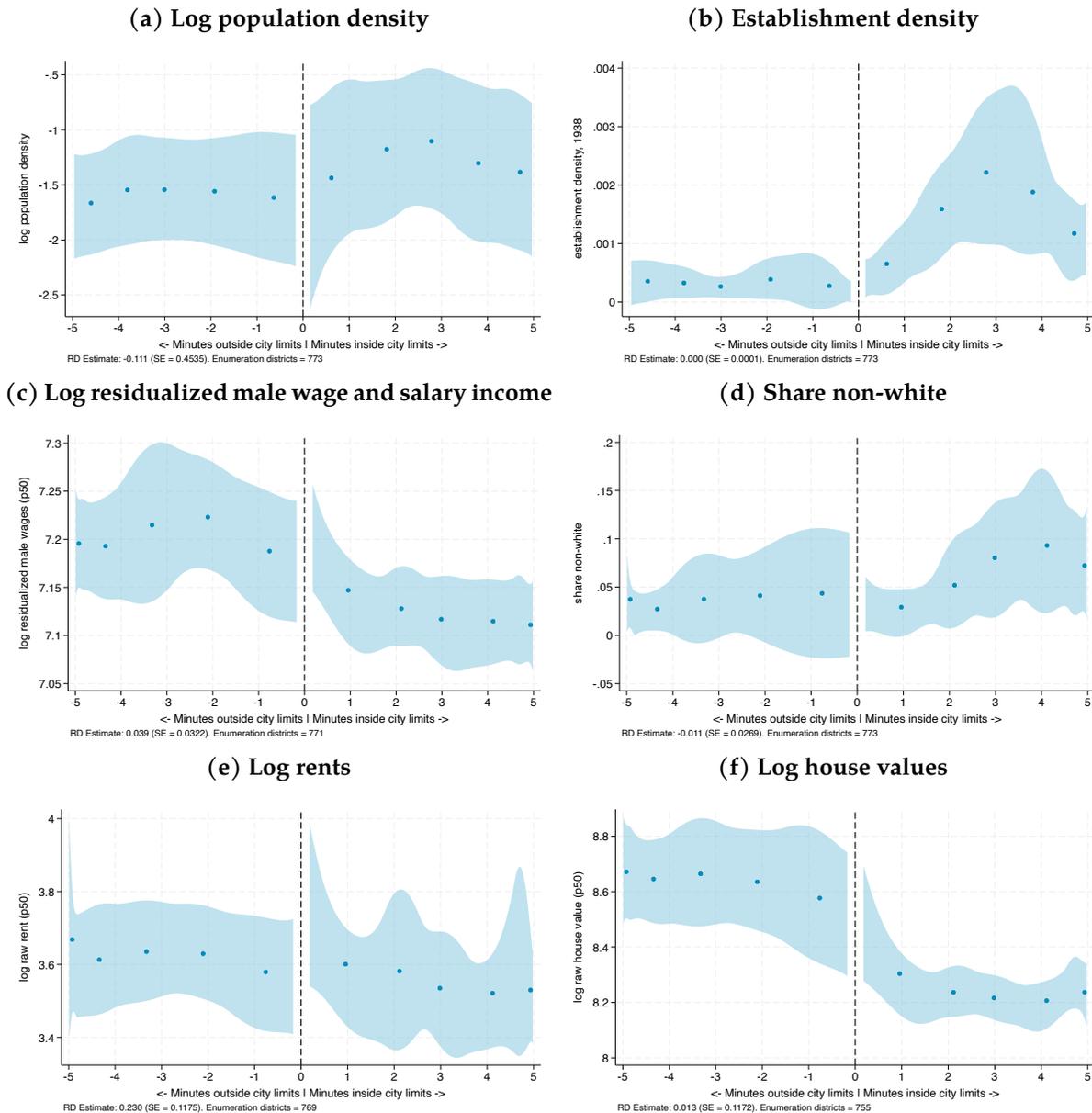
Note: The figures display the CDF of (reduced form) RD coefficients of changes in log commuting shares for 2,500 simulated borders for Philadelphia County. To construct the simulated city limits, we draw straight lines at random (such that there is tangency with the actual Philadelphia County border in at least one point), and compute straight-line distance between census block centroids and the simulated border. We assign negative values to census tracts that lie to the northwest of the simulated border. We then estimate the spatial RD design, controlling for log distance to City Hall. The RD estimates shown next to the vertical dashed lines are the reduced form coefficients from Table 2. The p-values are the fraction of estimates larger than those in Table 2 for 1960 and 1970, and smaller than those in Table 2 for 2003.

Figure A.24: Robustness: Gravity equation estimates for net-of-tax rates at different income percentiles



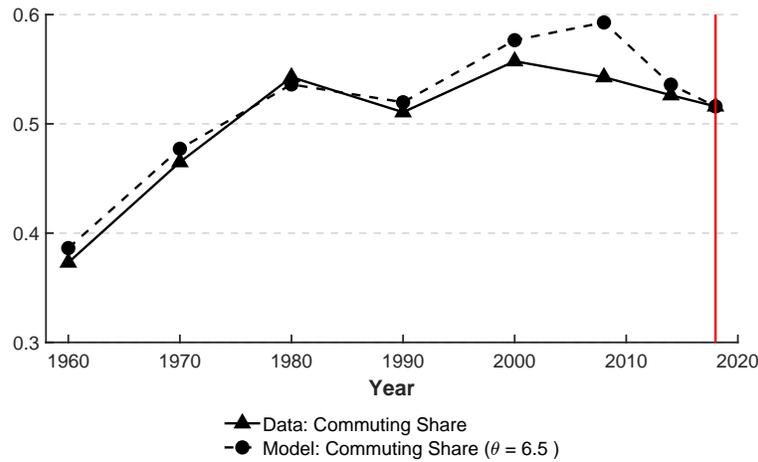
Note: The figures display robustness of the elasticities in Table 3 to different net-of-tax rates by wage and salary income percentiles in the Philadelphia Commuting Zone. The coefficients at the 50th percentile in Panel (a), (b), (c), and (d) correspond to the coefficients in Table 3 Columns (1), (2), (3), and (4). The point estimate and 95% CI are also displayed in gray dashed lines. Standard errors used for 95% CIs are clustered at the home municipality-by-work municipality level.

Figure A.25: RD on measures of male wage and salary income, rents, and house values around Philadelphia’s borders in 1940



Note: The figure displays RD estimates on log price measures for 1940 enumeration districts within a five-minute drive of Philadelphia’s city limits. Establishment density obtained from the 1938 Industrial Directory of Pennsylvania is not logged to include zeroes. Male wage residuals are those estimated for employed males ages 16 to 64 in Philadelphia and its four collar counties, and are orthogonal to indicators for educational attainment, an age quartic, indicators for race-by-Hispanic status, and an indicator for foreign-born status. The index is the simple average of the z-scores for each measure (using the mean and standard deviation in the enumeration districts outside of Philadelphia). Controls include latitude and longitude. Standard errors clustered at percentiles of distance from City Hall.

Figure A.26: Observed, model-predicted share of workers living and working in suburbs



Note: The figure displays observed and model-predicted commutes starting and ending outside of Philadelphia County as a share of all commutes in Bucks, Camden, Delaware, Gloucester, Montgomery, and Philadelphia counties. The commuting elasticity that minimizes the sum of squared deviations between the model's predictions for the suburb-to-suburb commuting share and the data is 6.5. We use LODES data for 2018, marked by the vertical red line, to estimate the model baseline. See Section 6.2.2 for assumed model parameter values for consumption and production.

Table A.1: Summary of Fréchet parameter and tax elasticity estimates

Study	Sample	Variation	Estimation	Estimate	
<i>Panel A: Fréchet parameter</i>					
Ahlfeldt et al. (2015)	Berlin city blocks	Division, reunification	QSM	6.83	
Monte et al. (2018)	U.S. counties	Million dollar plants	QSM	3.30	
Bryan and Morten (2019)	Indonesian regencies (migration)	–	Gravity equation	3.18	
Bryan and Morten (2019)	U.S. states (migration)	–	Gravity equation	2.69	
Heblich et al. (2020)	Greater London boroughs	Steam railway	QSM	5.25	
Owens et al. (2020)	Detroit census tracts	–	Gravity equation	4.62-8.34	
Allen and Arkolakis (2022)	Seattle grid cells	–	–	6.83	
Zárate (2022)	Mexico City municipalities	Subway	Gravity equation	3.11-4.66	
Kreindler and Miyauchi (2023)	Dhaka, Colombo cell phone towers	–	Gravity equation	1.57-5.00	
Severen (2023)	Los Angeles census tracts	LA Metro Rail	Gravity equation	2.18-2.90	
Tsivanidis (2024)	Bogotá city blocks	TransMilenio (BRT)	Gravity equation	3.40	
Franklin et al. (2024)	Addis Ababa neighborhoods	Public works program	Gravity equation	2.08	
<i>Panel B: Tax elasticity</i>					
				<i>Stock</i>	<i>Flow</i>
Moretti and Wilson (2017)	U.S. scientists (migration)	State income tax rate	Diff-in-diff	0.4	1.8
Agrawal and Foremny (2019)	Rich Spanish taxpayers (migration)	Region top income tax rate	IV	0.9	1.7
Martínez (2022)	Rich Swiss taxpayers (migration)	Canton income tax rate	Diff-in-diff	2.0	7.2
Rubolino and Giommoni (2023)	Italian taxpayers (migration)	Municipal income tax rate	Panel FE	1.2	2.2
Agrawal and Tester (2024)	U.S. pro golfers (labor supply)	State income tax rate	Panel FE	–	2.1
<i>Panel C: Elasticity w/r/t wage tax</i>					
				<i>Stock</i>	<i>Flow (θ)</i>
This paper	Census tracts, Phila. city limits	Wage tax rate	Spatial RD	–	6.39
This paper	Census tracts, Phila. region	Wage tax rate	Gravity equation	–	8.57
This paper	Census tracts, Phila. region	Wage tax rate	QSM	0.86	6.50

Note: This table displays Fréchet spread parameter estimates (Panel A), migration elasticity of taxes (Panel B), and the commuting elasticity to the wage tax estimates in this paper (Panel C). We emphasize studies of within-city variation wherever possible, complemented by other studies of within-country variation referenced therein. QSM refers to quantitative spatial model. Gravity equations refer to regressing bilateral commuting flows on some measure of transportation costs, usually using PPML or IV-PPML methods. Allen and Arkolakis (2022) borrow estimate from Ahlfeldt et al. (2015). The median Fréchet parameter estimate in Panel A is 4.7. The median tax elasticities across Panel B are 1.05 (stock) and 2.1 (flow).

Table A.2: Geocode rates by county for Dun & Bradstreet establishments, 1970-2018

	All counties	Bucks County	Delaware County	Montgomery County	Philadelphia County
<i>Geocode rate, all years</i>					
Establishments (raw)	98.82	99.85	93.81	99.48	99.88
Employment counts	97.84	99.45	90.44	99.23	98.63
<i>1970 Geocode rate</i>					
Establishments (raw)	98.82	98.78	96.67	97.98	99.52
Employment counts	95.03	98.75	95.00	97.47	93.80
<i>1980 Geocode rate</i>					
Establishments (raw)	98.63	99.23	95.44	98.53	99.61
Employment counts	97.21	98.16	90.23	98.26	98.29
<i>1990 Geocode rate</i>					
Establishments (raw)	99.33	99.46	98.72	99.09	99.65
Employment counts	98.55	97.76	98.56	99.00	98.55
<i>2000 Geocode rate</i>					
Establishments (raw)	98.77	99.90	94.02	99.36	99.87
Employment counts	98.53	99.78	93.32	99.28	99.43
<i>2008 Geocode rate</i>					
Establishments (raw)	98.86	99.93	93.48	99.66	99.95
Employment counts	98.21	99.85	89.40	99.59	99.51
<i>2014 Geocode rate</i>					
Establishments (raw)	98.76	99.94	92.95	99.65	99.96
Employment counts	98.37	99.87	89.39	99.65	99.69
<i>2018 Geocode rate</i>					
Establishments (raw)	98.74	99.95	92.81	99.66	99.98
Employment counts	97.62	99.98	85.04	99.62	99.72

Note: For each of the counties in our spatial RD sample, we display geocode rates for select years by establishment or by establishments weighted by employee count. We consider an observation to be successfully geocoded if 1) the address is matched to a (non-administrative) point or street point, and 2) if the Dun & Bradstreet county and the geocoded county match.

Table A.3: Robustness: continuity of characteristics at city limits at baseline for different subsamples

	Subsample 1			Subsample 2			Subsample 3			Subsample 4			Subsample 5		
	Sub.	Phil.	Δ	Sub.	Phil.	Δ									
<i>Panel A: Workplace characteristics in 1938</i>															
Manufacturing plants per km ²	0.524	0.744	0.220	-0.194	0.741	0.935	0.038	0.725	0.686	0.496	1.178	0.683	0.015	-0.000	-0.015
Mfg employees per km ²	16	-245	-261	-4	10	14	43	42	-1	10	22	12	1	-0	-1
Mfg employees per km ² (\perp industry FEs)	11	-583	-594	-7	-0	7	16	66	50	59	132	73	1	-0	-1
<i>Panel B: Main residential characteristics in 1940</i>															
Log population per km ²	6.291	7.142	0.851	7.748	7.581	-0.167	5.617	5.816	0.199	6.516	6.919	0.403	4.404	3.821	-0.584
Log residential rents	3.446	4.507	1.062*	3.717	3.914	0.197	3.492	3.712	0.219	3.377	3.056	-0.321	3.050	3.202	0.152*
Log house values	8.755	8.500	-0.255	8.634	8.877	0.244	8.615	8.945	0.330	8.258	7.819	-0.440	8.277	8.490	0.213
Log male wage (\perp Mincer ctrls)	7.126	7.175	0.049	7.200	7.225	0.025	7.059	7.169	0.110	7.182	7.190	0.008	6.972	7.004	0.032
<i>Panel C: Other residential characteristics in 1940</i>															
Share population nonwhite	0.064	0.005	-0.058	0.006	0.021	0.015	0.014	0.026	0.012	0.074	0.215	0.141	0.054	0.037	-0.017
Share population foreign born	0.128	0.138	0.009	0.090	0.131	0.041**	0.110	0.172	0.062	0.086	0.104	0.018	0.099	0.077	-0.022
Household size	5.440	4.076	-1.364	4.002	3.883	-0.119	4.595	5.786	1.192	4.416	4.808	0.392	4.000	4.187	0.187
Homeownership rate	0.627	0.626	-0.001	0.529	0.436	-0.093*	0.668	0.506	-0.162**	0.514	0.429	-0.085	0.645	0.522	-0.123
Employment/population (males 16-64)	0.763	0.821	0.057***	0.800	0.772	-0.028	0.803	0.827	0.024	0.813	0.707	-0.107**	0.797	0.729	-0.068
Mfg emp/pop (males 16-64)	0.232	0.315	0.083**	0.178	0.177	-0.001	0.230	0.203	-0.027	0.341	0.273	-0.068	0.283	0.170	-0.113***
Share males 25-54 w/ some college	0.236	0.164	-0.072	0.253	0.280	0.027	0.271	0.246	-0.026	0.180	0.057	-0.123	0.093	0.240	0.147***

Note: The table presents robustness checks for 1 for different subsamples of our 1940 data. The subsamples are five clusters of enumeration districts along the city limits. The columns show the estimate for each variable at the suburbs' limits with Philadelphia (i.e., the estimate from the left), Philadelphia's limits with the suburbs (estimate from the right) and the discontinuity at the city limits. The RD sample consists of the 773 enumeration districts within a 5 minute drive of the city limits. Standard errors to obtain statistical significance are clustered at the grid cell level. Stars in the Δ columns denote statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.4: Specification check for pooled commuting elasticities

Base year Specification	1960 (To \mathcal{P})		1970 (To \mathcal{P})		2003 (To \mathcal{P})		2003 (From \mathcal{P})	
	Fuzzy RD	Fuzzy RD	Fuzzy RD	Fuzzy RD	Fuzzy RD	Fuzzy RD	Fuzzy RD	Fuzzy RD
Commuting elasticity (θ)	6.29 (1.36)	6.42 (1.35)	10.85 (1.94)	10.77 (1.88)	12.57 (2.58)	12.85 (2.58)	10.60 (5.12)	10.85 (5.04)
Controls								
Log dist from City Hall	✓	✓	✓	✓	✓	✓	✓	✓
White share	✓	✓	✓	✓	✓	✓		
College share		✓		✓		✓		
Log HH Income		✓		✓		✓		
Share estab. corp.-owned							✓	
Share empl. in corp.-owned								✓
Census Tracts	300	300	300	300	292	292	267	267
Clusters	100	100	100	100	99	99	97	97

Note: This table displays specification checks for the pooled commuting elasticity estimates in Table 2 when controlling for census tract white share of the working population, college share of the working age population, log household income, share of tract establishments that are corporate owned, and share of tract employees that work in corporate-owned establishments. Controls are measured in the base year. Specifications use the optimal bandwidth in [Calonico et al. \(2022\)](#). See Table 2 notes for further details. Standard errors clustered at the grid cell level are shown in parentheses.

Table A.5: Total tax revenue scenarios for counterfactual wage tax rates

Counterfactuals	No agglomeration	No agglomeration	Agglomeration
	Inelastic floorspace (1)	Elastic floorspace (2)	Elastic floorspace (3)
Total Revenue at 0% wage tax	38	39	42
Wage tax revenue	0	0	0
Property tax revenue	22	23	24
Sales tax revenue	16	17	18
Total Revenue at 1% wage tax	56	58	61
Wage tax revenue	19	19	21
Property tax revenue	21	22	24
Sales tax revenue	15	16	17
Total Revenue at 2% wage tax	73	74	78
Wage tax revenue	37	38	40
Property tax revenue	21	21	23
Sales tax revenue	15	15	16
Total Revenue at 3% wage tax	89	89	90
Wage tax revenue	54	54	55
Property tax revenue	21	21	21
Sales tax revenue	15	15	15
Total Revenue at 4% wage tax	105	103	103
Wage tax revenue	70	69	69
Property tax revenue	20	20	20
Sales tax revenue	14	14	14
Total Revenue at 5% wage tax	120	116	113
Wage tax revenue	86	84	82
Property tax revenue	20	19	19
Sales tax revenue	14	13	13
Total Revenue at 6% wage tax	133	126	122
Wage tax revenue	101	95	92
Property tax revenue	19	18	18
Sales tax revenue	13	13	12
Total Revenue at 7% wage tax	147	136	130
Wage tax revenue	115	107	102
Property tax revenue	19	18	17
Sales tax revenue	13	12	11
Total Revenue at 8% wage tax	158	145	136
Wage tax revenue	127	117	109
Property tax revenue	19	17	16
Sales tax revenue	13	11	11
Total Revenue at 9% wage tax	171	154	141
Wage tax revenue	140	127	116
Property tax revenue	18	16	15
Sales tax revenue	12	11	10
Total Revenue at 10% wage tax	183	163	144
Wage tax revenue	153	136	121
Property tax revenue	18	16	14
Sales tax revenue	12	10	9

Note: Values as a proportion of the sum of the wage tax, property tax, and sales tax revenue in 2018. See Section 7.2 for details.

B Data Construction

B.1 Census tract commuting data

1960 We use census tract tabulations for “Employed Population by Place of Work [from printed report]” from NHGIS (ID ds92 code B6Y). The raw tabulations contain no labels for place of work and are simply marked as “Area A”, “Area B”, etc. To match areas to identifiable places, we use scanned versions of the Census Published Tables (Table P-3) available on the IPUMS website (<https://usa.ipums.org/usa/voliii/pubdocs/1960/pubvols1960.shtml>). An example of the scanned versions is shown in Appendix Figure B.27 below. For Philadelphia, we keep commuting flows to Bucks, Delaware, Montgomery, and Philadelphia County. For placebo cities, we keep commuting flows to the central jurisdiction and to any county that has a land border with the central jurisdiction. The Brooklyn-Queens sample only considers commuting between Kings and Queens County. To crosswalk to 2010 census tract definitions, we overlay the 1960 census tract shapefile with the 2010 census tract shapefile in ArcGIS. We then calculate the share of the 1960 tract area that lies in each 2010 tract, which assumes equal population density. Finally, to apportion 1960 flows to 2010 tracts, we multiply 1960 commuting flows by the area shares and aggregate to the 2010 tract level.

Figure B.27: Detail of Census Published Tables for 1960 (Table P-3)

INSIDE SMSA	123	4 130
PHILADELPHIA (PHILADELPHIA CO.)	114	3 946
BUCKS COUNTY, PA	4
CHESTER COUNTY, PA	12
DELAWARE COUNTY, PA	51
MONTGOMERY COUNTY, PA	12
BURLINGTON COUNTY, N. J.	5	4
CAMDEN CITY, N. J.	74
REMAINDER OF CAMDEN COUNTY, N. J.	4	23
GLOUCESTER COUNTY, N. J.	4
OUTSIDE SMSA	21
PLACE OF WORK NOT REPORTED	356

Note: This figure is a detail of the scanned version of the Census Published Tables for 1960 (Table P-3). Rows represent the number of workers aged 16+ residing in a census tract by place of work. Columns are different census tracts whose codes are displayed at the top of the printed table (codes not shown in figure).

1970 We use census tract tabulations for “Place of Work” from NHGIS (ID ds99 code C2F). The raw tabulations contain no labels for place of work and are simply marked as “Place of work 1”, “Place of work 2”, etc. To match areas to identifiable places, we use scanned versions of the Census Published Tables (Table P-2) available on the IPUMS website (<https://usa.ipums.org/usa/voliii/pubdocs/1970/pubvols1970.shtml>). An example of the scanned versions is shown in Appendix Figure B.28 below. For Philadelphia, we keep commuting flows to Bucks, Delaware, Montgomery, and Philadelphia County. For placebo cities, we keep commuting flows to the central jurisdiction and to any county that has a land border with the central jurisdiction. The Brooklyn-Queens sample only considers commuting between Kings and Queens County. To crosswalk to 2010 census tract definitions, we overlay the 1970 census tract shapefile with the 2010 census tract shapefile in ArcGIS. We then calculate the share of the 1970 tract area that lies in each 2010 tract,

which assumes equal population density. Finally, to apportion 1970 flows to 2010 tracts, we multiply 1970 commuting flows by the area shares and aggregate to the 2010 tract level.

Figure B.28: Detail of Census Published Tables for 1970 (Table P-2)

Inside SMSA.....	3 005	563
Philadelphia, Pa. – central business district.....	182	26
Remainder of Philadelphia city (county), Pa.....	2 538	524
Delaware County, Pa.....	38	–
Montgomery County, Pa.....	87	6
Bucks County, Pa.....	95	–
Chester County, Pa.....	26	–
Camden County, N.J.....	32	7
Burlington County, N.J.....	7	–
Gloucester County, N.J.....	–	–
Outside SMSA.....	7	–
Place of work not reported.....	220	16

Note: This figure is a detail of the scanned version of the Census Published Tables for 1970 (Table P-2). Rows represent the number of workers aged 16+ residing in a census tract by place of work. Columns are different census tracts whose codes are displayed at the top of the printed table (codes not shown in figure).

1980 We combine two census tract tables for 1980. First, we identify commuting flows to out-of-state workplaces using “Place of Work – State and County Level” from NHGIS (ID ds107 code DG9). Second, we identify commuting flows to the central city of the SMSA with “Place of Work – SMSA Level” from NHGIS (ID ds107 code DHB). We deduct out-of-state commuting from the total flows to the SMSA. Because of the data structure, we do not consider Brooklyn-Queens in our placebo sample for 1980. To crosswalk to 2010 census tract definitions, we overlay the 1980 census tract shapefile with the 2010 census tract shapefile in ArcGIS. We then calculate the share of the 1980 tract area that lies in each 2010 tract, which assumes equal population density. Finally, to apportion 1980 flows to 2010 tracts, we multiply 1980 commuting flows by the area shares and aggregate to the 2010 tract level.

1990 We use bilateral commuting flows between census tracts from the Census Transportation Planning Package: Urban Element Part 3 (<https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/QFA8IS>). The 1990 CTPP was the third iteration of a Census program designed to inform the planning of transportation infrastructure, but it is the earliest for which residential and workplace shares can be analyzed for the same census tract. The first two iterations were the Urban Transportation Package, launched in 1970, and the Urban Transportation Planning Package, launched in 1980. These data had to be purchased from the Census by Metropolitan Planning Organizations. To our knowledge, these data sets are not available online, in the National Transportation Library, or in the FHWA Library and may be stored on “obsolete mainframe computer storage media” across several MPOs. See <https://www.trbcensus.com/olddata.html>. The data contains bilateral commuting flows between 1990 census tracts, subject to the data suppression rules discussed in the main text. We first transform the data to bilateral flows between 2010 census tract definitions. To crosswalk to 2010 census tract definitions, we overlay the 1990 census tract shapefile with the 2010 census tract shapefile in ArcGIS. We then calculate the share of the 1990 tract area that lies in each 2010 tract, which assumes equal population density. We multiply 1990

commuting flows by the area shares for both residence and workplace tracts, and then aggregate the bilateral data to 2010 census tract definitions. We then apply sample restrictions for workplace census tracts. For Philadelphia, we keep commuting flows to Bucks, Delaware, Montgomery, and Philadelphia County. For placebo cities, we keep commuting flows to the central jurisdiction and to any county that has a land border with the central jurisdiction. The Brooklyn-Queens sample only considers commuting between Kings and Queens County.

2000 We use bilateral commuting flows between census tracts from the Census Transportation Planning Package (CTPP) 2000 Part 3 (<https://rosap.nhtl.bts.gov/view/dot/49849>). The data contains bilateral commuting flows between 2000 census tracts, subject to the data suppression rules discussed in the main text. We first transform the data to bilateral flows between 2010 census tract definitions. To crosswalk to 2010 census tract definitions, we overlay the 2000 census tract shapefile with the 2010 census tract shapefile in ArcGIS. We then calculate the share of the 2000 tract area that lies in each 2010 tract, which assumes equal population density. We multiply 2000 commuting flows by the area shares for both residence and workplace tracts, and then aggregate the bilateral data to 2010 census tract definitions. We then apply sample restrictions for workplace census tracts. For Philadelphia, we keep commuting flows to Bucks, Delaware, Montgomery, and Philadelphia County. For placebo cities, we keep commuting flows to the central jurisdiction and to any county that has a land border with the central jurisdiction. The Brooklyn-Queens sample only considers commuting between Kings and Queens County.

2006-2010 We use bilateral commuting flows between census tracts from the Census Transportation Planning Package (CTPP) for the 2006-2010 ACS (<https://transportation.org/census-transportation-s-datasets/previous-datasets/2006-2010-5-year-ctpp/>). The data contains bilateral commuting flows between 2010 census tracts, subject to the data suppression rules discussed in the main text. For Philadelphia, we keep commuting flows to Bucks, Delaware, Montgomery, and Philadelphia County. For placebo cities, we keep commuting flows to the central jurisdiction and to any county that has a land border with the central jurisdiction. The Brooklyn-Queens sample only considers commuting between Kings and Queens County.

2012-2016 We use bilateral commuting flows between census tracts from the Census Transportation Planning Package (CTPP) for the 2012-2016 ACS (<https://transportation.org/census-transportation-s-datasets/previous-datasets/2012-2016-ctpp/>). The data contains bilateral commuting flows between 2010 census tracts, subject to the data suppression rules discussed in the main text. For Philadelphia, we keep commuting flows to Bucks, Delaware, Montgomery, and Philadelphia County. For placebo cities, we keep commuting flows to the central jurisdiction and to any county that has a land border with the central jurisdiction. The Brooklyn-Queens sample only considers commuting between Kings and Queens County.

2002-2019 We use Version 7 of the LEHD Origin-Destination Employment Statistics (<https://lehd.ces.census.gov/data/>). The Origin-Destination data contains bilateral commuting flows between 2010 census blocks for each year between 2002 and 2019, subject to the added noise discussed in the main text. We aggregate flows to 2010 bilateral census tract definitions. We also take three-year averages (considering zeroes) to construct flows for 2003, 2006, ..., 2018. For Philadelphia, we keep commuting flows to Bucks, Delaware, Montgomery, and Philadelphia County. For placebo cities, we keep commuting flows to the central jurisdiction and to any county that has a land border with the central jurisdiction. The Brooklyn-Queens sample only considers commuting between Kings and Queens County.

B.2 Drive times to city limits

We generate points, spaced 100 meters apart, along the boundary for Philadelphia County. To ensure continuity in commuting costs, we retain all points located on land but exclude points situated on bodies of water. We exclude these points because ArcGIS defaults to the nearest road when using the Generate Drive Time Trade Area tool for points located on water. In these cases, the nearest road may be far from the jurisdiction border, and including such points to calculate drive times would introduce measurement error to our running variable. Drive time surfaces, or isochrones, are obtained from these points in increments of 0.2 minutes up to 5 minutes. At an average speed of 24 mph, the sample spans 2 miles (road distance) on either side of the city limits. By overlaying census tracts on the 5-minute drive time isochrone, we obtain the running variable at the drive-time-surface-by-tract level. This means there are multiple observations per tract, which we take into account for inference. We also weight each drive-time-surface-by-tract observation by the area share that the surface occupies in the census tract.

B.3 Tax rates for bilateral commutes

We use historical statutory wage tax rates for each municipality from [Strumpf \(1999\)](#) and Pennsylvania's Municipal Statistics agency.⁴⁰ We assume that all income is earned through wages and salaries, account for all institutional details we describe in Section 2.1, and that workers pay the average federal income tax rate for the median wage and salary income in the Philadelphia CZ.

To calculate the wage and salary income percentiles, we use the 1960 5% sample, 1970 1% metro sample, 1980 5% sample, 1990 5% sample, 2000 5% sample, 2010 ACS 5-year sample, 2016 ACS 5-year sample, and combine the 2017, 2018, and 2019 ACS 1-year samples ([Ruggles et al., 2022](#)). Our universe consists of employed individuals ages 25 to 54 not in group quarters that report strictly positive, non-missing wage and salary earnings in the Philadelphia CZ. We use mini-PUMA-to-CZ crosswalks for 1960, as described in [Rose \(2018\)](#), to match observations to the Philadelphia CZ. We otherwise use geography crosswalks from [Autor and Dorn \(2013\)](#). We interpolate income percentiles for years between these samples to obtain yearly estimates for wage and salary income.

⁴⁰See https://munstats.pa.gov/Public/ReportInformation2.aspx?report=taxes_Dyn_Excel.

To assign tax rates to bilateral commutes, we consider home and work jurisdictions for four different cases. For commutes between Pennsylvania municipalities, we assume that workers pay the larger wage tax rate between the home and work jurisdictions, plus the flat PA income tax rate. For residents of PA that work in NJ, we assume that workers pay the local resident wage tax and the PA income tax. For residents of NJ that work in PA, we assume the worker is eligible for NJ tax credits paid to other jurisdictions, and therefore pays the larger amount between the average NJ income tax rate and the nonresident wage tax rate, per Schedule NJ-COJ (also known as Schedule A; <https://www.nj.gov/treasury/taxation/njit14.shtml>). For commutes that start and end in NJ, we assume the worker pays the average state income tax rate for the median wage and salary income in the Philadelphia CZ. State tax rates are set to 0 for 1960 and 1970, as neither state taxed income in those years. We use the NBER-TAXSIM model to compute state and federal tax rates for the median wage and salary income in the Philadelphia Commuting Zone, which we compute using Census/ACS data. To allow for some dynamic adjustment to changes in the wage tax, we average the net-of-tax rates in years t , $t - 1$, and $t - 2$ for each base and end year.

B.4 1938 Industry Directory

We digitize the 1938 Industrial Directory of the Commonwealth of Pennsylvania. Assembled by the state statistical bureau, this directory contains the name, location of plant, location of main office, employment count, and industrial sector for thousands of industrial plants. The location of plants in and near Philadelphia is given by an address string, which we geocode using ArcMap. If the string only denotes the name of a suburban township, we assign the township centroid to that plant. We then use plant latitude and longitude to assign plants to 1940 enumeration districts, which we describe below. We use consistent industry classifications from [Fiszbein et al. \(2020\)](#), who map strings listing industries from this time period into IPUMS ind1950 codes. We use employment counts for the universe of plants in Philadelphia, Bucks, Montgomery, and Delaware counties and use PPML to estimate industry fixed effects for plant size.

B.5 1940 Census and enumeration districts

We use microdata from the full-count 1940 Census ([Ruggles et al., 2021](#)) and merge to place identifiers from [Berkes et al. \(2023\)](#) to identify the set of enumeration districts in 26 municipalities surrounding Philadelphia. Enumeration districts are the rough historical equivalent of census tracts. For these municipalities, we georeference digitized enumeration district maps from the National Archives Catalog and build a shapefile linkable to the census microdata. We then combine this GIS data with a shapefile for 1940 enumeration districts within Philadelphia from [Shertzer, Walsh, and Logan \(2016\)](#). Finally, we overlay the drive time isochrone to the enumeration districts, and obtain a sample of 773 enumeration districts within a 5-minute drive of the Philadelphia County border.

B.6 Dun & Bradstreet (1969-2023)

We use yearly Dun & Bradstreet Private Company Listings from 1969 to 2023. This data set contains uniquely-identified establishments' addresses, SIC industries, corporate ownership statuses, employment counts, and sales estimates. We geocode establishments using the Census Geocoding API and ArcMap, achieving an establishment geocode match rate of 98.82%. Appendix Table A.2 shows geocode rates by year and county. To measure employee count and sales, we use establishment employment and sales estimates provided directly in the raw data. Because we cannot measure wages directly, we instead compute log sales per worker and, weighting establishments by employee counts, residualize with state-by-year-by-4-digit-SIC-industry fixed effects. We then compute the median residual per census tract, again weighting by establishment employee count.

B.7 Infutor data (1990-2019)

We use individual address histories from Infutor to measure migration flows within the U.S. The data includes residential addresses, moving dates, names and basic demographic information. As established in [Diamond, McQuade, and Qian \(2019\)](#), this data set is well suited to measure domestic migration flows and has coverage for the overwhelming majority of U.S. adults. We do not observe place of work information in Infutor data.

Census rent and house value We use census tract tabulations for median contract rent and house value for each census year and ACS sample in our commuting data. In using our area-based crosswalks, we take spatially-weighted averages to compute values at the 2010 census tract level across time, and deflate using the CPI. We then take log changes between census years.

B.8 Imputation of historical land values

We use tract-level land value estimates for 2018 from [Davis et al. \(2021\)](#), and use tract characteristics that we can consistently observe through time to find the best predictors of land values for that year. We use a LASSO linear regression of log land values and consider a second-degree polynomial of logged characteristics, including interactions, to select the predictors. Using the coefficient for the selected variables, we then project land values at each point in time for our sample of census tracts in the Philadelphia metro area. The selected predictors are log acres, the square of log acres, the product of log acres and log median house value, log rent, the square of log household income, and the square of log median house value. The post-LASSO R^2 equals 0.86. Our imputation approach implies that changes in land value across time come exclusively from changes in values for the underlying predictors. Our measure for rents (R_i) is the predicted land value for a 1 acre lot.

B.9 Predicted bilateral commute flows for estimation year

Following [Redding \(2024\)](#), we predict commuting flows using bilateral commuting flow data for the estimation year, the commuting cost matrix for 2017 ([McDonald, 2017](#)), and indicators for cross-

county and cross-state flows. We estimate the model using the 2018 LODES sample as the base.

C Workplace commuting probabilities

We condition (1) on workplace j to obtain expressions for workplace shares, and integrate residences to get the share of workers commuting from Philadelphia in a workplace in Philadelphia,

$$\lambda_{Cj|j \in C} = \frac{(1 - \tau_C^R)^\theta \Phi_{j \in C}^C}{(1 - \tau_C^R)^\theta \Phi_{j \in C}^C + (1 - \tau_C^L)^\theta (1 - \zeta)^\theta \Phi_{j \in C}^S} \quad (\text{C.1})$$

and the share of workers commuting from Philadelphia in a workplace in the suburbs,

$$\lambda_{Cj|j \in S} = \frac{(1 - \tau_C^R)^\theta (1 - \zeta)^\theta \Phi_{j \in S}^C}{(1 - \tau_C^R)^\theta (1 - \zeta)^\theta \Phi_{j \in S}^C + \Phi_{j \in S}^S}$$

where terms such as $\Phi_{j \in \mathcal{K}}^{\mathcal{H}} \equiv \sum_{i \in \mathcal{C}} \left[\frac{B_i^R}{P_i^\alpha Q_i^{1-\alpha} d_{ij}} \right]^\theta$ are gross firm market access variables which integrate residential amenities, consumption good prices, rents, and commuting costs across all neighborhoods in municipality \mathcal{H} for a workplace located in municipality \mathcal{K} . Echoing the results for residential shares, taxes and border effects can pull out of these market access terms since they do not vary between neighborhoods within municipality \mathcal{H} .

Assuming Philadelphia's workplace and residence tax rates are similar (i.e., $\tau_C^R \approx \tau_C^L$), (C.1) becomes

$$\lambda_{Cj|j \in C} = \frac{\Phi_{j \in C}^C}{\Phi_{j \in C}^C + (1 - \zeta)^\theta \Phi_{j \in C}^S}.$$

With this expression in hand, we can construct changes in log commuting shares from Philadelphia as in Section 3.2.2. The outcome takes the following form for a Philadelphia neighborhood,

$$\Delta \ln \lambda_{Cj,t|j \in C} = \ln \frac{\lambda_{Cj,t|j \in C}}{\lambda_{Cj,t_0|j \in C}} = \underbrace{\ln \left(\frac{\Phi_{j,t}^C}{\Phi_{j,t_0}^C} \right)}_{\text{change in FMA from } C} - \underbrace{\ln \left(\frac{\Phi_{j,t}^C + (1 - \zeta)^\theta \Phi_{j,t}^S}{\Phi_{j,t_0}^C + (1 - \zeta)^\theta \Phi_{j,t_0}^S} \right)}_{\text{change in total FMA}}$$

and the next form for a suburban neighborhood,

$$\Delta \ln \lambda_{Cj,t|j \in S} = \underbrace{\theta \ln \left(\frac{1 - \tau_{C,t}^R}{1 - \tau_{C,t_0}^R} \right)}_{\text{wage tax variation}} + \underbrace{\ln \left(\frac{\Phi_{j,t}^C}{\Phi_{j,t_0}^C} \right)}_{\text{change in FMA from } C} - \underbrace{\ln \left(\frac{(1 - \tau_{C,t}^R)^\theta (1 - \zeta)^\theta \Phi_{j,t}^C + \Phi_{j,t}^S}{(1 - \tau_{C,t_0}^R)^\theta (1 - \zeta)^\theta \Phi_{j,t_0}^C + \Phi_{j,t_0}^S} \right)}_{\text{change in total FMA}},$$

which removes variation from time-invariant border effects from the changes in FMA to Philadelphia.

C.1 Continuity assumptions for FMA terms

We assume that two objects are continuous in travel time at the county border: 1) the expectation of log FMA from Philadelphia at any point in time ($\ln \Phi_{j,t}^C$), and 2) the expected change in total FMA inclusive of border effects and taxes.

Continuity in the expectations of log FMA to Philadelphia can be stated as

$$\lim_{x \uparrow 0} \mathbf{E}[\ln \Phi_{j \in \mathcal{S}, t}^C | X_j = x] = \lim_{x \downarrow 0} \mathbf{E}[\ln \Phi_{j \in \mathcal{C}, t}^C | X_j = x] = \phi_t^C,$$

which says that, standing at the city limits, average amenity- and commuting-cost-adjusted (log) rents across Philadelphia vary smoothly if measured in the tract one step to the right, or in the tract one step to the left.

Continuity in the expectation of changes in log (total) FMA can be stated as

$$\lim_{x \uparrow 0} \mathbf{E}[\Delta \ln \Phi_{j \in \mathcal{S}, t} | X_j = x] = \lim_{x \downarrow 0} \mathbf{E}[\Delta \ln \Phi_{j \in \mathcal{C}, t} | X_j = x] = \Delta \phi_t,$$

As in (9), continuity must be tested empirically through various proxies.

C.2 Mapping to RD estimator

Following an analogous procedure to Section 3.3 yields the following expression,

$$\begin{aligned} \delta_{RD} &\equiv \lim_{x \uparrow 0} \mathbf{E}[\Delta \ln \lambda_{\mathcal{C}, j, t} | j \in \mathcal{C} | X_j = x] - \lim_{x \downarrow 0} \mathbf{E}[\Delta \ln \lambda_{\mathcal{C}, j, t} | j \in \mathcal{S} | X_j = x] \\ &= \left(\phi_t^C - \phi_{t_0}^C - \Delta \phi_t \right) - \left[\theta \ln \left(\frac{1 - \tau_{\mathcal{C}, t}^R}{1 - \tau_{\mathcal{C}, t_0}^R} \right) + \left(\phi_t^C - \phi_{t_0}^C - \Delta \phi_t \right) \right] \\ &= -\theta \ln \left(\frac{1 - \tau_{\mathcal{C}, t}^R}{1 - \tau_{\mathcal{C}, t_0}^R} \right), \end{aligned}$$

which links empirical estimates from our RD design to the model parameter and the observed *resident* wage tax rates.

D Panel Gravity Equation and Structural Approach

We derive alternative estimators for the elasticity of commuting to the net-of-tax rate (θ). The exposition closely follows [Heblich, Redding, and Sturm \(2020\)](#).

D.1 Panel Gravity Equation

The model implies the following population mobility condition with the rest of the United States:

$$\bar{U} \left(\frac{L^C + L^S}{L^{USA}} \right) = \Gamma \left(\frac{\theta - 1}{\theta} \right) \left\{ \sum_{k \in \text{CUS}} \sum_{l \in \text{CUS}} [B_{kl} w_l (1 - \tau_{kl}) / d_{kl}]^\theta [P_k^\alpha Q_k^{1-\alpha}]^{-\theta} \right\}^{\frac{1}{\theta}} \quad (\text{D.2})$$

where \bar{U} is expected utility in the wider U.S. economy, $L^C + L^S / L^{USA}$ is the probability that a worker chooses a residence-workplace pair in the Philadelphia CZ, and $\Gamma(\cdot)$ is the gamma function.

From (1) and (D.2), we can write the following gravity equation for the log commuting probability at each point in time:

$$\ln \lambda_{i,j,t} = \alpha_{j,t} + \gamma_{i,t} + \theta \ln(1 - \tau_{i,j,t}) + \theta \ln d_{i,j,t} + \theta(1 - \zeta_{i,j}) + \xi_{i,j,t}$$

where $\alpha_{j,t}$ denotes workplace-by-time fixed effects, $\gamma_{i,t}$ captures residence-by-time fixed effects and common expected utility in the Philadelphia CZ, and $\xi_{i,j,t}$ accounts for idiosyncratic shocks to bilateral amenities and commuting costs. We use this expression as the basis for (11), which allows for a more flexible specification for the time-varying commuting cost and the border effect terms.

D.2 Structural Approach

This approach requires us to think about neighborhoods simultaneously functioning as residences and workplaces. Since we reserve sub-indices i and j for residences and workplaces, we adopt a neutral placeholder subindex k here. In this approach, we assume that both the border effect (ζ_{ij}) and workplace amenities (B_j) are time-invariant.

We assume that a final good is produced in each neighborhood k with Cobb-Douglas technology and three factors: commercial floor space (H_k^L), physical capital (M_k), and labor (L_k). Cost minimization implies that payments to each of these factors are constant shares of revenue (X_k):

$$w_k L_k = \beta^L X_k, \quad q_k H_k^L = \beta^H X_k, \quad r M_k = \beta^M X_k, \quad \beta^L + \beta^H + \beta^M = 1$$

which implies payments for commercial floor space are proportional to workplace income

$$q_k H_k^L = \frac{\beta^H}{\beta^L} w_k L_k \quad (\text{D.3})$$

On the other hand, commuter market clearing implies that neighborhood resident income per capita (v_k) is a weighted average of wages in all locations, where the weights are the commuting

probabilities conditional on residence in (3):

$$v_k = \sum_j \lambda_{kj|k} w_j \quad (\text{D.4})$$

Combining (D.3) and (D.4) yields the land market clearing condition, which states that total income from the ownership of land equals the sum of payments for residential (H_k^R) and commercial (H_k^L) floor space use,

$$Q_k = q_k(H_k^R + H_k^L) = (1 - \alpha) \left[\sum_j \lambda_{kj|k} w_j \right] R_k + \frac{\beta^H}{\beta^L} w_k L_k, \quad (\text{D.5})$$

where R_k denotes residents. This equation can be written in exact hat algebra form (Dekle, Eaton, and Kortum, 2008) as

$$\hat{Q}_{kt} Q_{kt} = (1 - \alpha) \hat{v}_{kt} v_{kt} \hat{R}_{kt} R_{kt} + \frac{\beta^H}{\beta^L} \hat{w}_{kt} w_{kt} \hat{L}_{kt} L_{kt} \quad (\text{D.6})$$

where $\hat{x}_{kt} = x_{kt}/x_{kt_0}$. In turn, the relative change in neighborhood income per capita ($\hat{v}_{kt} v_{kt}$) is

$$\hat{v}_{kt} v_{kt} = \sum_j \left[\frac{\lambda_{kjt|k} \hat{w}_{jt}^\theta (1 - \widehat{\tau_{kjt}})^\theta}{\sum_j \lambda_{kjt|k} \hat{w}_{jt}^\theta (1 - \widehat{\tau_{kjt}})^\theta} \hat{w}_{jt} w_{jt} \right] \quad (\text{D.7})$$

while the relative change in workplace employment ($\hat{L}_{kt} L_{kt}$) is

$$\hat{L}_{kt} L_{kt} = \sum_k \left[\frac{\lambda_{kjt|k} \hat{w}_{jt}^\theta (1 - \widehat{\tau_{kjt}})^\theta}{\sum_j \lambda_{kjt|k} \hat{w}_{jt}^\theta (1 - \widehat{\tau_{kjt}})^\theta} \hat{R}_{kt} R_{kt} \right] \quad (\text{D.8})$$

since time-varying residence amenities (B_{it}) have canceled out from the numerators and denominators from these fractions. Substituting (D.7) and (D.8) into (D.6) yields the combined land and commuter market clearing condition,

$$\begin{aligned} \hat{Q}_{kt} Q_{kt} = & (1 - \alpha) \sum_j \left[\frac{\lambda_{kjt|k} \hat{w}_{jt}^\theta (1 - \widehat{\tau_{kjt}})^\theta}{\sum_j \lambda_{kjt|k} \hat{w}_{jt}^\theta (1 - \widehat{\tau_{kjt}})^\theta} \hat{w}_{jt} w_{jt} \right] \hat{R}_{kt} R_{kt} \\ & + \frac{\beta^H}{\beta^L} \hat{w}_{kt} w_{kt} \sum_k \left[\frac{\lambda_{kjt|k} \hat{w}_{jt}^\theta (1 - \widehat{\tau_{kjt}})^\theta}{\sum_j \lambda_{kjt|k} \hat{w}_{jt}^\theta (1 - \widehat{\tau_{kjt}})^\theta} \hat{R}_{kt} R_{kt} \right]. \end{aligned} \quad (\text{D.9})$$

Suppose we observe the initial equilibrium values of the commuting probabilities conditional on residence, workplace employment, resident employment, wages, and average per capita income by residence. Suppose also that we observe relative changes in residents and land values between years t_0 and t . Given these observed variables and known values for changes in the wage tax rate, (D.9) provides a system of N equations that determines unique values for the N unknown relative

changes in wages in each location (and, hence, residents, employment, and bilateral commuting flows).

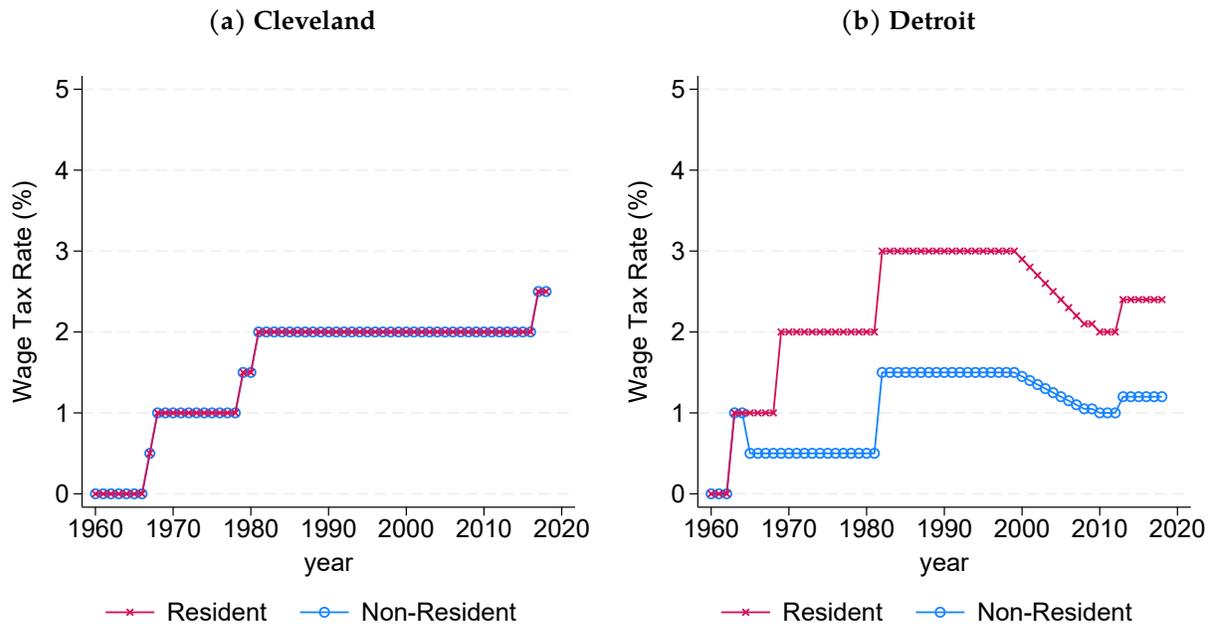
We hypothesize a value of θ and use (D.9) to obtain the predicted share of workers in the Philadelphia CZ that live and work in the suburbs in a given year t . As stated above, our only data requirement for that given year is that we also observe residents and land values in that year. We repeat this for a range of hypothesized values of θ . We then calibrate θ by minimizing the squared deviations between the observed and predicted share of workers in the Philadelphia CZ that live and work in the suburbs (which we only observe starting 1960).

E Wage tax variation in other cities

Several cities adopted a wage tax after Philadelphia. The first major city to impose a wage tax outside of Philadelphia was Toledo, OH in 1946. St. Louis and Kansas City, MO enacted wage taxes in 1948, where the resident and nonresident rates have remained at 1% ever since. New York City had a wage tax between 1966 and 1999, when the nonresident rate was abolished. Cleveland passed its wage tax ordinance in 1966, and began taxing workers the following year. Detroit did so in 1962.

In this analysis, we consider the wage tax variation in Cleveland and Detroit. A straightforward implementation of our spatial RD design is not possible in Toledo, due to its relatively small size, nor in Cincinnati or Columbus, due to their multiple enclaves and highly complex jurisdiction borders. Data availability on wage tax rates through time is also limited. Appendix Figure E.29 displays the variation in wage tax rates for Cleveland and Detroit.

Figure E.29: Wage tax variation for Cleveland and Detroit



Note: The figure displays the wage tax variation for Cleveland and Detroit. Source for Cleveland: City of Cleveland Central Collection Agency, <http://ccatax.ci.cleveland.oh.us/?p=effective>. Source for Detroit: City of Detroit Code of Ordinances, https://library.municode.com/mi/detroit/codes/code_of_ordinances/389797?nodeId=PTIVDEC0_CH44TA_ARTIICIINTA.

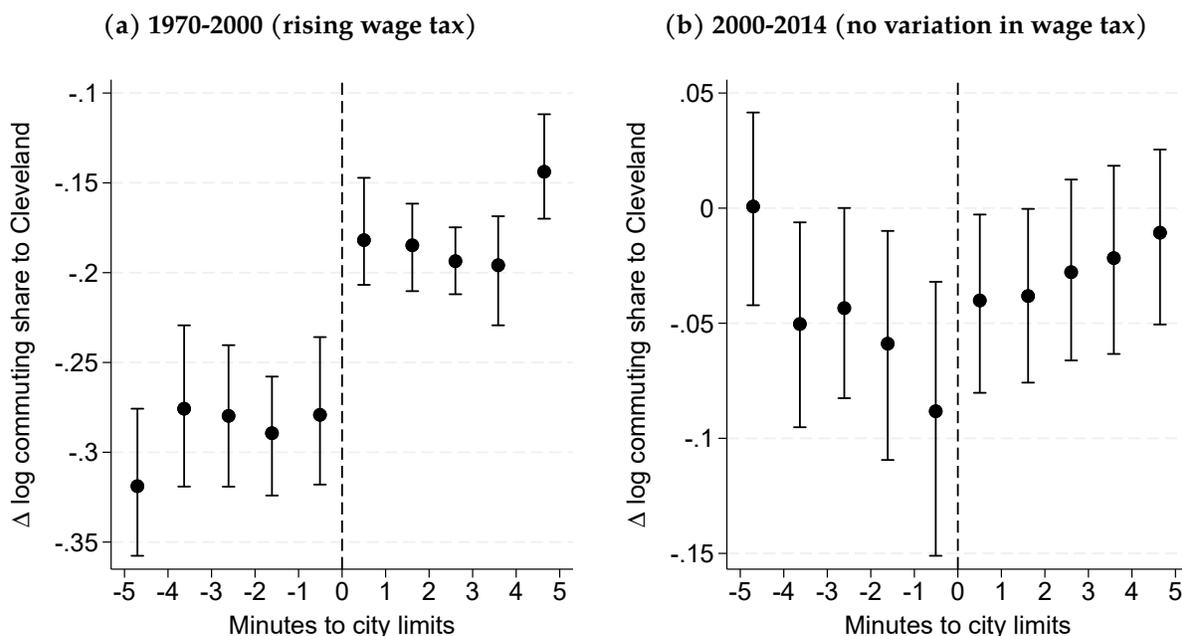
E.1 Cleveland

Cleveland's wage tax increased steadily between its introduction in 1967 and 1981, and remained at 2% until 2016. Cleveland's resident and nonresident wage tax rates have always been equivalent. We analyze the 1 percentage point increase in the nonresident wage tax rate between 1970 and 2000. We also analyze the period between 2000 and 2014 to test whether there is any discontinuity

in periods in which the wage tax didn't vary.

Appendix Figure E.30 shows our results. In Panel (a), we detect a positive discontinuity for the log change in commuting shares to Cleveland between 1970 and 2000. This is consistent with our sign prediction in Equation 10. Using our fuzzy RD design, the implied commuting elasticity here is 9.3 (SE = 2.99), which is similar in magnitude to the estimates we obtain for Philadelphia in Table 2. In Panel (b), we estimate a small, non-statistically significant discontinuity for the log change in commuting shares to Cleveland between 2000 and 2014. The null result for this time period is also consistent with our framework.

Figure E.30: RD plots for change in log commuting share to Cleveland



Note: The figure displays RD plots of changes in log commuting share to Cleveland for census tracts within a five-minute drive of the city limits. City tracts correspond to positive values along the x-axis. Panel (a) shows the outcome between 1970 and 2000. Panel (b) does the same for 2000-2014. Standard errors are clustered at the grid cell level.

E.2 Detroit

Detroit implemented its city wage tax in 1962. The municipalities surrounding Detroit have never had wage taxes, but Detroit's two enclave cities, Highland Park and Hamtramck, enacted wage tax ordinances in 1965. In this context, we consider three types of wage tax variation for Detroit.

The first case is the most complex since it involves the resident and nonresident wage tax rates moving in opposite directions. Detroit began taxing residents and nonresidents at a rate of 1% in 1962. In 1968, this rate *halved* to 0.5% for nonresidents, but *doubled* to 2% for residents. This variation in opposite directions effectively meant that, in relative terms, the city appeared to subsidize suburban commuting into the city in 1968. In other words, city workers had a higher incentive to relocate close to the city limits in 1970.

The second case is a more straightforward one, similar to our main analysis between 1960 and 2000 for Philadelphia. By the year 2000, the wage tax rates had stabilized in Detroit. The resident and nonresident wage tax rates had been at about 3 and 1.5% for nearly two decades. In 2000, the resident rates were 2% in Highland Park and 1% in Hamtramck. Relative to 1960, then, there would have been a penalty to commute into the city from the suburbs, but not from either of these two enclaves.

The third case is one of falling wage tax rates in Detroit. Between 2000 and 2014, the nonresident wage tax rate fell modestly from 1.45% to 1.2%. We have no clear data on the historical wage tax rates for the enclaves of Highland Park and Hamtramck, but available information appears to suggest the tax rates did not change during this period.⁴¹

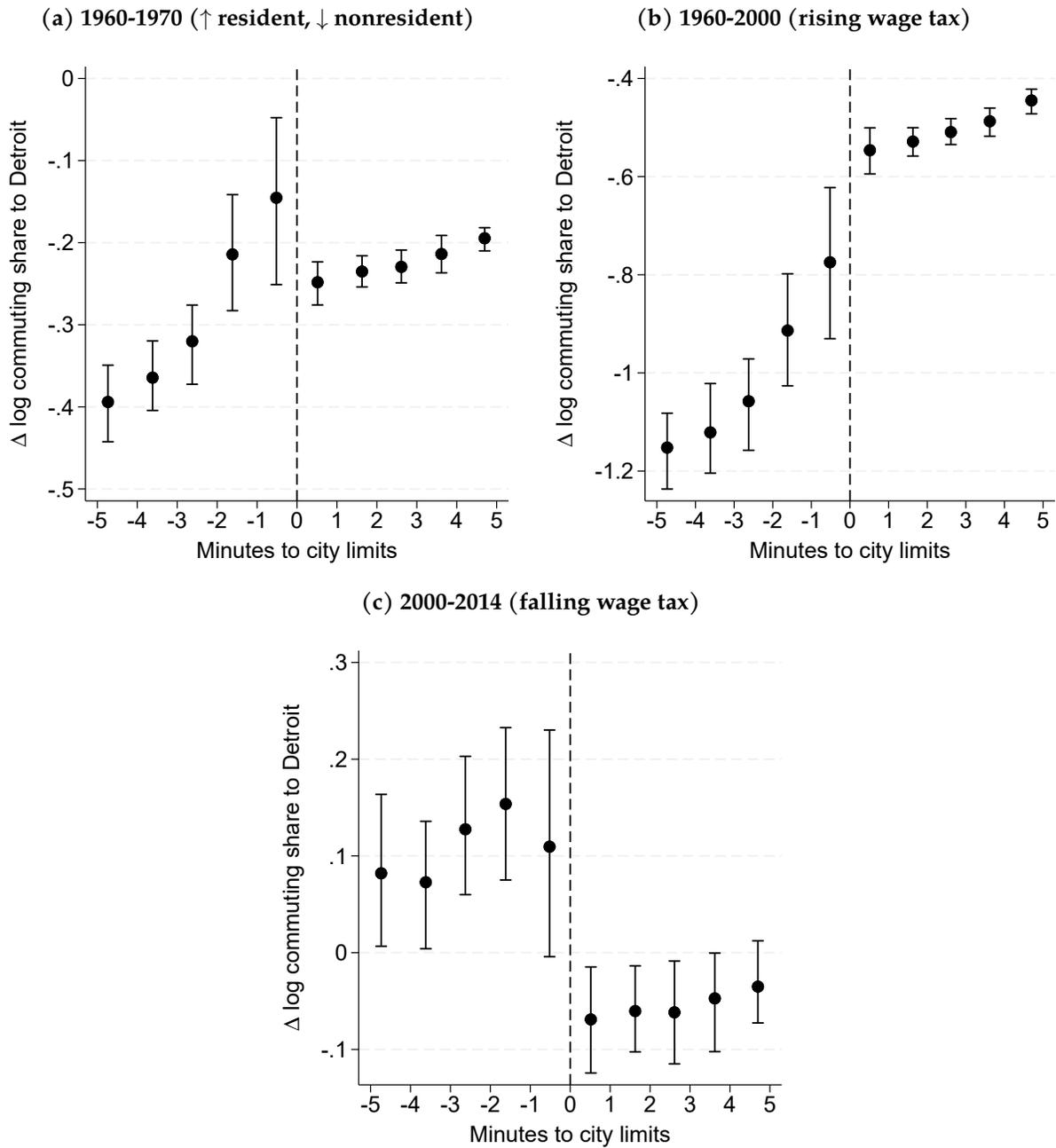
Appendix Figure E.31 applies our spatial RD design to the log change in commuting shares to Detroit for each of these three episodes. Consistent with our intuition for the first episode (1960-1970), we obtain a *negative* discontinuity for the change in commuting shares to Detroit in Panel (a), where the variation of the wage tax in opposite directions increased the incentives to commute into the city from the suburbs. Indeed, the slope to the left of the threshold reflects this unusual variation. The implied commuting elasticity is 10.1 (SE = 4.17), which is noisier but still consistent with the magnitudes we obtain in Table 2.

In the second episode in Panel (b), where suburban workers faced an increasing penalty to commute into Detroit between 1960 and 2000, we now obtain a positive discontinuity. This is consistent with our predictions from Equation 10. The implied commuting elasticity here is 10.6 (SE = 6.4), which is imprecisely estimated but still consistent with our main estimates.

In the modestly falling wage tax rate episode between 2000 and 2014 in Panel (c), we obtain a negative discontinuity. Because the change in Detroit's nonresident wage tax rate is so small (0.25 *p.p.*), we estimate a large commuting elasticity of 56.9 (SE = 25.022). Given the magnitude of the standard error, the 95% CI for this estimate nevertheless includes the elasticities we document above.

⁴¹See <https://www.michigan.gov/taxes/questions/iit/accordion/general/what-cities-impose-an-income-tax>.

Figure E.31: RD plots for change in log commuting share to Detroit



Note: The figure displays RD plots of changes in log commuting share to Detroit for census tracts within a five-minute drive of the city limits. City tracts correspond to positive values along the x-axis. Panel (a) shows the outcome between 1960 and 1970. Panels (b) and (c) do the same for the 1960-2000 and 2000-2014 periods. Standard errors are clustered at the grid cell level.