Monopsony in Space:
Commuting & Labor Market Power *

Abdelrahman Amer

November 14, 2025

[PLEASE CLICK HERE FOR LATEST VERSION]

Abstract

This paper studies the role of commuting costs in shaping labor market power and the allocation of workers to firms. I build, identify, and estimate a two-sided labor market matching model that features strategic interactions in wage setting, commuting costs, and residential choice. I use the model to study the direct and distributional consequences of a commuting cost shock due to a subway expansion in Vancouver. Empirically, workers who gained improved access to the subway network experienced an increase in earnings by 1.5-2% relative to workers with no change in access. Using the estimated model, I show that the expansion improved access to more productive firms for workers in affected areas, but increased competition for high-productivity jobs for workers elsewhere. Neighborhoods with improved access experienced an 8% drop in labor market concentration. These results show that improvements in access to firms for some neighborhoods can generate adverse spillover effects for others due to higher competition amongst workers. To underscore the role of differential job access in shaping labor market power, I show that 10-15% of the spatial variation in wage markdowns can be explained by the non-uniform access to firms within a commuting zone.

*I am extremely indebted to Kory Kroft, Ismael Mourifié, Nate Baum-Snow, and Clémentine Van Effenterre for unwavering guidance and support. I also thank Samy Amer, Lady Bolongaita, Michael Boutros, Grant Benjamin, Yanyou Chen, Victor Couture, Stephan Heblich, Mahmood Haddara, Gueorgui Kambourov, Marlène Koffi, Alexandre Lehoux, Peter Morrow, Sobia Jafry, Stephen Tino, Aloysius Siow, and Román Zárate, for helpful discussions; and seminar participants at University of Toronto SWEAT seminar series, University of Toronto Empirical Microeconomics seminar series, and University of Toronto Urban seminar series. In loving memory, of Jojo Zhang. I acknowledge the use of the Toronto Research Data Center for the completion of this research. All errors are my own.
Department of Economics, University of Toronto; e-mail: abdelrahman.amer@mail.utoronto.ca

1 Introduction

Attractive residential locations need not coincide with desirable employment opportunities. To find good jobs, workers need to commute. Commuting, however, can be burdensome, and workers are willing to sacrifice higher pay for a shorter trip to work (Le Barbanchon et al., 2021).[1] In addition, if a worker's next best employment option requires a longer commute, they may choose to stay at their current employer even when facing a marginal drop in their wage (Caldwell and Danieli, 2024). Such imperfect substitutability across differentially located jobs gives firms market power to set wages below their competitive level and affects the allocation of workers to firms (Hotelling, 1929; Salop, 1979).

Understanding the role of commuting costs as a source of labor market power can offer new insights on the welfare and distributional consequences of policies lowering workers' travel burden. However, studying the effects of commuting shocks on equilibrium labor market outcomes is challenging. Quantitative urban models, although are well-suited to study effects of commuting shocks across space, assume perfectly competitive labor markets, and therefore the number of geographically accessible jobs plays no role in shaping workers' trade-offs (Ahlfeldt et al., 2015; Heblich et al., 2020; Tsivanidis, 2022).[2] Models of job differentiation, although allow firms to set wages, cannot isolate shocks to commuting costs from other idiosyncratic preferences (Card et al., 2018; Azar et al., 2022; Lamadon et al., 2022). Both wage-setting power and independent commuting preferences are necessary to isolate the effect of commuting shocks on equilibrium labor market outcomes.

This paper examines the role of commuting costs in shaping labor market power and the allocation of workers to firms by incorporating firms' wage-setting behavior into a quantitative spatial model with heterogeneous firms. In equilibrium, a lower commuting cost increases the substitutability across differentially located jobs, incentivizes workers to match with more distant employers, and reduces firms' labor market power. I study the direct and distributional consequences of a shock to commuting costs through the lens of a subway expansion in Vancouver, and use the model to highlight the spillover effects of increased competition for high-paying jobs across space.

In preparation for the 2010 Winter Olympics, the Canada Line was added to the Vancouver rapid transit network (SkyTrain) to connect the airport with the downtown area. This allowed new neighborhoods to access the SkyTrain network and had a significant impact on travel times.[3]

[1] Le Barbanchon et al. (2021) use survey data to show that unemployed workers are less likely to accept employment offers at more distant jobs. Using layoff shocks Duan et al. (2022) estimate a job ladder model with commuting costs, and show that due to the trade-off between wages and commuting, wage losses alone fail to account for the total cost of job loss. This highlights the trade-off between pay and commuting that workers face.

[2] The role of the number of jobs in shaping agglomeration forces was first studied by (Helsley and Strange, 1990). Moretti and Yi (2024) show that laid-off workers fare better in labor markets with a greater number of jobs, conditional on labor market tightness.

[3] Herzog (2025) shows the Canada Line had a causal effect on transit ridership due to improved travel times around

Leveraging Canadian Employer-Employee data, with granular information on workers' residential locations, and station-level coordinates, I use an event study design to show that earnings for workers who gained access to the subway network increased by 1.5% three years after the expansion. Worker-level panel data allows me to follow workers over time, thereby alleviating compositional change issues inherent in neighborhood-level analyses.

I find a higher wage effect for job switchers compared to stayers. This offers suggestive evidence that under lower commuting costs, workers reallocate to higher-paying jobs. The partial equilibrium analysis nets out general equilibrium effects and is therefore silent on the spillover effects of the expansion across space. General equilibrium effects are important when changes to commuting costs for a set of neighborhoods can influence labor market outcomes by increasing competition for jobs across space.

To assess the distributional effects of the subway expansion, I incorporate commuting costs into a general equilibrium model of the labor market, featuring strategic interactions in wage setting, residential choice, and rich firm heterogeneity. In equilibrium, workers of different observable types choose a neighborhood-employer pair to balance housing costs and residential amenities, with compensation packages, net of commuting costs. Compensation packages include wages, deterministic non-wage amenities, and stochastic idiosyncratic preferences. Neighborhoods are characterized by an inelastic housing supply and exogenous residential amenities.

Firms face an upward-sloping labor supply curve due to horizontal differentiation by commuting costs and other stochastic preferences. By allowing workers' residential locations to affect workers' labor supply behavior through heterogeneous commuting costs, the model implies a concentration index that varies across neighborhoods. This allows me to assess the distributional effect of shocks to commuting costs on labor market concentration within and across cities.

Turning to identification, there are two main blocks of parameters to identify. First, labor supply parameters, are summarized by workers' stochastic preferences over unobservable non-wage amenities, and their disutility from commuting. Second, the set of parameters representing firms' heterogeneous production technologies. To identify workers' stochastic preferences over non-wage amenities, separate from commuting, I use the quasi-labor supply function from Berry (1994) at the neighborhood-firm level. The identification challenge stems from the potential correlation between systematic preferences for unobserved amenities, other than commuting, and wages. To overcome this challenge, I follow Lamadon et al. (2022) and Chan et al. (2024) by using an internal instrumental variable approach, assuming that innovations to firm productivity are persistent, while deterministic non-wage amenities are transitory.

To identify workers' commuting disutility, I show that the model implies a gravity equation

previously congested areas. In addition, Chernoff and Craig (2022) show that the introduction of the new line benefited lower-income workers in newly connected neighborhoods.

between workers' place of residence and firms. Although using gravity to identify commuting preferences is standard, unlike other spatial models, traveling to a higher productivity location does not fully determine workers' deterministic payoff, as longer commutes can be compensated by working for higher paying firms within a given destination. Neighborhood-level data is therefore insufficient to identify the semi-elasticity of commuting flows with respect to commuting distance. To address this identification challenge, I control for differences in firms' deterministic utility and leverage within firm variation in labor supply with respect to distance to identify the commute flows elasticity. Combining the commute flows elasticity with the given stochastic preference distribution parameters allows me to back out workers' commuting disutility.

On the demand side, to identify the elasticity of substitution across workers with different observable types, I use the elasticities implied by the labor supply parameters and invert firms' first-order condition for firm profit maximization. Given the elasticity of substitution and firm revenue from the balance sheet data, I identify other parameters associated with firms' production technology by inverting the production function under a normalization assumption on within-firm productivity differences.

I estimate the model using the Canadian Employer-Employee Dynamics Database (CEEDD). The CEEDD covers the entire population of individuals and businesses with taxable income from 2001 to 2019. I leverage the longitudinal nature of the data to estimate workers' and firms' primitives. I also combine the data with a rental price index at the neighborhood level using the 2016 population Census for Canada, to recover the model-implied housing supply. The worker-firm matched data contains balance sheet data, and coordinates on place of residence, and workplace.[4] This allows me to separately identify stochastic preferences from commuting costs. Finally, the available firm balance sheet data allow for the estimation of heterogeneous firm production functions.

Estimated parameters reveal an average firm-specific labor supply elasticity of 7.5, with an average wage markdown of 84% below the marginal product of labor. There is substantial heterogeneity across age groups. Workers are willing to pay between 1 to 2.7% of their annual earnings to reduce their commuting distance by 1 km. Older female workers have a 20% higher willingness to pay for a shorter commute, relative to same same-aged men. On the firm side, labor demand estimation reveals significant heterogeneity in firms' total factor productivity, and a sharp decrease in returns to firm size, with an average returns to scale of 0.29.

Using the estimated firm-level productivity parameters, I show that workers with longer commutes sort to higher productivity firms. An increase in commuting distance by 1 km is associated

[4] Data includes worker and firm (enterprise) location coordinates. For multi-establishment firms, only the location of their corresponding enterprise headquarters is reported. To overcome this issue, I impute the location using workers' residential choice. More detail on this in Section 2

with a 1% increase in firm total factor productivity (TFP). The relationship between firm TFP and commuting is stronger for female workers, who have a higher elasticity by 0.47 percentage points. Showing that female workers require matching higher productivity firms when commuting longer distances validates the willingness-to-pay estimates.

I simulate the estimated model in the absence of the Canada Line using 2016 data for the entire economy. Simulation results show that workers in neighborhoods near the Canada Line experienced a 1% increase in earnings due to the presence of the new subway line, while workers in neighborhoods located on the preexisting network and non-connected areas faced a drop in earnings due to the expansion. There was little change in average neighborhood markdowns due to workers' resorting to more productive firms, which offer higher wages, but lower markdowns on average. Workers in neighborhoods with access to the Canada Line resorted to more productive firms, as the average TFP of a given worker's employer in this area increased by by 5%. Simultaneously, workers in other areas resorted to less productive firms. This provides evidence that workers on the Canada Line resorted to higher productivity firms, while crowding out workers from other neighborhoods, underscoring the role of spatial competition for high-paying jobs in shaping the distributional consequences of a transit-induced commuting cost shock.

Investigating the effects of the shock on labor market concentration, I find that the average generalized concentration index (GCI) dropped by 8.6% in neighborhoods on the Canada Line, with modest increases in other areas. The effects on concentration are driven by changes in shares on the extensive margin, as well as changes in firm shares within a given neighborhood. Lower commuting costs improves an area's access to firms, while simultaneously reducing the value of the outside option, non-employment. As a result, employment in neighborhoods on the Canada Line increased by 8.4%, and decreased by 0.65% and 0.4% in areas with no subway, and pre-connected areas, respectively. This provides evidence that the improved employment accessibility on the Canada Line induced more competition for workers in other areas on the extensive margin too. Note that unaffected neighborhoods are larger and have bigger populations, explaining the differences in the magnitude of the estimated effects. In addition, due to the improved accessibility, the population on Canada Line neighborhoods increased by 3.75%, resulting in an increase in average rents by 8.5%. The increase in population is driven by changes in accessibility to firms in treated areas.

Overall, simulation results highlight that increased competition for high-productivity jobs benefited areas on the Canada Line, but crowded out workers in other areas from high-paying jobs. This is also true on the extensive margin, as employment increased in neighborhoods with access to the Canada Line. These results highlight the distributional consequences of asymmetric commuting cost shocks, and extend the findings in Manning and Petrongolo (2017) and Chapelle and Ubeda (2025) regarding increased extensive margin competition between workers across space to the intensive margin of firm productivity. [5] The effects on markdowns are less clear, as workers in affected areas resorted to higher productivity firms, who also have lower markdowns on average. Note that markdowns are needed on average for differences in firm productivity to be translated into differences in worker earnings across neighborhoods.

To further show the effects of asymmetric changes in commuting costs on competition amongst workers for high-paying jobs, I simulate the model under the counterfactual scenario of uniform distance between workers and firms. Simulation results show that differential firm access explains 10-15% of the variation in markdowns across space, and 40% of the spatial variation in the marginal revenue product of labor. By construction, the variation in both equilibrium objects across space is due to spatial variation in worker-to-firm sorting. These results highlight the role of heterogeneous commuting costs in shaping spatial variation in market power.

Related Literature My paper contributes to three literatures. First, it relates to the fast-growing literature on local monopsony power (Manning and Petrongolo, 2017; Heise and Porzio, 2022; Datta, 2023; Pérez et al., 2022).[6] I contribute to this literature by analyzing how asymmetric changes in commuting costs across space affect the spatial distribution of labor market concentration and the sorting of workers to firms. Manning and Petrongolo (2017) show that transit shocks increase spatial competition for jobs amongst the unemployed. My results show that such a shock increases competition for jobs at higher-paying firms, not only the extensive margin. Conceptually, I extend the role of commuting frictions in Heise and Porzio (2022) to models of job differentiation with rich firm heterogeneity, and granular geographies to capture *local* labor market power via a neighborhood-specific concentration index. This is motivated by the empirical evidence on the highly local nature of workers' labor markets (Manning and Petrongolo, 2017; Monte et al., 2018; Le Barbanchon et al., 2021; Caldwell and Danieli, 2024).[7]

Second, this paper adds to the growing literature on imperfect competition (Manning, 2003; Staiger et al., 2010; Card et al., 2018; Lamadon et al., 2022; Kroft et al., 2025). By nesting job differentiation into a quantitative spatial model, this paper investigates how an *observable* horizontal amenity, commuting costs, can be a source of labor market power; separately from other *unobservable* preferences. Understanding what shapes workers' idiosyncratic preferences is key to studying the sources of labor market power. Recent evidence by Humlum et al. (2025) show that

[5] Amior and Manning (2018) show that the unemployment rate is a sufficient statistic for a neighborhood's local welfare. This is true when firms are homogeneous and markdowns are constant with respect to market shares. With heterogeneous firms, or variable markdowns, the GCI is needed to capture a neighborhood's access to employers (Chan et al., 2024).

[6] Datta (2023) investigates how changes to commuting costs shape labor supply elasticities for a given firm across its establishments in London, but abstracts from workers' location choice, and excludes vertical non-pay amenities.

[7] In Canada, 40% of workers are employed within 5 km of their place of residence (Statistics Canada, 2006). In the US, 70% of US workers are employed within their county of residence (Monte et al., 2018). In Germany, 73% of workers work within 5 miles of their residential location (Caldwell and Danieli, 2024).

systematic firm-wide amenities explain only 14% of the amenity penalty associated with moving to a higher paying job, the rest is attributable to match effects in pay and preferences. Although this evidence stresses the importance of horizontal preferences in shaping labor market power, there remains limited evidence on what characteristics differentiate employers, while simultaneously allowing them to set wages (Mas, 2025; Kline, 2025). This paper underscores the role of commuting costs in shaping these heterogeneous preferences.[8] Finally, the model-implied concentration index at a granular location level is closely tied to the outside option index constructed by Caldwell and Danieli (2024), who also allow for commuting distance to shape workers' employment options. However, they do not allow for firm wage-setting power, a feature of the labor market that distorts the allocation of workers to firms.

Third, my paper contributes to the literature on quantitative urban models (Ahlfeldt et al., 2015; Monte et al., 2018; Tsivanidis, 2022; Heblich et al., 2020; Zárate, 2022; Baum-Snow and Han, 2024). Conceptually, by relaxing the assumption of perfect competition, and allowing for rich heterogeneity across firms, my framework can be used to study the effect of spatially heterogeneous shocks (e.g. labor demand) on worker to firm sorting across space. This generalizes the framework in Monte et al. (2018), who study how spatial shocks affect local employment elasticities, but do not consider the role of labor market power. With regard to identification, leveraging matched worker-firm data with information on residential and workplace locations, this paper offers a new method for identifying commuting disutility under labor market power and firm heterogeneity. In particular, the use of within-firm variation in labor supply with respect to commuting distance, combined with lagged firm productivity shocks, allows me to separately identify the distribution of stochastic preferences from commuting disutility. [9]

The remainder of the paper is structured as follows. Section 2 describes the Canadian Employer-Employee Dynamic Database. Section 3 provides evidence on the effects of the Vancouver Subway expansion. The model is presented in Section 4. Section 5 discusses identification and estimation. Section 6 presents model counterfactual simulation results, and Section 7 concludes.

2 Data

I use the Canadian Employer-Employee Dynamics Database (CEEDD). The data serves two purposes: First, it allows leveraging the Canada Line expansion in Vancouver to provide reduced-form evidence. Second, to estimate the model outlined in Section 4 on the effects of changes in commuting costs, due to subway access, on labor market outcomes. The CEEDD covers the universe of

[8] Recent work by Bils et al. (2025) argues that workers' unobservable comparative advantage across firms shapes firms' upward sloping labor supply. In my framework, observable commuting costs can be interpreted as an observable characteristic that shapes such a comparative advantage.

[9] Le Barbanchon et al. (2021) use workers' stated preferences for commuting and job choice to estimate workers' willingness to pay for a shorter commute. I implement a revealed preference approach using gravity equations.

individuals and firms in Canada from 2001 to 2019. The database integrates multiple data sources: One, the T4-slips linked with the Record of Employment (T4ROE) includes data on *annual* earnings from the T4 employment slips (similar to W-2 forms in the US), consisting of any form of employment pay, such as wages, commissions, or bonuses. Employment slips are linked with the Record of Employment to form the T4ROE dataset. In a given year, if a worker is linked to multiple employers, due to holding multiple jobs or job switches, I select the job with the highest earnings. Second, the T1 Personal Master File (T1PMF) provides demographic information such as age, gender, and most importantly for this paper, the exact geo-coordinates of workers' residential location (Longitude & Latitude). This information will be used to calculate workers' distance to the nearest subway station and will also be used, as described below, to impute *establishment* locations, allowing the estimation of the commute flows elasticity with respect to distance to the workplace. Finally, the database includes the National Accounts Longitudinal Microdata File (NALMF), which contains firms' balance sheet data and NAICS industry codes. Importantly for this paper, it includes information on firms' revenue, which is crucial for identifying firms' heterogeneous production function parameters. Separate from the CEEDD, I use the 25% sample of the 2016 Canadian Population Census at the Census Tract (CT) level to obtain housing costs, which are combined with residential characteristics in a hedonic regression to construct a rental price index for 6300 CTs, in 50 Census Metropolitan Areas (CMAs) & Census Agglomerations (CAs). The third dataset comprises the coordinates of all subway stations that form the Vancouver SkyTrain network, along with their corresponding opening dates. Combining this with workers' residential coordinates allows me to compute the distance to the nearest subway station for all workers in Vancouver in any given year.

I restrict the model estimation and analysis to individuals residing in one of the 50 Census Metropolitan Areas (CMAs) for which housing price indices are available from the Census. This mainly consists of dropping individuals residing in the territories and other smaller Census Agglomerations (CAs). Moreover, those outside the age range of 25-60 are also excluded. I also drop individuals with missing gender, age, or residential location geo-coordinates. The CEEDD is derived from tax records and therefore lacks information on hours worked, hourly wage, or full-time indicators. To address this issue, I follow Li et al. (2023) and further narrow the sample down to full-time equivalent (FTE) workers, defined as those earning at least $18,000 in 2016 dollars.

On the firm side, firms in the public sector (NAICS 91), education (NAICS 61), or healthcare sectors (NAICS 62) are excluded from the analysis. I also only keep firms with a total revenue of at least $50,000 and those with a wage bill lower than 80% of their total revenue. Moreover, I only keep firms that have at least 5 full-time workers, as defined above. Regarding firm location, although I observe the exact geo-coordinates of workers' residential location, firms are defined by their Enterprise ID from the Business Registry, which means only the headquarters' location information is available. This is an issue for firms with multiple establishments. Although they represent around 10

To check the validity of this imputation, I compute the deviation (Euclidean distance) between the imputed and actual location for single establishment firms for which their precise location is reported in NALMF. The mean deviation is approximately 23 km; however, this distribution is highly skewed, with a median of only 3 km. One reason for this issue is that some enterprises, including single establishments, often report their auditor's city even when all of their workers are in another location. To further validate the establishment location imputation, I compare the derived commuting to workplace distribution with that of the 2006 Canadian Census. The bottom row of Table A-1 presents the commuting distance to the workplace in Canada for the 2006 population census. Comparing these figures with the commuting distribution in Table 1 shows that the derived commuting distances in my data replicate that of the census quite closely. For example, Table 1 shows that 21% of workers commute between 5 and 10 km, whereas this value is 23% in A-1. Panel B of Table 1 presents summary statistics at the *enterprise* level. It shows that the average number of workers within a given enterprise is around 60, and 10% of enterprises are multi-establishment units. Given worker and firm locations, I define distance traveled to work as the Euclidean distance between workers' residential location and the establishment's imputed location. I keep worker-establishment pairs with at most 75 km, which consists of more than 95% of matches in any given year.

Table 1 presents summary statistics at the worker and enterprise level for the constructed sample for the period of 2001-2019. All dollar values are in 2016 dollars. Workers earn on average $ 67,000. There is a gender gap in earnings of approximately 32%. Note that I drop workers in the public and health sectors, where women are overrepresented. Around 40% of workers reside in Ontario, the largest Canadian province, 25% in Quebec, and 10% in British Columbia. Focusing on commuting patterns, approximately 38% of workers have a workplace that is within 5 km of their place of residence. The high share of workers who work close to their place of residence is consistent with the findings of Caldwell and Danieli (2024) who show that 73% of German workers had a workplace within 5 miles of their place of residence, pointing towards the localized nature of the labor market.

Finally, I define a *neighborhood* to be a set of clustered Census Tracts (CTs) within a CMA. I group 6200 CTs into 1034 neighborhoods across Canada. Clustering happens at the CMA level, such that each neighborhood is nested within a CMA. I drop clusters where the minimum distance between the neighborhood centroid and any CT in the cluster is greater than 20 km. This ensures the removal of remote areas with a sparse set of neighborhoods. This aggregation allows me to estimate and compute the model on a granular number of neighborhoods with sufficient variation in distance between origin-destination pairs. Descriptives at the neighborhood level are reported in Panel C of Table 1.

I combine the CEEDD with station-level data in Vancouver to study the effects of a subway expansion on labor market outcomes. Station-level data is sourced from citylines API. I retrieve the stations' longitude and latitude, as well as their opening year. The station-level data is then linked to its corresponding CT in the CEED. I discuss sample and data construction for this empirical exercise in Section 3. Importantly, the CEEDD contains granular residential location data. Combining this with station data allows me to construct annual changes in distance to the nearest station for all workers.

3 Effect of Subway Access on Labor Market Outcomes

3.1 Vancouver SkyTrain Expansion

I first present reduced-form evidence on the effect of reduced commuting costs, as a result of gaining access to a subway station, on workers' earnings and travel distance to work. The empirical analysis will focus on the Metro Vancouver area, leveraging the expansion of the SkyTrain transit network in 2009 into new neighborhoods. The Vancouver SkyTrain is the longest automated rapid transit system in the world. It includes three lines connecting downtown Vancouver to suburbs such as Surrey, Burnaby, Richmond, and Coquitlam. Prior to 2009 the SkyTrain had only two lines; The Expo Line, built in 1986, and later extended in 1990 and 1994, and runs approximately 22 km through Vancouver, Burnaby and New Westminster; And the Millenium Line, for which operations began in 2002, and spans 25 km from west Vancouver to the skirts of Coquitlam in the east, before looping back again to the downtown area. As illustrated by Figure 1, prior to the Canada Line construction, there was no rapid rail connecting Richmond in the southwest with downtown Vancouver in the north by crossing the Fraser River. Further, prior to 2016, there was no rapid rail connecting Coquitlam in the northeast with the downtown area.

In preparation for the 2010 Winter Olympics, the Canada Line became a priority to connect the airport with the downtown area. As illustrated in Figure 1, the expansion entailed the construction of a new line connecting downtown Vancouver in the North with Richmond in the south across the Fraser River. Plans for the new line were finalized in 2003, construction started in 2005, and the line was operational ahead of schedule in 2009. This new line runs approximately 20 km with 16 stations from Vancouver International Airport to Waterfront Station. As of 2010, the line served approximately 100,000 riders on a daily basis (Sroka, 2021).[10] Although both Coquitlam and Richmond were in need of a subway connection with the downtown area, funds and efforts were directed towards the Canada line to host the Winter Olympics. This resulted in plans to extend the

[10]More details on the SkyTrain network and ridership available here

Millennium line eastward into Coquitlam being postponed, and did not start until 2016.[11]

The Canada Line expansion serves as an ideal quasi-exogenous shock to workers' travel costs for several reasons. One, the goal of the expansion was to better prepare Vancouver to host the 2010 Winter Olympics. Therefore, one of the primary goals of the line was to connect the downtown area with the airport, rather than targeting a specific neighborhood. I leverage this institutional feature, along with panel data on residents of the Metro Vancouver area, to identify the effect of gaining access to the subway network on earnings and commuting distance. Treated workers will be those in the Richmond area, and control units will span the eastern parts of the Metro area, including residents of Coquitlam who were waiting for the Millennium Line extensions, as well as sections of Surrey across the Fraser River. The individual panel data allow me to rule out residential sorting and other compositional changes that could potentially lead to differences in earnings and work behavior for treatment and control neighborhoods, as is usually a threat when assessing the effects of subway expansion using spatial panel data only. Individual-level panel data also allows me to follow individuals over time independently of changes in place of residence. Two, the expansion had a considerable impact on travel time. Prior to the Canada Line, public transit users traveling between Richmond and downtown Vancouver had to rely on the *98 B-Line bus*. During rush hour, the trip can take between 50-60 minutes (Sroka, 2021; Herzog, 2025). The improvement in public transit travel time is reflected in mode choice for workers residing in Richmond, the Census Subdivision most affected by the new line. According to Statistics Canada Population Census, in 2006, only 11% of workers in Richmond used public transit as their primary mode of travel to work. In 2011, after the introduction of the Canada Line, approximately 20% did so. As illustrated in Figure A-2, this represented a steeper increase in public transit use compared to the rest of Metro Vancouver. Although not causal, this provides suggestive evidence that the new line reduced travel time for public transit users to the extent of changing the travel mode distribution. This is indeed supported by the causal evidence found by Herzog (2025), who shows that the Canada Line *causally* increased transit use by 4.6 percentage points, and the increase is due to reduced driving to work. Using the 2016 General Transit Feed Specification (GTFS) file Herzog (2025) simulates the time savings resulting from the new Canada Line by replacing the new rail line with an express bus to compute historical travel times. Simulations show that the Canada line reduced travel time by almost 7-10 minutes as a lower bound for any route that passes along the Canada Line corridor, representing a decrease in travel time of approximately 15-25% across different times of the day. These reductions in travel time will be used to simulate the counterfactual model in the absence of the Canada line to quantify and zoom in on the role of how gaining better access to firms mediates the observed treatment effects.

[11] This article discusses the political and municipal landscape prior to this extension, and importantly shows that the Millennium line extension was delayed as a result of the Canada line in preparation for the Winter Olympics.

3.2 Empirical Strategy

I leverage SkyTrain station-level data, which includes station geo-coordinates and opening year, to combine with workers' residential locations in the CEEDD and calculate individual-year-level distance to the nearest subway station. The exact residential and station coordinates allow precise measurement of changes in distance to the subway station at a granular geographic level. I use these coordinates to calculate the percentage change in distance to the nearest subway station during the 2009 expansion. As illustrated by Figure 2, residents of the Richmond area in the south experienced the largest drop in distance to nearest subway station - *treatment exposure*- after the construction of the Canada line. On the other hand, residents of eastern Surrey and Coquitlam did not gain access to a subway station until the end of 2016 when the Millennium Line was extended into Coquitlam. Table 2 shows that the average change in distance to nearest subway station was around 70% conditional on experiencing a non-zero change in subway access.

The goal is to estimate the *level* effect of gaining subway access on earnings and distance to workplace. As such, the estimand of interest is a weighted average of the Average Treatment Effect on the Treated (ATT), where weights are assigned across the treatment exposure distribution. Following Callaway et al. (2024), the standard Parallel Trends assumption is sufficient to estimate this *level* effect of gaining subway access. I define treated workers as those with a non-zero change in distance to their nearest subway station between 2008 and 2009, a set of workers identified as gaining access to a subway network, to estimate the ATT given below.

$$ATT^{\mathrm{loc}} := \mathbb{E}[\Delta y | \%\Delta dist < 0] - \mathbb{E}[\Delta y | \%\Delta dist = 0] \tag{1}$$

Although it is possible to run a two-way fixed effects model with a continuous exposure measure, Callaway et al. (2024) shows that such a specification identifies a weighted average of treatment effects across exposure groups, but suffers from having negative weights and does not integrate up to one. To address this issue, I identify the above ATT by defining treatment T_i equal to one if a worker experienced *some* drop in distance to nearest subway station, and zero otherwise. Another benefit of discretizing treatment into two groups is that it offers a simple and transparent pre-trends test (Sun and Shapiro, 2022). With a continuous treatment definition, a stronger pre-trends assumption is needed to rule out selection into treatment intensity. However, our goal is to estimate the effect of the *level* effect of gaining access to a subway, not the marginal increase in access. I also present results using percentage change quintiles as treatment bins to test whether workers who experienced a sharper change in distance experienced a greater treatment effect. Results from this specification present estimates for a weighted ATT within each quintile.

I limit the analysis to workers whose nearest subway station was more than 3 km away in

2008. This rules out comparisons between workers who reside close to the Millennium line (earlier treated), which started operating in 2002, and workers treated later due to the Canada Line construction. This is illustrated in Figure (2) where the gray area represents neighborhoods not included in either treatment or control groups. The time period of interest is 2005-2015, during which no other subway expansion took place; however, this does not rule out changes in travel costs due to time variation in congestion, changes in bus routes, or other infrastructure projects. Note that the Expo Line was expanded in 2016 to connect Burnaby with Coquitlam. This extension meant that a part of the control group eventually gained access to the SkyTrain network. This extension was planned since the 1990s, but was set aside for political reasons and to redirect funds for the Canada line in preparation for the Olympics. Restricting the time period of interest to 2005-2015 to be prior to the Millennium line extension.

Descriptives Figure 2 illustrates the spatial distribution of treated and control neighborhoods at the Census Tract (CT) level. Comparing this map with Figure 1 reveals that areas with the highest treatment intensity are those located along the Canada line route. Note that neighborhoods on the Expo and Millennium lines are excluded from the sample. Table 2 presents summary statistics at the worker-year level by treatment status before and after the expansion for the estimation sample. Panel A shows that prior to the expansion, treated workers had higher average earnings- $78,000$ relative to $68,000$ for their control counterparts, who are closer to the downtown area, are more likely to work in professional industries, and less likely to be in manufacturing or construction. They are, however, employed at firms with lower total wage bills. The *level* differences between both groups do not preclude labor market outcomes from being on parallel trends prior to the expansion. This will be tested using the dynamic event study specifications below, with $t = -1$ normalized to 0, and with standard errors clustered at the individual level.

To estimate the ATT in equation (1), I run the following dynamic event study specification:

$$Y_{igt} = \alpha_i + \alpha_{gt} + \sum_{\tau=-4,\tau\neq-1}^{\tau=6} \beta_\tau \mathbb{1}\{\tau = t - 2009\} T_i + \varepsilon_{igt} \tag{2}$$

Where Y_{igt} is the outcome for individual i in demographic group g, at time t. I include individual fixed effects α_i, and demographic group by calendar year fixed effects given by α_{gt}. Demographic groups are defined as the interaction of gender and age bins. There are two age bins, 25-40 and 40-59. The dummy variable T_i takes the value 0 or 1 depending on treatment status as outlined above. The parameters of interest are β_τ, which capture the treatment effect relative to the time of Canada Line construction.

Worker-level panel data allows me to rule out sorting of workers into neighborhoods with better subway access as the main mechanism. Although residential sorting will respond to changes in the transit network, following *initially* affected workers allows me to rule it out. This is similar to Pérez et al. (2022) who studied how the expansion of the Santiago transit system affected workers in closer proximity to the new stations.

3.3 Results

Figure 3 presents the estimated coefficients from equation (2). The results indicate an increase in earnings by approximately 1.5% three years following the expansion. This effect persists and remains at almost 2% until 2015, six years after the expansion. The estimated coefficients prior to the expansion indicate that the earnings of the treatment and control groups were on parallel trends, lending validity to the empirical design and the choice of treatment definition. Specifically, it provides evidence that the Canada Line expansion was not correlated with prior trends in outcomes in treatment and control areas relatively. Additionally, Figure A-3 presents the effects of going from the bottom to the top quartile from a TWFE specification under a continuous treatment definition. The estimated coefficients show that an increase in the absolute value of the reduction in distance to the nearest subway, from the lowest to the highest quartile, increases earnings by 0.6% as a result of the SkyTrain expansion. Although this specification leverages variation in proximity to a new subway station, offering richer variation, it suffers from the negative weights problem as detailed by Callaway et al. (2024). Overall, the observed effect on earnings could be driven by several mechanisms. First, improved lower commuting costs may have incentivized workers to move to more productive, farther away firms that were initially very costly to commute to. It is also possible that stayers benefited from increased productivity due to the less burdensome commute. Finally, it is possible that firms along the route of the new line had a demand shift, especially those in the service sector. It is challenging to empirically isolate the role of each channel. My model-based simulations will allow me to isolate the effect of improved access to firms on changes in workers' labor market concentration and reallocation channels, while holding all other productivity-related variables constant.

Given these empirical challenges, I provide some suggestive evidence that an important mechanism through which workers benefited from this expansion is the ability to move to more productive employers located farther away. Figure 4 shows that treated workers are 4 percentage points more likely to change their 2008 employer 3 years after the expansion. This represents approximately a 10% increase in the probability of switching employers relative to the control group. This provides some suggestive evidence that switching employers is a primary channel through which the effect on wages is manifested. To better illustrate this, I split the estimation by job movers and stayers. Movers are defined as workers who changed their 2008 primary employer at any point between 2009 and 2010. Figure 5 illustrates results from estimating equation 2 separately by movers and stayers. Estimated coefficients show that movers experienced a larger gain in earnings due to the new line, reaching 2.5% 3 years after the expansion. Confidence intervals point to statistically significant differences in earnings gains between movers and stayers. Figure **??** shows estimated coefficients under a continuous treatment definition, split by movers and stayers. For movers, transitioning from the bottom to the top quartile of treatment is associated with a 0.9% increase in earnings three years after the expansion. For stayers, this value is halved. The difference between the two groups is statistically significant. It is important, however, to note that job switching is endogenous, and splitting the sample by this margin can lead to spurious conclusions. For example, if workers at firms hit by a contemporaneous negative productivity shock were more likely to leave, then this would incorrectly attribute the higher wage gains for job switchers to the subway expansion. To provide more suggestive evidence that the subway expansion alleviated the commuting burden of affected workers and allowed them to move to farther employers, I show that job switchers, defined as those who switched employers between 2009 and 2010, had employers located farther away compared to job switchers in the control group. Figure 7 presents estimated coefficients from regressing distance to workplace on time and treatment dummies for the sample of job switchers. Estimates show that job switchers in the treated group worked farther away than their control counterparts, who also switched employers between 2008 and 2010. This corroborates the hypothesis that treated workers faced lower commuting costs, allowing them to travel farther away, provided they switched employers.

Time savings due to shorter commuting time is another margin through which workers who gain access to a subway station might benefit. For example, with less time commuting, workers are able to work longer hours or may be more productive for the same number of hours. To gauge the importance of this mechanism, I estimate treatment effects by treatment intensity quantile, where the lowest quantile comprises the set of treated workers who experienced the largest change in distance to the nearest subway station (e.g., a new station opened one block away). The rationale is that if time savings are the main drivers of the results in Figure 3, then those living closer to a new subway station should experience a higher gain in their earnings. This is possible due to the availability of accurate station-level coordinate data. Figure 8 presents the results from a static Difference-in-Difference with treatment intensity by quantile. The treatment effect for workers in the first quantile of treatment intensity (group with the highest change in distance to nearest station) is 4%, and declines gradually, reaching 3% for the third to fifth quantiles. This suggests that time savings from shorter commute times play a role in explaining the effect on earnings in Figure 3. However, the treatment effect on earnings does not vary across treated workers in the second to fifth quantiles of treatment intensity, and is relatively stable at 2%. These results suggest that time savings due to shorter commute time plays a non-negligible role in explaining the effect of subway access on earnings, however it is not sufficient to explain the effect on earnings for those who had a lower change in access to subway, nor to explain why amongst those workers there is no evidence that workers closer to the new subway station had a higher treatment effect.

The results presented from this empirical exercise are comparable to the findings by Pérez et al. (2022) who show that the expansion of Santiago's Subway system increased monthly earnings for workers closer to the new subway stations by more than 1% four years after the expansion. Moreover, the results are qualitatively in line with the findings of Caldwell and Danieli (2024) who show that the introduction of a new high-speed rail connecting the small city of Montabaur in Germany, with its relatively larger neighbors, Cologne and Frankfurt, increased workers' outside options due to the reduction in travel time. They do not, however, examine the effect on wages resulting from this shock. This is in line with the job-level specification, suggesting that changes in outside options affected incumbent workers without the need to commute further away for a better match. The effect on earnings is also consistent with the findings of Bütikofer et al. (2024) who show that Swedish workers living in close proximity to a new bridge (Öresund bridge) connecting Malmö with Copenhagen experienced an increase in wages by 13% eight years after the opening of the bridge. The larger magnitude of their estimates points to the large difference between the two cities. A difference that is likely lower within the neighborhoods of the Metro Vancouver area. Nevertheless, the results presented here are consistent with their findings; better market access increases workers' earnings.

The empirical evidence suggests that gaining access to the subway network has a positive impact on workers' earnings. The transit shock has several implications. First, by reducing commuting costs, it increases workers' effective number of employers, which jointly affects their labor supply elasticities and their allocation to firms. Second, it is possible that treatment effects are explained by other productivity-related factors, such as improved worker productivity resulting from a less burdensome commute. Finally, another concern is that untreated workers could potentially benefit from reduced congestion as a result of the subway expansion. This would violate the SUTVA assumption, indicating that spillovers could be a threat to the identification of treatment effects. To address these empirical concerns and isolate the role of improved access to more productive firms from other explanations, I build on Chan et al. (2024) to develop a two-sided labor market matching model with firm wage-setting power and commuting costs. The model nests a new classical monopsony framework into a quantitative spatial model, providing a tractable framework to assess the effect of changes in commuting costs on worker-firm matching and changes in labor supply elasticity.

4 Model

4.1 Environment

Setup At each point in time t, there is a mass of M_t workers $i \in \mathcal{I}$ who are heterogeneous in their type $k \in \mathcal{K} := \{1, \ldots, K\}$. Worker type is exogenous and can be thought of as a set of observable demographics. There is a finite number of cities $c \in C := \{1, \ldots, C\}$, where each city $c \in C$ has a finite set of neighborhoods denoted by $z \in \mathcal{Z}(c)$. Each neighborhood z has an *exogenous supply of housing* denoted by H^s_{zt}. Each worker type k has workers with mass m_{kt} such that $\sum_{\mathcal{K}} m_{kt} = M_t$ in each period. On the other hand, there is a finite number of heterogeneous firms $j \in \mathcal{J} := \{1 \ldots J\}$. Firms are *exogenously* distributed across the finite neighborhoods $\mathcal{Z} := \bigcup_c \mathcal{Z}(c)$.[12] Each firm belongs to some sector $g \in \mathcal{G} := \{1 \ldots G\}$ defined by the 2-digit NAICS code of the firm. Firms produce a set of perfectly substitutable freely tradable goods using different types of labor l_k as the only input into the production function $F_j(l_{\cdot j})$.[13]

Decisions Workers take their type k as given, and choose which neighborhood z to reside in, and which employer j to work for. This decision takes into account that attractive residential locations need not overlap with productive ones. Thus, workers make choices to balance housing costs, wages, and commute distance.[14] For example, a worker with a long commute is compensated for such disamenity either by higher wages and \or lower housing costs. Firms, on the other hand, take their productivity, amenities, and location $o(j)$ as given, and make hiring decisions by setting wages for each type of worker k. Firms are vertically differentiated by wages and systematic amenities, and horizontally differentiated due to worker-specific variation in distance to workplace, and other idiosyncratic preferences. Conditional on worker type, k, the horizontal differentiation components induce an upward-sloping firm-specific labor supply curve. Firm-specific labor augmenting TFPs depend on worker type, k, but are invariant to workers' locations. Therefore, firms only account for type-specific aggregate labor supply. The location of the workforce matters to the extent that it affects changes in the marginal worker labor supply elasticity. *The implicit assumption here is that employers can only observe the average labor supply curve for a given group* k, and therefore cannot discriminate by worker residential neighborhood, z. When making hiring decisions, firms

[12]This is equivalent to allowing firms to choose their location but are assumed to be ex-ante homogenous, and only observe their productivity following location decision. Some recent papers allow for worker-firm sorting across space within a wage-setting framework but assume that firms are either atomistic or ex-post homogenous (Bamford, 2021; Hong, 2024; Lindenlaub et al., 2024). Assuming firms are atomistic ex-ante excludes the Robinson (1933) economy of company towns, and other geographically localized forms of monopsony power as discussed by Manning (2003).

[13]A firm does not need to employ all types of worker k. Firm production function outlined in Section 4.3

[14]There is evidence that straight-line distance is a good proxy for travel time. Using data from the New York State Department of Transportation, Phibbs and Luft (1995) finds a correlation of 0.98 between straight-line distance and travel times; the number drops to 0.82

for distances below 15 miles, such that travel time per mile is generally larger than for long distances.

consider how a change in its wage schedule would affect not only the quantity of workers it attracts, but also from where, in relation to its location, it attracts those extra workers. If firms cast a wide geographical net in terms of hiring, the marginal worker would be overly sensitive to wage changes, thereby increasing the average firm-specific labor supply elasticity and, consequently, wages.

Strategic Interactions & Local Labor Markets I do not take a stance on the geographical borders of labor markets, but rather allow them to be pouros to the extent of workers' commuting costs. That is, regardless of a given worker's residential location she can always choose from the full menu of employers, but all else equal, further work commutes have lower average utility. This differs from other papers in the literature that assume labor markets are geographically segmented. Labor market concentration indices derived from such models do not capture the true set of outside options when workers can travel beyond their defined segment.[15] As such, firms compete for workers from *every* residential location, z, and aggregate *firm-location* specific labor supply curves to obtain production-relevant *firm-specific* labor supply for each type k. To sum up, due to commuting costs, firms are horizontally differentiated not only by idiosyncratic preferences but also by distance from the worker's specific residential location. This means changes in firms' wages differentially affect labor supply across neighborhoods.

Given the highly granular nature of geographic competition, firms are not assumed to be atomistic. In each neighborhood z, firms compete in an oligopsonistic fashion with other firms in their own sector, g, and other employers. This builds on Chan et al. (2024) with the main difference being that worker types can vary by their residential location, z, which is unobserved by the firm, and type k, which is observed. In other words, the model is a smoothed version of company towns in Robinson (1933), where, in the limit, a large degree of geographic isolation or sufficiently high commuting costs can lead to a single monopsonist. From a policy perspective, to see why assuming away a specific definition of a local labor market is appealing, note the following two examples. First, if antitrust authorities screen mergers based on measures such as HHI, then obtaining the *right* set of outside options for workers is crucial, and worker mobility must be taken into account. Second, to assess the wage effects of improved transit, we must counterfactually allow workers to travel to previously infeasible locations. Models with segmented labor markets can only speak to this topic across commuting zones, yet many infrastructure projects are within.

[15]For example, in the US Azar et al. (2020) define a labor market as the intersection of a Commuting Zone (CZ) with 6-digit occupational code, while Lamadon et al. (2022) defines them as the intersection of CZs with 2-digit NAICs codes. There are a few exceptions that allow labor markets to be geographically porous (Pérez et al., 2022; Datta, 2023; Schubert et al., 2024)

4.2 Workers' Problem

4.2.1 Preferences

In each period, t, worker i of skill k first draws a preference shock $\boldsymbol{\nu}_{it}$ over neighborhoods $\mathcal{Z}$. Then, after making their residential decision, they receive a preference shock over the set of all possible employers $\mathcal{J}$ given by the vector $\boldsymbol{\epsilon}_{it}$. They receive this shock over the set of all employers regardless of their residential location. This timing assumption is similar to Quantitative Spatial Models, with the exception that I allow workers to choose their residential location and employer, as opposed to only their work location in an Armington model of production (Tsivanidis, 2022; Couture et al., 2024; Baum-Snow and Han, 2024). The workers' indirect utility function is given by the following: [16]

$$U_{izjt} = \ln\left(\frac{a_{zt}}{r_{zt}^{\gamma}}\right) + \frac{1}{\varphi}\nu_{izt} + \underbrace{\ln w_{kjt} + \ln u_{zkjt} - \eta_k d_{z(i),o(j)} + \frac{1}{\beta_k}\epsilon_{ijt}}_{\textit{worker i compensation net of commuting costs at firm j}} \tag{3}$$

After $\boldsymbol{\nu}_{it}$ is revealed, workers make a location decision given the group k specific amenity, a_{zkt}, from residing in neighborhood z, and the unit price of housing, r_{zt}. Where γ is the share of expenditure on housing by households, and φ is a scale parameter. The stochastic residential preference shocks ν_{izt} follow a nested logit distribution with a *homogenous* within nest correlation parameter σ given by the following distribution:

$$F(\nu_{izt}) = \exp\left\{\sum_{c\in C}\left(\sum_{z\in\mathcal{Z}(c)} e^{-\frac{\nu_{izt}}{\sigma}}\right)^{\sigma}\right\} \tag{4}$$

Given residential location choices, idiosyncratic labor market preferences, ϵ_{ijt}, are revealed, and workers make employment decisions given the following. One, the type k specific vector of posted wages by firms, w_{kjt}. Two, the type-location specific vector of non-pecuniary benefits, u_{zkjt}. This denotes the non-pay amenities for type k workers residing in neighborhoods z from working at firm j. Importantly, this is assumed orthogonal to commuting distance. Finally, given residential location, z, workers consider the commuting distance to each possible employer j, as denoted by $d_{z(i),o(j)}$. Workers can also decide to remain non-employed, in which case they receive w_{k0t}, and do not commute by definition. Average utility from nonemployment is normalized to zero. The

[16] This is derived from Cobb-Douglas preferences over housing and consumption, with housing expenditure share given by γ

parameter β_k measures the importance of wages relative to non-pay amenities. A higher value of β_k means employers are less horizontally differentiated, increasing workers' labor supply elasticity, and reducing firms' labor market power. The parameter η_k measures workers' marginal rate of substitution between commuting distance and wages. A higher value implies workers require higher compensation to be indifferent between their current job and the next best alternative. In each period, labor market stochastic shocks, ϵ_{ijt}, denote a worker's idiosyncratic preference to work at firm j and follow a nested logit with the following distribution function:

$$G(\epsilon_{ijt}) = \exp\left\{\sum_{g\in\mathcal{G}}\left(\sum_{j\in J(g)} e^{-\frac{\epsilon_{ijt}}{\lambda_{kg}}}\right)^{\lambda_{kg}}\right\} \tag{5}$$

Each nest, $\mathcal{J}(g)$, is defined to contain all employers in an industry $g \in \mathcal{G}$. The type-industry specific parameters $\lambda_{kg} := \sqrt{1 - corr(\epsilon_{ikj}, \epsilon_{ikl})}$ measure the degree of independence between preference shocks for firms within the same nest. A higher value implies more independence and less correlation. As $\lambda_{kg} \to 1$ firms within a nest are at their maximal degree of horizontal differentiation, and are perceived as independent alternatives. Workers have relatively inelastic labor supply elasticities, and firms' market power is high. On the other hand, as $\lambda_{kg} \to 0$ they are perceived as perfect substitutes, eliminating any within nest horizontal differentiation. This means that, within a nest, any small change in wages is sufficient to cause workers to move to another employer, implying higher labor supply elasticities and lower firm market power. The parameters $\beta_k, \lambda_{kg}, \eta_k$ govern firm-specific labor supply elasticity with respect to wages, as will be shown in section (4.2.2)

In this model, firm-specific labor supply curves are upward sloping due to horizontal differentiation arising from two sources. First, the idiosyncratic preference term, which captures the stochastic preference over amenities. Second, which is a key contribution of this paper, is the worker-firm specific distance term. As the firm expands along its labor supply curve, it attracts workers from different neighborhoods with higher wages, who are now more willing to commute. Note that an important assumption is that employers can only observe the *average labor supply curve for a given group k*, and therefore cannot discriminate by worker residential neighborhood, z. Due to disutility from commuting, workers who commute further are more sensitive to wage changes, reducing the firms' ability to markdown wages. An empirical question is to what extent workers' sensitivity to commuting is sufficient to offset wage-setting power (Mas, 2025).

4.2.2 Labor Supply & Residential Sorting

In each period, t, workers observe the vector of neighborhood amenities $\{a_{zkt}\}$, housing rents $\{r_{zt}\}$, the set of firms $\mathcal{J}$ and firms' compensation packages denoted by $\{w_{kjt}, u_{kjt}, u_{zkg}\}$. Given these values, workers then draw their idiosyncratic preference shocks over neighborhoods, $\boldsymbol{\nu}_{izt}$, and choose where to locate by forming expectations over idiosyncratic labor market preference shocks. Given their residential choice and labor market preferences over employers, $\boldsymbol{\epsilon}_{ijt}$, workers choose which firm to supply their labor to. I solve the workers' problem backwards by deriving firm-specific labor supply, *given residential locations*, then I show how workers make these residential locations by forming expectations over labor market outcomes to derive residential choice probabilities. Worker i of type k, living in city c and neighborhood z, chooses amongst employers $j \in \mathcal{J}$ to maximize $\tilde{U}_{ijt}$ as denoted by (3)

$$\tilde{U}_{ijt} = \underbrace{\ln w_{kjt} + \ln u_{zkjt} - \eta_k d_{z(i),o(j)}}_{=:v_{zkjt}} + \frac{1}{\beta_k}\epsilon_{ij} \tag{6}$$

Workers who choose to remain in non-employment have a utility given by

$$U_{ik0t} = \ln w_{k0t} + \frac{1}{\beta_k}\epsilon_{ik0} \tag{7}$$

where w_{k0t} are unemployment benefits to be exogenously set. I assume that commuting distance when unemployed is zero. Note the assumption that non-employment average utility is location invariant.

Neighborhood-Firm specific Labor Supply Following from (6), and given the nested logit assumption on preferences, the share of workers from neighborhood z of type k working at firm j from nest $\mathcal{J}(g)$ is given by:

$$s_{zkjt} := s_{zkgt} \underbrace{\frac{\exp\left(\frac{\beta_k v_{zkjt}}{\lambda_{kg}}\right)}{I_{zkgt}}}_{s_{zkj|gt}} \tag{8}$$

Where s_{zkgt} is the share in industry g

$$s_{zkgt} = \frac{I_{zkgt}^{\lambda_{kg}}}{\sum_f I_{zkft}^{\lambda_{kf}}} \quad \text{and} \quad I_{zkft} := \sum_{l \in \mathcal{J}(f)} \exp\left(\frac{\beta_k v_{zkjt}}{\lambda_{kg}}\right)$$

The labor supply elasticity of workers in neighborhood z of type k to firm j is given by: [17]

$$\mathcal{E}_{zkjt} := \beta_k \left[\frac{1}{\lambda_{kg}} + \left(1 - \frac{1}{\lambda_{kg}}\right) s_{zkj|gt} - s_{zkjt} \right] \tag{9}$$

Equation (9) denotes several mechanisms. One, as $\beta_k \to \infty$ horizontal differentiation between firms vanishes and workers' labor supply to any particular firm becomes infinitely elastic, eroding firms' market power. Similarly, as $\lambda_{kg} \to 0$ firms are viewed as perfect substitutes within each nest, increasing the perceived number of different alternatives for workers, and increasing firm-specific labor supply elasticity.

At the other extreme, as $\lambda_{kg} \to 1$ we are in the logit case and elasticities can be written as $\beta_k \left(1 - s_{zkjt}\right)$. Two, firms with high labor market shares from neighborhood z face lower labor supply elasticities *from that location* since $\left(1 - \frac{1}{\lambda_{kg}}\right) s_{zkj|gt} - s_{zkjt} \leq 0$. Whether the neighborhood invariant elasticity labor supply is lower, too, depends on the geographical distribution of the firms' workforce. If all workers are from neighborhood z, then the change is one-to-one. Otherwise, it depends on whether the change in the neighborhood-firm specific elasticity (marginal) is lower than the within-firm elasticity (average). To sum, firms affect workers' labor supply elasticities by offering better compensation packages, but this differentially impacts workers in different locations due to commuting costs, which means the neighborhood-specific shares s_{zkjt} do not change uniformly.

Residential Sorting & Housing Demand Since workers make residential choices before observing idiosyncratic labor market preferences, residential choice is shaped by

residential amenities, housing costs, *and expected labor market compensation*. Given the distributional assumptions, the latter is given by

$$\begin{aligned} RMA_{zkt} &= \mathbb{E}\left[\max_{j \in \mathcal{J}} \{v_{zkjt} + \epsilon_{ij}\} \right] \\ &= \ln \left[e^{\beta_k v_{zk0t}} + \sum_g \left(\sum_{j \in N_g} e^{\frac{\beta_k v_{zkjt}}{\lambda_{kg}}} \right)^{\lambda_{kg}} \right] \end{aligned} \tag{10}$$

The term RMA_{zkt} summarizes access to employment opportunities from residential neighborhood z, for workers of type k at time t. The last equality follows from the distributional assumption on idiosyncratic labor market preferences (McFadden et al., 1978). Therefore, after drawing loca-

[17] Labor supply elasticity conditional on location is derived in Appendix B.2

tion preferences, workers choose a neighborhood z to maximize the following:

$$max_{z \in \mathcal{Z}} \underbrace{\ln a_{zkt} - \gamma \ln r_{zt} + RMA_{zkt}}_{y_{zkt}} + \frac{1}{\varphi} \nu_{zit} \tag{11}$$

Given the distributional assumption in (4), the share of workers residing in neighborhood z is given by:

$$\mu_{zkt} := \pi_{kct} \underbrace{\frac{\exp\left(\frac{\varphi y_{zkt}}{\sigma}\right)}{\Phi_{ckt}}}_{\pi_{zk|ct}} \tag{12}$$

where

$$\pi_{zkct} = \frac{\Phi^{\sigma}_{ckt}}{\sum_{c'} \Phi_{c'kt}} \quad \text{and} \quad \Phi_{ckt} = \sum_{z' \in \mathcal{Z}(c)} \exp\left(\frac{\varphi y_{z'kt}}{\sigma}\right)$$

The number of workers residing in neighborhood z of type k at time t is therefore given by $\ell_{zkt} := m_{kt}\mu_{zkt}$. Note that since firms are not assumed to be atomistic, a change in a single firm's wage schedule can affect the population distribution of a given neighborhood. This allows for potential population declines following localized labor demand shocks, for example, the demographics in factory or mining towns can drastically change after plant closures (Bound and Holzer, 2000; Notowidigdo, 2011; Krueger, 2017). This example shows that such an effect depends on employer size in a given city or neighborhood. Formally, define ζ_{zkjt} to be the population elasticity for type k workers in neighborhood z with respect to firm j wages at time t. Given the probabilities in (12) and the residential market access term in (10) this elasticity has the following form: [18]

$$\zeta_{zkjt} = \varphi \beta_k s_{zkjt} \left[\frac{1}{\sigma} + \left(1 - \frac{1}{\sigma}\right)\pi_{zk|ct} - \pi_{zkct}\right] \tag{13}$$

Finally, given employer choice for all residents in the neighborhood z, demand for housing has the following form:

[18] Derivation in Appendix B.2

$$H^d_{zt}(r_{zt}) = \frac{\gamma}{r_{zt}} \underbrace{\sum_{k \in \mathcal{K}} m_{kt} \mu_{zkt} \sum_{j \in \mathcal{J}} w_{kjt} s_{zkjt}}_{\text{Total Earnings in neighborhood } z := E_{zt}} \tag{14}$$

The Cobb-Douglas preference specification implies that aggregate housing demand in a given neighborhood z is a fraction of total earnings received by residents.

4.3 Firms' Problem

The other side of the labor market is characterized by a set of heterogeneous firms $\mathcal{J}$ and a wage-posting framework. The location of each firm j is assumed exogenous, and is given by $o(j) \in \mathcal{Z}$. In each period, firms observe the labor supply curve of each type k (*firm-specific* labor supply), but cannot observe the residential location of each worker. In other words, firms only observe the aggregate of equation (8) across neighborhoods for each type k. Therefore, for a given firm j, the labor supply curve of workers with type k is given by $\ell_{kjt} = \sum_{z \in \mathcal{Z}} \ell_{zkt} s_{zkjt}$. Firms also observe their production function $F_{jt}(\ell_{.j})$, where $\ell_{.j} = (\ell_{1jt} \dots \ell_{kjt})$ and an exogenous output quantity, Q_{jt}. Firms then decide on the number of workers ℓ_{kjt} to hire by setting *type-specific wages*, w_{kjt}, to minimize their wage bill.[19] Formally, in each period t, firm j sets wages $\{w_{kjt}\}$ to minimize their wage bill, taking as given other firms' wages, and type-specific labor supply functions. Assuming that output quantity Q_{jt} is exogenously given, the firm solves the following:

$$\begin{aligned} &\min_{w_{kj}} \sum_{k \in \mathcal{K}} w_{kjt} \ell_{kjt} \\ s.t.\ &F^j\left(\ell_{.jt}\right) \geq Q_{jt} \\ &\ell_{kjt} = \sum_{z \in \mathcal{Z}} \ell_{zkt} s_{zkjt} \quad \forall (k, j) \in \mathcal{K} \times \mathcal{J} \\ &w_{kjt} \geq 0 \end{aligned} \tag{15}$$

For each firm j, the production function follows the functional form below:

$$F_{jt}(\ell_{.j}) = \theta_{jt} \left(\sum_{k \in C_j} \gamma_{kjt} \ell^{\rho}_{kjt} \right)^{\frac{\alpha_{jt}}{\rho}} \qquad 0 < \alpha_{jt} \leq 1 \tag{16}$$

[19]The assumption of exogenous quantities follows from assuming firm output to be a set of perfectly substitutable, freely tradable goods.

The parameter θ_{jt} is a hicks-neutral firm-specific TFP. The parameters γ_{kjt} are labor augmenting TFP shocks, and govern firm relative productivity of workers. More importantly, they allow for sorting based on match-specific production complementarities between firms and worker types. The parameter α_{jt} represents firm-specific returns to scale. Finally, *homogenous* ρ captures the degree of substitutability between workers of different types k. A value of $\rho < 1$ indicates that workers are complementary in production. Note that in the presence of diminishing returns, all else equal, bigger firms have a lower marginal product of labor and therefore offer lower wages. A firm with better amenities, which is able to attract relatively more workers, will offer lower wages. This means that wages have a compensating differential component. Note that this is not the case under a constant returns to scale production function, where only worker productivity enters into wage determination.

Since the firm does not observe the residential location of each worker, the relevant labor supply elasticity for the firm's wage-setting problem is the weighted average of elasticities across all neighborhoods, denoted by $\mathcal{E}_{kjt}$. Where the weights correspond to the share of workers from each neighborhood *within* the firm, ω_{zkjt}. Thus the firm specific labor supply elasticity of type k to firm j is given by:[20]

$$\mathcal{E}_{kjt} = \sum_{z \in \mathcal{Z}} \omega_{zkjt} \left(\mathcal{E}_{zkjt} + \zeta_{zkjt} \right) \tag{17}$$

Equipped with the labor supply elasticity, the first order conditions of problem (15) yield the following Lerner condition for each firm j:

$$w_{kjt} = \underbrace{\lambda_{jt}}_{\text{shadow price of output}} \underbrace{\frac{\partial F^j(\ell_{.j})}{\partial l_{kjt}}}_{MPL_{kjt}} \times \underbrace{\frac{\mathcal{E}_{kjt}}{\mathcal{E}_{kjt} + 1}}_{mdn_{kjt}} \tag{18}$$

Equation (18) yields the standard Lerner condition for wages as a markdown $mdn_{kjt} \leq 1$ below the marginal product of labor, which is the product of the shadow price of output λ_{jt} and the marginal product of labor. Note that, unlike models of monopsonistic competition, labor supply elasticities, $\mathcal{E}_{kjt}$, are a function of labor supply parameters β_k, λ_{kg}, as well as labor market shares s_{kjt} (Berger et al., 2022; Chan et al., 2024). Further, horizontal differentiation due to commuting distance means elasticities are shaped by disutility from commuting, η_k, and the geographical distribution of workers relative to the firm. For example, if for a given firm workers live in close

[20]Full derivation of below equation is given in Appendix B

proximity, then labor supply is relatively inelastic, as little compensation for commuting is needed. However, if a given worker lives further away, they need to be compensated for their commute, and will be more sensitive to wage changes. Since the firm can only set wages given the average elasticity of its workforce, this raises the wage set by the firm for all workers. This is akin to screening models where workers' outside options are unobserved to the firm, even after conditioning on observable types (Kline, 2025; Mas, 2025). The case of company towns, as in Robinson (1933) is reached when residential locations and employers are sufficiently geographically isolated. In this case, labor supply elasticities are low due to the limited number of employers within commuting distance, and firms can, to a large degree, act as a single monopsonist within a given geographic area.

4.4 Equilibrium

In each period t, given neighborhood and firm fundamentals $\{\Theta_{\mathcal{Z}_t}, \Theta_{\mathcal{J}_t}\}$ and the distribution of workers $\{m_{kt}\}$, where the former denote housing supply, residential amenities, firm specific non-pay amenities, and firm location, an equilibrium is defined by an allocation of workers to neighborhood-firm pairs $\{m_{z,j|kt}\}$

and a set of wages $\{w_{kjt}\}$ and rental prices $\{r_{zt}\}$ such that in each period:

1. Workers choose neighborhood and employer pair (z, j) to maximize (3)
2. Given their locations,and workers' labor supply curves, wages $\{w_{kjt}\}$ solve firms' problem given by (15)
3. Rental prices $\{r_{zt}\}$ and wages $\{w_{kjt}\}$ are such that housing demand, as given by (14), equals housing supply in each neighborhood
4. Consistency & population constraints:
    (a) Firms' optimal labor inputs are consistent with the implied spatial distribution

$$s_{kjt}(\vec{w};\vec{r}) = \sum_{z\in\mathcal{Z}} \overbrace{\frac{m_{z,j|kt}}{\sum_j m_{z,j|kt}}}^{s_{zkjt}} \underbrace{\sum_j m_{z,j|kt}}_{\mu_{zk}}$$

    (b) Population constraints: $\sum_{j\in\mathcal{J}} s_{kjt} + s_{k0t} = m_{kt}$ and $\sum_{z\in\mathcal{Z}} \mu_{zkt} = m_{kt}$ $\forall k \in \mathcal{K}$

Equilibrium existence follows the fixed-point argument in Chan et al. (2024) Appendix B. Uniqueness, however, is only guaranteed under a set of sufficient conditions. Assuming the only component that horizontally differentiates firms is the unobservable idiosyncratic preferences, then under any log-concave distribution for workers' stochastic, for example, Type I Extreme Value, we have a unique equilibrium whenever. This is proved in Chan et al. (2024) showing that a unique equilibrium exists whenever the super-labor supply elasticity, and the cross super labor supply elasticities satisfy the gross substitutes property. Which holds under log-concavity. However, when there are two sources of horizontal differentiation, observable commuting costs, and stochastic preferences, the sum of these two need not be log concave. And additional restrictions need to be imposed to satisfy the assumptions in Chan et al. (2024). I provide a more detailed derivation of the semi-super elasticity in Appendix B.5. Conceptually, uniqueness is guaranteed when the firm's labor supply is concave, meaning that workers are more inelastic as the firm grows. However, with commuting costs, the firm grows by hiring more workers from a given neighborhood, as well as workers living farther away. The latter requires more compensation due to their longer commute. As such, firm growth changes the composition of workers, and therefore the average labor supply elasticity.

4.5 Residential Market Access & Concentration Index

A novel feature of nesting classical monopsony into a quantitative spatial model is a generalized concentration index at the neighborhood-type level. Building on Chan et al. (2024), and given the spatial distribution of workers, the generalized concentration index for neighborhood z and workers of type k is given by

$$GCI_{zkt} = \left[\prod_{g\in\mathcal{G}}\left(\exp\left\{\sum_{j\in\mathcal{J}(g)} s_{zkt|g}\ln s_{zkj|gt}\right\}\right)^{s_{zkgt}\lambda_{kg}}\right]\times\left[\exp\left\{\sum_{g\in\mathcal{G}} s_{zkg}\ln s_{zkg}\right\}\right] \tag{19}$$

If workers from a given neighborhood are all employed at one employer, the concentration index equals 1. Importantly, the concentration index is at the neighborhood level, offering a model-based measure of labor market power at granular spatial units.

Importantly, there is a tight connection between the neighborhood-level concentration GCI_{zkt} and the neighborhood-level residential market access RMA_{zkt}. More precisely, the residential market access term can be decomposed into two components. One, the "matching component" which captures the systematic level of utility for workers in a given neighborhood from matching with a set of employers. Second, is the "concentration component" which captures the extent of labor market concentration at the neighborhood level. More precisely

$$RMA_{zkt} = \sum_{j \in \mathcal{J}} v_{zkjt} s_{zkjt} - \ln GCI_{zkt} \tag{20}$$

Note that changes to firm productivity, commuting costs, or any other fundamentals change both concentration and the matching component simultaneously. This happens due to changes in sorting and firm wage-setting responses. All else equal, labor market concentration acts as a population shifter and therefore has an inverse relationship with housing rents. The model therefore offers a microfoundation for the empirical relationship between housing costs and labor market concentration documented by Kahn and Tracy (2024).

5 Identification & Estimation

In this section I outline the identification & estimation strategy used to recover labor supply and demand parameters. I start by describing the identification of labor supply preferences, which is broken down into the identification of the distribution function $G(.)$, which governs workers' idiosyncratic preferences, as summarized by β_k, λ_{kg}. Equipped with these parameters, I then outline how to identify workers' marginal disutility from commuting η_k using a modified gravity equation. Given the identification of labor supply parameters, the Lerner condition in equation (18) allows the recovery of the marginal product of labor, which, along with variation in the ratio of worker types, is used to identify the substitution parameter ρ. With this in hand, other firm-specific labor demand parameters are identified. Residential demand parameters will be calibrated outside the model. I calibrate residential choice preferences using parameter estimates from Baum-Snow and Han (2024) and Couture et al. (2024). Speficially, I set $\varphi = 3$ and $\sigma = 0.35$. This takes into account the mapping from the Frechet distribution in those papers to a nested logit, with a homogenous nesting parameter σ

5.1 Labor Supply Parameters

5.1.1 Idiosyncratic Scale & Correlation Terms: β_k , λ_{kg}

The identification strategy of the idiosyncratic labor supply preferences distribution $G(.)$ follows Chan et al. (2024), with the main difference being in the definition of worker types to include their residential locations. Following Berry (1994) the quasi-supply function of equation (8) is defined as the log ratio of the share of type k workers, residing in neighborhood z, who work at firm j at time t, s_{zkjt}, relative to their neighbors of the same type who are in non-employment, s_{zk0t}. The quasi-labor supply function is given by:[21]

[21] Derivations for the quasi-labor supply function are in Appendix B.1

$$\ln \frac{s_{zkjt}}{s_{zk0t}} = \beta_k \ln \frac{w_{kjt}}{w_{k0t}} + (1-\lambda_{kg}) \ln s_{zkj|gt} - \beta_k \eta_k d_{z,o(j)} + \ln u_{zkjt} \tag{21}$$

The identification challenge in estimating equation (21) is the correlation of earnings, neighborhood-specific inside shares, and potentially distance to workplace, $d_{z,o(j)}$, with the unobserved amenities term. More precisely, an increase in the unobserved non-wage amenities simultaneously changes the offered wage, and increases the labor supply for the firm, leading to an inconsistently lower estimate of β_k. To address this, I identify the model parameters using an internal instrumental variable approach following Lamadon et al. (2022). This entails varying wages through productivity changes, while holding non-pay amenities *relatively* constant.

Assumption 1. Labor augmenting TFP γ_{kjt} can be decomposed into a type-specific and firm-type specific components, where the latter follows an AR(1) process. That is:

1a) $\gamma_{kjt} = \bar{x}_{kt} x_{kjt}$

1b) The firm-type specific component, x_{kjt} ,follows an AR(1) process with persistence parameter δ of the following form:

$$\ln x_{kjt} = \delta \ln x_{kjt-1} + \bar{\xi}_k + \bar{\xi}_{kjt}$$

Assumption 2. The firm-type specific amenities $\ln u_{zkjt}$ follows an MA(1) process

In words, Assumptions 1 and 2 ensure that *changes* in amenities are transitory relative to *changes* in labor-augmenting productivity. Note that these two assumptions do not preclude the *levels* of amenities and productivity from being correlated. The differences in their transitory nature mean that after taking sufficient differences over time, recent changes in productivity-related variables (e.g. revenue and labor supply) are correlated with changes in wages, as they are a function of labor productivity, but are uncorrelated with changes in amenities, as they are transitory (Lamadon et al., 2022; Chan et al., 2024). Further, taking long differences of (21) eliminates the time invariant distance to workplace $d_{z,o(j)}$. The parameter η_k will be estimated using a Gravity equation, as will be outlined in the next section.

Proposition 1. Under assumptions 1 and 2 *sufficient* long differences satisfy the relevance and exclusion restrictions *whenever the underlying composition of workers' unobservable skill is independent of changes in firm-level productivity shocks over the same time period.*

1. relevance: $Cov(\gamma_{kjt+2} - \gamma_{kjt-3}, x_{kjt} - x_{kjt-1}) \neq 0$

2. exclusion: $Cov(\ln u_{zkjt+2} - \ln u_{zkjt-3}, x_{kjt} - x_{kjt-1}) = 0$

Proposition 1 implies that changes over a long enough peirod in any productivity-related variables, such as wages and shares $s_{zkj|gt}$, are affected by short-period changes in other productivity-related variables. However, changes in amenities over the same period, are orthogonal to such productivity shocks. Since quasi-labor supply is at the neighborhood-firm level, an important assumption is that neighborhood unobservable skill composition does not change over the same period over which the long difference is applied, conditional on type k. As such, at each period t take the difference in the quasi-labor supply of each (z, k) type of workers between periods $t + 2$ and $t + 3$ to obtain the following:

$$\Delta_{2,3} \ln \frac{s_{zkjt}}{s_{zk0t}} = \beta_k \Delta_{2,3} \ln \frac{w_{kjt}}{w_{k0t}} + (1 - \lambda_{kg})\Delta_{2,3} \ln s_{zkj|gt} + \Delta_{2,3} \ln u_{zkjt} \tag{22}$$

Where $\Delta_{2,3}h_t = h_{t+2} - h_{t-3}$. For instruments to be valid, they must be correlated with long differences in earnings, inside share, but not with long differences in amenities terms. Applying Proposition 1 this holds for any productivity-related variables. In practice, equation (22) will be estimated using 2SLS where the instruments to be used are one-period changes in log revenue, neighborhood-specific log inside share, and *firm-level* sum of the shares for other types employed by the firm. These instruments are valid whenever the firm-level productivity shocks have a persistent effect on wages and shares, but not correlated with *changes* in long-term amenities.

5.1.2 Marginal Commuting Disutility: η_k

To identify workers' commuting disutility, I leverage within-firm variation in the supply of workers *across* neighborhoods relative to variation in commuting distance to the firm. The exponential decay of the number of workers relative to distance to firm identifies commuting disutility for each type k under the assumption that firms are *spatially connected*, that is, there is no one-to-one mapping from neighborhoods to firms. This is akin to the mobility assumption in AKM models for worker and firm wage effects. This is analogous to gravity equations used in the Quantitative Spatial Models literature, where variation in labor supply to origin destinations is used to identify disutility from commuting.

(Ahlfeldt et al., 2015; Tsivanidis, 2022; Dingel and Tintelnot, 2020). The main difference, however, is that having already identified the parameters β_k, λ_{kg} from the previous section, I can therefore separately identify commuting disutility from other labor supply parameters without the need to use the variation in earnings across space as in Ahlfeldt et al. (2015). To have a gravity equation that governs the flow of workers from neighborhoods to firms, I impose the following separability assumption on the non-pay amenities term $\ln u_{zkjt}$.

Assumption 3. Conditional on distance to workplace, non-pay amenities $\ln u_{zkjt}$ take the following separable form:

$$\ln u_{zkjt} = \ln u_{kjt} + \ln u_{zkg(j)t} + \ln e_{zkjt} \tag{23}$$

where

$$\mathbb{E}[e_{zkjt}|d_{z,o(j)}, \Gamma_{kjgt}, \Gamma_{zkgt}, k] = 1$$

In addition to permitting the identification of commuting disutility using gravity, Assumption 3, along with the previously identified labor supply parameters, allows recovering firm-specific amenities $\ln u_{kjt}$ from firm fixed effects in the gravity equation. Further note that the model does not exactly fit the data. After estimating the model, I will carry out validation checks with respect to commuting flows and the wage distribution to show that the baseline simulation of the model fits the data well.

Proposition 2. Whenever Assumption 3 holds, the share of workers at neighborhood z of type k employed by firm j at time t, takes the following form:[22]

$$s_{zkjt} = s_{zkj|gt} s_{zkgt} = \exp\left\{\tilde{\eta}_{kg} d_{z,o(j)} + \Gamma_{kjgt} + \Gamma_{zkgt}\right\} \tilde{e}_{zkjt} \tag{24}$$

where

$$\tilde{\eta}_{kg} := \frac{\beta_k \eta_k}{\lambda_{kg}}$$
$$\Gamma_{kjgt} := \frac{\beta_k \left(\ln w_{kjt} + \ln u_{kjt}\right)}{\lambda_{kg}}$$
$$\Gamma_{zkgt} := \ln\left(\frac{u_{zkgt}^{\lambda_{kg}}}{I_{zkg}}\right)$$

That is, a regression of log commute probabilities between each neighborhood-firm pair on firm and neighborhood fixed effects, plus commuting distance, recovers an estimate of the parameter bundle $\frac{\beta_k \eta_k}{\lambda_{kg}}$. This parameter corresponds to the semi-elasticity of commuting flows with respect to distance.[23] *The identifying assumption is that, conditional on firm and neighborhood fixed effects,*

[22] Derivation of the model implied gravity equation is given in Appendix B.4

[23] This is equivalent to the parameter $\kappa\epsilon$ in Ahlfeldt et al. (2025). I do not report estimated values of this parameter, but they range between -0.06 and -0.09

the unobserved factors that affect commuting in the error term are independent of Euclidean distances between locations. Figure 9 illustrates the identification argument graphically; the elasticity of commuting flows with respect to distance is identified from the covariance between the number of workers and distance to different neighborhoods within each firm. In other words, conditional on firm productivity, the decay in the number of workers with distance identifies workers' willingness to pay for commuting. The faster the decay of residualized log workers as distance increases, the greater the commuting costs. Note that even if workers of different types are observed to have the same commuting distribution, information on employer productivity is needed to identify the shape of their indifference curves. This revealed preference approach differs from Le Barbanchon et al. (2021), who has stated preferences on reservation wages and maximum travel distance, and is closer to the empirical method proposed by Card et al. (2024), who show that workers with higher commuting constraints should be observed to match with higher paying firms to compensate for their higher disutility.

I estimate equation (24) using separate flow-weighted commuting gravity regressions with firm and neighborhood fixed effects in 2016 for each (k, g) tuple. An empirical challenge in estimating commuting gravity equations is the presence of zero bilateral commuting flows in the data. This is especially the case when the number of firms or neighborhoods is large, creating a sparse commuting matrix. Although we have a continuous measure of workers, and the logit distribution is unbounded from above, implying there should always be a positive measure of neighborhood-firm pairs, there are zero flows in the data. One way to rationalize this is to assume infinitely prohibitive commuting costs between neighborhood-firm pairs with zero flows. As shown by Dingel and Tintelnot (2020) and Redding (2024) this is equivalent to the *calibrated shares* approach in International Trade (Eaton et al., 2012), and has the disadvantage of ruling out zero flows becoming positive. For example, if no workers from the west of a commuting zone are observed to work for a firm in the east, even under a reasonable commuting distance, any counterfactual will continue to assume these workers are constrained to their commuting radius given by the data, potentially underestimating welfare implications of policies such as transit infrastructure improvement, commuting subsidies, or changes in work arrangements such as work-from-home. Instead, I assume the zero flows in the data are due to having a finite sample from the continuous model, that is, the population (sample) in each neighborhood is relatively small compared to the number of potential employers. In which case, the parameters of the model can be estimated using the Pseudo-Poisson Maximum Likelihood (PPML) estimator (Silva and Tenreyro, 2006; Correia et al., 2020). Although fixed-effect estimators in non-linear models potentially suffer from an Incidental Parameter Problem (IPP), the PPML estimator with two fixed effects is asymptotically unbiased (Fernández-Val and Weidner, 2016). After recovering the fixed effects from regressions in equation (24) I use the labor supply parameters to recover $\ln u_{kjt}$. I outline the exact execution of this step in the estimation procedure.

5.2 Labor Demand Parameters

In this section, I demonstrate how to use the estimated labor supply elasticities to identify firm production functions. Conditional on having firm-level balance sheet data, where firm revenue is observed, the method to identify firms' production function can be split into two subparts. One, conditional on having recovered the elasticity of substitution parameter, ρ, other firm-specific parameters can be identified via inverting the firms' FOC in Equation (18). Two, absent endogeneity concerns, the substitution parameter is identified from variation in the ratio of worker types relative to changes in their marginal product. Under perfectly competitive labor markets, the latter coincides with the wage ratio. However, due to wage markdowns, this is no longer the case, and the marginal product of labor must first be identified.

5.2.1 Elasticity of Substitution: ρ

From Equation (18) we know that the wage is a markdown times the marginal revenue product of labor. Since labor supply parameters have been identified, firm-specific labor supply elasticity can be constructed using (4) and used to define the marginal revenue product of k-type workers at firm j as:

$$\tilde{w}_{kjt} := \frac{\mathcal{E}_{kjt} + 1}{\mathcal{E}_{kjt}} w_{kjt} \tag{25}$$

Fixing workers of type $k, h \in \mathcal{K}$ the firm first order condition in (18), along with the production function functional form assumption 1 imply that

$$\ln \frac{\tilde{w}_{kjt}}{\tilde{w}_{hjt}} = (\rho - 1) \ln \frac{\ell_{kjt}}{\ell_{hjt}} + \ln \frac{\tilde{\gamma}_{kjt}}{\tilde{\gamma}_{hjt}} \tag{26}$$

Running OLS on (26) will provide biased estimates, since labor ratios are correlated with labor-augmenting TFP. Using the identification assumption 2, substitute in the innovation terms, we get the following estimation equation:

$$\ln \frac{\tilde{w}_{kjt}}{\tilde{w}_{hjt}} = \zeta_{kht} + (\rho - 1) \ln \frac{\ell_{kjt}}{\ell_{hjt}} + \delta(\rho - 1) \ln \frac{\ell_{kjt-1}}{\ell_{hjt-1}} + \delta \ln \frac{\tilde{w}_{kjt-1}}{\tilde{w}_{hjt-1}} + \zeta_{khjt} \tag{27}$$

I estimate (27) using OLS after controlling for lagged labor marginal revenue product ratios for all possible types (k, h) within a given firm. That is, after correcting for markdowns using the estimated labor supply parameters, the inverse of the elasticity of substitution is identified from the covariance of wage and factor input ratios.

5.2.2 Other Parameters

Equipped with the substitution parameter, I impose the following normalization assumptions to identify $k-$type specific labor augmenting TFP γ_{kjt} as well as the hicks-neutral TFP term θ_{jt}

Assumption 4. Normalize the labor-augmenting productivity terms for a given firm j and assume firms have no market power on the product market.

4a) Normalize the labor augmenting TFP at each firm such that:

$$\sum_{k \in C_{jt}} \gamma_{kjt} = 1$$

4b) Output market is perfectly competitive $\Rightarrow \lambda_{jt} = P_{jt}$

Proposition 3. Under Assumption 4 the following holds (Chan et al., 2024)

1. Given an estimate of ρ firm-specific labor augmenting TFP is identified from the cross-section by:

$$\gamma_{kjt} = \frac{1}{\sum_{h \in C_{jt}} A_{khjt}}$$

   Where

$$A_{khjt} = \frac{\tilde{w}_{kjt}^{-1} \ell_{kjt}^{\rho-1}}{\tilde{w}_{hjt}^{-1} \ell_{hjt}^{\rho-1}}$$

2. To back out the returns to scale parameter, invert the firms' FOC (18) to obtain:

$$\alpha_{jt} = \frac{\rho}{R_{jt}} \times \sum_{k \in C_{jt}} \frac{\tilde{w}_{kjt}}{\rho \ell_{kjt}^{\rho-1}} \times \sum_{k \in C_{jt}} \gamma_{kjt} \ell_{kjt}^{\rho}$$

3. Finally Hicks-neutral TFP identified up to scale

$$P_{jt}\theta_{jt} = \frac{R_{jt}}{\left[\sum_{k \in C_{jt}} \gamma_{kjt} \ell_{kjt}^{\rho}\right]^{\frac{\alpha_{jt}}{\rho}}}$$

5.3 Estimation Procedure

1. For each k: Estimate labor supply parameters by running 2SLS on equation (22). Use as instruments one-period changes in log revenue *at the enterprise level*, from firms' balance sheet data in the NALMF, and one-period changes in log neighborhood-type-firm shares s_{zkjt}. [24]

2. For each (k, g) tuple: Estimate the gravity equation (2) using PPML with firm and neighborhood by industry fixed effects. Given the estimated $\tilde{\eta}_{kg}$ and the estimated β_k, λ_{kg} recover η_k. [25]

3. From equation (2), given values of β_k, λ_{kg} , and firm fixed effects, the firm specific amenities $\ln u_{kjt}$ could be recovered from the following equation

$$\ln u_{kjt} = \lambda_{kg} \underbrace{\Gamma_{kjgt}}_{\textit{Firm FEs}} - \beta_k \ln w_{kjt}$$

4. Given estimated parameters and firm fixed effects, recover $\ln u_{zkgt}$ from neighborhood specific employment log odds using the following equation

$$\ln\left(\frac{s_{zkgt}}{s_{zk0t}}\right) = \ln u_{zkgt} + \lambda_{kg} \ln\left(\sum_{j\in} \exp\left\{\frac{\beta_k \ln w_{kjt} + \ln u_{kjt}}{\lambda_{kg}}\right\}\right)$$

5. I use establishment specific shares s_{zkjt}, residential distribution, π_{zkct}, the estimated labor supply preference paramters $\{\beta_k, \lambda_{kg}\}$, and the calibrated residential choice parameters $\{\varphi, \sigma\}$ to construct $\mathcal{E}_{zkjt}$, ζ_{zkjt} and the appropritate weights to construct $\mathcal{E}_{kjt}$ as presented in equation (17).

6. Given the constructed $\mathcal{E}_{kjt}$, I recover marginal revenue product of labor $\tilde{w}_{kjt}$, and estimate the elasticity of substitution parameter ρ, and persistence parameter δ as outlined by estimation equation (27). Given the estimated ρ apply proposition (3) to recover $\{\gamma_{kjt}, \alpha_{jt}, \tilde{\theta}_{jt}\}$

7. Construct RMA_{zkt} as outlined in Equation (10) for each neighborhood-type pair (z, k). Given these values, use residential choice probabilities π_{zkct}, and neighborhood-specific housing rental indices to back out neighborhood-type-specific amenities $\ln a_{zkt}$ using choice probabilities. The amenities are normalized with respect to the neighborhood with the lowest population across all 50 CMAs

[24] I run (22) with dummy variables for each industry to obtain industry specific λ_{kg}

[25] In practice, I recover multiple η_k since PPML is run at the industry level. I take the average value after multiplying and dividing by $\eta_k g$ and β_k, respectively.

5.4 Estimation Results

5.4.1 Labor Supply & Demand Primitives

Estimates of group k specific labor supply parameters are presented in Table 3. As discussed in section 4.2.1, a lower β_k and higher λ_{kg} contribute to higher horizontal differentiation across employers, leading to more monopsony power and therefore lower markdowns. Table 3 shows that younger workers have higher marginal utility of earnings, and therefore view employers as relatively more substitutable than older workers. The within-nest correlation parameter λ_{kg} is estimated to be around 0.3, and relatively stable across groups. On the other hand, the estimate of η_k shows significant heterogeneity across age and gender. Estimated values indicate that older male workers require an extra 2.2% increase in annual earnings to commute a further 1 km to work. Older female workers require a 2.6% increase in annual earnings, exhibiting a gender gap in commuting disutility of almost 20%. This is not true however, for younger workers, where independent of gender, compensation required for an additional one kilometer of commuting is around 0.9% of annual earnings. The estimated gender gap is close to that reported by Le Barbanchon et al. (2021), however workers in my sample exhibit much higher commuting disutility. This is expected given that these are full-time worker, whereas the sample in Le Barbanchon et al. (2021) consisted of unemployed workers only. Further, the estimated values of η_k are close to the values estimated by Ahlfeldt et al. (2025) who estimate origin-destination level gravity equations using commuting flows and travel time on a sample of full-time workers in Denmark. My estimates are close to theirs, and exhibit the same life-cycle pattern. Assuming average annual earnings of $84,000$ for this group, this amounts to approximately \$1848 of compensation per year for an increase in distance to work by one kilometer. Assuming a work-year of 250 days, this corresponds to an increase of \$7.40 per day. Further, assuming a one kilometer increase in travel distance corresponds to an increase in travel time by 3.4 minutes each way for a total of 6.8 minutes, older male workers require an increase in compensation corresponding to 1.55 times their hourly wage. For older female workers, with an elasticity of 2.7% their compensating differential for an extra 1 km of travel is 1.83 times the hourly wage. Together, these results indicate that younger workers are more wage sensitive and are less sensitive to commuting costs. On the other hand, older workers view employers as more differentiated, are less sensitive to wage changes, and put more weight on commuting distance when choosing between jobs. Furthermore, among older workers, there is a 20% gender gap in commuting disutility. This demonstrates that a primary advantage of including explicit commuting costs in worker preferences is identifying the role of commuting to work in explaining differences in job choice relative to other job attributes.

Table 5 shows the estimated labor demand parameters. Panel A shows the distribution of the firm-specific production function parameters $\left(\theta_{jt}, \gamma_{kjt}, \alpha_{jt}\right)$. The average value of the returns to scale parameter across all firms is around 0.29, exhibiting high decreasing returns and significant heterogeneity across firms; with a 90-10 ratio of 8.14. The labor augmenting labor productivity γ_{kjt} is a within-firm relative productivity measure and should be compared across groups within a given firm. Finally, the average $\ln\left(\tilde{\theta}_{jt}\right)$ is approximately 14.9 and varies significantly across firms. These values are close to firm production function estimation by Tino (2024), who applies a similar production functions estimation strategy for a set of Canadian firms using the CEEDD. Finally, estimates of elasticity of substitution in Panel B of Table 5 at 0.91 exhibits high levels of substitution across worker types. This result is consistent with estimates from Chan et al. (2024) and Volpe (2024), who, although allow for heterogeneity across worker types and firms, respectively, find an average EoS close to 0.9 with little variation across types and\or firms. Finally, estimates of δ at 0.73 validate the high level of persistence in productivity shocks in Assumption 1.

5.4.2 Elasticities & Markdowns

Table 4 provides the employment-weighted average of firm labor supply elasticity for all workers, and by group k. Panel A shows that the average employment-weighted labor supply elasticity is around 7.5. This value is slightly higher than estimated values in the literature, within the range of 3-7. To better understand why this is the case, note that workers living further away from the firm are more elastic. This can be seen from the neighborhood-firm specific labor supply elasticity in equation (9), where the shares are decreasing with distance to employers. Therefore, if the marginal worker lives farther away from the firm, labor supply elasticity at the k-level will be higher. This reflects a key mechanism of the model; proximity to the workforce affects the slope and intercept of firms' labor supply curves. To quantify the role of commuting in shaping firm labor supply elasticities, I simulate the model under a counterfactual scenario of no commuting costs, showing that commuting explains 5% of the reported employment-weighted elasticity. The corresponding aggregate markdowns are 84.7%, indicating that workers receive on average 85% of their marginal revenue product of labor. These averages hide significant heterogeneity across worker groups and firms. Panels B-D show that younger workers are more elastic and earn on average 91% of their marginal revenue product of labor, compared to only 79.3% for older workers. Note that although there is a gender gap in commuting costs for older workers, this does not necessarily show up in the average markdowns. To see why, note that there are two forces at play. One, more productive firms offer lower markdowns, since, independent of a worker's neighborhood, they have higher labor market shares at any level of geographical aggregation. Two small firms with low aggregate shares are more likely to attract workers from nearby neighborhoods, and therefore also offer low markdowns. That is, what shapes a firm's markdown is the *average shares* from its workforce neighborhoods. To better see this, note that under a fixed spatial distribution of workers, the firm-specific labor supply elasticity could be rewritten as:

$$\mathcal{E}_{kj} = \beta_k \left[\frac{1}{\lambda_{kg}} + \left(1 - \frac{1}{\lambda_{kg}}\right) \tilde{s}_{zkj|g} - \tilde{s}_{zkj} \right] =: \mathbb{E}[\mathcal{E}_{zkj}|j,k]$$

where

$$\tilde{x} = \sum_z x\mathbb{P}(z|j,k)$$

This equation nests the case where each firm is restricted to hire from a given CMA. Demonstrating that explicitly modelling spatial linkages between locations renders the proposed framework amenable at any aggregation level. Another implication of this equation is that within-firm variation in where workers live affects firms' labor supply elasticity, highlighting that workforce composition can shape wage markdowns. This result resembles the findings in Volpe (2024) who show that unobserved variation across workers in their relative valuation of wage and non-wage amenities allows even low productivity firms to have low wage markdowns, as they attract less wage-sensitive workers. The framework proposed here assumes workers differ in their marginal rate of substitution between wage and non-wage amenities due to idiosyncratic observable differences in commuting costs.

5.4.3 Commuting & Firm Productivity

If workers trade off wages against commuting time in making job choices, it should be the case that in travelling farther away, workers match with more productive firms. Table 6 presents results from regressing log firm TFP, $\ln(\tilde{\theta}_{jt})$ on workers' commuting distance.

Note that the recovered $\ln(\tilde{\theta}_{jt})$ are independent of the estimated η_k, since only β_k and λ_{kg} are needed to recover the firm-specific productivity parameters. Permitting this exercise to serve as a qualitative validation test for the estimated willingness to pay parameters. Estimates in column (1) show that an increase in commuting distance by 1 km is associated with an increase in firm TFP by 1%. With an average commuting distance of 11 km, this corresponds to an elasticity of 0.1. Within CMA estimates in column (3) are similar but slightly lower. Columns (2) and (4) of Table 6 show that the firm TFP commuting premium is higher for women, validating the estimated willingness to pay for shorter commute estimates η_k. This is because workers with higher disutility from commuting, in this case, women, require higher compensation to travel farther. An explanation for this is that female workers are more spatially mismatched. That is, they live farther away from high-paying jobs. This is possible if married couples bargain over residential location, particularly when female bargaining power is lower. These results are in line with Card et al. (2024), who use AKM fixed effects to show that longer commutes are associated with higher firm premiums. Note that with variable firm markdowns, firm premia capture both firm-specific markdowns and firm productivity.

6 Counterfactuals

I simulate the model using an underrelaxed Jacobi iteration algorithm to solve for equilibrium wages, rental prices, and allocation to firms and neighborhoods using recovered housing supply, firm production parameters, and given population for 2016 only. Refer to Appendix C for the detailed simulation algorithm. Note that the model does not fit the data exactly by construction. This is due to the separability assumption in equation (23). The simulated model matches the data well, recovering a variance of log wages of 0.21, which is close to that of the data of 0.22. The simulated commuting distribution also matches the 2016 commuting distribution closely.

6.1 SkyTrain Simulation

To better understand the underlying mechanisms behind the SkyTrain expansion effect on earnings with respect to changes in worker-firm sorting, labor supply elasticities, and markdowns, I simulate the model for the *entire* economy in 2016 under a range of changes to commuting costs to emulate the absence of the Canada Line. Note that commuting costs in the model between neighborhood z and firm j enter parametrically as $\kappa_{z,j} := e^{\frac{-\eta_k \beta_k d_{z,o(j)}}{\lambda_{kg}}}$. Specifically, I simulate the model under the following commuting costs parameterization

$$\kappa_{z,j} = \begin{cases} Ae^{\frac{-\eta_k \beta_k d_{z,j}}{\lambda_{kg}}} & \text{if } z \in CanadaLine \text{ and } o(j) \in SkyTrain \setminus CanadaLine \\ e^{\frac{-\eta_k \beta_k d_{z,j}}{\lambda_{kg}}} & \text{otherwise} \end{cases}$$

That is, I reduce commuting costs by a factor A between neighborhoods that gained access to the SkyTrain network and any other neighborhood that was on the SkyTrain prior to 2009. This corresponds to the dark blue and gray areas in Figure 2, respectively. Note that areas that did not gain access to the subway network could still be affected due to migration, employment responses, or reallocation between workers and firms across space. These correspond to the areas in yellow. I run simulations for multiple values of A to simulate the economy in the absence of the Canada Line and compare it to the simulated model in 2016. Importantly, I simulate the model for the entire economy, not the Vancouver area only. This allows for migration across CMAs as a result of the Vancouver area offering a better balance between housing costs, firm access, and commuting costs.

I report the percentage change in outcomes between the baseline economy, with the Canada Line, and the counterfactual economy, which takes as given the spatial distribution of firms in 2016, and assumes the Canada Line is removed. Outcomes are reported separately for three regions. First, neighborhoods that never had access to the SkyTrain network, corresponding to the yellow regions in Figure 2; Second, regions that have access to the Millenium or Expo lines, corresponding to the gray area in Figure 2; Finally, neighborhoods that gained access to the SkyTrain network due to the Canada Line, corresponding to the dark blue areas, and mainly consists of the suburb of Richmond.

Table 7 reports the percentage change in labor market outcomes, population and market concentration indices for $A = 0.7$ corresponding to a 30% reduction in commuting costs between neighborhoods on the Canada Line and any other neighborhood where the Millennium and Expo Lines are located. First, average wages for workers on the Canada Line are 1% higher in the presence of the Canada Line. At the same time, workers in other neighborhoods experienced a drop in earnings by 0.2 % for workers not on the SkyTrain network, and 0.3% for workers on the Expo & Millennium lines. Across all neighborhood types, however, there is little change in markdowns. Instead, wage effects seem to be driven by the resorting of workers on the Canada Line to more productive firms. To see that, note that under the Canada Line, workers in affected neighborhoods experienced an increase of 5% in average employer TFP, as captured by $\tilde{\theta}_{jt}$. However, firm productivity is not captured by $\tilde{\theta}_{jt}$ only, as firms differ in their returns to scale as well as their labor augmenting productivity. To assess how worker-firm sorting changes in presence of the Canada Line given the rich firm heterogeneity, I follow Chan et al. (2024) to assess how the composite term given by $\tilde{\theta}_{jt}\alpha_{jt}\sum_{k\in\mathcal{K}}\left(\gamma_{kjt}\ell_{kjt}^{\rho}\right)^{\frac{\alpha_{jt}}{\rho}-1}$ changes for each neighborhood type. Table 7 shows that workers on the Canada Line experienced an increase in this composite term by 1.7 %, indicating they resorted to more productive firms. On the other hand, workers at other neighborhoods resorted to less productive firms with an average drop in the composite term by 0.34 %. This indicates that workers on the Canada Line resorted to more productive firms while crowding out workers from other areas. This shows that the transit projects could have negative spillovers on workers in faraway neighborhoods due to increased firm access to other areas. This result is consistent with the findings of Manning and Petrongolo (2017) who show that improved transportation between an area of high and low unemployment can affect locals in high employment areas due to the stronger competition for jobs from new applicants. I further extend their findings by showing that competition between workers across neighborhoods also occurs on the intensive margin of firm productivity, as well as employment. One way to rationalize the change in worker firm sorting with little change in markdowns is to note that markdowns are at the firm-type level, not the neighborhood level. It is therefore possible that neighborhoods affected by a transportation shock reallocate to more productive firms that have lower markdowns for *all* workers.

Further, neighborhoods on the Canada Line experienced a population increase of 3.75% , whereas other neighborhoods saw a drop in their population. Note that the baseline population on the Canada Line is lower than in other areas, indicating that workers changed their place of residence to benefit from the increased accessibility. Moreover, Canada Line neighborhoods experienced an increase in employment by 8%, while other neighborhoods saw a slight drop in employment. This is in line with workers on the Canada Line crowding out workers from other areas, not only on the intensive margin of firm types, but also on the extensive margin of employment. Consistent with the increase in population, employment, and wages, and given the fixed housing supply assumption, average rents increased on the Canada Line neighborhoods by 8.5%, whereas rents dropped by 0.7% and 0.5% at neighborhoods not on the SkyTrain and Old station areas, respectively.

Turning to welfare and labor market concentration, a key insight from this simulation is the fall in concentration for neighborhoods on the Canada Line by 8.6%. There are two drivers of this result. First is the drop in the share of non-employed for workers in these neighborhoods. As presented in Equation 19, all else constant,if the share of workers in some group $g \in \mathcal{G}$, including nonemployment drops, the concentration index falls. This is assuming the newly employed workers have the same probabilities of sorting to firms. The second driver is the reallocation of workers across firms, presented by the increase in the employment-weighted average of productivity of employers. The drop in labor market concentration in neighborhoods on the Canada Line was accompanied by a modest increase in concentration in other neighborhoods. Areas not on the SkyTrain network, and neighborhoods on Expo & Millennium lines experienced an increase in concentration index by 0.5% and 0.08% respectively. Although the spillovers in concentration across neighborhoods are modest, the results show that reducing labor market concentration in a set of neighborhoods can be at the detriment of crowding out and increasing concentration in other areas. It is important to note, however, that there is no directional relationship between the generalized concentration index and wages; instead, both are determined in equilibrium by the allocation of workers to firms. Equation 20 indicates an inverse relationship between residential market access RMA_{zkt} and concentration $\ln GCI_{zkt}$ at the neighborhood level. Table 7 exhibits this relationship by indicating an increase in market access for neighborhoods on the Canada Line by 3%. While neighborhoods not on the SkyTrain network, and those on the Expo & Millennium lines experienced a drop in market access by 0.2 % and 0.13% respectively. The passthrough from concentration index to market access is not one-to-one due to changes in equilibrium shares s_{zkjt} and systematic utility v_{zkjt} , due to changes in wages. The increase in population, employment, wages, and the fall in concentration index all contribute to the increased housing rents experienced by neighborhoods on the Canada Line. It is not clear, however, which factor is the primary contributor.

Overall, the results show that the Canada Line benefited workers in neighborhoods who gained access to the subway network. This was in tandem with negative labor market consequences for workers from other neighborhoods who were crowded out of high-productivity jobs, and culminating in tougher competition for jobs on the intensive and extensive margins.

6.2 Heterogeneous Firm Access & Spatial Variation In Markdowns

I use the model to examine the role of the geography of firms in shaping wage markdowns across space. To do so, I simulate the model under the counterfactual scenario of uniform distance between workers and firms, regardless of residential location. More precisely, I set commuting distance to 20 km for any worker-firm pair that is within a 75 km radius of a given neighborhood. This is not necessarily a welfare-improving shock since workers lose the option to work close to home; however, within a given area, it removes differential access to firms across neighborhoods.

Table 8 presents the change in variance of mean wages, markdowns, and marginal product across space for the four different groups of workers. Results show that within a given CMA, the geography of firms explains a sizable share of the spatial variation in mean wages across neighborhoods. Focusing on male workers between the ages of 40 and 60, the firm geography of firms explains 16.8% of the variation in wages across neighborhoods. This rises to 42% within CMAs. The role of firm geography is more important in explaining earnings variation across space for female workers in the same age group. This is due to their higher disutility from commuting; as such, firm geography plays a more important role in shaping their earnings distribution. The effect on wages is primarily mediated through a reduction in the variance of the marginal product of labor. This suggests that worker-firm geography plays a significant role in shaping labor supply. Heterogeneous commuting costs explain around 40% of the variation in earnings across neighborhoods within a CMA. Although the firm's productivity is substantial in explaining the wage effect, the variance in markdowns across space also shrinks when commuting costs are equalized. The variance in markdowns drops by 16 % on average across neighborhoods. The effects within and across CMA are not significantly different. Both the effects on markdowns and marginal product are mediated through changes in the allocation of workers and firms. These results indicate that heterogeneous commuting costs account for a significant portion of the variance in markdowns across space, underscoring the importance of equalizing firm access to mitigate spatial inequality in labor market power.

7 Conclusion

Commuting costs are a classical source of market power. Alleviating workers' commuting burden increases the effective number of potential employers and therefore improves workers' outside options. I leverage a subway expansion in Vancouver to show that workers who benefited from lower commuting costs experienced an increase in earnings relative to those who did not gain access. Empirical evidence, although suggestive, cannot rule out changes in workers' and firms' productivity due to the shock, and is silent on the general equilibrium effects of the shock across space. To investigate the general equilibrium effects of the subway expansion, I build, identify, and estimate a two-sided labor market matching model that features strategic interactions in wage setting, commuting costs, and residential choice. The model extends the framework in Chan et al. (2024) to assess the role of commuting in shaping equilibrium outcomes, and extends the framework in Caldwell and Danieli (2024) to investigate worker-to-firm sorting in the presence of wage-setting behavior.

Simulation results show that workers in areas close to the subway expansion experienced an increase in earnings due to sorting into higher-productivity firms and faced lower labor market concentration. This came at the expense of workers from other areas, who faced higher competition for high-paying jobs and a slight increase in labor market concentration. To provide further evidence on the role of differential access to jobs in shaping the sorting of workers to firms and influencing wage markdowns, I demonstrate that heterogeneous commuting costs can explain a significant portion of the spatial variation in wage markdowns and the marginal product of labor. Underscoring the role of horizontal commuting costs in shaping labor market power via sorting in equilibrium.

The proposed framework can be used to study the labor market consequences of any change in commuting costs, regardless of its source. For example, it could be used to better understand how work-from-home arrangements affect the equilibrium allocation of workers to firms and the resulting wage-setting power (De Fraja et al., 2022). More generally, the proposed model can be used to study the effect of any spatially heterogeneous shock on labor market power while allowing for flexible commuting patterns (Monte et al., 2018).

References

Ahlfeldt, Gabriel M, Ismir Mulalic, Caterina Soto Vieira, and Daniel M Sturm, "The Geography of Life: Evidence from Copenhagen," 2025.

_ , Stephen J Redding, Daniel M Sturm, and Nikolaus Wolf, "The economics of density: Evidence from the Berlin Wall," *Econometrica*, 2015, *83* (6), 2127–2189.

Amior, Michael and Alan Manning, "The persistence of local joblessness," *American Economic Review*, 2018, *108* (7), 1942–1970.

Azar, José A, Steven T Berry, and Ioana Marinescu, "Estimating labor market power," Technical Report, National Bureau of Economic Research 2022.

Azar, José, Ioana Marinescu, Marshall Steinbaum, and Bledi Taska, "Concentration in US labor markets: Evidence from online vacancy data," *Labour Economics*, 2020, *66*, 101886.

Bamford, Iain, "Monopsony power, spatial equilibrium, and minimum wages," *Unpublished paper*, 2021.

Barbanchon, Thomas Le, Roland Rathelot, and Alexandra Roulet, "Gender differences in job search: Trading off commute against wage," *The Quarterly Journal of Economics*, 2021, *136* (1), 381–426.

Baum-Snow, Nathaniel and Lu Han, "The microgeography of housing supply," *Journal of Political Economy*, 2024, *132* (6), 1897–1946.

Berger, David, Kyle Herkenhoff, and Simon Mongey, "Labor market power," *American Economic Review*, 2022, *112* (4), 1147–1193.

Berry, Steven T, "Estimating discrete-choice models of product differentiation," *The RAND Journal of Economics*, 1994, pp. 242–262.

Bils, Mark, Bariş Kaymak, and Kai-Jie Wu, "Robinson Meets Roy: Monopsony Power and Comparative Advantage," Technical Report, National Bureau of Economic Research 2025.

Bound, John and Harry J Holzer, "Demand shifts, population adjustments, and labor market outcomes during the 1980s," *Journal of labor Economics*, 2000, *18* (1), 20–54.

Bütikofer, Aline, Katrine V Løken, and Alexander Willén, "Building bridges and widening gaps," *Review of Economics and Statistics*, 2024, *106* (3), 681–697.

Caldwell, Sydnee and Oren Danieli, "Outside Options in the Labour Market," *Review of Economic Studies*, 2024, *91* (6), 3286–3315.

Callaway, Brantly, Andrew Goodman-Bacon, and Pedro HC Sant'Anna, "Difference-in-differences with a continuous treatment," Technical Report, National Bureau of Economic Research 2024.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline, "Firms and labor market inequality: Evidence and some theory," *Journal of Labor Economics*, 2018, *36* (S1), S13–S70.

_ , Jesse Rothstein, and Moises Yi, "Reassessing the Spatial Mismatch Hypothesis," in "AEA Papers and Proceedings," Vol. 114 American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203 2024, pp. 221–225.

Chan, Mons, Kory Kroft, Elena Mattana, and Ismael Mourifié, "An empirical framework for matching with imperfect competition," Technical Report, National Bureau of Economic Research 2024.

Chapelle, Guillaume and Morgan Ubeda, "Competing for opportunity: Transport infrastructures and localized unemployment," Technical Report, HAL 2025.

Chernoff, Alex and Andrea N Craig, "Distributional and housing price effects from public transit investment: Evidence from Vancouver," *International Economic Review*, 2022, *63* (1), 475–509.

Correia, Sergio, Paulo Guimarães, and Tom Zylkin, "Fast Poisson estimation with high-dimensional fixed effects," *The Stata Journal*, 2020, *20* (1), 95–115.

Couture, Victor, Cecile Gaubert, Jessie Handbury, and Erik Hurst, "Income growth and the distributional effects of urban spatial sorting," *Review of Economic Studies*, 2024, *91* (2), 858–898.

Datta, Nikhil, "The measure of monopsony: the labour supply elasticity to the firm and its constituents," 2023.

Dingel, Jonathan I and Felix Tintelnot, "Spatial economics for granular settings," Technical Report, National Bureau of Economic Research 2020.

Duan, Yige, Oskar Jost, and Ramona Jost, "Beyond lost earnings: the long-term impact of job displacement on workers' commuting behavior," 2022.

Eaton, Jonathan, Samuel S Kortum, and Sebastian Sotelo, "International trade: Linking micro and macro," Technical Report, National bureau of economic research 2012.

Fernández-Val, Iván and Martin Weidner, "Individual and time effects in nonlinear panel models with large N, T," *Journal of Econometrics*, 2016, *192* (1), 291–312.

Fraja, Gianni De, Jesse Matheson, Paul Mizen, James Rockey, Shivani Taneja, and Gregory Thwaites, "Remote Work and Compensation Inequality," *Available at SSRN 4962603*, 2022.

Heblich, Stephan, Stephen J Redding, and Daniel M Sturm, "The making of the modern metropolis: evidence from London," *The Quarterly Journal of Economics*, 2020, *135* (4), 2059–2133.

Heise, Sebastian and Tommaso Porzio, "Labor misallocation across firms and regions," Technical Report, National Bureau of Economic Research 2022.

Helsley, Robert W and William C Strange, "Matching and agglomeration economies in a system of cities," *Regional Science and urban economics*, 1990, *20* (2), 189–212.

Herzog, Ian, "Does Demand for Public Transit Reflect Sorting or Substitution?," https://www.dropbox.com/scl/fi/rr66wiq9vhq2j5dmnq24u/Skytrain_and_Mode_Choice_March2025.pdf March 2025. Accessed: 2025-08-12.

Hong, Guangbin, "Two-sided sorting of workers and firms: Implications for spatial inequality and welfare," in "in" CLEF, Canadian Labour Economics Forum 2024.

Hotelling, Harold, "Stability in Competition," *The Economic Journal*, 1929, *39* (153), 41–57.

Humlum, Anders, Mette Rasmussen, and Evan K Rose, "Firm Premia and Match Effects in Pay vs. Amenities," Technical Report, National Bureau of Economic Research 2025.

Kahn, Matthew E and Joseph Tracy, "Monopsony in spatial equilibrium," *Regional Science and Urban Economics*, 2024, *104*, 103956.

Kline, Patrick M, "Labor market monopsony: Fundamentals and frontiers," 2025.

Kroft, Kory, Yao Luo, Magne Mogstad, and Bradley Setzler, "Imperfect competition and rents in labor and product markets: The case of the construction industry," *American Economic Review*, 2025, *115* (9), 2926–2969.

Krueger, Alan B, "Where have all the workers gone? An inquiry into the decline of the US labor force participation rate," *Brookings papers on economic activity*, 2017, *2017* (2), 1.

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler, "Imperfect competition, compensating differentials, and rent sharing in the US labor market," *American Economic Review*, 2022, *112* (1), 169–212.

Li, Jiang, Benoit Dostie, and Gaëlle Simard-Duplain, "Firm pay policies and the gender earnings gap: the mediating role of marital and family status," *ILR Review*, 2023, *76* (1), 160–188.

Lindenlaub, Ilse, Ryungha Oh, and Michael Peters, "Spatial Firm Sorting and Local Monopsony Power," 2024.

Manning, Alan, "The real thin theory: monopsony in modern labour markets," *Labour economics*, 2003, *10* (2), 105–131.

_ **and Barbara Petrongolo**, "How local are labor markets? Evidence from a spatial job search model," *American Economic Review*, 2017, *107* (10), 2877–2907.

Mas, Alexandre, "Non-wage amenities. NBER Working Paper 33643," 2025.

McFadden, Daniel et al., "Modelling the choice of residential location," 1978.

Monte, Ferdinando, Stephen J Redding, and Esteban Rossi-Hansberg, "Commuting, migration, and local employment elasticities," *American Economic Review*, 2018, *108* (12), 3855–3890.

Moretti, Enrico and Moises Yi, "Size matters: Matching externalities and the advantages of large labor markets," Technical Report, National Bureau of Economic Research 2024.

Notowidigdo, Matthew J, "The incidence of local labor demand shocks," Technical Report, National Bureau of Economic Research 2011.

Pérez, Jorge, Felipe Vial, and Román Zárate, "Urban transit infrastructure: Spatial mismatch and labor market power," 2022.

Phibbs, Ciaran S and Harold S Luft, "Correlation of travel time on roads versus straight line distance," *Medical care research and review*, 1995, *52* (4), 532–542.

Redding, Stephen J, "Quantitative urban economics," 2024.

Robinson, Joan, *The economics of imperfect competition*, Springer, 1933.

Salop, Steven C, "Monopolistic competition with outside goods," *The Bell Journal of Economics*, 1979, pp. 141–156.

Schubert, Gregor, Anna Stansbury, and Bledi Taska, "Employer concentration and outside options," *Available at SSRN 3599454*, 2024.

Silva, JMC Santos and Silvana Tenreyro, "The log of gravity," *The Review of Economics and statistics*, 2006, pp. 641–658.

Sroka, Robert, "Mega-events and rapid transit: evaluating the Canada line 10 years after vancouver 2010," *Public Works Management & Policy*, 2021, *26* (3), 220–238.

Staiger, Douglas O, Joanne Spetz, and Ciaran S Phibbs, "Is there monopsony in the labor market? Evidence from a natural experiment," *Journal of Labor Economics*, 2010, *28* (2), 211–236.

Statistics Canada, "2006 Census of Population," 2006. Catalogue no. 97-550-XWE, Ottawa.

Sun, Liyang and Jesse M Shapiro, "A linear panel model with heterogeneous coefficients and variation in exposure," *Journal of Economic Perspectives*, 2022, *36* (4), 193–204.

Tino, Stephen, "Labor Market Power, Firm Productivity, and the Immigrant-Native Pay Gap," 2024.

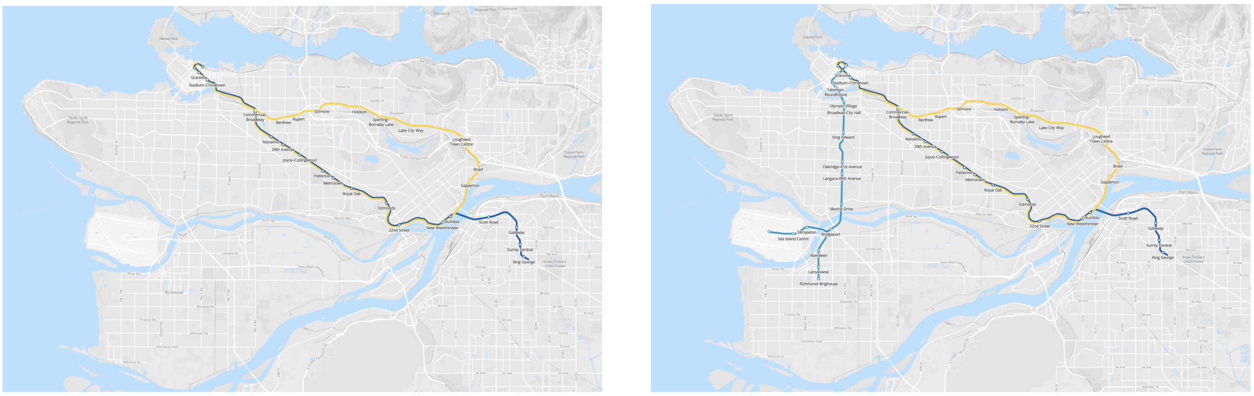
Tsivanidis, Nick, "Evaluating the impact of urban transit infrastructure: Evidence from bogota's transmilenio," *Unpublished manuscript*, 2022, *18.*

Volpe, Oscar, "Job Preferences, Labor Market Power, and Inequality," Technical Report, Discussion paper, Working Paper 2024.

Zárate, Román D, "Spatial misallocation, informality, and transit improvements," *Development Research*, 2022.

Figures

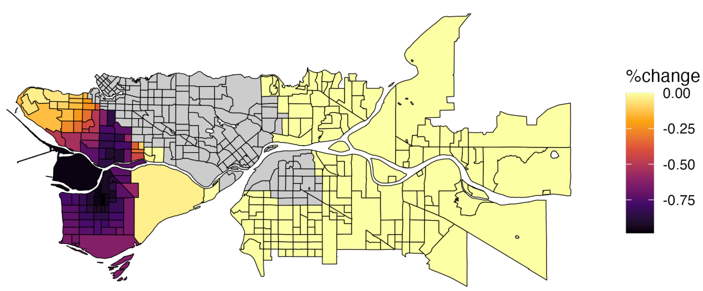
Figure 1: Vancouver SkyTrain Maps Pre and Post Canada Line expansion



(a) Pre-Expansion (b) Post-Expansion

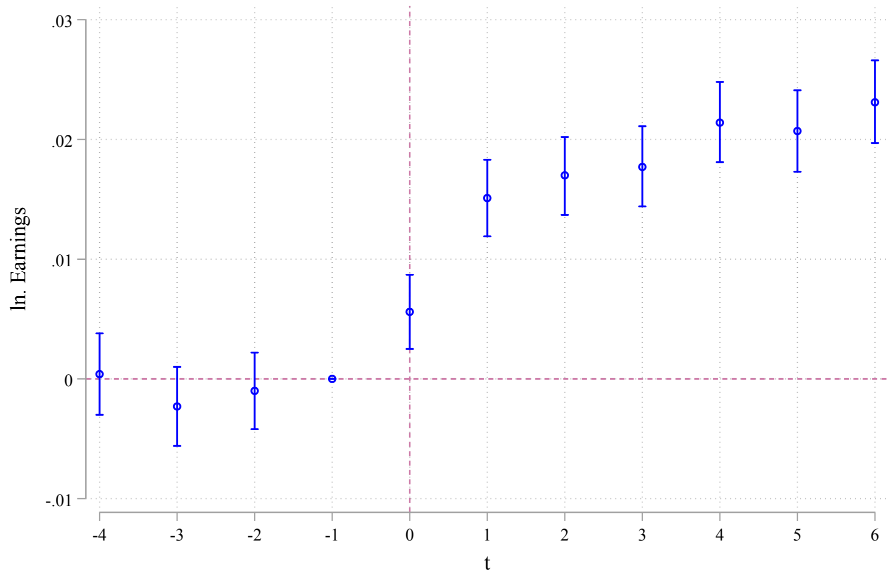
Notes: Panel A demonstrates the Vancouver SkyTrain in 2008, prior to the Canada Line expansion. At this point, the SkyTrain system includes the Expo and Millennium lines. Built in 1986 and 2002, respectively. Panel B illustrates the Canada line expansion in light blue. The new line connects the airport in the west and Richmond in the south, with downtown Vancouver across the Fraser River.

Figure 2: Richmond & Rest of Metro Vancouver Public Transit Shares



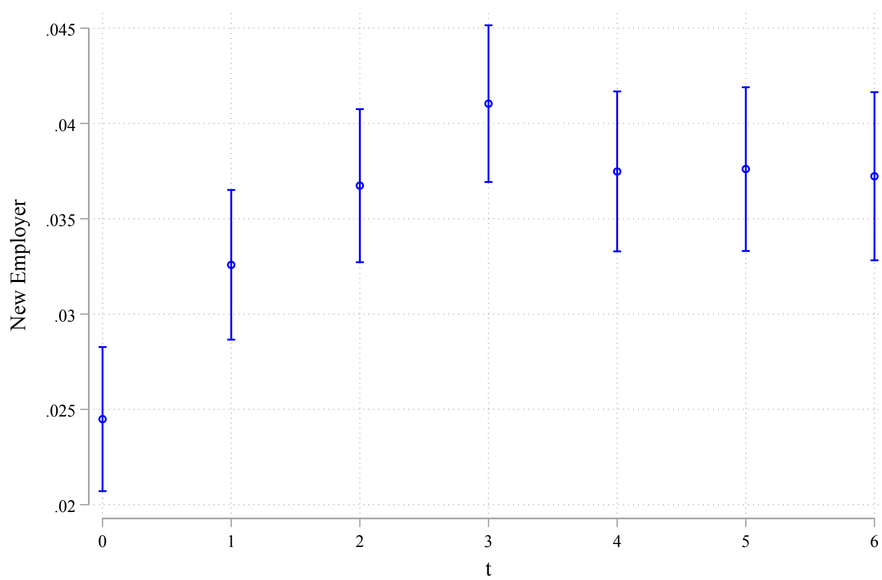
Notes: This map illustrates treatment and control areas for the Vancouver SkyTrain expansion estimation. Darker blocks reflect a higher change in distance to the nearest station during the expansion. Lighter colors reflect lower change. Neighborhoods in yellow represent the control group for the empirical strategy. Areas in gray are neither in the treatment nor control group, as they overlap with the Millennium and Expo lines.

Figure 3: Effect of SkyTrain Expansion on Earnings



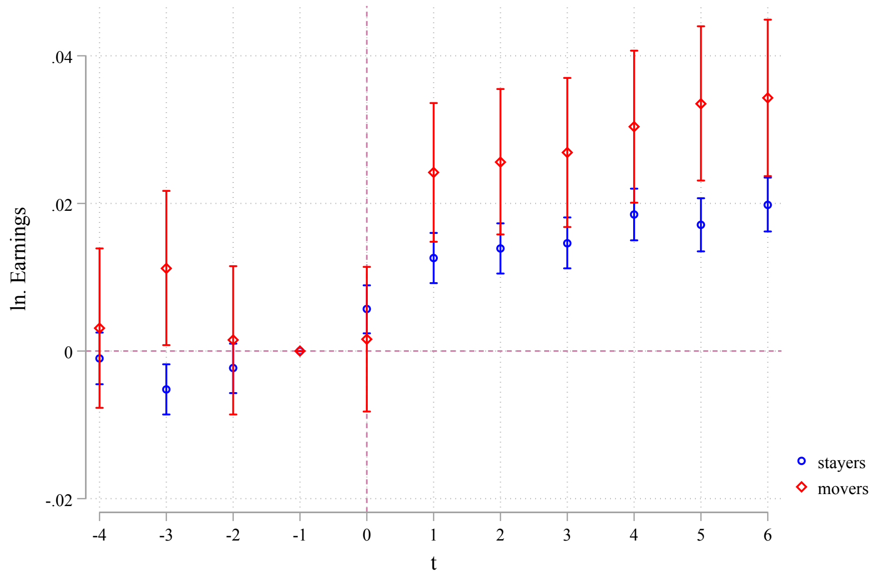
Notes: This figure plots estimates from equation (2) on an unbalanced sample of Vancouver residents. The time period is from 2005 to 2015, and coefficients are estimated relative to period $t = -1$, which corresponds to the year 2008, one year before the expansion. Specification includes individual fixed effects, and demographic group (gender × age bin) by year fixed effects.

Figure 4: Effect of SkyTrain Expansion on Probability of Switching Employers



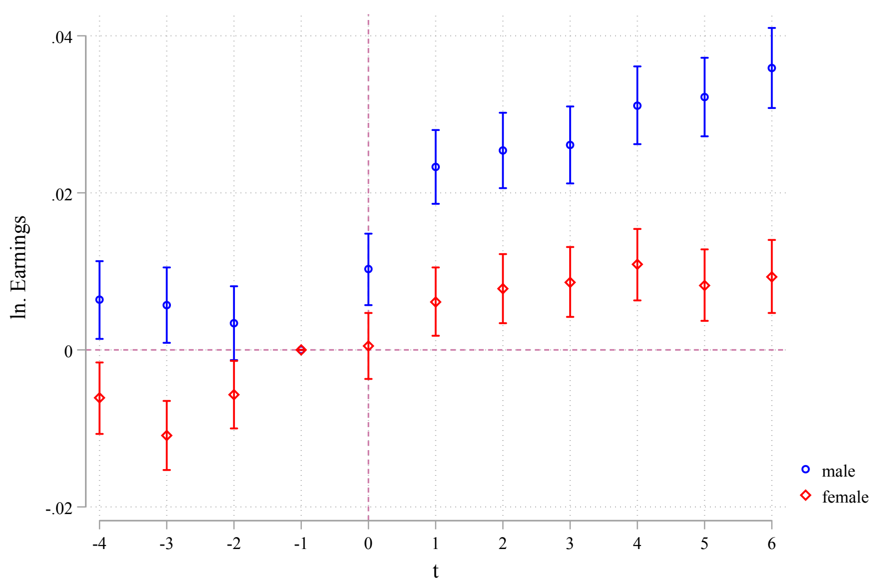
Notes: This figure plots yearly difference in probability of in switching 2008 primary employer between treatment and control groups. Specifcally, plotted coefficients are estimates from the following regression $D_{ikt} = \sum_{t=0}^{t=6} \beta_k \mathbb{I}\{k - 2009 = t\} \times T_i + \alpha_{kt} + \epsilon_{ikt}$, where D_{ikt} is a dummy indicating worker i has a new employer relative to 2008.

Figure 5: Effect of SkyTrain Expansion on Earnings By Job Switchers & Stayers



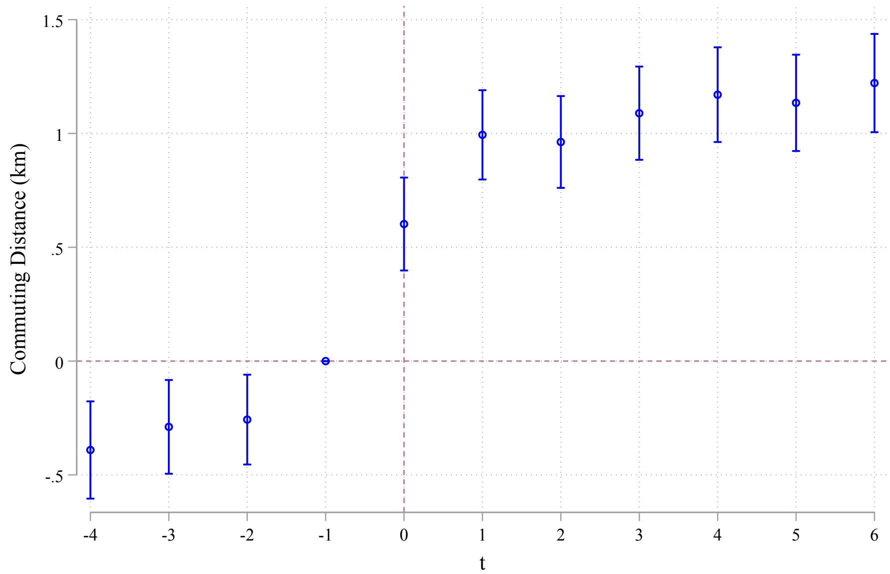
Notes: This figure plots estimates from (2) separately for job switchers and stayers. The estimation period spans from 2005 to 2015, using an unbalanced sample of Vancouver residents. The coefficients are estimated relative to period $t = -1$, which corresponds to the year 2008, one year prior to the expansion. Specification includes individual fixed effects, and demographic group (gender × age bin) by year fixed effects.

Figure 6: Effect of SkyTrain Expansion on Earnings By Gender



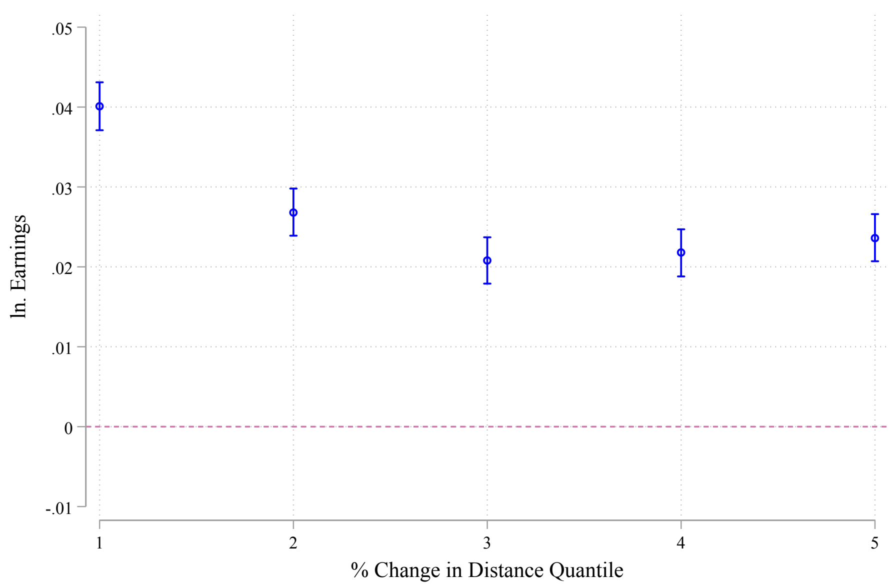
Notes: This figure plots estimates from (2) for individuals who had at least one job switch after the expansion, with the outcome variable being commuting distance, measured as the Euclidean distance between the worker's and firm's locations. The estimation period spans 2005-2015, using an unbalanced sample of Vancouver residents, and coefficients are estimated relative to period $t = -1$, one year prior to the expansion. Specification includes individual fixed effects, and demographic group (gender × age bin) by year fixed effects.

Figure 7: Effect of SkyTrain Expansion on Commuting Distance (km)



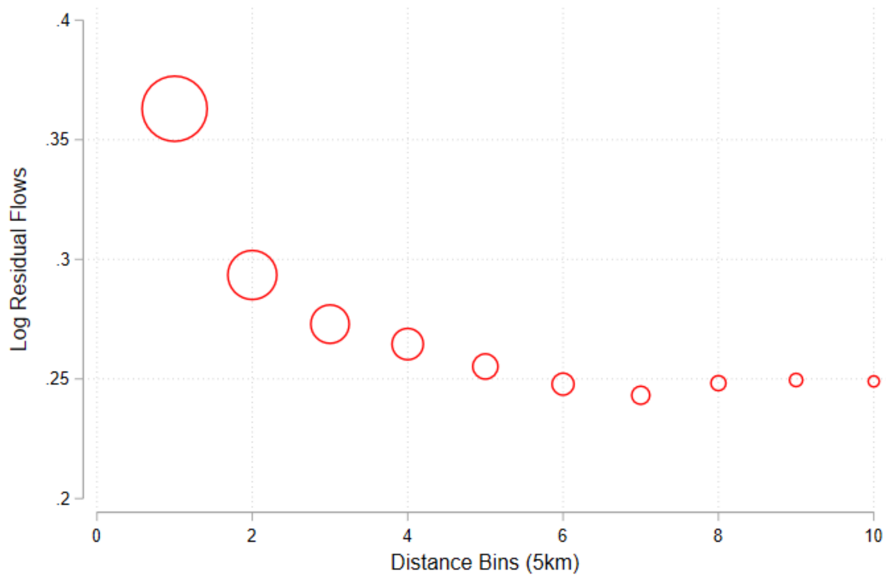
Notes: This figure plots estimates from (2), with the inclusion of firm× year fixed effects to account for potential wage gains due to changes in firm productivity. The estimation period is from 2005 to 2015, on an unbalanced sample of Vancouver residents, and the coefficients are estimated relative to period $t = -1$, which corresponds to the year 2008, one year before the expansion. Specification includes individual fixed effects, and demographic group (gender × age bin) by year fixed effects.

Figure 8: Effect of SkyTrain Expansion on Annual Earnings By Treatment Intensity



Notes: This figure presents estimates from a static Diff-in-Diff specification with multiple treatment intensities. Treatment intensity is defined by the percentage change in distance to the nearest station quantiles. The first quantile is the group that experienced the largest drop in distance (i.e, those with a station very close to the place of residence). The fifth quantile is workers who had a positive, but low decrease in distance to the nearest station. The control group is workers who had no change in distance to the nearest station, and live more than 3km away from any subway station. Specification is $Y_{it} = \alpha_i + \alpha_t + \sum_{d=1}^{5} \beta_k \mathbb{I}\{\%\Delta \in Q_d\}\text{Post}_t \times \textit{Treated}_i + \epsilon_{it}$ Where Q_d are the specified quantiles.

Figure 9: Residualized Number of Workers by Commuting Distance Bins



Notes: This figure plots the residualized number of workers at a given firm against 5 km travel distance bins. Residual flows are obtained from regressing the number of workers from neighborhood z of type k working at firm j on firm fixed effects using 2016 neighborhood-type-firm level data. I run the following $\ln n_{zkjt} = \alpha_{jt} + \epsilon_{zkjt}$. Note that firms hiring from only one neighborhood are dropped from this regression. However, zero hires are taken into account when using PPML to estimate equation (24).

Tables

Table 1: Sample Summary Statistics 2001-2019

	All	Male, 25-39	Male 40-59	Female, 25-39	Female,40-59
Panel A: Workers					
Earnings	67,000	64,000	84,000	49,000	57,000
Age	42	33	50	33	50
Share Female	0.36	-	-	-	-
Share Ontario	0.43	0.41	0.43	0.44	0.46
Share Quebec	0.25	0.25	0.26	0.24	0.25
Share B.C.	0.11	0.11	0.11	0.12	0.12
Share Manuf.	0.25	0.25	0.31	0.16	0.21
Share Hospitality	0.051	0.044	0.033	0.079	0.071
Share Finance	0.057	0.043	0.03	0.095	0.095
Share Construction	0.086	0.13	0.11	0.031	0.031
Share Other Professional.	0.20	0.21	0.16	0.28	0.21
Share commute ≤ 5	0.38	0.39	0.39	0.38	0.38
Share commute $\in (5, 10]$	0.21	0.21	0.21	0.22	0.21
Share commute $\in (10, 15]$	0.13	0.13	0.13	0.13	0.13
Share commute $\in (15, 20]$	0.083	0.082	0.085	0.082	0.082
Share commute $\in (20, 25]$	0.056	0.054	0.058	0.054	0.055
Share commute $\in (25, 30]$	0.04	0.038	0.041	0.039	0.039
Share commute > 30	0.10	0.10	0.10	0.093	0.09
Panel B: Enterprises					
PD7 Number Of Employees	60				
Share Multi establishment	0.10				
Number of k types	3.19				
Aggregate Shares s_{kj}	2×10^{-6}				
Neighborhood Specific Share s_{zkj}	0.00056				
Log. Revenue	15.13				
Log. Value Added	13.97				
Log. Total Payroll	13.75				
Distance to Actual Location	23.25				
Panel C: Neighborhoods					
Population	10,000				
Rental Price Index	1.1				
Log. Neighborhood Total Earnings	19.4				
Share Ontario	0.41				
Share Quebec	0.24				
Share British Columbia	0.13				
Average within CMA share $\pi_{zk\|c}$	0.05				
No. Individual-year observations.	81,000,000	24,000,000	28,000,000	13,000,000	16,000,000
No. Unique Individuals	15,000,000	5,000,000	4,000,000	3,000,000	3,000,000
No. Enterprise-year observations	2, 000,000				
No. Unique Enterprises	316,000				
No. Unique Neighborhoods	1,034				

Notes: This table presents summary statistics for the constructed sample used to estimate the model parameters. Panel A presents worker-level summary statistics for all workers, as well as separately by worker type. Panel B includes Descriptives at the *enterprise* level. Panel C presents descriptives for the constructed *neighborhoods*. Observation numbers are rounded to the nearest million in compliance with Statistics Canada confidentiality terms.

Table 2: Vancouver SkyTrain Expansion Summary Statistics by Treatment Status

Panel A: Pre-Expansion		
	Control	Treatment
Age	42.85	42.09
Share Female	0.47	0.49
Earnings	68,000	78,000
Log. Enterprise Total payroll	16.83	16.49
Distance to nearest station 2008	12.38	6.94
Commuting Distance	15.55	15.26
Res. Distance to CBD	32.17	10.39
Share Manuf.	0.12	0.086
Share Hospitality	0.035	0.054
Share Finance	0.051	0.065
Share Construction	0.066	0.037
Share Other Professional.	0.16	0.22
No. Individual-year observations.	547,000	341,000
No. Unique Individuals	165,000	109,000
Average Change in Distance (%)	0	−70
Panel B: Post-Expansion		
	Control	Treatment
Age	45.95	44.68
Share Female	0.48	0.5
Earnings	70,000	79,000
Log. Enterprise Total payroll	17.00	16.67
Distance to nearest station 2008	12.29	6.96
Commuting Distance	15.71	14.45
Res. Distance to CBD	32.54	10.28
Share Manuf.	0.099	0.072
Share Hospitality	0.034	0.055
Share Finance	0.049	0.062
Share Construction	0.072	0.042
Share Other Professional.	0.15	0.21
No. Individual-year observations.	888,000	588,000
No. Unique Individuals	170,000	116,000
Average Change in Distance (%)	0	−71

Notes: This table presents averages and shares pre- and post-SkyTrain expansion by treatment status for the estimation sample. Treatment is defined as whether or not an individual had *some* drop in distance to the nearest station between 2008 and 2009. Panel A reports summary statistics prior to the expansion (2005-2008), and Panel B reports summary statistics post-expansion (2009-2015). Res. distance to CBD refers to as the crow flies distance between Downtown Vancouver and a worker's place of residence in a given year. Earnings and log. Enterprise's total payroll is in 2016 dollars. The category *Other Professional* refers to workers in industries with 2-digit NAICS codes between 53 to 56. This includes Real Estate agents, managers, professional services, and administrative support. For confidentiality, the number of observations and average earnings are rounded to the nearest thousand.

Table 3: Labor Supply Parameters by Group k

Labor Supply Parameters		Male, 25-39	Male, 40-60	Female, 25-39	Female,40-60
Marginal Utility of Earnings	β_k	2.818	1.182	3.229	1.354
Within-Nest Independence Parameter	λ_{kg}	0.298	0.323	0.327	0.360
Marginal Commuting Disutility	η_k	-0.00865	-0.0223	-0.00901	-0.0265

Notes: This table presents labor supply parameters for each group k. Estimated β_k and λ_{kg} are estimated via 2SLS using the long-differnce equation (22). The instruments used are one-period changes in log revenue, one-period changes in neighborhood-firm shares, and one-period changes in establishment shares of workers of other types, excluding those of type k. Equation (22) is estimated separately for each type k. The reported λ_{kg} are employment-weighted industry averages of all industry-specific estimates from equation (22). The parameter η_k is estimated using 2016 data on commuting flows between neighborhoods and firms. It is estimated using a commuting-flow weighted PPML regression separately for each (k, g) pair using equation (24). This results in 84 separate estimates. Each estimate is multiplied by the industry corresponding $\frac{\lambda_{kg}}{\beta_k}$ to back out η_k. The reported parameter is the employment-weighted industry averages for each group k.

Table 4: Labor Supply Elasticities & Markdowns By Group k

Panel A: All		mean	median	p10	p90
Labor Supply Elasticity	$\mathcal{E}_{kjt}$	7.497	5.713	2.912	15.134
Markdown	$\frac{\mathcal{E}_{kjt}}{\mathcal{E}_{kjt}+1}$	0.847	0.851	0.744	0.938
Panel B: Male, 25-39		mean	median	p10	p90
Labor Supply Elasticity	$\mathcal{E}_{kjt}$	10.749	11.687	6.517	15.047
Markdown	$\frac{\mathcal{E}_{kjt}}{\mathcal{E}_{kjt}+1}$	0.907	0.921	0.867	0.938
Panel C: Male, 40-60		mean	median	p10	p90
Labor Supply Elasticity	$\mathcal{E}_{kjt}$	4.207	4.314	2.276	5.369
Markdown	$\frac{\mathcal{E}_{kjt}}{\mathcal{E}_{kjt}+1}$	0.793	0.812	0.695	0.843
Panel D: Female, 25-39		mean	median	p10	p90
Labor Supply Elasticity	$\mathcal{E}_{kjt}$	12.634	12.054	8.372	17.706
Markdown	$\frac{\mathcal{E}_{kjt}}{\mathcal{E}_{kjt}+1}$	0.918	0.923	0.893	0.947
Panel E: Female, 40-60		mean	median	p10	p90
Labor Supply Elasticity	$\mathcal{E}_{kjt}$	4.395	4.162	2.631	5.713
Markdown	$\frac{\mathcal{E}_{kjt}}{\mathcal{E}_{kjt}+1}$	0.795	0.806	0.725	0.851

Notes: This table presents elasticity and markdown distributions for all worker types, and each type k separately. Reported estimates are at the firm-level as defined by Equation (17), and equation (18) respectively. Mean elasticity and markdowns are *employment-weighted*. This calculation is based on the period 2001-2019, and weights are at the (k, j, t) level. The median, p10, and p90 are unweighted establishment-type level distribution percentiles for the period 2001-2019.

Table 5: Labor Demand Parameter Estimates & Distribution

Panel A: Distribution of Other Production Function Parameters		mean	median	p10	p90
Labor Productivity	γ_{kjt}	0.310	0.284	0.159	0.495
Returns to Scale	α_{jt}	0.289	0.240	0.076	0.578
Total Factor Productivity	$\ln\left(\tilde{\theta}_{jt}\right)$	14.912	14.73	13.296	16.770
Panel B: Estimated Parameters From Equation (27)			Est.Value		
Labor Substitution Parameter	ρ		0.9103		
Persistence of Labor Productivity	δ		0.734		

Notes: This table reports the distribution of production function parameters from equation (16) at the (k, j, t) level for γ_{kjt} and at the (j, t) level for α_{jt} and $\ln(\tilde{\theta}_{jt})$. The estimated parameters ρ and δ are from the estimating equation (27) using OLS. Given ρ, other firm-specific production function parameters are recovered as outlined in the estimation procedure.

Table 6: Commuting & Firm TFP

	$\ln(\tilde{\theta}_{jt})$			
	(1)	(2)	(3)	(4)
Commuting Distance (km)	0.011***	0.0091***	0.009***	0.0077***
	(0.00002)	(0.00002)	(0.00002)	(0.00002)
Commuting Distance × Female		0.0047***		0.0045***
		(0.00004)		(0.00004)
Type × Year FEs	Yes	Yes	No	No
Type k × CMA × Year FEs	No	No	Yes	Yes
Indv FEs	Yes	Yes	Yes	Yes

Notes: This table presents estimated coefficients from regressing log TFP as estimated by $\tilde{\theta}_{jt}$ on commuting distance in kilometers. Distance is measured as the Euclidean distance between neighborhood centroids. All specifications include individual fixed effects. Column (1) includes worker k type by year fixed effects. Column (2) is similar but includes an interaction with a female dummy. Columns (3) and (4) include type k by year by CMA fixed effects.

Table 7: Model Simulation of SkyTrain Expansion

	No Station	Millenium & Expo Line	Canada Line
Wages, Elasticities & Markdowns			
Average Wage	-0.20 %	-0.30%	1.0 %
Average Elasticity $\mathcal{E}_{zkj}$	-0.13 %	-0.11%	-0.5 %
Average Markdown	-0.02 %	-0.02%	-0.08 %
Population & Employment			
Total Population	-0.39%	-0.14%	3.75%
Total Employment	-0.65%	-0.30%	8.44%
Sorting			
Average Composite term	-0.17%	-0.34 %	1.7%
Average $\tilde{\theta}_{jt}$	-0.08%	-0.58%	5.12%
Welfare, Rents, Market Access & Concentration			
Average utility	-0.11%	-0.07%	0.45%
Average Rents	-0.70%	-0.49%	8.50%
Average RMA_{zk}	-0.22%	-0.13%	3.07%
Average GCI_{zk}	0.57%	0.08%	-8.60 %

Notes: This table presents the percentage change between the baseline model in 2016 and the counterfactual economy in the absence of the Canada Line. The no Canada Line economy is simulated by increasing commuting costs by a factor of 0.7 between neighborhoods on the Canada Line and other neighborhoods on the SkyTrain network. Column (1) corresponds to neighborhoods not on the SkyTrain network. Column (2) corresponds to neighborhoods near the Millennium & Expo lines, and column (3) corresponds to neighborhoods on the Canada Line. Labor market outcomes in wages, elasticities, markdowns, composite term, and $\tilde{\theta}_{jt}$ are employment-weighted averages. Rents, RMA_{zk} and GCI_{zk} are simple averages across (z,k) within each group.

Table 8: Model Simulation of Uniform Commuting Costs

Panel A: Within CMA	$\% \Delta \mathbb{V}ar(w_{zk})$	$\% \Delta \mathbb{V}ar(mdn_{zk})$	$\% \Delta \mathbb{V}ar(mrpl_{zk})$
Male, 25-39	-40.68	-15.23	-38.46
Male, 40-60	-42.34	-16.40	-37.13
Female, 25-39	-44.88	-12.24	-42.35
Female, 40-60	-52.53	-11.65	-46.47
Panel B: Across CMA	$\% \Delta \mathbb{V}ar(w_{zk})$	$\% \Delta \mathbb{V}ar(mdn_{zk})$	$\% \Delta \mathbb{V}ar(mrpl_{zk})$
Male, 25-39	2.68	-16.12	1.37
Male, 40-60	-16.79	-14.07	-18.34
Female, 25-39	-20.50	-10.19	-19.38
Female, 40-60	-38.78	-10.65	-35.04

Notes: This table presents the percentage change in variance of mean wages, markdowns, and marginal product across neighborhoods. Panel A focuses on residual variance within a given Census Metropolitan Area, which is measured by regressing the underlying quantities on CMA fixed effects. Panel B considers differences across neighborhoods and CMAs. The percentage change compares the baseline model for 2016 with the counterfactual economy assuming all workers face the same commuting distance of 20 km to any employer within a 75 km radius of their chosen place of residence.

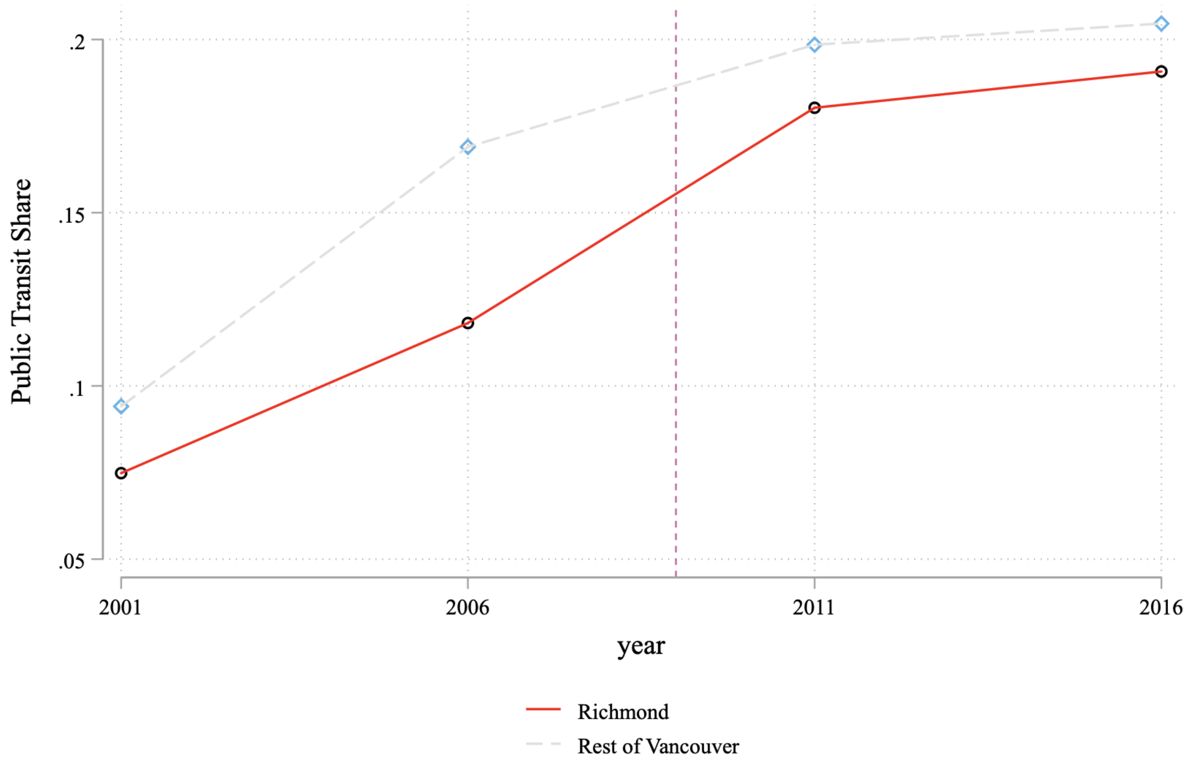
Appendix A - Additional Figures

Figure A-1: Commuting Distribution for Canada & GTA municipialities

			% of Workers by Commuting Distance						
CMA	Municipality	Employed Labourforce	Less than 5 km	5 to 9.9 km	10 to 14.9 km	15 to 24.9 km	25 km or more	Median Distance km	ST Mode%
Barrie	Barrie, CY	55,810	43%	18%	3%	3%	33%	6	10%
Barrie	Innisfil, T	12,555	11%	8%	18%	13%	49%	24	4%
Barrie	Springwater, TP	7,255	14%	19%	14%	22%	31%	16	4%
Barrie, CMA		75,625	35%	17%	6%	7%	35%	9	8%
Hamilton	Burlington, CY	74,340	32%	18%	11%	14%	26%	10	12%
Hamilton	Grimsby, T	10,440	24%	8%	6%	26%	36%	21	8%
Hamilton	Hamilton, C	207,120	33%	25%	13%	12%	16%	8	16%
Hamilton, CMA		291,905	33%	23%	12%	13%	19%	8	15%
Oshawa	Clarington, MU	33,940	17%	18%	14%	20%	31%	15	6%
Oshawa	Oshawa, CY	59,660	39%	22%	6%	8%	24%	7	13%
Oshawa	Whitby, T	50,070	23%	17%	9%	7%	44%	15	13%
Oshawa, CMA		143,680	28%	20%	9%	10%	33%	11	12%
Toronto	Ajax, T	41,355	24%	10%	5%	19%	42%	21	17%
Toronto	Aurora, T	21,350	26%	12%	5%	22%	35%	19	12%
Toronto	Brampton, CY	188,995	21%	24%	20%	20%	15%	11	13%
Toronto	Caledon, T	24,990	17%	7%	8%	29%	39%	21	4%
Toronto	E.Gwillimbury, T	9,410	16%	19%	13%	10%	43%	16	6%
Toronto	Georgina, T	17,595	17%	4%	7%	21%	52%	27	6%
Toronto	Halton Hills, T	25,290	26%	6%	13%	28%	27%	17	8%
Toronto	King, TP	8,100	10%	11%	13%	27%	39%	21	7%
Toronto	Markham, T	111,650	22%	24%	17%	23%	14%	11	17%
Toronto	Milton, T	26,255	26%	5%	14%	26%	29%	17	8%
Toronto	Mississauga, CY	294,340	24%	25%	17%	20%	14%	10	19%
Toronto	Mono, T	2,915	24%	14%	8%	8%	46%	18	4%
Toronto	New Tecumseth, T	11,790	32%	6%	7%	11%	45%	19	9%
Toronto	Newmarket, T	34,655	41%	9%	3%	11%	36%	11	11%
Toronto	Oakville, T	71,450	27%	15%	8%	17%	32%	15	18%
Toronto	Orangeville, T	12,250	50%	2%	1%	3%	44%	7	9%
Toronto	Pickering, CY	39,755	20%	10%	12%	26%	33%	18	16%
Toronto	Richmond Hill, T	68,820	19%	19%	18%	27%	17%	13	16%
Toronto	Toronto, C	1,016,825	34%	28%	18%	14%	6%	8	43%
Toronto	Uxbridge, TP	7,620	21%	7%	5%	13%	54%	27	7%
Toronto	Vaughan, CY	103,665	24%	26%	16%	27%	8%	10	11%
Toronto	W.Gwillimbury, T	10,815	20%	10%	19%	11%	40%	15	7%
Toronto	W-Stouffville, T	9,995	20%	10%	11%	29%	30%	18	8%
Toronto, CMA		2,160,020	29%	23%	16%	18%	14%	9	28%
	Canada	**13,041,190**	**36%**	**23%**	**13%**	**14%**	**14%**	**8**	**19%**

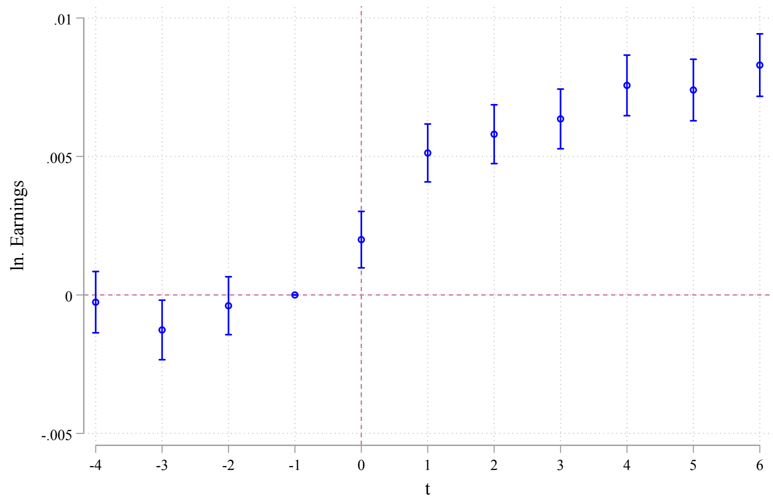
Source: 2006 Census

Figure A-2: Richmond & Rest of Metro Vancouver Public Transit Shares



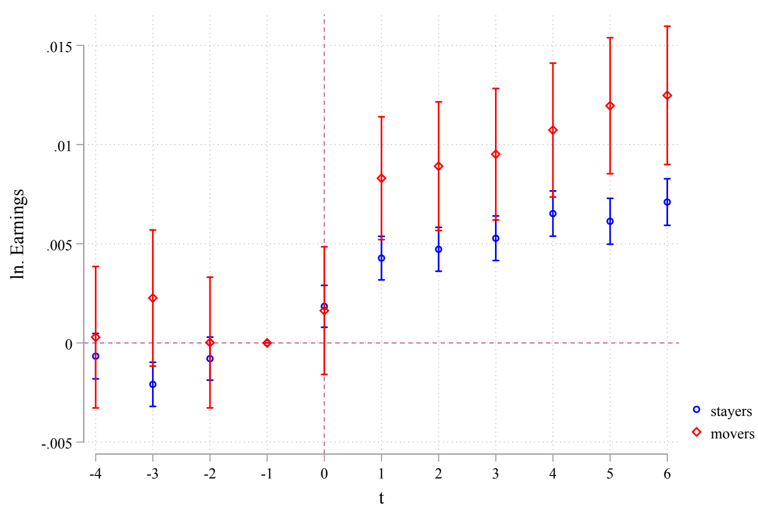
Notes: This Figure plots the share of public transit users in Richmond (red) and rest of Metro Vancouver area (light). Summary Statistics are from the long form 25% sample of the Canadian Census of Poulation for the corresponding years.

Figure A-3: Effect of SkyTrain Expansion on Earnings With Continuous Treatment



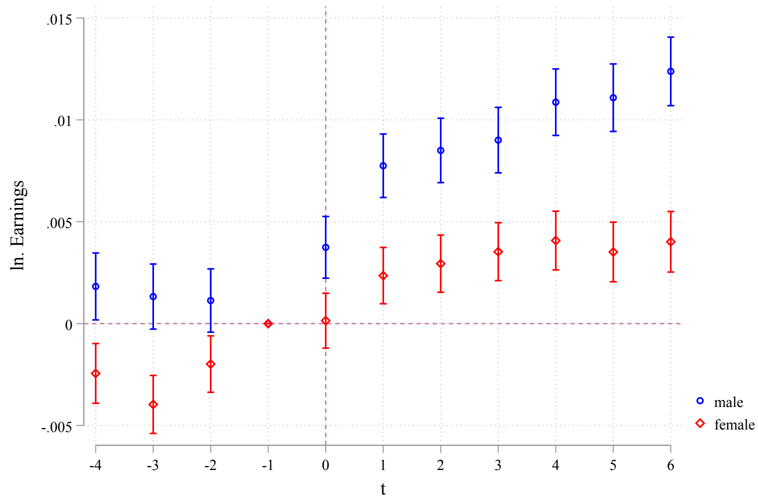
Notes: This figure plots estimates of the effect of going from p25 to p75 in the treatment exposure distribution after estimating the following TWFE specification: $Y_{it} = \alpha_i + \alpha_t + \sum_{d=1}^{5} \beta_k \%\Delta Dist \text{ Post}_t + \epsilon_{it}$. Time period is from 2005-2015, and coefficients are estimated relative to period $t = -1$, which corresponds to the year 2008, one year before the expansion. Specification includes individual fixed effects, and demographic group (gender $\times$ age bin) by year fixed effects.

Figure A-4: Effect of SkyTrain Expansion on Earnings With Continuous Treatment By Job Switchers & Stayers



Notes: This figure plots estimates of the effect of going from p25 to p75 in the treatment exposure distribution after estimating the following TWFE specification: $Y_{it} = \alpha_i + \alpha_t + \sum_{d=1}^{5} \beta_k \%\Delta Dist \text{ Post}_t + \epsilon_{it}$. Time period is from 2005-2015, and coefficients are estimated relative to period $t = -1$, which corresponds to the year 2008, one year before the expansion. Specification includes individual fixed effects, and demographic group (gender $\times$ age bin) by year fixed effects.

Figure A-5: Effect of SkyTrain Expansion on Earnings With Continuous Treatment By Gender



Notes: This figure plots estimates of the effect of going from p25 to p75 in the treatment exposure distribution after estimating the following TWFE specification: $Y_{it} = \alpha_i + \alpha_t + \sum_{d=1}^{5} \beta_k \% \Delta Dist\ \text{Post}_t + \epsilon_{it}$. The time period is from 2005 to 2015, and coefficients are estimated relative to period $t = -1$, which corresponds to the year 2008, one year before the expansion. Specification includes individual fixed effects, and demographic group (gender $\times$ age bin) by year fixed effects.

Appendix B - Model Appendix

B.1 Neighborhood z Labor Supply to Firm j

Following the notation from (8). Recall that

$$s_{zkj} = \frac{\exp\left(\frac{\delta_{kj}+\mu_{zkj}}{\lambda_{kg}}\right) I_{zkg}^{\lambda_{kg}-1}}{\sum_f I_{zkf}^{\lambda_{kf}}} \tag{A.1}$$

The probability of remaining in non-employment is given by $s_{zk0} = \frac{\exp(\delta_{k0})}{\sum_f I_{zkf}^{\lambda_{kf}}}$

$$\begin{aligned}
\Rightarrow \frac{s_{zkj}}{s_{zk0}} = \frac{\exp\left(\frac{\delta_{kj}+\mu_{zkj}}{\lambda_{kg}}\right) I_{zkg}^{\lambda_{kg}}}{\exp\left(\delta_{k0}\right) I_{zkg}} = \frac{s_{zkj|gt} I_{zkg}^{\lambda_{kg}}}{\exp(\delta_{k0})} &= s_{zkj|g} \frac{\exp\left(\delta_{kj}+\mu_{zkj}\right)}{\exp(\delta_{k0})} \left[\frac{I_{zkg}}{\exp\left(\frac{\delta_{kj}+\mu_{zkj}}{\lambda_{kg}}\right)}\right]^{\lambda_{kg}} \\
&= \frac{\exp\left(\delta_{kj}+\mu_{zkj}\right)}{\exp(\delta_{k0})} s_{zkj|g}^{1-\lambda_{kg}} \\
&= \frac{w_{kj}^{\beta_k} u_{kj} \exp(\eta_k d_{z,z(j)})}{w_{k0}^{\beta_k}} s_{zkj|g}^{1-\lambda_{kg}} \\
&\Rightarrow \ln \frac{s_{zkjt}}{s_{zk0t}} = \beta_k \ln \frac{w_{kjt}}{w_{k0t}} + (1-\lambda_{kg}) \ln s_{zkj|gt} - \tilde{\eta}_k d_{z,z(j)} + \ln u_{zkjt}
\end{aligned}$$

The second equality follows from noting that the probability of choosing firm j conditional on choosing nest N_g is given by

$$s_{zkj|g} = \frac{\exp\left(\frac{\delta_{kj}+\mu_{zkj}}{\lambda_{kg}}\right)}{\underbrace{\sum_{j' \in N_g} \exp\left(\frac{\delta_{kj'}+\mu_{zkj'}}{\lambda_{kg}}\right)}_{I_{zkg}}}$$

B.2 Neighborhood z Labor Supply *Elasticity* to Firm j

Labor supply elasticity between a given neighborhood firm pair depends on the sum of elasticity conditional on workers' spatial allocation with population elasticity of the neighborhood with respect to the firm's wages.

Elasticity Conditional on Workers' Spatial Allocation This section derives equation (9) from the model. This is the firm labor supply elasticity conditional on workers' spatial distribution. First note that the share of workers from neighborhood z working for firm j is the probability of working in nest g where $j \in N_g$ and the conditional probability of choosing this firm. We can rewrite (8) as the following

$$s_{zkj} = \underbrace{\frac{\exp\left(\frac{\delta_{kj}+\mu_{zkj}}{\lambda_{kg}}\right)}{I_{zkg}}}_{s_{zkj|g}} \times \overbrace{\frac{I_{zkg}^{\lambda_{kg}}}{\sum_f I_{zkf}^{\lambda_{kf}}}}^{s_{zkg}} \tag{A.2}$$

$$= s_{zkj|g} s_{zkg} \tag{A.3}$$

Where $I_{zkf} := \sum_{l \in N_f} \exp\left\{\frac{\delta_{kl}+\mu_{zkl}}{\lambda_{kf}}\right\}$. Taking logs on both sides of the above equation we get

$$\ln s_{zkj} = \frac{\beta_k}{\lambda_{kg}} \ln w_{kj} + \frac{1}{\lambda_{kg}} \ln u_{kj} - \frac{\eta_k}{\lambda_{kg}} d_{z,z(j)} + (\lambda_{kg} - 1) \ln I_{zkg} - \ln\left(\sum_f I_{zkf}^{\lambda_{kf}}\right)$$

Assuming firms behavior affects the wage index in their local labor market, and also globally means that

$$\frac{\partial \ln I_{zkg}}{\partial \ln w_{kj}} \neq 0$$

$$\frac{\partial \ln\left(\sum_f I_{zkf}^{\lambda_{kf}}\right)}{\partial \ln w_{kj}} \neq 0$$

Deriving the first of these two equations we have

$$\begin{aligned} \frac{\partial \ln I_{zkg}}{\partial \ln w_{kj}} &= \frac{\partial \ln\left(\sum_{j' \in N_g} w_{kj'}^{\frac{\beta_k}{\lambda_{kg}}} u_{kj'}^{\frac{1}{\lambda_{kg}}} \exp(-\eta_k d_{z,j'})\right)}{\partial w_{kj}} w_{kj} \\ &= \left[\frac{\frac{\beta_k}{\lambda_{kg}}}{w_{kj}} \frac{w_{kj}^{\frac{\beta_k}{\lambda_{kg}}} u_{kj}^{\frac{1}{\lambda_{kg}}} \exp(-\frac{\eta_k}{\lambda_{kg}} d_{z,j})}{\sum_{j' \in N_g} w_{kj'}^{\frac{\beta_k}{\lambda_{kg}}} u_{kj'}^{\frac{1}{\lambda_{kg}}} \exp(-\frac{\eta_k}{\lambda_{kg}} d_{z,j'})}\right] w_{kj} \\ &= \frac{\beta_k}{\lambda_{kg}} s_{zkj|g} \end{aligned} \tag{A.4}$$

The last line follows from (A.2). Deriving the second of these equations follows a similar line of reasoning

$$
\begin{aligned}
\frac{\partial \ln\left(\sum_f I_{zkf}^{\lambda_{kf}}\right)}{\partial \ln w_{kj}} &= \frac{\partial \ln\left(\sum_f I_{zkf}^{\lambda_{kf}}\right)}{\partial w_{kj}} w_{kj} \\
&= \frac{\partial I_{zkg}^{\lambda_{kg}}}{\partial w_{kj}} \frac{w_{kj}}{\sum_f I_{zkf}^{\lambda_{kf}}} \\
&= \lambda_{kg} I_{zkg}^{\lambda_{kg}-1} \frac{\partial I_{zkg}}{\partial w_{kj}} \frac{w_{kj}}{\sum_f I_{zkf}^{\lambda_{kf}}} \\
&= \lambda_{kg} I_{zkg}^{\lambda_{kg}-1} \times I_{zkg} \times \frac{\partial \ln I_{zkg}}{\partial \ln w_{kj}} \times \frac{1}{\sum_f I_{zkf}^{\lambda_{kf}}} \\
&= \lambda_{kg} \frac{\beta_k}{\lambda_{kg}} s_{zkg} s_{zkj|g} \\
&= \beta_k s_{zkj}
\end{aligned}
\tag{A.5}
$$

The last two equalities follow from the previous derivation and (A.2). Putting these together we get that

$$
\begin{aligned}
\frac{\partial \ln s_{zkj}}{\partial \ln w_{kj}} &= \underbrace{\frac{\beta_k}{\lambda_{kg}}}_{\textit{monopsonistic (LMS)}} \underbrace{+ (\lambda_{kg} - 1)\frac{\ln I_{zkg}}{\ln w_{kj}}}_{\textit{oligopsonistic (BHM)}} - \underbrace{\frac{\ln\left(\sum_f I_{zkf}^{\lambda_{kf}}\right)}{\ln w_{kj}}}_{\textit{aggregate competition (CKMM)}} \\
&= \frac{\beta_k}{\lambda_{kg}} + (\lambda_{kg} - 1)\frac{\beta_k}{\lambda_{kg}} s_{zkj|g} - \beta_k s_{zkj} \\
&= \beta_k \left[\frac{1}{\lambda_{kg}} + \left(1 - \frac{1}{\lambda_{kg}}\right) s_{zkj|g} - s_{zkj}\right] \\
&= \mathcal{E}_{zkj}
\end{aligned}
\tag{A.6}
$$

Elasticity of Neighborhood z population to Firm j Since firms' are non-atomisitc and can set wages to influence worker's residential choice, the labor supply elasticity facing a firm is the sum of population elasticity, and labor supply elasticity conditional on location. Following the same argument as above, and taking into account the distributional assumption in ν_{it} in (4) we can show that

$$
\zeta_{zkjt} = \varphi \beta_k s_{zkjt} \left[\frac{1}{\sigma} + \left(1 - \frac{1}{\sigma}\right) \pi_{zk|ct} - \pi_{zkct}\right]
\tag{A.7}
$$

B.3 Aggregate Labor Supply to Firm j

In this section I show how to aggregate up to get the firm-specific, neighborhood invariant, labor supply and its elasticity for a given type. I show that the latter is an average of the elasticities from all neighborhoods.

Aggregation

$$
\begin{aligned}
\ell_{kj} &= \sum_{z \in \mathcal{Z}} s_{zkj} \ell_{zk} \\
&= w_{kj}^{\frac{\beta_k}{\lambda_{kg}}} u_{kj}^{\frac{1}{\lambda_{kg}}} \sum_{z \in \mathcal{Z}} \exp\left(-\frac{\eta_k}{\lambda_{kg}} d_{z,o(j)}\right) \frac{I_{zkg}^{\lambda_{kg}-1}}{\left(\sum_f I_{zkf}^{\lambda_{kf}}\right)} \ell_{zk}
\end{aligned}
$$

Labor supply to firm j is inversely proportional to a population weighted average of distance between the firm from and potential workforce. This indicates that businesses that are located further away from their residential locations, e.g. those in different neighborhoods, stand to experience a larger labor supply shift as a result of remote work introduction.

Elasticity $\mathcal{E}_{kj}$ First, let μ_{zk} be the share of workers of type k who reside in neighborhood z. Then firm specific labor supply is given by

$$
s_{kj} = \sum_{z \in \mathcal{Z}} s_{zkj} \mu_{zk}
$$

Leveraging the linearity of the summation operator we have

$$
\begin{aligned}
\mathcal{E}_{kj} := \frac{\partial s_{kj}}{\partial w_{kj}} \frac{w_{kj}}{s_{kj}} &= \sum_{z \in \mathcal{Z}} \frac{\partial s_{zkj}}{\partial w_{kj}} \frac{w_{kj}}{s_{kj}} \mu_z \\
&= \sum_{z \in \mathcal{Z}} \frac{\partial s_{zkj}}{\partial w_{kj}} \frac{w_{kj}}{s_{zkj}} \frac{s_{zkj}}{s_{kj}} \mu_{zk} \\
&= \sum_{z \in \mathcal{Z}} \frac{\partial \ln s_{zkj}}{\partial \ln w_{kj}} \frac{s_{zkj}}{s_{kj}} \mu_{zk} \\
&= \sum_{z \in \mathcal{Z}} \left(\mathcal{E}_{zkj} + \zeta_{zkjt}\right) \underbrace{\frac{s_{zkj}}{s_{kj}} \mu_{zk}}_{\omega_{zkj}} \\
&= \sum_{z \in \mathcal{Z}} \left(\mathcal{E}_{zkj} + \zeta_{zkjt}\right) \omega_{zkj}
\end{aligned}
$$

The third equality follows from multiplying and dividing each term in the summation by the respective s_{zkj}. Recall that

$$\omega_{zkj} = \frac{\mathbb{P}(z, j|k)}{\mathbb{P}(j|k)} = \frac{\text{\# workers } (z, k, j)}{\text{\# workers } (k, j)}$$

This proves that firm specific labor supply of type k is a weighted average of the neighborhood to firm specific labor supplies. The weights take into account that the sum of log is not equal to log of sums.

B.4 Model Implied Gravity Equation

$$\ln\left(\frac{s_{zkjt}}{s_{zkj0t}}\right) = \beta_k \ln\left(\frac{w_{kjt}}{w_{k0t}}\right) + (1 - \lambda_{kg}) \ln\left(s_{zkj|gt}\right) - \tilde{\eta}_k d_{zo(j)} + \ln u_{zkjt}$$

Assumption 5. Location-type-firm specific amenities take have the following separable form

$$u_{zkjt} = u_{kjt} \times u_{zkt} \times \epsilon_{zkjt}$$

and the residual error term satis

$$\mathbb{E}[\epsilon_{zkjt}|z, j, k, t] = 1$$

The second part of this assumption allows the estimation of η_k using PPML as shown by shown Silva and Tenreyro (2006). Equation (A.1) combined with the separability assumption implies

$$\begin{aligned}
n_{zkjt} &= \exp\left(\frac{\eta_k d_{zj}}{\lambda_{kg}}\right) \exp\left(\frac{\beta_k \ln \frac{w_{kj}}{w_{k0}} + \ln u_{zkjt}}{\lambda_{kg}}\right) \frac{I_{zkg}^{\lambda_{kg}}}{\sum_f I_{zkf}^{\lambda_{kf}}} L_{zk} \\
&= \epsilon_{zkjt} \exp\left(\frac{\eta_k d_{zj}}{\lambda_{kg}}\right) \exp\left(\frac{\beta_k \ln \frac{w_{kj}}{w_{k0}} + \ln u_{kj}}{\lambda_{kg}}\right) \frac{\exp\left(\frac{\ln u_{zk}}{\lambda_{kg}}\right) \tilde{I}_{zkg}^{\lambda_{kg}}}{\sum_f \tilde{I}_{zkf}^{\lambda_{kf}}} L_{zk} \\
&= \epsilon_{zkjt} \times \exp\left(\frac{\eta_k d_{zj}}{\lambda_{kg}}\right) \times \exp\left(\ln FE_{kj}\right) \times \exp\left(\ln FE_{zk}\right) \\
&= \epsilon_{zkjt} \times \exp\left(\frac{\eta_k d_{zj}}{\lambda_{kg}} + \ln FE_{kj} + \ln FE_{zk}\right)
\end{aligned}$$

$$\begin{aligned}
FE_{kj} &= \frac{\beta_k \ln \frac{w_{kj}}{w_{k0}} + \ln u_{kj}}{\lambda_{kg}} \\
\Rightarrow \ln u_{kj} &= \lambda_{kg} FE_{kj} - \beta_k \ln \frac{w_{kj}}{w_{k0}}
\end{aligned}$$

Given $\ln u_{kj}$ we can use the log odds of employment relative to non-employment to back out $\ln u_{zkt}$ in the following manner. First, define $s_{zk\emptyset}$ to be the the probability of employment at any firm. Given the separability assumption, this could be written as

$$
\begin{aligned}
I_{zkf} &= \exp\left(\frac{\ln u_{zk}}{\lambda_{kf}} \sum_{j\in N_f} \left(\frac{\beta_k \ln \frac{w_{kj}}{w_{k0}} + \ln u_{kj} + d_{zj}}{\lambda_{kg}}\right)\right)^{\lambda_{kf}} \\
&= u_{zk} \exp\left[\sum_{j\in N_f} \left(\frac{\beta_k \ln \frac{w_{kj}}{w_{k0}} + \ln u_{kj} + d_{zj}}{\lambda_{kg}}\right)\right]^{\lambda_{kf}}
\end{aligned}
$$

$$
s_{zk\emptyset} = \frac{u_{zk} \sum_f \exp\left[\sum_{j\in N_f} \left(\frac{\beta_k \ln \frac{w_{kj}}{w_{k0}} + \ln u_{kj} + d_{zj}}{\lambda_{kg}}\right)\right]^{\lambda_{kf}}}{1 + \sum_f I_{zkf}^{\lambda_{kf}}}
$$

$$
\frac{s_{zk\emptyset}}{s_{zk0}} = u_{zk} \sum_f \exp\left[\sum_{j\in N_f} \left(\frac{\beta_k \ln \frac{w_{kj}}{w_{k0}} + \ln u_{kj} + d_{zj}}{\lambda_{kg}}\right)\right]^{\lambda_{kf}}
$$

B.5 Equilibrium Uniqueness

The goal is to derive the firm-level super-elasticity with respect to its own wages and those of other firms. I will demonstrate that it is a composite of the average super elasticity and the within-firm variance in elasticities. The same applies to the cross-super elasticity.

Useful Properties

1. Define μ_{zk} to be the share of workers of type k living in neighborhood z
2. Recall that the firm level elasticity was the weighted average of elasticities across all neighborhoods
$$
\mathcal{E}_{kj} = \sum_{z\in\mathcal{Z}} \mathcal{E}_{zkj}\omega_{zkj}
$$
where
$$
\omega_{zkj} = \frac{\mathbb{P}(z,j|k)}{\mathbb{P}(j|k)} = \frac{\#\text{ workers }(z,k,j)}{\#\text{ workers }(k,j)} = \frac{s_{zkj}\mu_{zk}}{\sum_{z'} s_{z'kj}\mu_{z'k}} = \mathbb{P}(z|j,k)
$$

3. So we have that

$$\mathcal{E}_{kj} = \sum_{z \in \mathcal{Z}} \mathcal{E}_{zkj} \frac{s_{zkj} \mu_{zk}}{s_{kj}}$$

4. Recall too that

$$\mathcal{E}_{zkj} = \beta_k \left[\frac{1}{\lambda_{kg}} + \left(1 - \frac{1}{\lambda_{kg}}\right) s_{zkj|g} - s_{zkj} \right]$$

5. Combining 2 and 3 implies the firm-level productivity could be written as the following:

6. $$\mathcal{E}_{kj} = \beta_k \left[\frac{1}{\lambda_{kg}} \left(1 - \frac{1}{\lambda_{kg}}\right) \tilde{s}_{zkj|g} - \tilde{s}_{zkj} \right] := \mathbb{E}[\mathcal{E}_{zkj}|j,k]$$

   where

$$\tilde{x} = \sum_z x \mathbb{P}(z|j,k) =: \mathbb{E}[x|j,k]$$

Equipped with this property, I will now derive the semi super-elasticity

$$\begin{aligned}
\frac{\partial \mathcal{E}_{kj}}{\partial \ln w_{kj}} &= \sum \frac{\partial}{\partial \ln w_{kj}} \left(\frac{\mathcal{E}_{zkj} s_{zkj} \mu_{zk}}{s_{kj}} \right) \\
&= s_{kj}^{-1} \left[\sum_z \frac{\partial \mathcal{E}_{zkj}}{\partial \ln w_{kj}} s_{zkj} \mu_{zk} + \sum_z \frac{\partial s_{zkj}}{\partial \ln w_{kj}} \mathcal{E}_{zkj} \mu_{zk} \right] - \frac{\partial s_{kj}}{\partial \ln w_{kj}} s_{kj}^{-2} \sum_{z \in Z} \mathcal{E}_{zkj} s_{zkj} \mu_{zk} \\
&= s_{kj}^{-1} \left[\sum_z \frac{\partial \mathcal{E}_{zkj}}{\partial \ln w_{kj}} s_{zkj} \mu_{zk} + \sum_z \frac{\partial s_{zkj}}{\partial \ln w_{kj}} \mathcal{E}_{zkj} \mu_{zk} \right] - \mathcal{E}_{kj}^2 \\
&= \sum_z \frac{\partial \mathcal{E}_{zkj}}{\partial \ln w_{kj}} \frac{s_{zkj} \mu_{zk}}{s_{kj}} + \sum_z \left(\mathcal{E}_{zkj}^2 \frac{s_{zkj} \mu_{zk}}{s_{kj}} \right) - \mathcal{E}_{kj}^2 \\
&= \underbrace{\mathbb{E}\left[\frac{\partial \mathcal{E}_{zkj}}{\partial \ln w_{kj}} \middle| j,k \right]}_{<0} + \underbrace{\mathbb{V}ar\left[\mathcal{E}_{zkj} \middle| j,k \right]}_{>0}
\end{aligned}$$

The third equality follows from noting that:

$$\frac{\partial s_{kj}}{\partial \ln w_{kj}} s_{kj}^{-1}$$

And

$$s_{kj}^{-1} \sum_{z \in Z} \mathcal{E}_{zkj} s_{zkj} \mu_{zk}$$

are two different ways of writing down the firm-level elasticity term. The fourth equality follows from multiplying and dividing the second term in the square bracket by s_{zkj}. The fifth equality follows from the definition of conditional expectation and variance.

The neighborhood-level superelasticities are negative, as in a standard nested logit framework. This could be viewed as the intensive margin of elasticity adjustment. However, since firms could source labor from any neighborhood, the within-firm distribution of elasticities also matters, as presented by the second moment. This is the extensive margin, as firms could hire workers from further away neighborhoods as they increase the offered wages. When firms offer higher wages, there are two effects. One, workers from a given neighborhood become more inelastic as the firm grows and increases its market share. This allows the firm to re-adjust and offer lower wages until it converges to some middle point. Two, when a demand shock hits the firm, it changes the workforce composition. So that when firms offer higher wages, it attracts workers from further away neighborhoods who are more sensitive to wage changes due to commuting. This increases firm-level elasticity and reduces the firm's wage-setting power. Through this process, wage changes within a firm can become self-reinforcing, leading to multiple equilibria. The question is now: *What are the restrictions that would allow there to be a unique equilibrium? I will now derive the variance term in the above equation more explicitly.*

$$
\begin{aligned}
\mathbb{V}ar\left[\mathcal{E}_{zkj}\middle| j,k\right] &= \sum_z \left(\mathcal{E}_{zkj} - \mathcal{E}_{kj}\right)^2 \frac{s_{zkj}\mu_{zk}}{s_{kj}} \\
&= \sum_z \left[\left(1 - \frac{1}{\lambda_{kg}}\right)\left(s_{zkj|g} - \tilde{s}_{kj|g}\right) - \left(s_{zkj} - \tilde{s}_{zkj}\right)\right]^2 \frac{s_{zkj}\mu_{zk}}{s_{kj}} \\
&= \left(1 - \frac{1}{\lambda_{kg}}\right)^2 \sum_z \left(s_{zkj|g} - \tilde{s}_{kj|g}\right)^2 \frac{s_{zkj}\mu_{zk}}{s_{kj}} + \sum_z \left(s_{zkj} - \tilde{s}_{kj}\right)^2 \frac{s_{zkj}\mu_{zk}}{s_{kj}} \\
&\qquad + 2\left(1 - \frac{1}{\lambda_{kg}}\right) \sum_z \left(s_{zkj|g} - \tilde{s}_{kj|g}\right)\left(s_{zkj} - \tilde{s}_{kj}\right) \frac{s_{zkj}\mu_{zk}}{s_{kj}} \\
&= \left(1 - \frac{1}{\lambda_{kg}}\right)^2 \mathbb{V}ar\left(s_{zkj|g}\middle| j,k\right) + \mathbb{V}ar\left(s_{zkj}\middle| j,k\right) - 2\left(1 - \frac{1}{\lambda_{kg}}\right) Cov\left(s_{zkj|g}, s_{zkj}\middle| j,k\right)
\end{aligned}
$$

The second equality uses the alternative definition of firm-level elasticity in bullet point (5) from the useful properties section. The rest follows the definitions of variance and covariance. This could be further simplified by noting the following:

$$
\begin{aligned}
\mathbb{V}ar(s_{zkj}|j,k) &= \mathbb{V}ar(s_{zkj|g}s_{zkg}|j,k) \\
&= \mathbb{V}ar(s_{zkj|g}|j,k)
\end{aligned}
$$

Conditional on being at firm j, the probability of working in industry g is one. As the industry of the firm is implicitly being conditioned on. The same rationale also implies the following:

$$Cov(s_{zkj|g}, s_{zkj}|j,k) = Cov(s_{zkj|g}, s_{zkj|g}s_{zkg}|j,k) = \mathbb{V}ar(s_{zkj|g}|j,k)$$

So the last equation from the previous derivations implies the following:

$$\mathbb{V}ar\left[\mathcal{E}_{zkj}\middle|j,k\right] = \frac{1}{\lambda_{kg}^2}\mathbb{V}ar\left(s_{zkj|g}\middle|j,k\right) \tag{A.8}$$

And the super firm level labor supply elasticity could then be written as

$$\frac{\partial \mathcal{E}_{kj}}{\partial \ln w_{kj}} = \underbrace{\mathbb{E}\left[\frac{\partial \mathcal{E}_{zkj}}{\partial \ln w_{kj}}\middle|j,k\right]}_{\textit{intensive margin}} + \underbrace{\frac{1}{\lambda_{kg}^2}\mathbb{V}ar\left(s_{zkj|g}\middle|j,k\right)}_{\textit{extensive margin}} \tag{A.9}$$

Appendix C - Simulation Algorithm

I simulate the model following the Jacobi algorithm over wages and rental prices. The algorithm makes use of two primary equations. One, firm best response function, and housing market clearing condition.

Firm best response functions

$$w_{kjt} = \frac{\mathcal{E}_{kjt}}{\mathcal{E}_{kjt}+1} P_{jt}\theta_{jt} \sum_{k\in\mathcal{K}} \left(\gamma_{kjt}\ell_{kjt}^{\rho}\right)^{\frac{\alpha}{\rho}-1} \tag{A.10}$$

Housing Market Clearing:

$$r_{zt} = \frac{\gamma}{H_{zt}^{s}} \underbrace{\sum_{k\in\mathcal{K}} m_{kt}\mu_{zkt} \sum_{j\in\mathcal{J}} w_{kjt}s_{zkjt}}_{\text{Total Earnings in neighborhood } z := E_{zt}} \tag{A.11}$$

The housing clearing follows from Equation (14), and setting $H_{zt}^{d}(r_{zt}) = H_{zt}^{s}$

Algorithm 1: Under-relaxed Jacobi Algorithm

Initialize vector of wages $\mathbf{w_{kj}}^{0}$ for all $\mathcal{K}\times\mathcal{J}$ and rents $\mathbf{r}_{z}^{0}$ for all locations $z \in \mathcal{Z}$
repeat
  Use equations (10) and (12) to generate neighborhood shares μ_{zk}
  Given μ_{zk}, use equation (8) to generate s_{zkj}
  Given s_{zkj}, use equation (9) to generate $\mathcal{E}_{zkj}$. And equation (13) to generate ζ_{zkjt}
  Given $s_{zkj}, \mu_{zk}, \mathcal{E}_{zkj}, \zeta_{zkjt}$ use equation (17) to generate $\mathcal{E}_{kj}$
  Given s_{zkj}, μ_{zk} sum to get $l_{kj} = \sum_{z} s_{zkjt}\mu_{zk}\ell_{k}$ and calculate $P_{jt}\theta_{jt}\sum_{k\in\mathcal{K}}\left(\gamma_{kjt}\ell_{kjt}^{\rho}\right)^{\frac{\alpha}{\rho}-1}$
  Given $s_{zkj}, \mu_{zk}, \mathbf{w_{kj}}^{new}$ construct E_{zt} in equation (A.11)
  Use equation (A.10) to define $\mathbf{w_{kj}}^{new}$, and equation (A.11) to define $\mathbf{r_{z}}^{new}$
  Update wages using $\mathbf{w_{kj}}^{n+1} = \lambda\mathbf{w_{kj}}^{new} + (1-\lambda)\mathbf{w_{kj}}^{n}$
  Update rents using $\mathbf{r_{z}}^{n+1} = \lambda\mathbf{r_{z}}^{new} + (1-\lambda)\mathbf{r_{z}}^{n}$
until $\|\mathbf{w_{kj}}^{n+1} - \mathbf{w_{kj}}^{n}\|_{max} < \epsilon$ and $\|\mathbf{r_{z}}^{n+1} - \mathbf{r_{z}}^{n}\|_{max} < \epsilon$
