# Urban Transit Infrastructure and Inequality: The Role of Access to Non-Tradable Goods and Services*

Brandon Joel Tan†

December 2020

**Abstract**

With 68% of the world population projected to live in urban areas by 2050, mass transit networks are expanding faster than ever before. But how are the economic gains from such expansions being shared between low- and high-income workers? Existing research focuses on the role of commuting to work (Tsivanidis 2019; Balboni et al. 2020), however much of urban travel is related to the consumption of non-tradable goods and services (retail, F&B, personal services etc.). Since low-income workers are overwhelmingly employed in these non-tradable sectors, changes in consumption travel patterns in response to a transit expansion leads to a spatial re-organization of low-income jobs in the city which has important implications for inequality. This paper develops an urban spatial model with heterogeneous worker groups and incorporating travel to consume non-tradable goods and services. We estimate our model using detailed farecard and administrative data from Singapore to quantify the impact of the Downtown Line (DTL). We find large welfare gains for high-income workers, but near zero gains for low-income workers. All workers benefit from improved access to consumption opportunities, but low-income non-tradable sector jobs move to less attractive workplaces. Abstracting away from consumption travel results in a five-fold underestimation of the inequality effects and failure to capture the spatial re-organization of low-income jobs in the city.

*We thank Pol Antras, Nick Buchholz, Edward Glaeser, Nathan Hendren, Adam Kapor, Jakub Kastl, Gabriel Kreindler, Kate Ho, Marc Melitz, Christopher Neilson, Stephen Redding, and a number of other colleagues, conference and seminar participants for helpful comments and suggestions. Responsibility for results, opinions and errors is the authors' alone.

†Department of Economics, Harvard University, btan@g.harvard.edu

# 1 Introduction

The world is expanding public transit networks faster than ever before. Each year, more than a trillion dollars is invested in building and expanding transportation infrastructure (Lefevre, Leipziger, and Raifman 2014). In particular, there has been a surge in urban mass transit systems in response to the continued rapid densification of cities. Worldwide, mass transit carried 53 billion passengers in 2017 with nearly 650 transit lines covering 14,000 kilometers (UITP 2018).[1] Just between 2015 and 2017, roughly 1,900 kilometers of new track was put into service, and the International Association of Public Transport projects that more than 200 new transit lines will open across the world between 2018 and 2023 (UITP 2018). With 68% of the world population projected to live in urban areas by 2050, governments will continue spend vast sums on mass transit (United Nations 2018). As inequality is increasingly becoming a major issue across cities around the world, policymakers must consider how the economic gains from public transit expansion are shared between low- and high-income workers. Existing research focuses on differential access to employment opportunities across worker groups emphasizing the cost of travel to work (Tsivanidis 2019; Balboni et al. 2020). However, a large share of urban trips are unrelated to work and are made for the consumption of non-tradable goods and services such as restaurants, coffee shops, retail stores, salons, and cinemas. According to the National Household Travel Survey, 31% of travel miles are made for work purposes, while 33% of travels miles are made for shopping or meals and another 26% for social or recreational purposes.[2] Additionally, non-tradable consumption makes up over 50% of total household expenditure (Department of Statistics 2018).[3]

Accounting for travel related to the consumption of non-tradables is critical for evaluating the impact of public transit expansion on inequality. First, low- and high-income workers face differential reductions in travel costs from consumption trips and differential changes in access to consumption opportunities. Second, in general equilibrium, workers trade off rents with

[1] See Figure A1.
[2] See Appendix Figure A2 (Department of Transportation 2017).
[3] See Appendix Figure A3.

*both* access to employment and consumption, shaping residential patterns by worker group across the city. Third, where workers choose to consume non-tradables determines the spatial distribution of non-tradable sector jobs. These urban non-tradable sectors overwhelmingly employ low-income workers - such as waiters, dishwashers, salespeople, cashiers, manicurists, and ticket takers etc. According to Autor and Dorn (2013), low income jobs have been rapidly moving into non-tradable services since the 1980s, and the trend is only expected to intensify with the decline of low-skill intensive manufacturing.[4] In response to changes in transit access, jobs may move to locations which are less (or more) productive, that pay lower (or higher) wages, or that require longer (or shorter) commutes by worker group. This re-organization of commercial activity across space changes each worker group's expected access to employment, or expected income net of commuting, differentially.

In this paper, we develop an urban spatial model with heterogeneous worker groups, low- and high-income, which incorporates both travel to work and to consume non-tradable goods and services. The internal structure of the city is modeled as a set of neighborhoods with a transportation network defined as the bilateral travel times between locations. Worker groups differ in their preferences over residential and consumption locations, productivities over workplaces and sectors, consumption patterns, travel costs, and elasticities. First, workers choose where to live trading off rents, residential amenities, and access to employment and consumption opportunities by type. Then, workers choose where to work based on type-specific wages, match-specific productivities and commute distances, and choose where to consume non-tradables based on prices, type-specific idiosyncratic consumption amenities, and travel distances. Non-tradable and tradable sectors have different input requirements over commercial floor space and labor provided by each worker group, with the non-traded sector being much more intensive in low-income labor. Market clearing in the non-traded market implies that consumption travel patterns drive the demand for non-traded production, and hence also the demand for low-income workers, in each location. The model also

[4] See Appendix Figure A4.

accommodates agglomeration forces and spillovers in endogenous residential amenities by type as a force towards segregation.

We use the model to study the impact of the Downtown Line (DTL) in Singapore. We exploit farecard data on the universe of public transportation trips linked to individuals over a three year period before and after the line opening to distinguish travel patterns across worker groups and locations. We also use detailed administrative spatial data on land use, rents, expenditures and employment for structural estimation. The DTL is the longest underground and automated rapid mass transit line in Singapore, designed to connect the northwest and east to the center of the city. We estimate a spatial differences-in-differences specification and find that the DTL increased nearby housing prices by 5% over a four year period. Event study estimates show that neighborhoods with a DTL station experienced a 20% in increase in trips after the line opening. However, there were large differences in how low- and high-income workers adjusted their travel patterns in response to the DTL across locations.

We estimate the model using data from before the DTL opening. Across worker types, we find significant differences in theoretically-consistent measures of residence, workplace and consumption location attractiveness, and validate our measures with external data on amenities. Model-implied average wages by worker type and residence are also highly correlated with that of administrative data. Low-income workers have larger travel elasticities than high-income workers, and elasticities are larger than for consumption than for work travel across both types. Spillovers in residential amenities are larger for low-income workers, and residential elasticities are slightly larger for high-income workers. In the pre-DTL equilibrium, low-income workers commute shorter distances than high-income workers, but make similar distanced consumption trips. High-income workers live near the center of the city with the DTL directly serving many of these neighborhoods, while low-income workers live away from the city center. High-income workers spend more of their overall expenditures on non-tradables than low-income workers.

We quantify the welfare and inequality effects of the DTL by undertaking counterfactuals using exact-hat methods popularized by Dekle, Eaton, and Kortum (2008). Using changes in travel times as observed in the fare card data before and after the opening of the line, we find that the predictions of the model are highly correlated with post-DTL data on travel flows, residential patterns and non-tradable firm entry. We find that the DTL improves welfare for high income workers by 3.15%. However, low income workers experience near zero net benefits. Although access to consumption opportunities increases for both groups, non-tradable production and low income jobs move to less attractive locations for low-income workers offsetting those gains. The inequality effects are driven by two main mechanisms. First, the DTL disproportionately improves access for many high-income areas in which the line directly serves. Second, more consumers consume non-tradables near the city center in response to new access via DTL. Thus, non-tradable low-income jobs move to the center of the city, while low income workers live away from the center. We find that average commute time increased 1.5% for low income workers, while high income workers saw a 1% reduction. Abstracting away from travel to consume non-tradables results in a five-fold underestimation of the inequality effects of the DTL, estimating that both worker groups benefit but with a slightly larger share of the gains going to high-income workers. This is a result of failing to capture the spatial re-organization of low-income non-tradable sector jobs. Aggregate welfare gains are also underestimated by about 40%, ignoring gains in access to consumption opportunities for both groups.

Last, we conduct a decomposition exercise to isolate the role of various differences across worker groups impact inequality. First, low-income workers have lower expenditures on non-tradables than high-income workers. Thus, they experience smaller gains in welfare from improved access to consumption opportunities. Second, since low-income workers have lower travel costs than high-income workers, they experience smaller gains from reductions in travel time. Third, higher travel elasticities imply that low-income workers are more able to substitute to attractive work and consumption locations in equilibrium, thus benefit less when travel time falls. Last, a more greater dispersion in residential preferences relative to high-income workers implies that low-income workers are less able to move to take advantage of improved access in other neighborhoods.

This paper contributes directly to the literature focused on studying the impact of transportation infrastructure within cities (McDonald et al. 1995; Gibbons and Machin 2005; Baum-Snow and Kahn 2005; Glaeser, Kahn, and Rappaport 2008; Billings 2011; Donaldson and Hornbeck 2016; Severen 2018; Donaldson 2018; Brooks and Lutz 2019; Gupta, Van Nieuwerburgh, and Kontokosta 2020; Heblich, Redding, and D. Sturm 2020). The most closely related papers to ours are Tsivanidis (2019) and Balboni et al. (2020) which study the Bus Rapid Transit (BRT) systems focusing on inequality in Colombia and Tanzania respectively. Both papers use quantitative spatial models restricted to workplace commuting and estimate small inequality effects. Our paper emphasizes the importance of incorporating the role of travel costs to access non-tradable goods and services in evaluating the distributional welfare effects of transit expansions. We find similarly small inequality effects of the DTL line abstracting away from consumption trips in our model, but find large effects when they are incorporated.

Our paper also contributes to several other strands of literature. The first is the growing body of work on quantitative spatial models more generally (Redding and D. M. Sturm 2008; Allen and Arkolakis 2014; Ahlfeldt et al. 2015; Redding 2016; Allen, Arkolakis, and Li 2016; Caliendo et al. 2018; Desmet, Nagy, and Rossi-Hansberg 2018; Monte, Redding, and Rossi-Hansberg 2018; Adao, Arkolakis, and Esposito 2019; Allen and Arkolakis 2019). Our paper is closely related to and builds on methods from Miyauchi, Nakajima, and Redding (2020) who explore the role of consumption access and agglomeration in the Greater Tokyo metropolitan area. We are the first paper to consider a quantitative spatial model with heterogeneous workers incorporating both travel for work and consumption. This paper is also related to the literature on the value of urban amenities (Roback 1982; Blomquist et al. 1988; Glaeser, Kolko, et al. 2001; Albouy et al. 2016; Cragg and Kahn 1997; Bayer et al.

2009, Fan, Guthrie, and Levinson 2016; Diamond 2016; Almagro and Domínguez-Iino 2020). We demonstrate the important role of non-tradable retail amenities in shaping inequality and the spatial distribution of economic activity in the city. We are also able to quantify differences in preferences over amenities across worker groups. Last, our paper is related to the literature on inequality in cities (Brueckner et al. 1999, Glaeser, Resseger, and Tobio 2009; Su et al. 2018; Fogli and Guerrieri 2019; Couture et al. 2019). We show that changes in travel access can have large implications for urban inequality.

The rest of the paper proceeds as follows. Section 2 discusses the context of Singapore and the Downtown Line, as well as the data. Section 3 presents reduced form results and facts which motivate our structural model. Section 4 develops the model, while Section 5 estimates it. Section 6 quantifies the impact of the Downtown Line. We conclude in Section 7.

# 2 Background and Data

## 2.1 Context: Singapore and the Downtown Line

Singapore is an island city-state in Southeast Asia. With a population of 5.5 million inhabitants, Singapore is the third densest country in the world, and is among the densest cities (World Bank 2019).[5] There exists significant inequality with a Gini coefficient of 46.4, ranking only behind Hong Kong and the United States among high-income countries.

The population in Singapore is heavily reliant on public transportation, which is composed of buses, light-rail networks and mass-rail networks. The Singapore government heavily restricts the supply of cars with a Vehicle Quota System. According to the 2019 Worldwide Cost of Living Survey carried out by the Economist Intelligence Unit, Singapore "remains the most expensive place in the world to buy and run a car." As of 2018, the price of a

[5] There are highly restrictive immigration policies with strict foreign worker quotas, and limited out and in migration (Ministry of Manpower 2020).

Volkswagen Golf 1.4 is US $110,479.80, 5 times more expensive than in the United States (US$21,845). Additionally, the city's Electronic Road Pricing System imposes large tolls for driving to encourage the use of public transportation. Tolls are highest during commuting hours – a car trip from the northwest to the center of the city 8am and 9am on a weekday is tolled up to 15 dollars. As a result only 10% of households in Singapore own a car, and far fewer drive to work. With the public transit system carrying over 4 million passengers per day, public transportation is the primary means of travel for the population.[6] Approximately 60% of trips are made on buses, as opposed to rail lines.

The Downtown Line is the longest underground and mass rapid transit line in Singapore. At 41.9 kilometres (26.0 miles) and with 34 operation stations, the line runs from Bukit Panjang station in the north-west to Expo station in the east via the Central Area, see Figure A5. The line was first announced on 23 October 2001, and was built in three phases. The first phase opened in December 2013, and was composed of 6 stations from Bugis to Chinatown station within the Downtown area. The second phase from Bukit Panjang to Rochor station, which linked the north-west to the center of the city, opened in December 2015. The third and final phase from Fort Canning to Expo station, connecting the east to the city of the city, opened in October 2017. The line, at a cost of 15.5 billion USD, is considered the government's most ambitious public transit project to date (The Straits Times 2017). The purpose of the line was to provide the north-west and eastern areas a direct link to the center of the city, and to alleviate congestion on various other rail and bus routes.

## 2.2 Data

This sub-section provides an overview of the datasets used in the analysis.

Our primary source of data is public transit fare card data (EZ-Link) from the Land Transport Authority of Singapore.[7] We observe all trips made by public transit (mass rail,

[6] Appendix Table A1 provides a summary of the mass-rail lines in Singapore.

[7] We present a heat map visualization of the data in Appendix Figure A6.

light rail or bus) linked to a individual's fare card[8]. Our data set covers one week every quarter between June 2015 and June 2018, and three full months between Dec 2015 and February 2016. The longer quarterly data set allows to observe changes in transit patterns from before the opening of Phase 2 of the Downtown Line to after the opening of Phase 3. The full three months of data captures the period directly before and after the opening of Phase 2 of the Downtown Line. We observe a total of over a billion trips. For each trip, we observe the origin, destination, and the start and end time of the trip. Each individual in our data set is categorized into eight groups: Adult, Low-Income Worker, Primary Student, Secondary Student, Tertiary Student, Senior Citizen, Military, and Person with Disabilities. In our analysis, we focus the Adult and Low-Income Worker categories. Low-Income Workers are those who earn a monthly salary below the 25th percentile ($2000 SGD prior to 2020). Low Income Workers receive up to a 25% subsidy on their fare costs. The Central Provident Fund automatically determines eligibility based on tax filings and sends an individual application package informing workers of their eligibility, details of the scheme, and how to apply for the subsidized card (Ministry of Transport 2014). We link all bus and train stops to subzones as delineated by the Urban Redevelopment Authority which is the smallest geographic unit across our datasets. Singapore is divided into 323 subzones with a median size of 1,229,894 square meters. These are contained within larger spatial units including 55 planning areas and 5 planning regions.

We use our fare card data to generate work and consumption travel probabilities conditional on residential subzone. We restrict our data set to 3 million fare cards with at least 20 trips over our panel.[9] We include only Adult and Low-Income fare cards in order to capture the working population, dropping students, people with disabilities, those in the military, and the elderly. First we identify each individual's residence as the modal first origin and last destination of the day, where each person typically starts or ends the day. Next, we identify each individual's workplace as the modal destination during the morning rush hour

[8] These are called EZ-Link cards.

[9] We get similar probabilities with different trip thresholds for inclusion in our data set.

(5am to 11 am) and origin during evening rush hour (3pm to 11pm). Finally, we classify all remaining trips as consumption trips.

We also use data from several administrative sources. The General Household Survey (GHS) 2015 from the Department of Statistics provides detailed information on population, employment, income, and demographics at the subzone level. The GHS is conducted in between the Population Censuses which are conducted once in ten years covering a wider range of topics. Most data is compiled from administrative records across multiple sources, and additional information not available from administrative sources is collected from a sample survey of over 30,000 households. The Labor Force Survey 2018 provides data on employment and wages by industry. The Household Expenditure Survey 2018 provides detailed data on household and worker expenditures by income bracket. We classify itemized expenditures on goods and services into tradables and non-tradables. The REALIS dataset from the Urban Redevelopment Authority provides data at the subzone level on residential and commercial land use (broken down by sector) as well as average rent per square meter by commercial and residential land. Housing and Development Board (HDB) transaction data provides the universe of HDB flat sales with information on price, address, flat size, and number of rooms between 1999 and 2019.

Finally, we also collect spatial data on amenities from various sources. We have two cross-sectional data sets on the universe of licensed food establishment in Singapore from the Singapore Food Authority for 2015 and 2018. We have data on all supermarkets and hawker centers in Singapore from the National Environment Agency.[10] Ministry of Education provides data on all schools. Singapore Land Authority provides data on all parks and community clubs. Ministry of Health provides data on all clinics. We geocode all data and link each address or coordinates to subzones

[10]We present a map of the data in Appendix Figure A7.

# 3 Reduced Form Results

In this section, we present reduced form results on the impact of the Downtown Line and motivating facts which guide and motivate our structural model. In all our analysis, we define low-income workers as those earning below the 25-percentile ($2000 SGD or $1500 USD which is less than half of median earnings)[11] and high-income workers as those earning above, consistent with our farecard data set. Our unit of analysis is the subzone, henceforth a neighborhood.

## 3.1 Residential patterns

Low and high income groups have different residential patterns across the city. Figure 1 plots the share of high-income workers that live in each neighborhood in 2015. High income workers primarily cluster near the center of the city (eg. Bukit Timah, Tanglin) with only some concentration in certain neighborhoods near the coasts of the island (eg. East Coast Park). On the other hand, low income workers live further from the city center, with many living towards to north. Figure 1 also shows that the Downtown Line runs through some of the neighborhoods with the highest concentration of high-income workers in Singapore, clustered in or just west of the city center. The many low income residents in the north are not directly served by the Downtown Line. This suggests that high-income workers may be disproportionately benefit from the DTL from greater direct service.

## 3.2 Employment patterns

We find that neighborhoods which employ many low income workers are intensive in non-tradables sector production. Figure 2 plots the relationship between the share of commercial land used by non-tradable sectors and the share of workers that are low-income. We find a strong positive correlation as expected. Labor Force Survey data shows that 51% of workers

[11] Singapore's median income in 2019 was $4,563.

in the non-tradables sector are low income, while making up only 25% of the labor force, see Table 1. Specifically, 61% of employees in the food and accommodation, 49% in retail, and 41% in personal services are low-income. Only 18% of workers in the tradables sector are low income. Appendix Figure A8 plots the share of high-income workers that work in each neighborhood. This suggests that the spatial distribution of non-tradable jobs is highly important for low-income workers.

### 3.3 Consumption patterns

Using itemized Household Expenditure Survey data broken down by income group, we find that high-income consumers spend more on non-tradables than low-income consumers. Figure A9 shows that high-income workers spend 67% of their income on non-tradable goods and services, 14% on tradables and the remaining on housing. In contrast, low-income workers spend just 57% of their income on non-tradable goods and services and 20% on tradables. This implies that improvements in access to consumption opportunities will improve welfare more for high-income than low-income workers.

### 3.4 Travel patterns

We find that low income workers commute shorter distances at a median of 25 minutes, than high income workers at a median of 30 minutes. However, consumption trips are more similar in distances across income groups and shorter than workplace trips at around 23 minutes for high-income workers and 22 minutes for low-income workers. Figure 3 plots the travel time distributions by worker group and type of travel. Low-income workers make slightly more, 5% more, trips on average than high-income workers. Consumers also make more trips on the weekend than weekdays. Figure A10 presents the average number of daily trips by worker group and weekday vs weekend.

### 3.5 Impact of the DTL on Housing Prices

We find that the Downtown Line had a large impact on residential prices.

First, we use a spatial differences-in-differences framework. We exploit the timing of the announcement of Downtown Line (DTL) alignment on July 15, 2008.[12] Homebuyers cannot anticipate precisely where the DTL will be constructed before the announcement, but will incorporate information on distance to DTL line stations into prices after. Figure 4 plots the log difference between residential prices within 0 to 1 km of a DTL station and prices between 1 to 5km of a DTL station over time.[13] The price difference trend is flat before the announcement. After the alignment announcement, prices within 1km on the DTL begin increasing relative to those between 1 to 5km. Table A3 presents regression estimates of the following specification:

$$\log(Price_{it}) = \beta_0 + \beta_1 Post_t + \beta_2 Close_i + \beta_3 Post_t * Close_i + \alpha_i + \epsilon_{it} \quad (1)$$

where $i$ is an apartment building - number of rooms pair[14] and $t$ is the quarter. $Price_{it}$ is the price of flat $i$ in quarter $t$, $Post_t$ equals one for all quarters after the announcement and zero otherwise, and $Close_i$ equals one for all flats within 1 km of a DTL station and 0 or all flats between 1 and 5 km of a DTL station. $\alpha_i$ is a fixed effect for the apartment building-number of rooms pair, and $\epsilon_{it}$ is an error term. We estimate that over 4 years, residential prices increase by 4.84%, all else equal.

We also consider the relationship between prices and distance from the DTL over time. We estimate the following specification:

$$\log(Price_{it}) = \tau_t + \sum_w \gamma_w * Distance_i * \mathbb{1}\{t = w - T\} + \epsilon_{it} \quad (2)$$

where $Distance_i$ is the distance in kilometers from the nearest DTL stop. We restrict

[12] Here we focus on phase 2.

[13] We use HDB data on flats for prices.

[14] We observe addresses but not the unit number, and observe the number of rooms in the flat.

the sample to include apartments within 5km of a DTL stop. $\mathbb{1}\{t = w - T\}$ is an indicator for prices $w$ years relative to the announcement on quarter $T$. We plot the coefficients $\gamma$ in Figure 5. Consistent with our previous result, time trend of the relationship between prices and distance to the nearest DTL station is flat before the announcement and decreases over time after the announcement. Prices are higher the closer flats are to the DTL after the announcement.

Increased property prices indicate that residents highly value improvements in access from the DTL. These price increases reflects a combination of both improved access to work and consumption locations. We find that the price effects increase over time which suggests long-run dynamic responses with the re-organization of economic activity, motivating our general equilibrium structural model.

### 3.6 Impact of the DTL on Travel

We find a sharp uptick in travel to DTL subzones in response to the line opening. Figure 6 plots the daily volume of trips to or from subzones with and without a DTL station between December 2015 and February 2016 by low- and high-income workers, with Phase 2 of the DTL opening on December 27th 2015. We observe about a 20% increase in trips after the opening. Over all sub-zones, the volume of low-income trips increases by a smaller percentage than high-income trips. This suggests that high-income workers are taking greater advantage of the new DTL than low-income workers. We also find that low- and high-income workers adjust their travel patterns differently across neighborhoods. In Figure 7a and 7b, we plot the travel patterns by income group for two of the major locations on the DTL line separately, Bukit Panjang and Upper Bukit Timah. High-income workers increase their travel to and from Bukit Panjang by about 40% on average, while low-income workers increase their travel volume by about 20%. On the other hand, both worker groups increase their travel to and from Upper Bukit Timah by the same amount. This suggests that it is important for a model to account for differential travel response in evaluating welfare across worker groups.

Last, we plot changes in travel patterns over a longer time horizon in Figure 8. We find that the responses in travel flow are dynamic, growing over time. Again, this suggests that accounting for long-run responses with the re-organization of economic activity may be important, motivating our general equilibrium structural model.

### 3.7 DTL and Non-Tradable Production

Using two cross-sections of food establishment license data between 2015 (before the opening of Phase 2 and 3 of the DTL) and 2018 (after the opening), we find that there is a greater growth in food establishments between 2015 and 2018 in subzones with a DTL station. The number of food establishments in DTL locations increased by about 15% relative to non-DTL locations, see Appendix Table A2.

## 4 Quantitative Model

This section presents a model of the internal structure of a city with heterogenous workers, and travel choice over workplace and consumption locations.

### 4.1 Setup

We model a city as $n \in \mathbb{N}$ neighborhoods. Neighborhoods differ in their exogenous amenities, productivities, residential and commercial floor space, and the time it takes to commute to any other location. High- and low-income workers decide where to live, where to consume non-tradable goods and services, and where to work and in which sector. Each worker type has different preferences, productivities, wages, travel costs, and consumption shares. Utility is derived from the consumption of a tradable good, a non-tradable good, and residential floor space. There are two sectors $G \in \{T, N\}$, the tradable and non-tradable sectors. Firms are located across the city and produce using labor and commercial floorspace. Sectors differ in their demand for different worker types, with the non-tradable sector relying more low-income workers and the tradable sector relying more on the high-income workers. Demand for non-tradable sector production depends on where consumers choose to travel to consume non-tradables, while tradable goods are costlessly traded across the city. Demand for labor by worker type varies across the city based on the productivity of each sector in each location, commercial rents and demand for production. A competitive land sector supplies floor space using land and capital with constant returns to scale technology. In equilibrium, the price of floor space, wages, and prices of goods adjust to clear the goods, land and labor markets. We consider a closed city with an exogenous city population by worker group.[15]

## 4.2 Workers

The city is populated by different worker groups indexed by $\theta \in \{+,-\}$ (high- and low-income) with a fixed population $R^\theta$. A worker $\omega$ in group $\theta$ chooses a location $n$ in which to live, a location $i$ in which to work, and a location $j$ in which to consume non-tradable goods and services. We assume indirect utility is Cobb-Douglas and is evaluated according to:

$$u^\theta_{niGj}(\omega) = \frac{B^\theta_n b^\theta_n(\omega)}{Q_n^{\alpha^{H\theta}} {P^T_n}^{\alpha^{T\theta}}} \times \frac{w^\theta_{iG} a^\theta_{iG}(\omega)}{\exp(\kappa^\theta \tau_{ni})} \times \frac{s^\theta_j(\omega)}{{P^N_j}^{\alpha^{N\theta}} \exp(\kappa^\theta \tau_{nj})}, \tag{3}$$

where $0 < \alpha^{T,\theta}, \alpha^{H,\theta}, \alpha^{N,\theta} < 1$ and $\alpha^{T,\theta} + \alpha^{H,\theta} + \alpha^{N,\theta} = 1$.

Individuals derive utility from consumption of residential floorspace, consumption of a tradable good, and consumption of a non-tradable good. We accommodate different preferences, productivities, wages, travel costs and consumption shares by type. $B^\theta_n$ represents common residential amenities in neighborhood $n$ for workers of type $\theta$. Differences in residential preferences across groups will drive worker types to live in different neighborhoods. $Q_n$ is the price of residential floor space in neighborhood $n$. $P^T_n$ is the price of the tradable good in neighborhood $n$. We assume that the tradable good is traded without cost in the city so that $P^T_n = P^T = 1$, serving as our numeraire. $P^N_j$ is the price of the non-tradable

[15]Singapore is a small city-state with little in- and out-migration.

good in neighborhood $j$. $w^{\theta}_{iG}$ is the wage per efficiency unit in neighborhood $i$ in sector $G$ for workers of type $\theta$. Iceberg travel costs to work in neighborhood $i$ and consume non-tradables in neighborhood $j$, $\exp(\kappa^{\theta}\tau_{ni})$ and $\exp(\kappa^{\theta}\tau_{nj})$, increases with the time $\tau$ it takes to travel between neighborhoods. The parameter $\kappa^{\theta}$ controls the size of these travel costs by worker group. We expect differences in travel costs since low-income workers receive subsidies on travel. Higher travel costs imply larger gains from reductions in travel time. High- and low-income workers have different consumption shares over housing, $\alpha^{H,\theta}$, tradables, $\alpha^{T,\theta}$, and non-tradables, $\alpha^{N,\theta}$. These shares measure the relative importance of housing, tradable and non-tradable consumption for utility for people of type $\theta$. Larger consumption shares on housing imply larger implications of changes in rents for welfare. Larger consumption shares on non-tradales imply larger implications of access to consumption opportunities for welfare.

Worker $\omega$ has idiosyncratic draws for residential amenities in neighborhood $n$, $b^{\theta}_n(\omega)$, efficiency units of labor (productivity) in neighborhood $i$ and sector $G$, $a^{\theta}_{iG}(\omega)$, and consumption amenities in neighborhood $j$, $s^{\theta}_j(\omega)$). We assume that these idiosyncratic shocks are drawn from the following independent Frechet distributions:

$$\begin{aligned} F^{R\theta}_n(b) &= \exp(-T^{R\theta}_n b^{-\epsilon^{R\theta}}) \\ F^{L\theta}_{iG}(b) &= \exp(-T^{L_I\theta}_i T^{L_G\theta}_G a^{-\epsilon^{L\theta}}) \\ F^{S\theta}_j(b) &= \exp(-T^{S\theta}_j s^{-\epsilon^{S\theta}}) \end{aligned} \tag{4}$$

where all the shape parameters, $\epsilon$, are greater than 1 and all scale parameters, $T$, are greater than zero. The scale parameters, $T^{R\theta}_n, T^{L_I\theta}_i, T^{L_G\theta}_G, T^{S\theta}_j$ control the overall level of the draws for residential preferences, work location productivity, sector productivity, and consumption preferences respectively. We allow scale parameters to vary across worker groups, to capture differences in preferences and productivities over locations across types. We also allow shape parameters, $\epsilon^{R\theta}, \epsilon^{L\theta}, \epsilon^{S\theta}$, which control the dispersion of the distributions to differ across groups. A higher $\epsilon$ corresponds to a smaller dispersion. The sensitivity of choices to other variables such as travel costs is governed by the dispersion of preferences or productivity. When workers have similar matches in different locations (high $\epsilon$), choices are more sensitive to these other variables. Differences in heterogeneity across groups will be important in determining the incidence of travel costs, since it controls the extent to which individuals are willing to bear high travel costs to work or consume in a location.

**Timing.** Workers first choose where to live, and then choose where to work and where to consume non-tradables. First, individuals observe their idiosyncratic residential amenities draws for all neighborhood and chooses to live in some neighorhood $n$. Second, individuals observe their idiosyncratic productivities for all workplaces and idiosyncratic consumption preferences in all neighborhoods, then choose to work in some neighborhood $i$ and sector $G$, and to consume non-tradables in some neighborhood $j$. We solve the worker problem by backward induction.

### 4.2.1 Employment Location Decision

Having chosen where to live $n$, individuals draw a vector of match-productivities with firms across the city as in Equation 4. With these draws in hand, workers choose to work in the neighborhood that offers the highest income net of commute costs by type.

$$\max_{i,G} \frac{w_{iG}^{\theta} a_{iG}^{\theta}(\omega)}{\exp(\kappa^{\theta} \tau_{ni})}. \tag{5}$$

Properties of the Frechet distribution imply that the probability a worker of type $\theta$ who has made the choice to live in $n$ decides to work in $i$ and sector $G$ is given by the following gravity equation:

$$\lambda_{niG|n}^{L\theta} = \frac{T_i^{L_I\theta} T_G^{L_G\theta} [w_{iG}^{\theta} \exp(-\kappa^{\theta} \tau_{ni})]^{\epsilon^{L\theta}}}{\sum_l \sum_H T_l^{L_I\theta} T_H^{L_G\theta} [w_{lH}^{\theta} \exp(-\kappa^{\theta} \tau_{nl})]^{\epsilon^{L\theta}}} \tag{6}$$

Individuals are more likely to travel to work in a neighborhood that pays a high wage net of travel costs, as in the numerator, relative to those in all other locations, as in the denominator. The sensitivity of employment decisions to commute costs is governed by the dispersion of productivities across neighborhoods. Similar matches across workplaces in different locations and sectors (high $\epsilon$) imply that choices are more sensitive to commute costs. Differences in productivity heterogeneity across worker types is important in determining the incidence of travel costs, controlling the extent to which workers are willing to bear high commute costs to work in a neighborhood. Differences in productivity across neighborhoods and commute costs by type drives differences in work travel patterns across worker groups.

Expected income prior to drawing the vector of match productivities is directly related to the denominator in Equation 6 through

$$\mathbb{W}_n^\theta \triangleq \mathbb{E}_n^{L\theta}[v] = \Gamma\left(\frac{\epsilon^{L\theta}-1}{\epsilon^{L\theta}}\right)\left[\sum_l \sum_H T_l^{L_I\theta} T_H^{L_G\theta}[w_{lH}^\theta \exp(-\kappa^\theta \tau_{nl})]^{\epsilon^{L\theta}}\right]^{1/\epsilon^{L\theta}} \tag{7}$$

where $\Gamma(.)$ is the Gamma function. Intuitively, in locations with better *access to employment*, or access to locations with high expected income, by type, workers are better off.

#### 4.2.2 Consumption Location Decision

Having chosen where to live $n$, individuals draw a vector of idiosyncratic preference shocks with consumption locations across the city as in Equation 4. With these draws in hand, workers choose to consume non-tradables in the neighborhood with the best consumption amenities net of travel costs and the price of non-tradables.

$$\max_j \frac{s_j^\theta(\omega)}{P_j^{N^{\alpha^N}} \exp(\kappa\tau_{nj})} \tag{8}$$

Properties of the Frechet distribution imply that the probability a worker of type $\theta$ who has made choice to live in $n$ decides to consume non-tradables in $j$ is given by the following gravity equation:

$$\lambda_{nj|n}^{S\theta} = \frac{T_j^{S\theta}[P_j^{N^{\alpha^N}} \exp(\kappa\tau_{nj})]^{-\epsilon^{S\theta}}}{\sum_m T_m^{S\theta}[P_m^{N^{\alpha^N}} \exp(\kappa^\theta\tau_{nm})]^{-\epsilon^{S\theta}}} \tag{9}$$

Individuals are more likely to travel to consume in a neighborhoods with high consumption amenities net of the price of non-tradables and travel costs, as in the numerator, relative to those in all other locations, as in the denominator. The sensitivity of consumption decisions to travel costs is governed by the dispersion of preferences across neighborhoods. Similar amenities across consumption locations (high $\epsilon$) imply that choices are more sensitive to travel costs. Differences in preference heterogeneity across worker types is important in determining the incidence of travel costs, controlling the extent to which workers are willing to bear high travel costs to consume non-tradables in a neighborhood. Differences in preferences across neighborhoods and travel costs by type drives differences in consumption travel patterns across worker groups.

Expected utility from non-tradable consumption prior to drawing the vector of idiosyncratic preferences is directly related to the denominator in Equation 10 through

$$\mathbb{S}_n^\theta \triangleq \mathbb{E}_n^{S\theta}[\gamma] = \Gamma\left(\frac{\epsilon^{S\theta}-1}{\epsilon^{S\theta}}\right)\left[\sum_m T_m^{S\theta}[P_m^{N^{\alpha^N}} \exp(\kappa^\theta \tau_{nm})]^{-\epsilon^{S\theta}}\right]^{\alpha^{N\theta}/\epsilon^{S\theta}} \tag{10}$$

where $\Gamma(.)$ is the Gamma function. Intuitively, in locations with better *access to consumption*, or access to locations with low prices and high amenities, by type, workers are better off.

### 4.2.3 Residential Location Decision

In the first stage, individuals choose where to live to maximize their expected indirect utility after observing their idiosyncratic residential amenities draws across all neighborhoods. Workers of type $\theta$ solve the following problem

$$\max_n U_n^\theta(\omega) = \frac{B_n^\theta b_n^\theta(\omega)}{Q_n^{\alpha^H}} \mathbb{W}_n^\theta \mathbb{S}_n^\theta \tag{11}$$

Workers are attracted to locations with high residential amenities, low housing prices, high net incomes, and utility from consumption of non-tradables.

Properties of the Frechet distribution imply that the share of type-$\theta$ workers who live in neighborhood $n$ is

$$\lambda_n^{R\theta} = \frac{T_n^{R\theta}[P_n^{T^{-\alpha^T}} Q_n^{-\alpha^H} B_{n\theta} \mathbb{W}_{n\theta} \mathbb{S}_{n\theta}]^{\epsilon^{R\theta}}}{\sum_k T_k^{R\theta}[P_k^{T^{-\alpha^T}} Q_k^{-\alpha^H} B_{k\theta} \mathbb{W}_{k\theta} \mathbb{S}_{k\theta}]^{\epsilon^{R\theta}}} \tag{12}$$

Similar residential amenities across neighborhoods (high $\epsilon$) imply that choices are more sensitive to changes in access to employment, access to consumption, and rents. Differences in preferences across neighborhoods by type drives differences in residential patterns in the city across worker groups.

#### 4.2.4 Expected Utility

Since we consider a "closed city", expected utility from living in the city prior to drawing the vector of idiosyncratic preferences is directly related to the denominator in Equation 12 through

$$\bar{U}^{\theta} = \Gamma\left(\frac{\epsilon^{R\theta} - 1}{\epsilon^{R\theta}}\right) \left[\sum_k T_k^{R\theta}[P_k^{T^{-\alpha^T}} Q_k^{-\alpha^H} B_{k\theta} \mathbb{W}_{k\theta} \mathbb{S}_{k\theta}]^{\epsilon^{R\theta}}\right]^{1/\epsilon^{R\theta}}. \tag{13}$$

### 4.3 Firms

#### 4.3.1 Technology

There is a representative firm for the non-tradable and non-tradable sectors, $G \in \{N, T\}$, in each neighborhood $i \in \mathbb{N}$. Firms produce under perfect competition and constant returns to scale. Firms produce using a Cobb-Douglas technology over labor and commercial floor space with output

$$Y_{iG} = A_{iG} L_{iG}^{\beta^G} H_{iG}^{1-\beta^G} \tag{14}$$

where $\beta^G \in [0, 1]$, $H_{iG}$ is commercial floor space, and labor input is a Cobb-Douglas aggregate over each worker group's effective labor, $\tilde{N}_{iG\theta}$

$$L_{iG} = \prod_{\theta} \tilde{N}_{iG\theta}^{\beta^{G\theta}} \tag{15}$$

where $\sum_\theta \beta^{G\theta} = 1$. Each sector and neighborhood has different productivities, $A_{iG}$. Sectors differ in the intensity in which they use different types of workers $\beta^{G\theta}$. The tradable sector requires a higher share of high-income workers, while the non-tradable sector heavily relies on low-income workers. Thus, the spatial distribution of non-tradable versus tradable production is strongly related to that of low-income versus high-income jobs, impacting workers' access to employment across neighborhoods by type.

#### 4.3.2 Factor Demand

Perfect competition implies that profits are zero and that the price of each variety is equal to its marginal cost

$$P_i^G = A_{iG}{}^{-1} W_{iG}^{\beta^G} q_i^{1-\beta^G} \tag{16}$$

where $q_i$ is the price of commercial floor space in $i$ and

$$W_{iG} = w_{iG+}^{\beta^{G+}} w_{iG-}^{\beta^{G-}} \tag{17}$$

is the cost of labor for sector $G$ in location $i$ where $w_{iG\theta}$ is the wage per efficiency unit for a worker of type $\theta$. Wages are different across both sectors and locations, with each sector and location pair facing an upward-sloping supply function for effective units of labor for each worker type. Solving the firm's profit maximization problem implies that the demand for labor and commercial floorspace is

$$L_{iG} = \beta^G P_{iG} Y_{iG} / W_{iG} \tag{18}$$

$$H_{iG} = (1 - \beta^G) P_{iG} Y_{iG} / q_{iG} \tag{19}$$

$$\tilde{N}_{iG\theta} = \beta^{L\theta,G} L_{iG} W_{iG} / w_{iG\theta} \tag{20}$$

## 4.4 Land Market

Following Ahlfeldt et al. (2015), the land market equilibrium requires no-arbitrage between the commercial and residential use of floor space after the tax equivalent of land use regulations. The commercial price of floor space for both the tradeable and non-tradable sector is

$$q_i = \xi_i Q_i \tag{21}$$

where $\xi_i$ captures one plus the tax equivalent of land use regulations that restrict commercial land use relative to residential land use. We allow this wedge between commercial and residential floor prices to vary across neighborhoods.

Floor space is supplied by a perfectly-competitive developers using land, $K_i$, and capital, $M_i$, with constant returns scale:

$$H_i = M_i^{\mu_H} K_i^{1-\mu_H} \tag{22}$$

where $H_i$ is total floor space and $\mu_H$ is the share of land in floor space production. Therefore, the corresponding dual cost function for floor space is

$$Q_i = \mu_H^{-\mu_H}(1-\mu_H)^{-(1-\mu_H)}\mathbb{P}^{\mu_H}\mathbb{R}_i^{1-\mu_H} \tag{23}$$

where $\mathbb{P}$ is the common price for capital across all neighborhoods, and $\mathbb{R}_i$ is the price for land. Cost minimization implies that

$$Q_i = \mathbb{P} K_i^{\frac{\mu_H - 1}{\mu_H}} H_i^{\frac{1-\mu_H}{\mu_H}} \mu_H^{-1} \tag{24}$$

As non-tradable production and demand for floor space increases in response to increased consumption travel, increased rents may drive out tradable production and residents, changing commercial and residential spatial patterns across the city.

## 4.5 Market Clearing

**Land.** Demand for residential floor space is

$$H_n^R = \frac{\sum_\theta \alpha^{H,\theta} R^\theta \lambda_n^R \mathbb{W}_n}{Q_n} \tag{25}$$

where we sum over the housing expenditures for all the residents in neighborhood $n$ across worker groups, using the fact that expenditures on residential floor space is a constant share of income from Cobb-Douglas.

Market clearing for floor space requires that the total supply of floor space equals the total floor space demanded from both residents and firms in each neighborhood $j$.

$$H_j = H_j^R + \sum_{G \in \{N,T\}} H_{jG} \tag{26}$$

**Labor.** Using the commuting probabilities from Equation 6, the supply of workers to any location is found by summing over the number of residents who commute there

$$N_{iG\theta} = \sum_n \lambda_{niG|n}^{L\theta} \lambda_n^R R^\theta \tag{27}$$

Labor supply in the model takes a log-linear form that depends on two forces. First, more workers commute to destinations paying higher wages. Second, firms attract workers when they have better access to them through the commuting network. Individuals care about wages net of commute costs. Total effective labor supply to location is given by

$$\tilde{N}_{iG\theta} = \sum_n R_n \lambda_{ni|n}^{L\theta} \lambda_{niG|ni}^{L\theta} \bar{a}_{niG}^\theta \tag{28}$$

where $\bar{a}_{niG}^\theta$ is the average productivity of type-$\theta$ workers who live in $n$ and decide to work in

$i$. Using the properties of Frechet we have[16]

$$\bar{a}^{\theta}_{niG} = \Gamma(\epsilon^{L\theta})(T_i^{L_I\theta}T_G^{L_G\theta}/\lambda^{L\theta}_{niG|n})^{1/\epsilon^{L\theta}} \tag{29}$$

Market clearing requires that supply of effective labor in Equation 28 equals demand for effective labor in Equation 18 in each neighborhood and sector. Wages by worker type are endogenously determined by market clearing.

**Non-tradables.** In each neighborhood $j$, total production for non-tradables must equal total expenditures on non-tradables.

$$P_j^N A_{iN} L_{iN}^{\beta^N} H_{iN}^{1-\beta^N} = \sum_{n\in\mathbb{N}}\sum_{\theta} \alpha^{N,\theta}\lambda^{S\theta}_{nj|n}\lambda^R_n R^\theta \mathbb{W}_{n\theta} \tag{30}$$

Prices of non-tradables are endogenously determined to clear the market. We sum over the non-tradables expenditures for all the workers who travel from some neighborhood $n$ to consume in neighborhood $j$ across worker groups, using the fact that expenditures on non-tradables is a constant share of income from Cobb-Douglas. This market clearing condition implies that where workers decide to consume non-tradables is closely linked to where non-tradables are produced.

## 4.6 Externalities

### 4.6.1 Productivities

A location's productivity depends on both an exogenous component $\bar{A_{iG}}$ that reflects features independent of economic activity (e.g. access to roads, slope of land) as well as the endogenous density of employment in that location

$$A_{iG} = \bar{A}_{iG}(\sum_{\theta} N_{i\theta}/\bar{H}_j)^{\mu_A} \tag{31}$$

[16] $\bar{a}^{\theta}_{niG} = \mathbb{E}[a^{\theta}_{inG}|T_i^{L_I\theta}T_G^{L_G\theta}(w^{\theta}_{iG}\exp(-\kappa\tau_{ni}))^{\epsilon^{L\theta}} > T_l^{L_I\theta}T_H^{L_G\theta}(w^{\theta}_{lH}\exp(-\kappa\tau_{nl}))^{\epsilon^{L\theta}}, \quad \forall l, H]$

where $\bar{H}_j$ is the total units of land. The strength of agglomeration externalities is governed by the parameter $\mu_A$.

### 4.6.2 Amenities

Amenities in a neighborhood depend on an exogenous component $\bar{B}_n$ and a residential externality that depends on the endogenous ratio of high-skilled and low-skilled residents

$$B_n^\theta = \bar{B}_n^\theta (R_n^\theta / R_n^{-\theta})^{\mu_{U\theta}} \tag{32}$$

Endogenous amenities depend on demographic composition across income groups rather than the total density of residents, similar to Tsivanidis (2019). Workers are more willing to pay to live in type-specific amenity neighborhoods and by doing so increase type-specific amenities even more. Thus, this is an additional force towards segregation. We let the strength of residential externalities differ across worker groups and let the data speak to the relative strength of these forces in estimation.

## 4.7 Equilibrium

We now define the general equilibrium of the city.

**Definition** Given vectors of exogenous location characteristics $\{\bar{B}_n^\theta, \bar{A}_{iG}, \tau_{ij}, \bar{H}_j, \psi_i, \xi_i\}$, city group-wise populations $\{R^\theta\}$, and model parameters $\{\alpha, \beta, \kappa^\theta, T_n^{R\theta}, T_i^{L_I\theta}, T_G^{L_G\theta}, T_j^{S\theta}, \epsilon^{R\theta}, \epsilon^{L\theta}, \epsilon^{S\theta}, \mu_H, \mu_A, \mu_{U\theta}\}$, an equilibrium is defined as a vector of endogenous objects $\{q_i, w_{iG}^\theta, H_{iG}, \tilde{N}_{iG\theta}, \lambda_n^{R\theta}, \lambda_{niG|n}^{L\theta}, \lambda_{nj|n}^{S\theta}, \bar{U}^\theta\}$ such that:

1. **Labor Market Clearing**: The supply of labor by individuals in Equation 28 is consistent with demand for labor by firms in Equation 18.

2. **Floorspace Market Clearing**: The market for floorspace clears as in Equation 26, its price is consistent with residential populations 25 and commercial floorspace use 19, and total floor space is consistent with land developer optimality 24.

3. **Goods Market Clearing**: Non-tradable consumption matches non-tradable production in each neighborhood as in Equation 30.

4. **Closed City**: Populations add up to the city total.

## 4.8 Model Discussion

First, our model captures the direct distributional implications of changes in access to consumption, in addition to access to employment. Changes in transit affect access to consumption opportunities across worker types differently depending on where they (or where they want to) live, work and consume. If travel costs decrease between low (high) income neighborhoods and attractive consumption locations with lower prices and high consumption amenities which are type-specific, low (high) income worker groups will be better off. Access to consumption is more important to welfare for worker types with 1) larger expenditures on non-tradables, 2) larger dispersion of preferences over consumption locations (more inelastic), and 3) higher travel costs.

Next, consumption travel shapes the residential patterns of city across types. Workers now trade off access to non-tradables, along with access to employment and higher rents in their residential decisions. Worker types with less dispersion in residential amenities are better able to take advantage of both improved access to consumption and employment in other locations.

Last, incorporating travel to the consumption of non-tradables also has distributional implications for access to employment across worker groups. In response to changes in transit access, consumers change where they consume non-tradables, trading off travel costs and type-specific amenities and prices. This results in a change in relative factor demand across locations. Since non-tradable firms hire more low-income workers, low-income jobs move to where residents are traveling more for consumption and demanding more non-tradable production. High-income tradable jobs may also move away from locations with greater consumption travel, since commercial rents will rise in response to greater non-tradable production. Thus, low (high) income jobs could move to locations closer or further from where low (high) income workers live, locations with higher or lower wages for low (high) income workers, or locations which are more or less productive for low (high) income workers. This re-organization of commercial activity between the tradable and non-tradable sectors across space changes each worker group's expected access to employment, or expected income net of commuting, differentially and could increase or decrease inequality.

# 5 Estimation

This section estimates the model from Section 4. First, we estimate workplace and consumption preferences and access across neighborhoods using gravity specifications from the model across worker groups. Second, we estimate work and consumption travel costs and elasticities by type. Third, we estimate parameters which govern the dispersion of idiosyncratic residential amenities and residential spillovers by type. Fourth, we estimate residential amenities across neighborhoods across high- and low-income workers. Fifth, we show how the model can be inverted to obtain unobservable endogenous variables that rationalize the observed data as an equilibrium of the model. Last, we describe the calibration of remaining parameters to data or literature.

## 5.1 Estimating Consumption Preferences and Access

Taking logs of the expression for consumption travel flows in Equation 9 combined with the specification of travel costs $\phi^{S\theta} = \epsilon^{S\theta}\kappa^{\theta}$ provides a gravity equation relating variation in work travel flows to variation in travel times for each worker type $\theta \in \{+, -\}$

$$\ln \lambda^{S\theta}_{nj|n} = \eta^{S\theta}_j - \mu^{S\theta}_n - \phi^{S\theta}\tau_{nj} + u^{S\theta}_{nj} \tag{33}$$

where $\eta^{S\theta}_i$ and $\mu^{S\theta}_n$ are fixed effects, and $u^{S\theta}_{ni}$ is an unobserved component of consumption travel costs. We define $\tau$ as the average time between neighborhoods in the farecard data, and construct work travel shares from identified residences and consumption trips in the farecard data. We instrument for $\tau$ with the linear distance between neighborhoods. Columns 3 and 4 of Table 2 reports the results. We estimate a negative and statistically significant semi-elasticity of consumption travel flows with respect to travel time of $\hat{\phi}^{L+} = -0.070$ for high-income workers and $\hat{\phi}^{L+} = -0.094$ for low-income workers. We find that low-income workers are more sensitive to changes in travel time, with a larger gravity coefficient. In the next subsection, we will show that gravity coefficients are larger for consumption travel than work travel across both groups.

The consumption location fixed effect from estimation gives us measures of consumption attractiveness across neighborhoods

$$\exp(\hat{\eta}_j^{S\theta}) = T_j^{S\theta} P_{jN}^{-\alpha^N \epsilon^{S\theta}} \tag{34}$$

which captures average consumption amenities by type and prices of non-tradables in each neighborhood $i$. Figure 9a presents a scatter plot of consumption preferences by high- and low-income workers. We observe substantial differences in where low and high income groups prefer to consume non-tradables. In Table 8, we find that consumption attractiveness for both groups is highly correlated with various measures of retail amenities from external data. Preferences are positively correlated to the density of food establishments, supermarkets and clinics of in neighborhood. We find that low-income workers prefer locations where low-cost options make up a larger share of food establishments, such as hawker stalls and food courts, while the coefficient is insignificant at the 5% level for high-income workers.

The residence fixed effect from estimation gives us

$$\exp(\hat{\mu}_n^{S\theta}) = \sum_{m\in} T_m^{S\theta} P_m^{N-\alpha^N \epsilon^{S\theta}} \exp(-\epsilon^{S\theta} \kappa \tau_{nm}) \tag{35}$$

which allow us to recover access to consumption, $\mathbb{S}_n^{\theta}$, across neighborhoods given $\epsilon^{S\theta}$, capturing expected utility from consumption in each neighborhood $i$.

### 5.2 Estimating Workplace Preferences and Access

Taking logs of the expression for work travel flows in Equation 6 combined with the specification of travel costs $\phi^{L\theta} = \epsilon^{L\theta}\kappa^{\theta}$ provides a gravity equation relating the variation in work travel flows to variation in travel times for each worker type $\theta \in \{+, -\}$

$$\ln \lambda_{ni|n}^{L\theta} = \eta_i^{L\theta} - \mu_n^{L\theta} - \phi^{L\theta}\tau_{ni} + u_{ni}^{L\theta} \tag{36}$$

where $\eta_i^{L\theta}$ and $\mu_n^{L\theta}$ are fixed effects, and $u_{ni}^{L\theta}$ is an unobserved component of work travel costs. Again, we define $\tau$ as the average time between neighborhoods in the farecard data, and construct work travel shares from identified residences and workplaces in the farecard data. We instrument for $\tau$ with the linear distance between neighborhoods. Columns 1 and 2 of Table 2 reports the results. We estimate a negative and statistically significant semi-elasticity of work travel flows with respect to travel time of $\hat{\phi}^{L+} = -0.51$ for high-income workers and $\hat{\phi}^{L+} = -0.53$ for low-income workers. Again, we find a larger gravity coefficient for low-income workers, indicating that low-income workers are more sensitive to travel time. Comparing our estimates for work to that of consumption, we find smaller gravity coefficients for work travel than consumption travel across both groups.

The workplace fixed effect from estimation gives us measures of workplace attractiveness across neighborhoods

$$\exp(\hat{\eta}_i^{L\theta}) = \sum_G T_i^{L_I\theta} T_G^{L_G\theta} w_{iG}^{\epsilon^{L\theta}} \tag{37}$$

which captures average productivity and wages in each neighborhood $i$ by type. Figure 9b presents a scatter plot of workplace preferences by high- and low-income workers. We observe substantial differences in where low and high income groups prefer to work. In Table 4, we find that workplace attractiveness for both groups is highly correlated with employment density.

The residence fixed effect from estimation gives us

$$\exp(\hat{\mu}_n^{L\theta}) = \sum_{l \in \mathbb{N}} \sum_{G} T_i^{L_I\theta} T_G^{L_G\theta} w_{lG}^{\epsilon^{L\theta}} \exp(-\epsilon^{L\theta} \kappa^{\theta} \tau_{nl}) \tag{38}$$

which allow us to recover access to employment, $\mathbb{W}_n^{\theta}$, across neighborhoods given $\epsilon^{L\theta}$, capturing the expected wages net of commuting in each neighborhood $i$.

## 5.3 Parameter Estimation: Travel Elasticities and Costs

### 5.3.1 Dispersion of Workplace Idiosyncratic Productivity

Taking logs of the expression for expected income by worker type in Equation 7, we can estimate the workplace dispersion parameter, $\epsilon^{L\theta}$, by regressing the estimated residence fixed effects, $\mu_n^{L\theta}$, on the log of average residential income from observed administrative data by type.

$$\mu_n^{L\theta} = \epsilon^{L\theta} \log(\text{Residential Income}_n) + e_n \tag{39}$$

Table 5 presents our estimation results. We find a strong positive correlation between our model's estimate of expected income with that of non-targeted data. We show in Figure 10, similar to Miyauchi, Nakajima, and Redding (2020), that the relationship is log-linear as consistent with the model except for outliers in a few high-income neighborhoods, reflecting non-labor income at high income levels. in our preferred specification we drop the four richest neighborhoods. Thus, we estimate a workplace idiosyncratic dispersion of $\epsilon^{L+} = 2.912$ for high-income workers and $\epsilon^{L-} = 5.023$ for low-income workers. High-income workers have more inelastic work travel patterns than low-income workers. This is consistent with Tsivanidis (2019). Thus, high-income workers benefit more from reductions in travel costs as they are less able to substitute to more attractive work locations in equilibrium.

### 5.3.2 Dispersion of Idiosyncratic Consumption Amenities

Next, we can estimate the consumption travel dispersion parameter by worker type, $\epsilon^{S\theta}$, using the estimated semi-elasticity for consumption and work travel, $\phi^{S\theta}$ and $\phi^{L\theta}$:

$$\epsilon^{S\theta} = \epsilon^{L\theta}\phi^{S\theta}/\phi^{L\theta} \tag{40}$$

We estimate a consumption idiosyncratic dispersion of $\epsilon^{S+} = 3.00$ for high-income workers and $\epsilon^{S-} = 5.81$ for low-income workers, see Table 6. Again, High-income workers have more inelastic consumption travel patterns than low-income workers. Thus, high-income workers benefit more from reductions in travel costs as they are less able to substitute to more attractive consumption locations in equilibrium. Dispersion is larger for consumption trips and workplace trips.

### 5.3.3 Travel Costs

Finally, we can back out the travel cost parameter by worker type:

$$\kappa^{\theta} = \phi^{S\theta}/\epsilon^{S\theta} \tag{41}$$

We estimate a travel cost parameter of $\kappa^{+} = 0.018$ for high-income workers and $\kappa^{+} = 0.014$ for low-income workers, see Table 6. Low-income workers have smaller travel costs. The wedge ($0.014/0.018 = 0.778$) between travel costs is consistent with the 25% subsidy low-income workers receive on their fare costs.

## 5.4 Parameter Estimation: Dispersion of Idiosyncratic Residential Amenities and Externalities

We exploiting the fact that changes in market access induced by the Downtown line provide a shock to the supply of labor and residents across the city. We estimate the model in

2015 and 2018 before and after the opening of the DTL. By log-linearizing Equation 12 and substituting in Equation 32, we can estimate $\mu_{U\theta}$ and $\epsilon^{R\theta}$ with the following equation.

$$\Delta \ln \lambda_n^{R\theta} = \epsilon^{R\theta} \Delta \ln Q_n^{-\alpha^{H\theta}} \mathbb{W}_n^\theta \mathbb{S}_n^\theta + \epsilon^{R\theta} \mu_{U\theta} \Delta \ln (R_n^\theta / R_n^{-\theta}) + \Delta \ln e_{n\theta} \tag{42}$$

where $\Delta$ indicates the difference in estimated model unobservables or data before and after the line opening. We present the estimates in Table 7. We find that dispersion of idiosyncratic residential amenities is slightly larger for high-types at $\epsilon^{R+} = 1.48$ than low-types at $\epsilon^{R-} = 1.38$. This implies that high-income workers are more able to take advantage of improvements in access across neighborhoods from moving. We find stronger externalities for low-types at $\mu_{U-} = 0.45$ than high types at $\mu_{U+} = 0.22$), which are consistent with Tsivanidis (2019). This implies that there are stronger spillovers in endogenous amenities for low income workers, and hence a stronger force towards clustering in the same neighborhoods.

## 5.5 Estimating Residential Amenities

We estimate composite residential amenities as

$$\mathbb{B}_{n\theta} \triangleq B_n^\theta (T_n^{R\theta})^{1/\epsilon^{R\theta}} \tag{43}$$

We are able to recover composite residential amenities using known parameters, observed land prices and travel shares, and estimated measures of access from earlier:

$$\mathbb{B}_{n\theta} = \frac{Q_n^{\alpha^H} (\lambda_n^{R\theta})^{1/\epsilon^{R\theta}} \bar{U}^\theta}{\mathbb{W}_{n\theta} \mathbb{S}_{n\theta}} \Gamma \left( \frac{\epsilon^{R\theta} - 1}{\epsilon^{R\theta}} \right)^{-1} \tag{44}$$

In Table 8, we show that residential preferences are highly correlated with external data on amenities - specifically schools, community clubs, and parks. We find that high-income residents care more about community clubs, and schools overall. However, low-income residents care more about the presence of "neighborhood schools" which are public and non-legacy,[17] and parks.

### 5.6 Model Inversion

The model contains unobserved endogenous variables which are necessary to conduct our counterfactuals in Section 6. We are able to recover unique values of these variables that rationalize the observed data as a model equilibrium. First, although we observe the workplace commuting flows between each neighborhood by worker type, we do not directly observe the share of employment in each sector, non-tradable versus tradable, conditional on residential and workplace choice by worker type. However, we are able to combine data on tradable versus non-tradable land quantities with the model structure to solve for the share of employment in each sector by location and type. Second, although we observe residential land prices and commercial land quantities and rents from administrative data sources, we do not observe residential land quantities. However, we can use the model's structure to recover residential land quantities by dividing total housing expenditures for all the residents in neighborhood by residential land prices, as in Equation 25, as a function of observable data on residential wages and population by type. The following proposition formalizes this intuition.

**Proposition 1.** ***Sector Employment by Type and Neighborhood*** *(i) Given data on residence and work travel by type and commercial land quantities by sector* $\{\lambda^{L\theta}_{ni|n}, \lambda^{R\theta}_{n}, R^{\theta}, H_{nG}\}$ *in addition to model parameters, there exists a unique vector of neighborhood-sector employment shares conditional residence by type* $\{\lambda^{L\theta}_{niG|n}\}$*, that rationalizes the observed data as an equilibrium of the model.*

***Residential Quantities*** *(ii) Given data on residence and wages by type, and residential land prices* $\{\lambda^{R\theta}_{n}, R^{\theta}, \mathbb{W}^{\theta}_{n}, Q_n\}$ *in addition to model parameters, there exists a unique vector of residential land quantities,* $\{H^{R}_{n}\}$*, that rationalizes the observed data as an equilibrium of the model.*

[17] Wealthier families typically can send their children to elite and private schools where they have legacy status or which are more inexpensive to enroll in.

## 5.7 Parameters Calibrated to Exogenous Data

The consumption shares, $\alpha$, for low- and high-income groups (below and above the 25th percentile) are calibrated to match line-itemized expenditures data from the Household Expenditure Survey 2018 which provides data by income bracket. Overall labor share, $\beta^G$, by sector is matched to estimates from the Economic Survey of Singapore 2013. Labor shares by type, $\beta^{G\theta}$, are estimated using the average share of the wage bill paid to low- and high-income workers in Singapore by sector using the Labor Force Survey 2018. We assume a share of land in construction costs of $\mu = 0.25$ following Epple, Gordon, and Sieg (2010) and Combes, Duranton, and Gobillon (2019) Ahlfeldt et al. (2015). We consider production agglomeration parameters consistent with Rosenthal and Strange (2004), Melo et al. (2009), and Ahlfeldt et al. (2015) who estimate the elasticity of producitivity, $\mu_A$, to be between 3–8 percent.[18]

# 6 Impact of the Downtown Line

In this section, we undertake counterfactuals to estimate the impact of the Downtown Line on inequality and welfare.[19] First, we present our exact-hat approach. Second, we test our counterfactual predictions on ex-post data. Third, we present our main results. Last, we conduct a decomposition exercise.

## 6.1 Exact-hat

We undertake counterfactuals with an exact-hat approach as popularized by Dekle et al. (2008). Rather than estimating our model in terms of levels, we specify the model in terms of changes from the current equilibrium. We thereby finesse having to assemble proxies for various unobservables in our model such as the prices of non-tradables, neighborhood-

[18] We use $\mu_A = 0.03$ as our preferred specification to be conservative.

[19] We focus on Phase 2 and 3 of the line due to data availability which covers over 80% of the line.

sector wages per efficiency unit by type, and mean idiosyncratic draws etc. We let $\hat{x} = x'/x$ denote the relative change in the equilibrium endogenous variable $x$, where $x'$ is the value in the new equilibrium. We consider an exogenous change in travel times, $\Delta\tau_{ij} = \tau'_{ij} - \tau_{ij}$. Given model parameters $\{\alpha, \beta, \kappa^{\theta}, \epsilon^{R\theta}, \epsilon^{L\theta}, \epsilon^{S\theta}, \mu_H, \mu_A, \mu_{U\theta}\}$, assumed bilateral changes in travel times, $\{\exp(\Delta\tau_{ij})\}$, and endogenous variables in the initial equilibrium $\{\lambda^{S\theta}_{nj|n}, \lambda^{L\theta}_{ni|n}, \lambda^{R\theta}_{n}, H_{nG}, H^R_n, \mathbb{W}\}$, we solve for the counterfactual equilibrium with the following equations.

First, we re-write Equations 9, 6 and 12 to derive expressions for changes in residential, commuting and consumption probabilities by worker type.

$$\lambda^{\hat{L\theta}}_{niG|n} = \frac{\hat{w^{\theta}_{iG}}^{\epsilon^{L\theta}} e^{-\kappa^{\theta}_{L}\Delta\tau_{ni}\epsilon^{L\theta}}}{\sum_{l\in\mathbb{N}}\sum_{H}\lambda^{L\theta}_{nlH|n}\hat{w^{\theta}_{lH}}^{\epsilon^{L\theta}} e^{-\kappa^{\theta}_{L}\Delta\tau_{nl}\epsilon^{L\theta}}} \tag{45}$$

$$\lambda^{\hat{S\theta}}_{ni|n} = \frac{\hat{P_{iN}}^{-\alpha^{N,\theta}\epsilon^{S\theta}} e^{-\kappa^{\theta}_{S}\Delta\tau_{ni}\epsilon^{S\theta}}}{\sum_{l\in\mathbb{N}}\lambda^{S\theta}_{nl|n}\hat{P_{lN}}^{-\alpha^{N,\theta}\epsilon^{S\theta}} e^{-\kappa^{\theta}_{S}\Delta\tau_{nl}\epsilon^{S\theta}}} \tag{46}$$

$$\lambda^{\hat{R\theta}}_{n} = \frac{\left(\hat{B}^{\theta}_{n}\hat{\mathbb{S}}^{\theta}_{n}\hat{\mathbb{W}}^{\theta}_{n}\hat{Q}_{n}^{-\alpha^{H}}\right)^{\epsilon^{R\theta}}}{\sum_{l\in\mathbb{N}}\lambda^{R\theta}_{l}\left(\hat{B}^{\theta}_{l}\hat{\mathbb{S}}^{\theta}_{l}\hat{\mathbb{W}}^{\theta}_{l}\hat{Q}_{l}^{-\alpha^{H}}\right)^{\epsilon^{R\theta}}} \tag{47}$$

Next, we re-write Equations 25, 19, 26, and 27 to derive expressions for changes in land and labor quantities.

$$\hat{H}^{R}_{n} = \frac{\sum_{\theta}\alpha^{H,\theta}R^{\theta}\lambda'^{R}_{n}\mathbb{W}'_{n}}{\hat{Q}_{n}\sum_{\theta}\alpha^{H,\theta}R^{\theta}\lambda^{R}_{n}\mathbb{W}_{n}} \tag{48}$$

$$\hat{H}_{jG} = \frac{\hat{\hat{N}}_{iG\theta}\hat{w}_{iG\theta}}{\hat{q}_{j}} \tag{49}$$

$$\hat{H}_j = \frac{H_j^R \hat{H}_j^R + \sum_{G \in \{N,T\}} H_{jG} \hat{H}_{jG}}{H_j^R + \sum_{G \in \{N,T\}} H_{jG}} \tag{50}$$

$$\hat{\tilde{N}}_{iG\theta} = \frac{\sum_n (\lambda'^{L\theta}_{niG|n})^{1-\frac{1}{\epsilon^{L\theta}}} \lambda'^{R\theta}_n}{\sum_n (\lambda^{L\theta}_{niG|n})^{1-\frac{1}{\epsilon^{L\theta}}} \lambda^{R\theta}_n} \tag{51}$$

We re-write Equations 16, 24, 21, and 20 to derive expressions for changes in rents and wages.

$$\hat{P}_i^N = \frac{\sum_n \sum_\theta \alpha^{N,\theta} R^\theta \mathbb{W}'^\theta_n \lambda'^{S\theta}_{ni|n} \lambda'^{R\theta}_n}{\hat{A}_{iN} \hat{L}_{iN}^{\beta^N} \hat{H}_{iN}^{1-\beta^N} \sum_n \sum_\theta \alpha^{N,\theta} R^\theta \mathbb{W}^\theta_n \lambda^{S\theta}_{ni|n} \lambda^{R\theta}_n} \tag{52}$$

$$\hat{Q}_n = \hat{H}_i^{\frac{1-\mu}{\mu}} \tag{53}$$

$$\hat{q_n} = \hat{Q}_n \tag{54}$$

$$\hat{w}_{nG+} = \left( \frac{\hat{A}_{nG} \hat{P}_n^G}{\hat{q}_i^{1-\beta^G} \hat{w}_{nG-}^{\beta^G \beta^{G-}}} \right)^{\frac{1}{\beta^G \beta^{G+}}} \tag{55}$$

$$\hat{w}_{nG-} = \frac{\hat{\tilde{N}}_{iG+} \hat{w}_{iG+}}{\tilde{N}_{iG-}} \tag{56}$$

We re-write Equations 31 and 32 to derive expressions for changes in productivity and amenities by type

$$\hat{A}_i = \left( \frac{\sum_n \sum_H \sum_\theta R^\theta \lambda'^{R\theta}_n \lambda'^{L\theta}_{niH|n}}{\sum_n \sum_H \sum_\theta R^\theta \lambda^{R\theta}_n \lambda^{L\theta}_{niH|n}} \right)^{\mu_A} \tag{57}$$

$$\hat{B}_n^\theta = \left(\frac{\hat{\lambda}_n^{R\theta}}{\hat{\lambda}_n^{R-\theta}}\right)^{\mu_{U\theta}} \tag{58}$$

Last, we rewrite Equations 7 and 10 to derive expressions for changes in expected income and utility from consumption by type

$$\hat{\mathbb{W}}_n^\theta = \left[\hat{w}_{nG}^{\epsilon^{L\theta}} e^{-\kappa_L^\theta \Delta\tau_{nn}\epsilon^{L\theta}} / \lambda_{nnG|n}^{\hat{L\theta}}\right]^{1/\epsilon^{L\theta}} \tag{59}$$

$$\hat{\mathbb{S}}_n^\theta = \left[\hat{P}_n^{-\epsilon^{S\theta}} e^{-\kappa_S^\theta \Delta\tau_{nn}\epsilon^{S\theta}} / \lambda_{nn|n}^{\hat{S\theta}}\right]^{\alpha^{N\theta}/\epsilon^{S\theta}} \tag{60}$$

We solve equations 45 to 60 starting with an initial guess in each endogenous variable such that $\hat{x} = 1$, and update until convergence to an equilibrium. With the counterfactual changes in endogenous variables $\{\hat{B}_n^\theta, \hat{\mathbb{S}}_n^\theta, \hat{\mathbb{W}}_n^\theta, \hat{Q}_n\}$, the change in expected utility by type is:

$$\hat{\bar{U}}^\theta = \left(\sum_n \lambda_n^{R\theta}\left(\hat{B}_n^\theta \hat{\mathbb{S}}_n^\theta \hat{\mathbb{W}}_n^\theta \hat{Q}_n^{-\alpha^H}\right)^{\epsilon^{R\theta}}\right)^{1/\epsilon^{R\theta}} \tag{61}$$

We estimate the initial equilibrium using data from 2015. We use changes in travel times as observed in the fare card data from before the opening of the DTL in 2015 to after its opening in 2018. Figure 11 presents the distribution of the changes in travel time.

## 6.2 Testing Model Predictions

We test the predictions of the model on data from 2018.

First, we show that model predicted changes in travel patterns are highly correlated with observed post-DTL changes in the fare card data. In Figure 12b, we plot the reduced form relationship between predicted changes in consumption travel according to the model, $\lambda_{ni|n}^{\hat{S\theta}}$, and observed changes in consumption travel from the fare card data conditional on neighborhood of residence. We find a highly positive correlation, significant at the 1 percent level. In Figure 12a, we plot the reduced form relationship between predicted changes in work location travel according to the model, $\lambda^{L\hat{\theta}}_{niG|n}$, and observed changes in work location travel from the fare card data conditional on neighborhood of residence. Again, we find a highly positive correlation, significant at the 1 percent level. We conclude that the model strongly predicts changes travel patterns, specifically where people work and shop, in response to the Downtown Line.

Second, we show that model predicted changes in residential shares are highly correlated with observed post-DTL changes. In Figure 12c, we plot the reduced form relationship between predicted changes in residential shares according to the model, $\lambda^{\hat{R}\theta}_{n}$, and observed changes in residential shares from the data. We find a highly positive correlation, significant at the 1 percent level. We conclude that the model strongly predicts changes travel patterns, specifically where people work and shop, in response to the Downtown Line.

Last, we show that model predicted changes in non-tradables land use are correlated with observed post-DTL food establishment entry between 2015 and 2018. In Figure 12d, we plot the reduced form relationship between predicted changes in non-tradable floor space according to the model, $\hat{H}_{jN}$, and observed entry of food establishments from the data. We find a positive correlation which is significant. We conclude that the model has predictive power over changes in commercial activity in response to the Downtown Line.

## 6.3 Welfare and Inequality

### 6.3.1 Main Results

Column 1 of Table 9 presents the effect of the Downtown line on welfare broken down by access to employment and access to consumption by high- and low-income. We find that the DTL improves welfare, $\bar{U}^{\theta}$, for high income workers, but not for low income workers. Welfare increases by 3.15% for high-income workers, while welfare for low-income workers is stagnant at 0.04. This is because although consumption access increases for both groups, non-tradable production and low-income jobs move to less attractive locations. Access to consumption, $\mathbb{S}$, improves for both workers, with high income workers benefiting more at 1.96% than low-income workers at 1.55%. However, while access to employment or expected wages net of commuting costs, $\mathbb{W}$, increases for high-income workers by 0.77%, low-income workers experience a 1.65% decline. Overall, inequality increases by 3.11%.

There are two main mechanisms driving our results. First, the DTL directly disproportionately improves access for many high income areas. Figure 1 shows that the Downtown Line directly serves the many high income areas who cluster just west of the city center. This partially explains the larger improvements in both access to consumption and employment experienced by high-income workers relative to low-income workers. Second, in response to the DTL, a greater number consumers travel to consume non-tradables near the center of the city because of new access via the new line. Figure 13 plots the model predictions for changes in consumption trips across neighborhoods. We see a large increase in consumption trips made to the center of the city. As a result, non-tradable jobs which are disproportionately low income jobs move to the center of the city, while low income workers live far away from the center as seen in Figure 1. We find that average commute time increased 1.5% for low income workers, while high income workers saw a 1% reduction.

We also estimate an increase in segregation as measured by the dissimilarity index.[20].

### 6.3.2 Results with Only Work Travel

We see much larger inequality effects by including travel to consume non-tradables in the model. We shut down the non-tradable consumption channel by setting $\alpha^{N,\theta} = 0$, $\alpha^{T,\theta} = 1 - \alpha^{H,\theta}$ and $\lambda^{S\theta}_{ni|n} = 0$ for both worker groups, and re-estimate the model. Column 2 of Table 9 presents the estimation results. We find that the DTL improves welfare, $\bar{U}^{\theta}$, for both workers groups. Welfare increases by 1.44% for high-income workers, while welfare increases by 0.84% for low-income workers. Expected income net of commuting, or access to employment, increases by 1.43% for high-income workers and by 0.43% for low-income

[20]Dissimilarity equals $\frac{1}{2}\sum_n |\frac{R^+}{R^-} - \frac{R^+\lambda_n^{R+}}{R^-\lambda_n^{R-}}|$

workers. We do not capture improvements in access to consumption. Overall inequality only increases by 0.60% which is a five-fold underestimation.

The model without consumption trips underestimates welfare effects because it ignores the fact that workers value access to consumption in utility. It also severely underestimates the inequality effects of the DTL in two key ways. First, the model fails to capture the disproportionate gains in access to consumption for high-income workers. Second, this model misses the re-organization of non-tradable production across the city in response to changes in consumption travel, worsening access to employment for low-income workers.

#### 6.3.3 No Spillovers

We show that the results are similar without agglomeration or residential externalities by setting $\mu_A$ and $\mu_{U\theta}$ to zero. We present the results in Table 10. We find that low-income workers are slightly more worse off, and high-income workers experience slightly smaller gains. This is driven by smaller gains (and larger losses) in access to employment, or expected wages, due to a lack of agglomeration effects from greater density in the city center. We also see less segregation due to a lack of residential externalities.

### 6.4 Decomposition

In this section, we explore how differences in parameters across worker groups in the model impact the inequality effects of the Downtown line. Our baseline will be the simplified model with no spillovers as in Column 1 of Table 10, and we will isolate the impact of each parameter separately.

Low-income workers have smaller expenditures on non-tradables than high-income workers. We set low-income consumption shares, $\{\alpha^{N-}, \alpha^{T-}, \alpha^{H-}\}$, to be equal to high-income consumption shares, $\{\alpha^{N+}, \alpha^{T+}, \alpha^{H+}\}$, and re-estimate the counterfactual in Row 2 of Table 11. We find smaller impacts on inequality relative to baseline. This is driven by larger gains in access to consumption for low-income workers, where inequality in access to consumption now actually decreases. High-income workers experience greater gains in welfare from improved consumer access relative to low-income workers because they spend a greater share of their income on non-tradables. Increasing non-tradable consumption also slightly strengthens the force of non-tradable jobs moving to less attractive locations, slightly worsening access to employment but more than offset by improvements in access to consumption. However, setting consumption shares to be equal only mitigates the inequality effects of the DTL slightly, decreasing welfare inequality effects by just 0.2%, with the main mechanisms from the previous section still driving the results.

Low-income workers have smaller travel costs than high-income workers. We set low-income travel costs, $\kappa^{-}$, to be equal to that high-income workers (removing subsidies), $\kappa^{+}$, and re-estimate the counterfactual in Row 3 of Table 11. We find smaller impacts on inequality relative to baseline. Inequality effects are smaller across both access to consumption and employment for low-income workers. High-income workers experience greater gains in welfare from the DTL because larger travel costs result in larger gains from improvements in access. However, setting travel costs to be equal only mitigates the inequality effects of the DTL slightly, decreasing welfare inequality effects by just 0.2%, with the main mechanisms from the previous section still driving the results.

Low-income workers have higher travel elasticities, or smaller dispersion of idiosyncratic consumption amenities and productivities, across locations than high-income workers. We set low-income dispersion parameters for work and consumption, $\{\epsilon^{L-}, \epsilon^{S-}\}$, to be equal to that high-income workers, $\{\epsilon^{L+}, \epsilon^{S+}\}$ and re-estimate the counterfactual in Row 4 of Table 11. We find smaller impacts on inequality relative to baseline. Inequality effects are smaller across both access to consumption and employment for low-income workers. High-income workers experience greater gains in welfare from the DTL because lower travel elasticities imply that high-income workers are less able to substitute to more attractive work and consumption locations in equilibrium and thus benefit more when travel time falls. However, setting travel elasticities to be equal only mitigates the inequality effects of the DTL partially, decreasing welfare inequality effects by 0.79%, with the main mechanisms from the previous section still driving the results.

Finally, low-income workers have greater dispersion in idiosyncratic residential amenities across locations than high-income workers. We set the low-income residential dispersion parameter, $\epsilon^{R-}$, to be equal to that high-income workers, $\epsilon^{R+}$ and re-estimate the counterfactual in Row 5 of Table 11. We find smaller impacts on inequality relative to baseline. High-income workers experience greater gains in welfare from the DTL because higher residential elasticities imply that high-income workers are more able to move to take advantage of improved access in other neighborhoods. However, setting residential elasticities to be equal only mitigates the inequality effects of the DTL partially, decreasing welfare inequality effects by just 0.13%, with the main mechanisms from the previous section still driving the results.

# 7 Conclusion

This paper demonstrates that consumption trips have important implications for the inequality effects of public transit expansions. We develop an urban spatial model with heterogeneous worker groups, low- and high-income, which incorporates both travel to workplace and to consume non-tradable goods and services. We use the model to study the impact of the Downtown Line (DTL) in Singapore. We find that the DTL improves welfare for high income workers by 3%. Abstracting away from travel to consume non-tradables results in a five-fold underestimation of the inequality effects of the DTL, estimating that both worker groups benefit but with a slightly larger share of the gains going to high-income workers. This is a result of failing to capture the spatial re-organization of low-income non-tradable sector jobs. Aggregate welfare gains are also underestimated by about 40%, ignoring gains in access to consumption opportunities for both groups. Policy makers should consider the welfare implications of changes in consumption patterns, in addition to commuting patterns, as they design new mass-transit infrastructure in cities across the world.

# References


Adao, Rodrigo, Costas Arkolakis, and Federico Esposito (Feb. 2019). *General Equilibrium Effects in Space: Theory and Measurement*. Tech. rep. Cambridge, MA: National Bureau of Economic Research.

Ahlfeldt, Gabriel et al. (2015). "The Economics of Density: Evidence From the Berlin Wall". In: *Econometrica* 83.6, pp. 2127–2189.

Albouy, David et al. (Mar. 2016). "Climate amenities, climate change, and american quality of life". In: *Journal of the Association of Environmental and Resource Economists* 3.1, pp. 205–246.

Allen, Treb and Costas Arkolakis (Aug. 2014). "Trade and the Topography of the Spatial Economy". In: *The Quarterly Journal of Economics* 129.3, pp. 1085–1140.

— (Jan. 2019). *The Welfare Effects of Transportation Infrastructure Improvements*. Tech. rep. Cambridge, MA: National Bureau of Economic Research.

Allen, Treb, Costas Arkolakis, and Xiangliang Li (2016). "Optimal City Structure". In:

Almagro, Milena and Tomás Domínguez-Iino (2020). *Location Sorting and Endogenous Amenities: Evidence from Amsterdam **. Tech. rep.

Autor, David H. and David Dorn (2013). "The growth of low-skill service jobs and the polarization of the US Labor Market". In: *American Economic Review* 103.5, pp. 1553–1597.

Balboni, Clare et al. (2020). "Transportation, Gentrification, and Urban Mobility: The Inequality Effects of Place-Based Policies".

Baum-Snow, Nathaniel and Matthew E. Kahn (2005). "Effects of Urban Rail Transit Expansions: Evidence from Sixteen Cities, 1970-2000". In: *Brookings-Wharton Papers on Urban Affairs* 2005.1, pp. 147–206.

Bayer, Patrick et al. (2009). "Migration and hedonic valuation: The case of air quality". In: *Journal of Environmental Economics and Management* 58.1, pp. 1–14.

Billings, Stephen B. (Nov. 2011). “Estimating the value of a new transit option”. In: *Regional Science and Urban Economics* 41.6, pp. 525–536.

Blomquist, Glenn C et al. (1988). “New Estimates of Quality of Life in Urban Areas”. In: *American Economic Review* 78.1, pp. 89–107.

Brooks, Leah and Byron Lutz (July 2019). “Vestiges of transit: Urban persistence at a microscale”. In: *Review of Economics and Statistics* 101.3, pp. 385–399.

Brueckner, Jan K. et al. (1999). “Why is central Paris rich and downtown Detroit poor?: An amenity-based theory”. In: *European Economic Review* 43.1, pp. 91–107.

Caliendo, Lorenzo et al. (Oct. 2018). “The Impact of Regional and Sectoral Productivity Changes on the U.S. Economy”. In: *The Review of Economic Studies* 85.4, pp. 2042–2096.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon (July 2019). “The Costs of Agglomeration: House and Land Prices in French Cities”. In: *The Review of Economic Studies* 86.4, pp. 1556–1589.

Couture, Victor et al. (Aug. 2019). *Income Growth and the Distributional Effects of Urban Spatial Sorting*. Tech. rep. Cambridge, MA: National Bureau of Economic Research.

Cragg, Michael and Matthew E. Kahn (1997). “New Estimates of Climate Demand: Evidence from Location Choice”. In: *Journal of Urban Economics* 42.2, pp. 261–284.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum (July 2008). “Global rebalancing with gravity: Measuring the burden of adjustment”. In: *IMF Staff Papers* 55.3, pp. 511–540.

Department of Statistics (2018). “Household Expenditure Survey 2017/18”. In:

Department of Transportation (2017). “2017 National Household Travel Survey”. In:

Desmet, Klaus, Dávid Krisztián Nagy, and Esteban Rossi-Hansberg (June 2018). “The Geography of Development”. In: *Journal of Political Economy* 126.3, pp. 903–983.

Diamond, Rebecca (Mar. 2016). *The determinants and welfare implications of US Workers’ diverging location choices by skill: 1980-2000*.

Donaldson, Dave (Apr. 2018). “Railroads of the Raj: Estimating the impact of transportation infrastructure”. In: *American Economic Review* 108.4-5, pp. 899–934.

Donaldson, Dave and Richard Hornbeck (May 2016). “Railroads and American Economic Growth: A “Market Access” Approach *”. In: *The Quarterly Journal of Economics* 131.2, pp. 799–858.

Epple, Dennis, Brett Gordon, and Holger Sieg (June 2010). “A new approach to estimating the production function for housing”. In: *American Economic Review* 100.3, pp. 905–924.

Fan, Yingling, Andrew Guthrie, and David Levinson (2016). “Waiting time perceptions at transit stops and stations: Effects of basic amenities, gender, and security”. In:

Fogli, Alessandra and Veronica Guerrieri (Aug. 2019). “The End of the American Dream? Inequality and Segregation in US Cities”. In: *SSRN Electronic Journal*.

Gibbons, Stephen and Stephen Machin (Jan. 2005). “Valuing rail access using transport innovations”. In: *Journal of Urban Economics* 57.1, pp. 148–169.

Glaeser, Edward L., Matthew E. Kahn, and Jordan Rappaport (2008). “Why do the poor live in cities The role of public transportation”. In: *Journal of Urban Economics* 63.1, pp. 1–24.

Glaeser, Edward L., Jed Kolko, et al. (2001). “Consumer city”. In: *Journal of Economic Geography* 1.1, pp. 27–50.

Glaeser, Edward L., Matt Resseger, and Kristina Tobio (Oct. 2009). “Inequality in cities”. In: *Journal of Regional Science* 49.4, pp. 617–646.

Gupta, Arpit, Stijn Van Nieuwerburgh, and Constantine Kontokosta (Feb. 2020). *Take the Q Train: Value Capture of Public Infrastructure Projects*. Tech. rep. Cambridge, MA: National Bureau of Economic Research.

Heblich, Stephan, Stephen Redding, and Daniel Sturm (Nov. 2020). “The Making of the Modern Metropolis: Evidence from London*”. In: *The Quarterly Journal of Economics* 135.4, pp. 2059–2133.

Lefevre, Benoit, David Leipziger, and Matthew Raifman (2014). *The Trillion Dollar Question: Tracking Public and Private Investment in Transport*. Tech. rep. World Resources Institute.

McDonald, John F. et al. (1995). "The effect of anticipated transportation improvement on residential land values". In: *Regional Science and Urban Economics* 25.3, pp. 261–278.

Melo, Patricia C. et al. (2009). "A meta-analysis of estimates of urban agglomeration economies". In: *Regional Science and Urban Economics* 39.3, pp. 332–342.

Ministry of Manpower (2020). *S Pass quota and levy requirements*.

Ministry of Transport (2014). "Workfare Transport Concession Scheme". In:

Miyauchi, Yuhei, Kentaro Nakajima, and Stephen Redding (2020). *Consumption Access and Agglomeration: Evidence from Smartphone Data*. Tech. rep.

Monte, Ferdinando, Stephen Redding, and Esteban Rossi-Hansberg (Dec. 2018). *Commuting, migration, and local employment elasticities*.

Redding, Stephen (July 2016). "Goods trade, factor mobility and welfare". In: *Journal of International Economics* 101, pp. 148–167.

Redding, Stephen and Daniel M. Sturm (Dec. 2008). "The costs of remoteness: Evidence from German division and reunification". In: *American Economic Review* 98.5, pp. 1766–1797.

Roback, Jennifer (Dec. 1982). "Wages, Rents, and the Quality of Life". In: *Journal of Political Economy* 90.6, pp. 1257–1278.

Rosenthal, Stuart S. and William C. Strange (2004). "Evidence on the nature and sources of agglomeration economies". In: *Handbook of Regional and Urban Economics* 4, pp. 2119–2171.

Severen, Christopher (Mar. 2018). "Commuting, Labor, and Housing Market Effects of Mass Transportation: Welfare and Identification".

Su, Wensong et al. (Apr. 2018). "Regional inequality, spatial spillover effects, and the factors influencing city-level energy-related carbon emissions in China". In: *Journal of Geographical Sciences* 28.4, pp. 495–513.

The Straits Times (2017). “The Downtown Line: Soaring to new heights”. In:

Tsivanidis, Nick (2019). “Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogotá’s TransMilenio”.

UITP (2018). *World Metro Figures*. Tech. rep.

United Nations (2018). *Revision of World Urbanization Prospects*. Tech. rep.

World Bank (2019). “Population Estimates and Projections”. In:

# Figures and Tables

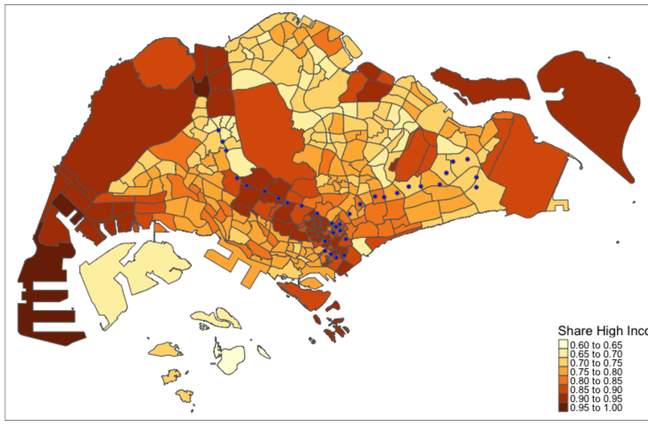



Figure 1: High-Income Residential Population Share and the Downtown Line

*Notes: This figure plots the share of high-income residents in each neighborhood, or subzone, according to farecard data in 2015. We define high-income residents as those above the 25th percentile. The Downtown line stations are plotted in blue.*

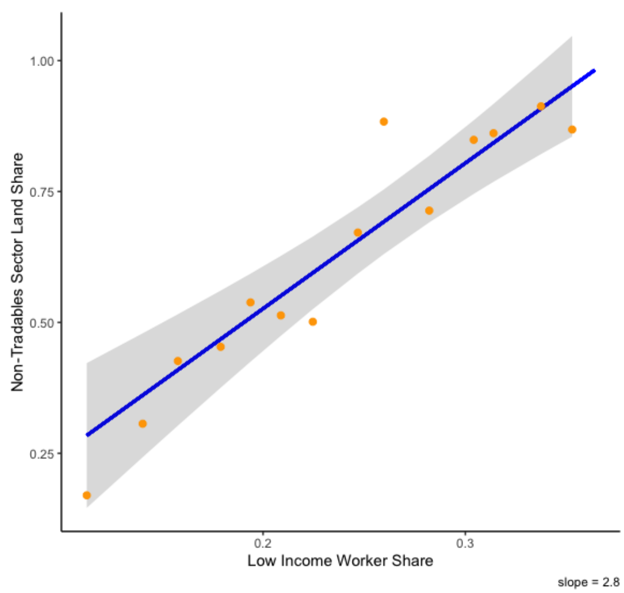
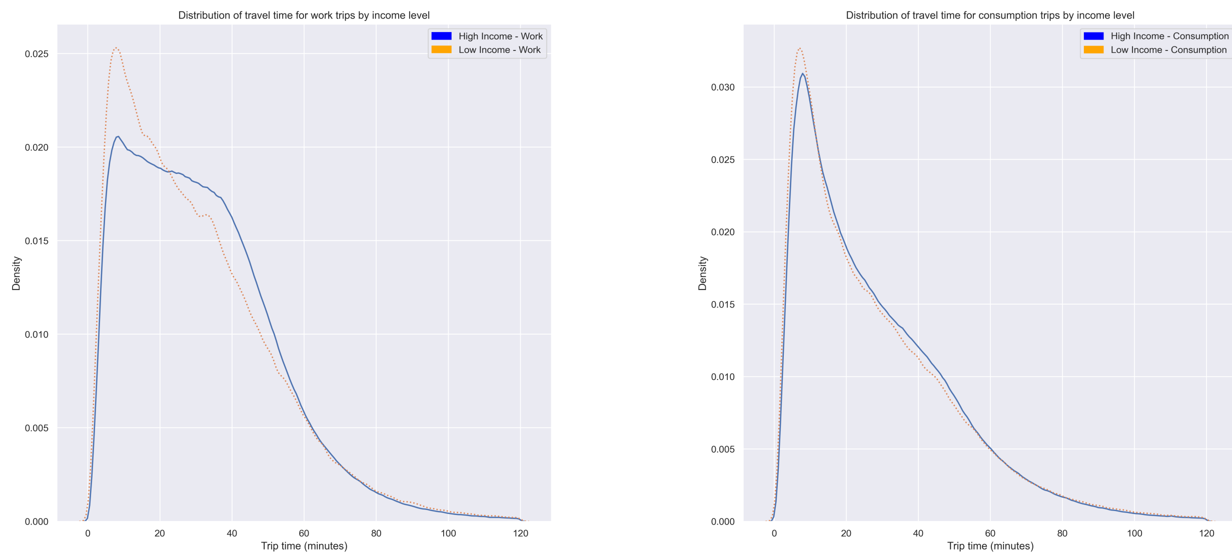



Figure 2: Relationship Between Low-Income Employment and Non-Tradables Commercial Land Share

*Notes: This figure presents a binned scatter plot of the share of low-income (below the 25th percentile) employment against the share of commercial land used by non-tradable sectors by neighborhood in 2015. Each point shows the mean of the variable with low-income employment share in the given bin.*




(a) Work (b) Consumption

Figure 3: Distribution of Travel Time

*Notes: This figure presents distribution of travel time for work in panel (a) and for consumption in panel (b) by high- and low-income workers (above and below the 25th percentile). Travel time is computed at the trip-level from farecard data pooled across all years.*

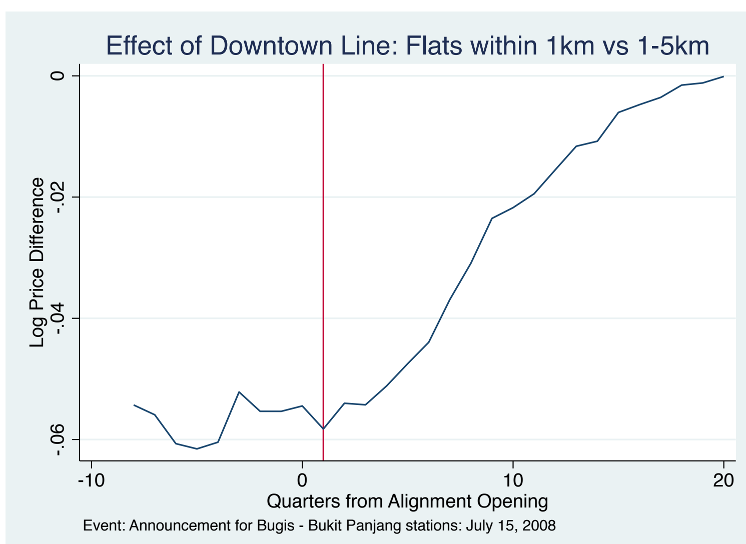



Figure 4: Spatial Differences-in-Differences Effect of the Downtown Line on Log Prices: Flats Within 1km vs 5km

*Notes: This figure plots the log difference in mean residential prices between flats within 1km and between 1-5 km of a Phase 2 DTL station in each quarter relative to the announcement date of the alignment of the line on 15 July 2008 (indicated with a red vertical line). Our data is a geocoded balanced panel of flats from Housing Development Board transaction data.*


Figure 5: Relationship between Prices and Distance to the Downtown Line over Time

*Notes: This figure plots the relationship between log prices and distance (in km) from the DTL over time. We estimate Equation 2 and plot coefficients, $\gamma$. We consider quarters from 2007 to 2011 in event time relative to the announcement date of the Phase 2 alignment of the line on 15 July 2008. We restrict our sample to addresses within 5km of a Phase 2 DTL station. Our data is a geocoded balanced panel of flats from Housing Development Board transaction data.*

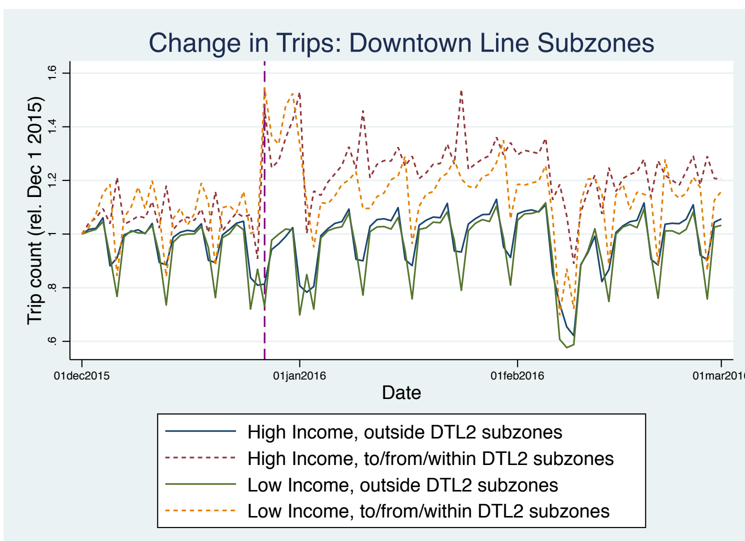
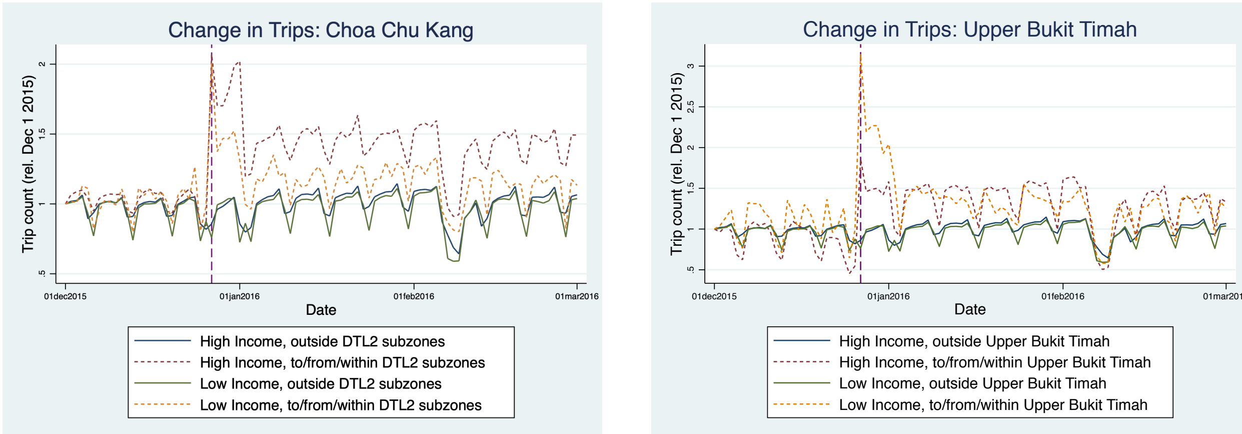



Figure 6: Impact of the Downtown Line on Travel by High- and Low-Income Workers

*Notes: This figure plots daily volume of trips to and from subzones with and without a DTL station over time by high- and low-income workers (above and below the 25th percentile) with farecard data between December 2015 and February 2016. We plot the number of trips relative to Dec 1 2015.*




(a) Choa Chu Kang

(b) Upper Bukit Timah

Figure 7: Impact of the Downtown Line on Travel by High- and Low-Income Workers: By Station

*Notes: This figure plots daily volume of trips to and from Choa Chu Kang in Panel (a) and Upper Bukit Timah in panel (b) compared to all other stations over time by high- and low-income workers (above and below the 25th percentile) with farecard data between December 2015 and February 2016. We plot the number of trips relative to Dec 1 2015.*

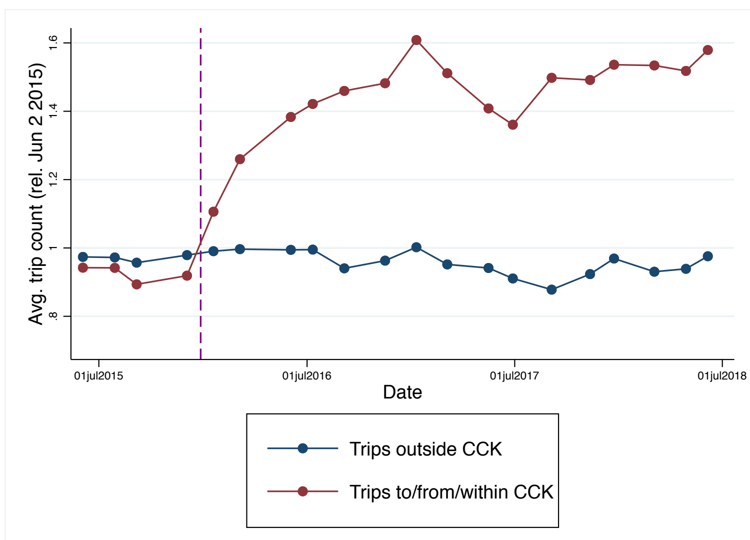
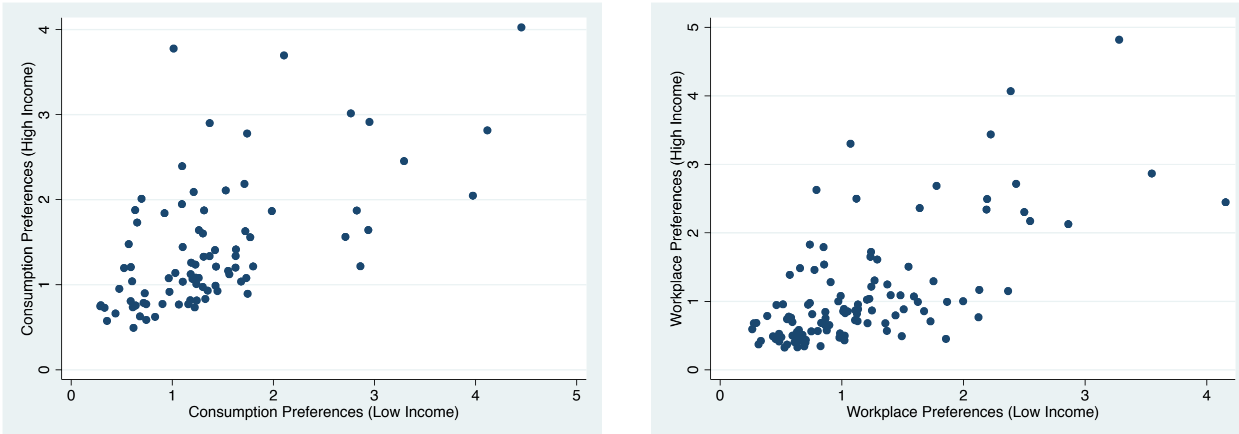



Figure 8: Impact of the Downtown Line on Long-run Travel: Choa Chu Kang (CCK) Stations

*Notes: This figure plots quarterly volume of trips to and from Choa Chu Kang and all other stations over time by high- and low-income workers (above and below the 25th percentile) with farecard data between July 2015 and July 2018. We plot the number of trips relative to June 2015.*




(a) Consumption (b) Work

Figure 9: Work and Consumption Location Attractiveness by High- and Low-Income Workers

*Notes: This figure plots consumption location attractiveness in Panel (a) and workplace attractiveness in Panel (b) as estimated in Section 5 by high- and low-income workers at the subzone level.*

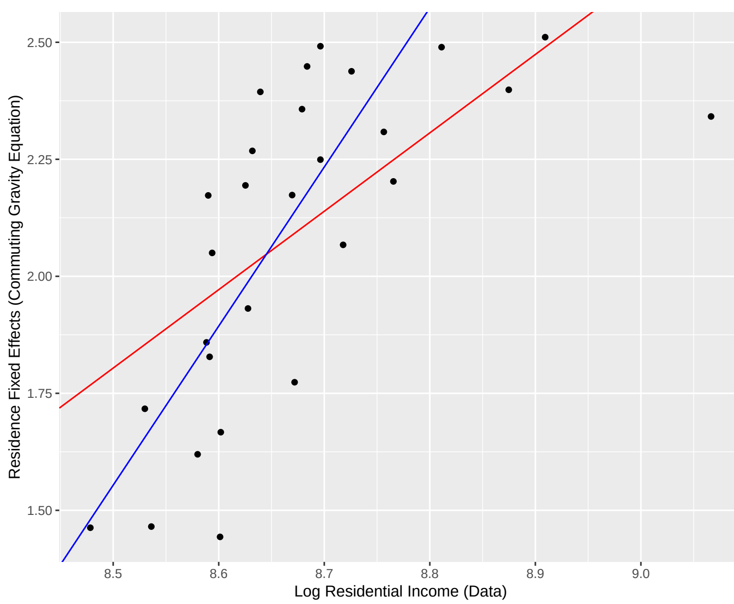



Figure 10: Expected Income: Data vs Model

*Notes: This figure presents a binned scatter plot of log model-estimated expected income by residence against log observed data on income by residence in 2015 from GHS. Each point shows the mean of the variable with observed data on income in the given bin. The red line is the linear fit across all neighborhoods, and the blue line is the linear fit dropping the top 4 highest income neighborhoods.*

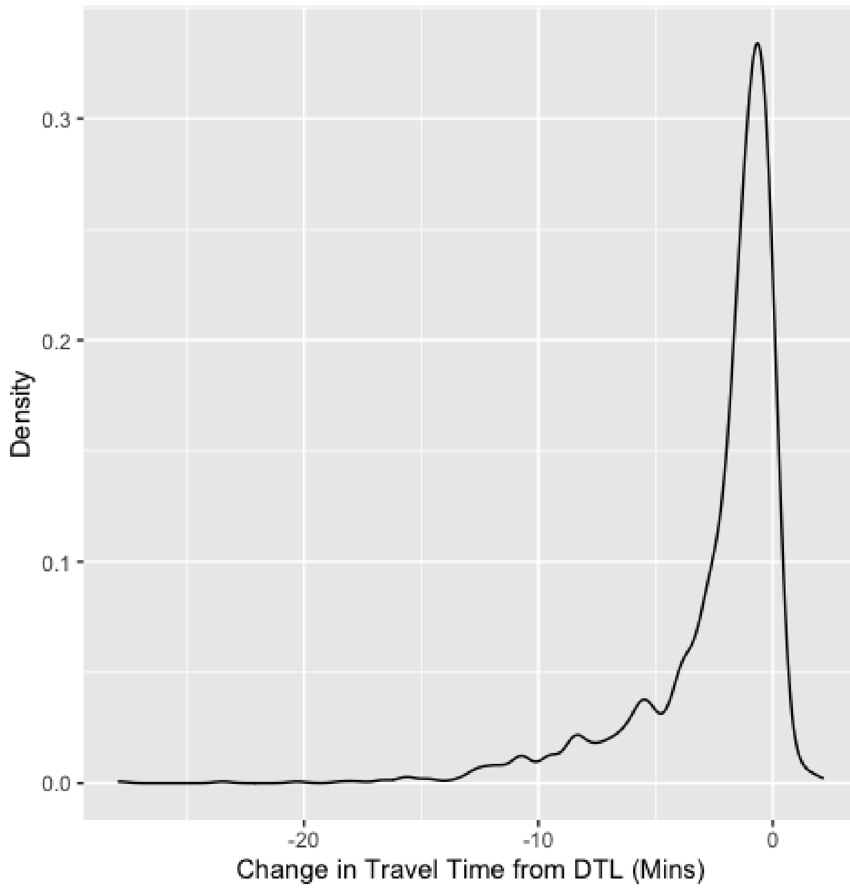
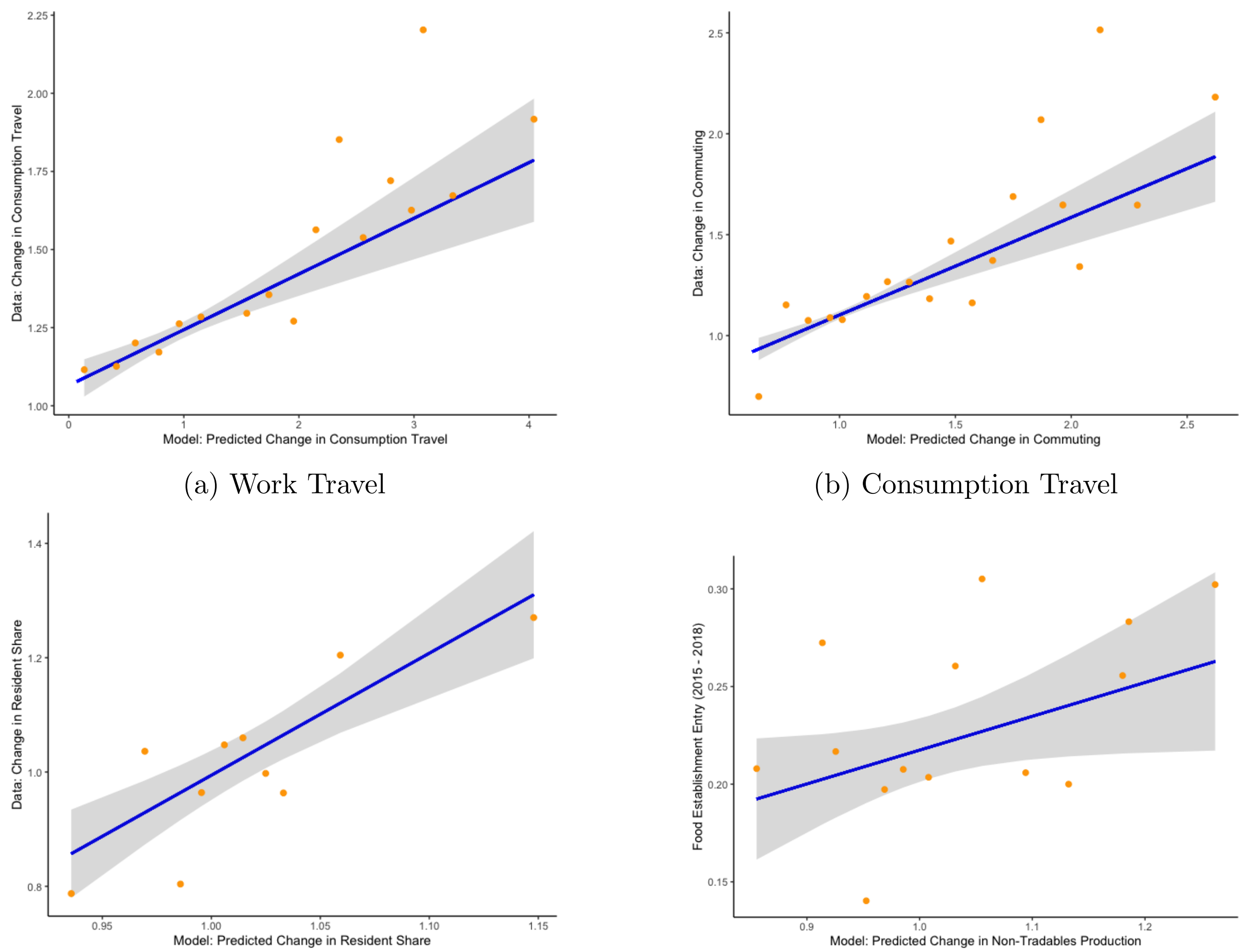



Figure 11: Changes in Bilateral Transit Time: Downtown Line

*Notes: This figure presents the distribution of the changes in bilateral travel times in minutes across all neighborhoods according to fare card between 2015 and 2018.*




(a) Work Travel

(b) Consumption Travel

(c) Residential Shares

(d) Food Establishment Entry

Figure 12: Model Counterfactual Prediction vs Ex-Post Data

*Notes: This figure presents binned scatter plots of predicted changes in variables from the counterfactual estimation in Section 6. against observed data. Each point shows the mean of the variable with predictions in the given bin. Panel (a) shows the relationship between predicted changes in bilateral work travel shares conditional on residence with observed changes between 2015 and 2018 computed from farecard data. Panel (b) shows the relationship between predicted changes in bilateral consumption travel shares conditional on residence with observed changes between 2015 and 2018 from farecard data. Panel (c) shows the relationship between predicted changes in residential shares by neighborhood with observed changes between 2015 and 2018 from farecard data. Panel (d) shows the relationship between predicted changes in the land quantities of the non-tradables sector and the entry rate of food establishments between 2015 and 2018 from Singapore Food Authority data by neighborhood.*

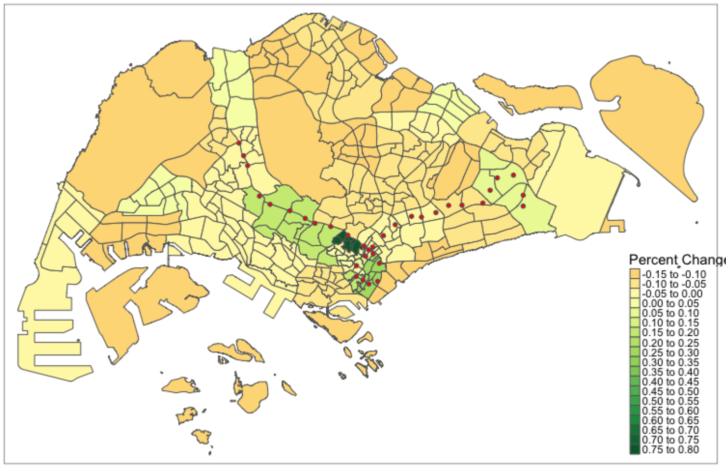



Figure 13: Model Prediction: Change in Consumption Trips

*Notes: This figure plots the predicted changes in total consumption trips in the counterfactual estimation in Section 6.*

Table 1: Labor Share of Low-Income Workers: Labor Force Survey 2018

| Industry | Under $2,000 | Over $2,000 |
|---|---|---|
| **Total** | **25%** | **75%** |
| | | |
| **Non-Tradables** | **51%** | **49%** |
| Retail Sales | 49% | 51% |
| Accommodation & Food Services | 61% | 39% |
| Personal Services | 41% | 59% |
| | | |
| **Tradables** | **18%** | **82%** |
| Construction | 23% | 77% |
| Manufacturing | 19% | 81% |
| Public Administration & Education | 13% | 87% |
| Real Estate Services | 24% | 76% |
| Financial & Insurance Services | 9% | 91% |
| Information & Communications | 10% | 90% |
| Administrative & Support Services | 51% | 49% |

*Notes: This table presents the share of employment over high- and low-income workers across sectors in Singapore from the Labor Force Survey 2018.*

Table 2: Gravity Estimation

| | Travel Share Conditional on Residence | | | |
|---|---|---|---|---|
| | Commuting (High Income) | Commuting (Low Income) | Consumption (High Income) | Consumption (Low Income) |
| | (1) | (2) | (3) | (4) |
| Travel Time (Minutes) | $-0.051^{***}$ | $-0.070^{***}$ | $-0.053^{***}$ | $-0.094^{***}$ |
| | (0.001) | (0.002) | (0.001) | (0.004) |
| Dest. FE | Yes | Yes | Yes | Yes |
| Origin FE | Yes | Yes | Yes | Yes |
| IV (geog. dist) | Yes | Yes | Yes | Yes |
| $R^2$ | 0.854 | 0.660 | 0.792 | 0.576 |
| Adjusted $R^2$ | 0.846 | 0.643 | 0.781 | 0.554 |

$^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01

*Notes: This table presents an estimation of Equation 33 in Columns 1 and 2, and Equation 36 in in Columns 3 and 4 for high- and low-income workers respectively. Destination and origin fixed effects are included and we instrument for travel time using linear distance. Each observation is a bilateral share of travel to a destination conditional on residence from 2015 farecard data.*

Table 3: Consumption Location Attractiveness: Data vs Model

| | *Log Leisure Attractiveness:* | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | High Income | Low Income | High Income | Low Income | High Income | Low Income | High Income | Low Income | High Income | Low Income |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Food Establishments | 0.386*** | 0.493*** | | | | | | | | |
| | (0.036) | (0.073) | | | | | | | | |
| Hawker Stalls / Food Establishments | | | 0.051 | 0.188** | | | | | | |
| | | | (0.043) | (0.076) | | | | | | |
| Food Court / Food Establishments | | | | | 0.080* | 0.295*** | | | | |
| | | | | | (0.043) | (0.075) | | | | |
| Supermarkets | | | | | | | 0.415*** | 0.586*** | | |
| | | | | | | | (0.035) | (0.071) | | |
| Clinics | | | | | | | | | 0.424*** | 0.733*** |
| | | | | | | | | | (0.035) | (0.067) |
| $R^2$ | 0.277 | 0.131 | 0.005 | 0.021 | 0.012 | 0.051 | 0.319 | 0.184 | 0.333 | 0.288 |
| Adjusted $R^2$ | 0.274 | 0.128 | 0.001 | 0.017 | 0.008 | 0.048 | 0.317 | 0.181 | 0.331 | 0.286 |
| Residual Std. Error | 0.625 | 1.275 | 0.732 | 1.294 | 0.729 | 1.274 | 0.607 | 1.235 | 0.600 | 1.154 |

*p<0.1; **p<0.05; ***p<0.01

*Notes: This table presents regression estimates of model-consistent measures of consumption attractiveness on various external data on retail amenities. We consider estimated high-income consumption attractiveness in odd columns and low-income consumption attractiveness in even columns as measured in Section 5. Each observation is a neighborhood/subzone. Columns 1 and 2 consider the normalized total number food establishments from Singapore Food Establishment data. Columns 3 and 4 consider the share of those food establishments which are hawker stalls. Columns 5 and 6 consider the share of those food establishments which are food courts. Columns 7 and 8 consider the normalized total number of supermarkets from National Environment Agency data. Columns 9 and 10 consider the normalized total number of clinics from Ministry of Health data.*

Table 4: Workplace Attractiveness and Employment Density by Worker Type

| | *Log Workplace Attractiveness* | |
|---|---|---|
| | High Income | Low Income |
| | (1) | (2) |
| Employment Density (Normalized) | 0.379*** | 0.228*** |
| | (0.053) | (0.056) |
| $R^2$ | 0.143 | 0.052 |
| Adjusted $R^2$ | 0.141 | 0.049 |
| Residual Std. Error | 0.927 | 0.975 |

*p<0.1; **p<0.05; ***p<0.01

*Notes: This table presents regression estimates of model-consistent measures of workplace attractiveness on employment density. We consider estimates of high-income workplace attractiveness in Column 1 and estimates of low-income workplace attractiveness in Column 2 as measured in Section 5. Each observation is a neighborhood/subzone. Employment density is normalized and is computed by dividing total employment (based on farecard data) and total square meters of the subzone.*

Table 5: Dispersion of Workplace Idiosyncratic Productivity by Worker Type: Estimation

| | *Parameter* | | | |
|---|---|---|---|---|
| | $\epsilon^{L+}$ | | $\epsilon^{L-}$ | |
| | (1) | (2) | (3) | (4) |
| Log Income (High Type) | 2.075*** | 2.912*** | | |
| | (0.482) | (0.641) | | |
| Log Income (Low Type) | | | 3.637** | 5.023*** |
| | | | (1.410) | (1.380) |
| Exclude Outliers | No | Yes | No | Yes |
| $R^2$ | 0.328 | 0.378 | 0.149 | 0.280 |
| Adjusted $R^2$ | 0.310 | 0.359 | 0.127 | 0.259 |
| Residual Std. Error | 0.384 | 0.380 | 0.527 | 0.489 |

*p<0.1; **p<0.05; ***p<0.01

*Notes: This table presents estimation results for Equation 39 in Section 5. The dependent variable is $\epsilon^{L\theta}$ multiplied by expected income by residence as estimated by the model, and the explanatory variable is observed income by residence from GHS 2015. We consider high-income workers (above 25th percentile) in Columns 1 and 2, and consider low-income workers (below 25th percentile) in Columns 3 and 4. We drop the four highest income neighborhoods in Columns 2 and 4.*

Table 6: Dispersion of Idiosyncratic Consumption Amenities and Travel Cost Parameters by High- and Low-Income

| **Parameter** | **Estimate** |
|---|---|
| $\epsilon^{S+}$ | 3.00 |
| $\epsilon^{S-}$ | 5.81 |
| $\kappa^{+}$ | 0.018 |
| $\kappa^{-}$ | 0.014 |

*Notes: This table presents estimates of parameters governing the dispersion of idiosyncratic consumption amenities in rows 1 and 2 from Equation 40, and travel costs in rows 3 and 4 from Equation 41. Rows 1 and 3 consider high-income workers (above-25th percentile), while rows 2 and 4 consider low-income workers (below-25th percentile).*

Table 7: Estimation of Parameters for Residential Dispersion of Idiosyncratic Residential Amenities and Residential Externalities by High- and Low-Income

| | *Dependent variable:* | |
|---|---|---|
| | $\Delta \ln \lambda_n^{R+}$ | $\Delta \ln \lambda_n^{R-}$ |
| | (1) | (2) |
| $\Delta \ln Q_n^{-\alpha^{H+}} \mathbb{W}_n^{+} \mathbb{S}_n^{+}$ | $1.481^{***}$ | |
| | (0.407) | |
| $\ln(R_n^{+}/R_n^{-})$ | $0.332^{***}$ | |
| | (0.085) | |
| $\Delta \ln Q_n^{-\alpha^{H-}} \mathbb{W}_n^{-} \mathbb{S}_n^{-}$ | | $1.380^{***}$ |
| | | (0.494) |
| $\ln(R_n^{-}/R_n^{+})$ | | $0.614^{***}$ |
| | | (0.085) |
| $R^2$ | 0.298 | 0.373 |
| Adjusted $R^2$ | 0.283 | 0.359 |
| Residual Std. Error | 0.237 | 0.245 |

$^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01

*Notes: This table presents estimates of Equation 42. Column 1 considers high-income workers (above-25th percentile), and Column 2 considers low-income workers (below-25th percentile). The dependent variable is the change in residential share by worker type between 2015 and 2018. The explanatory variables are changes in model estimates and observed data between 2015 and 2018. Each observation is a subzone/neighborhood.*

Table 8: Residential Amenities: Data vs Model

| | *Log Residential Amenities:* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | High Income | Low Income | High Income | Low Income | High Income | Low Income | High Income | Low Income |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Parks | $0.174^{***}$ | $0.220^{***}$ | | | | | | |
| | (0.054) | (0.058) | | | | | | |
| Community Clubs | | | $0.408^{***}$ | $0.400^{***}$ | | | | |
| | | | (0.059) | (0.073) | | | | |
| All Schools | | | | | $0.446^{***}$ | $0.399^{***}$ | | |
| | | | | | (0.062) | (0.081) | | |
| Neighborhood Schools | | | | | | | $0.196^{**}$ | $0.216^{**}$ |
| | | | | | | | (0.077) | (0.088) |
| High-Low Pop. Ratio Control | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.131 | 0.531 | 0.300 | 0.574 | 0.313 | 0.559 | 0.180 | 0.331 |
| Adjusted $R^2$ | 0.118 | 0.525 | 0.290 | 0.568 | 0.303 | 0.552 | 0.128 | 0.289 |
| Residual Std. Error | 0.870 | 0.945 | 0.781 | 0.901 | 0.773 | 0.917 | 0.456 | 0.518 |

$^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01

*Notes: This table presents regression estimates of model-consistent measures of residential amenities on various external data on amenities. We consider estimated high-income composite residential amenities in odd columns and low-income composite residential amenities in even columns as measured in Section 5. Each observation is a neighborhood/subzone. Columns 1 and 2 consider the normalized total number parks from Singapore Land Authority data. Columns 3 and 4 consider the normalized total number of community clubs from Singapore Land Authority data. Columns 5 and 6 consider consider the normalized total number of schools from Ministry of Education data. Columns 7 and 8 consider the normalized total number of "neighborhood schools" from and defined by Ministry of Education data. "Neighborhood schools" are public schools which are non-legacy and to not have gifted education programs.*

Table 9: Impact of the Downtown Line: Main Estimates

| | **(1) Full Model** | | **(2) Only Commuting** | |
|---|---|---|---|---|
| % Change in | Low Type | High Type | Low Type | High Type |
| Welfare, $\bar{U}$ | 0.04 | 3.15 | 0.84 | 1.44 |
| Access to Consumption, $\mathbb{S}$ | 1.55 | 1.96 | 0 | 0 |
| Access to Employment, $\mathbb{W}$ | -1.65 | 0.77 | 0.43 | 1.43 |
| Inequality | 3.11 | | 0.60 | |
| Segregation | 2.70 | | 1.79 | |

*Notes: This table presents results from the counterfactual estimation in Section 6. Column 1 presents results for the full model. Column 2 presents results abstracting away from consumption trips. Welfare, access to consumption and access to employment are defined by the model in Section 4. Inequality is the difference in percent change in welfare across high-income and low-income workers. Segregation is measured by the dissimilarity index.*

Table 10: Impact of the Downtown Line: No Spillovers

| | **(1) Full Model** | | **(2) Only Commuting** | |
|---|---|---|---|---|
| % Change in | Low Type | High Type | Low Type | High Type |
| Welfare, $\bar{U}$ | -0.84 | 2.09 | 0.55 | 1.32 |
| Access to Consumption, $\mathbb{S}$ | 1.63 | 1.80 | 0 | 0 |
| Access to Employment, $\mathbb{W}$ | -2.37 | 0.46 | 0.59 | 1.37 |
| Inequality | 2.93 | | 0.77 | |
| Segregation | -0.64 | | 0.59 | |

*Notes: This table presents results from the counterfactual estimation in Section 6 setting agglomeration and residential spillovers parameters to zero. Column 1 presents results for the full model. Column 2 presents results abstracting away from consumption trips. Welfare, access to consumption and access to employment are defined by the model in Section 4. Inequality is the difference in percent change in welfare across high-income and low-income workers. Segregation is measured by the dissimilarity index.*

Table 11: Impact of the Downtown Line: Decomposition of Parameters

| % Change in | **Welfare Inequality** | **Consumption Access Inequality** | **Workplace Access Inequalty** |
|---|---|---|---|
| Baseline | 2.93 | 0.17 | 2.83 |
| $\alpha^{-} = \alpha^{+}$ | 2.76 | -0.10 | 2.91 |
| $\kappa^{-} = \kappa^{+}$ | 2.76 | 0.07 | 2.77 |
| $\epsilon^{L/S-} = \epsilon^{L/S+}$ | 2.14 | 0.14 | 2.05 |
| $\epsilon^{R-} = \epsilon^{R+}$ | 2.80 | 0.17 | 2.70 |

*Notes: This table presents results from the counterfactual estimation in Section 6 with no agglomeration and residential spillovers setting various parameters across high- and low-income workers equal. Row 1 presents the baseline results as in Column 1 of Table 10. Row 2 sets low-income consumption shares to equal high-income consumption shares. Row 3 sets low-income travel costs to equal high-income travel costs. Row 4 sets low-income travel elasticities to equal high-income travel elasticities. Row 5 sets low-income residential elasticities to equal high-income residential elasticities. Welfare, access to consumption and access to employment are defined by the model in Section 4. Inequality is the difference in the percent change across high-income and low-income workers.*

# Online Appendix

# Appendix Figures and Tables

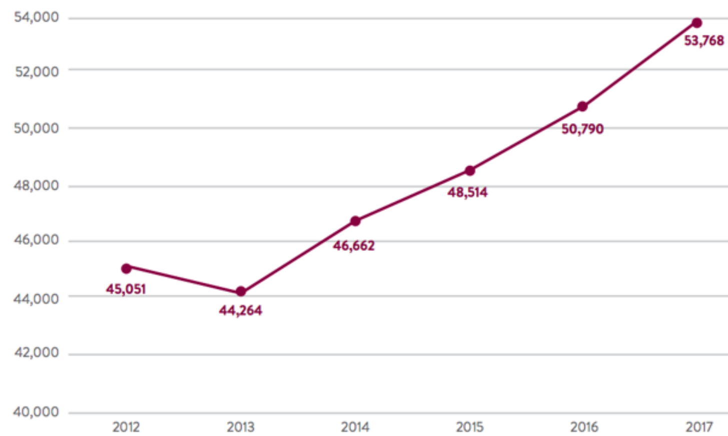



Figure A1: Mass-Transit Ridership over Time (in Millions)

*Notes: This figure presents the total number of mass-transit riders in millions over time between 2012 and 2017.*
*Source: UITP (2018).*

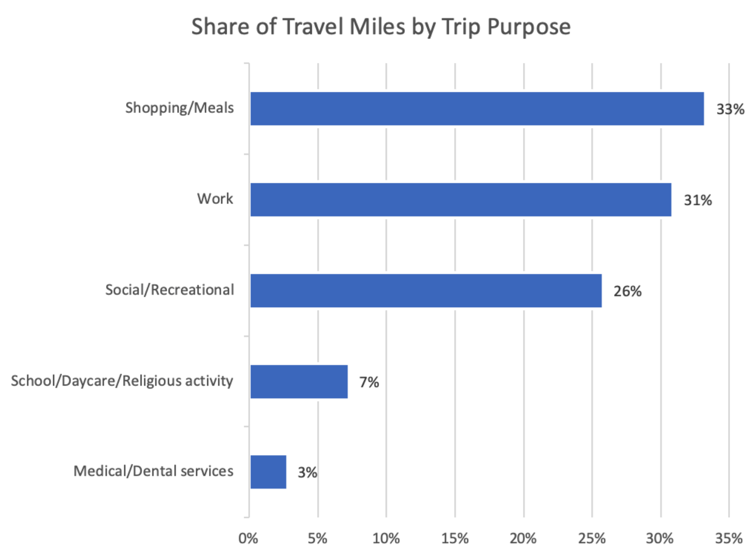



Figure A2: Purpose of Travel: National Household Travel Survey 2019

*Notes: This figure presents the share of travel miles classified by purpose of trip in 2017.*
*Source: Department of Transportation (2017)*

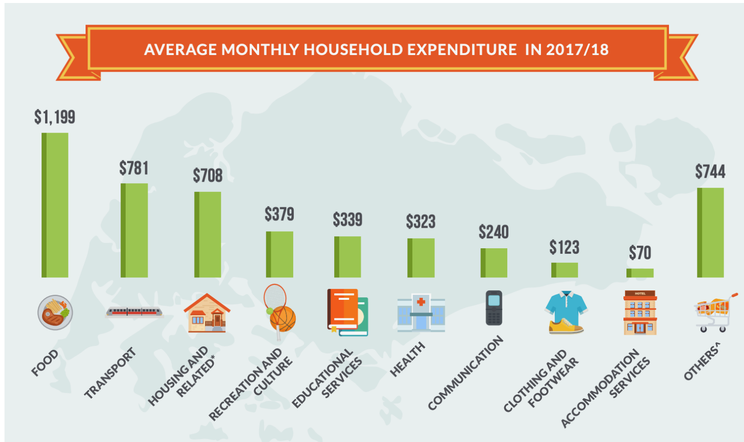
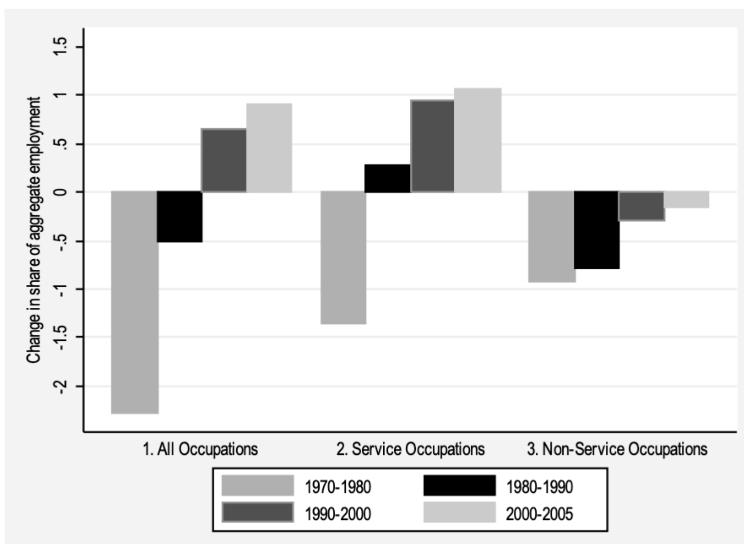



Figure A3: Household Expenditure Survey 2018: Singapore

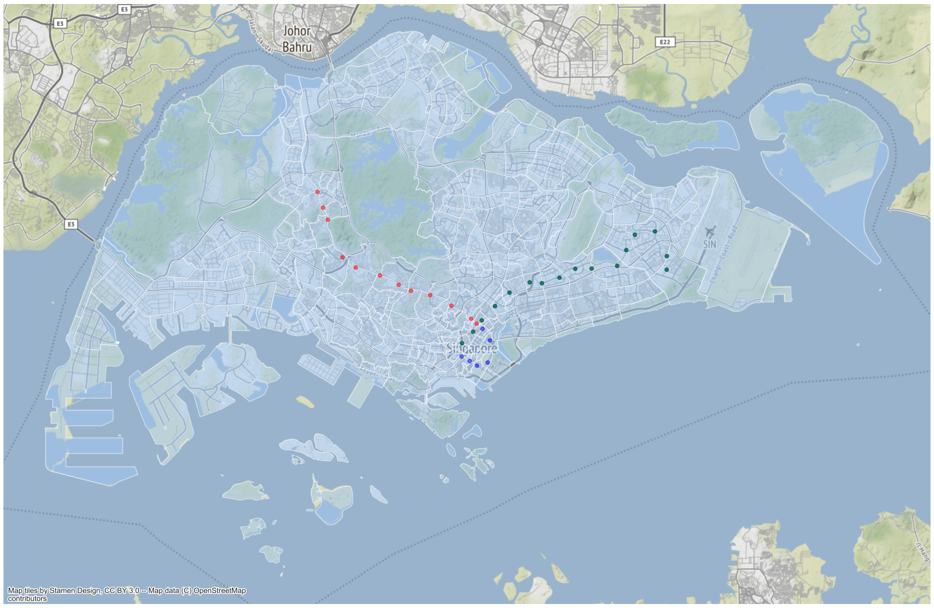
*Notes: This figure presents the average amount of household consumption by category in Singapore for fiscal year 2017-2018.*
*Source: Department of Statistics (2018)*




Change in Aggregate Employment Share by Decade 1970 through 2005 in
Occupations Comprising the Lowest Skill Quintile of Employment in 1980

Figure A4: Rise in Low-Income Non-tradable Service Sector Jobs: Autor and Dorn (2013)

*Notes: This figure presents the change in share of aggregate employment in low-income service and non-service occupations over time. Skill quintiles are measured by the mean occupational wage in 1980.*
*Source: Autor and Dorn (2013)*




Phase 1: Blue , Phase 2: Red , Phase 3: Green

Figure A5: The Downtown Line: Construction Phases

*Notes: This figure plots the Downtown Line stops by phase.*

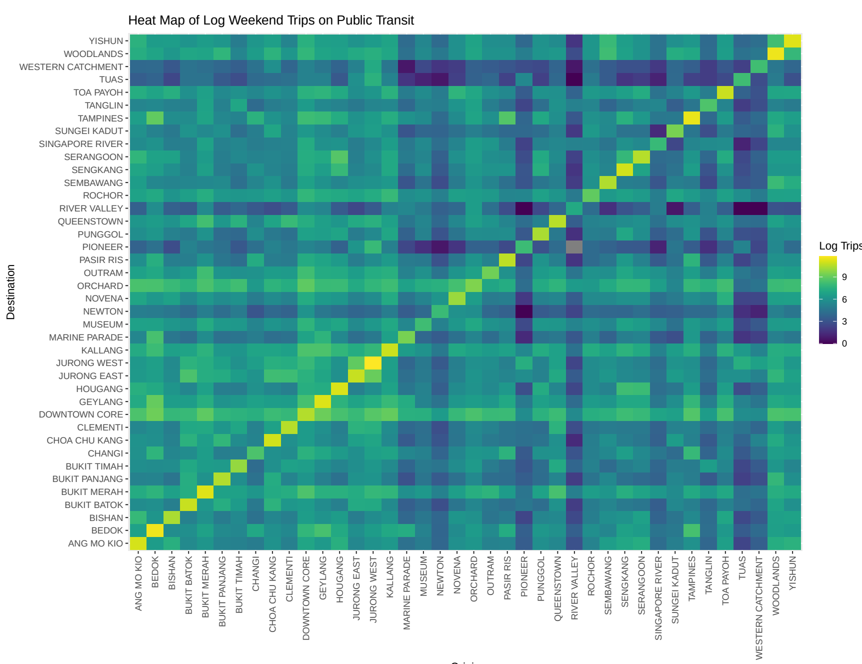



Figure A6: Farecard Data Heatmap

*Notes: This figure presents a heatmap of log farecard data weekend trips by planning area.*

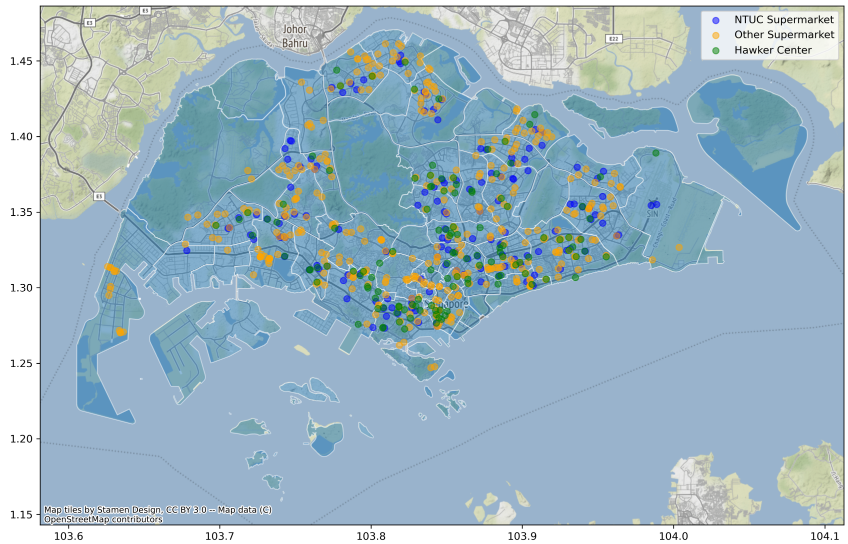



Figure A7: Amenities Data on Supermarkets and Hawker Centers

*Notes: This figure plots all supermarkets and hawker centers in Singapore from National Environment Agency data.*

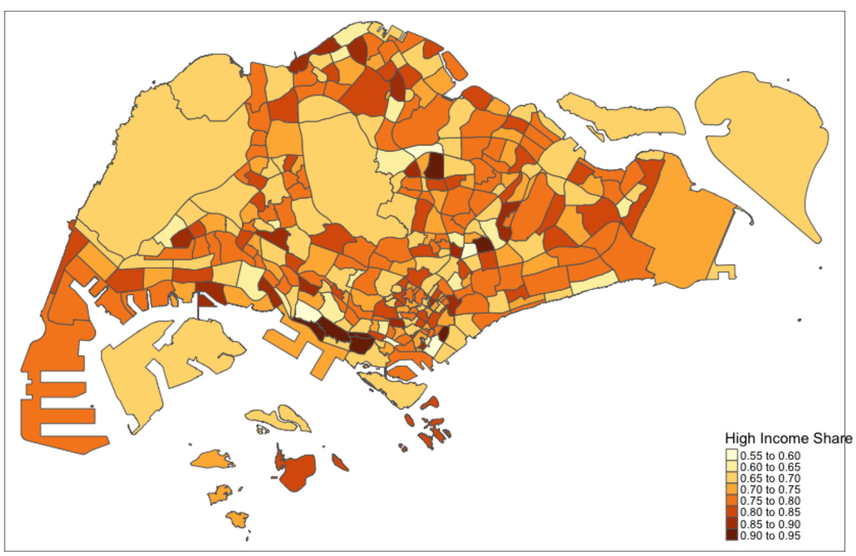



Figure A8: High-Income Employment Share

*Notes: This figure plots the high-income employment share in each subzone/neighborhood according to farecard data in 2015.*

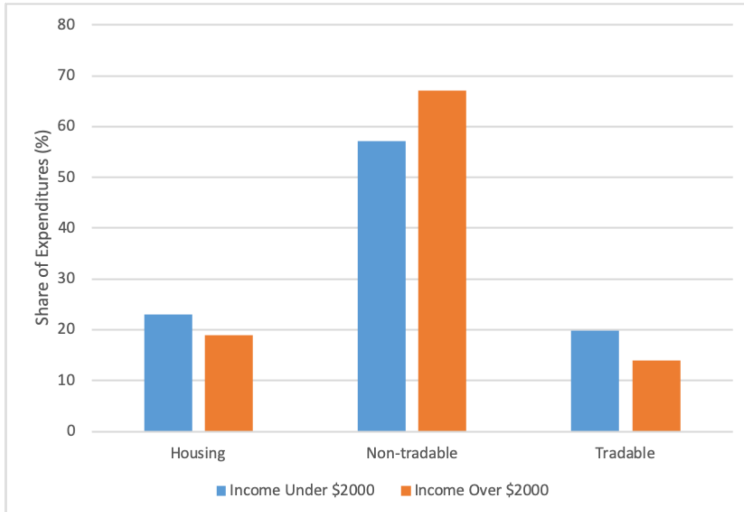



Figure A9: Expenditure Shares by High- and Low-Income Workers: Household Expenditure Survey 2018

*Notes: This figure presents the share of expenditures by housing, tradable goods, and non-tradable goods and services across high- and low-income groups in fiscal year 2017-2018.*

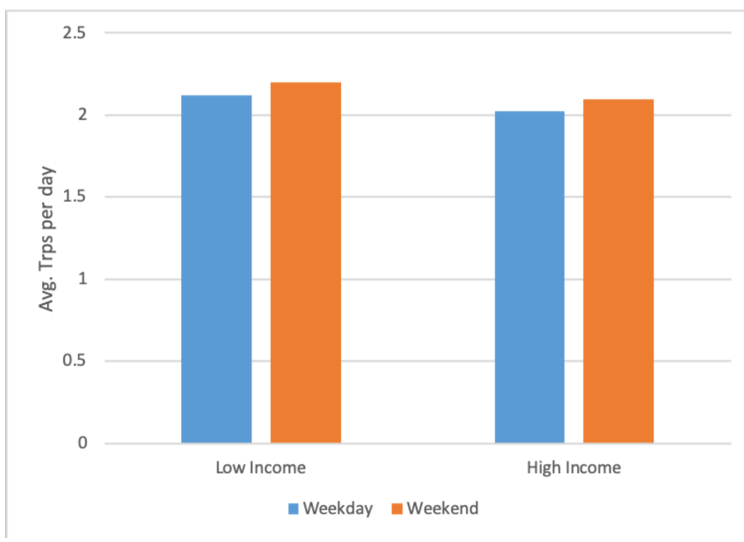



Figure A10: Daily Trips by High- and Low-Income Workers

*Notes: This figure presents the share average number of daily trips on weekdays and weekends by high- and low-income workers according to farecard data pooled over all years.*

Table A1: Summary of Mass Rail Lines in Singapore

| MRT line | Date announced | Date alignment announced | Open to public^ |
|---|---|---|---|
| North-South | 5/28/1982 | before 1983 | 11/7/1987 |
| East-West | 5/28/1982 | before 1983 | 11/7/1987 |
| North-East | 1/16/1996 | before 1997 | 6/20/2003 |
| Circle | 6/13/1998* | before 2002 | 5/28/2009 |
| Downtown | 6/14/2005 | 7/15/2008 | 12/22/2013 |
| Jurong Region | 5/9/2018** | 5/9/2018 | ~2026 |

- ^ date first section of line was opened
- * as Marina Line
- ** second announcement date (2001: "not enough demand")

*Notes: This table presents a summary of mass-transit rail lines in Singapore.*

Table A2: The Downtown Line and Food Establishment Entry: 2015 - 2018

| | **Downtown Line Subzones** | **Non-Downtown Line Subzones** |
|---|---|---|
| **% Change in Food Establishments** | 7.44% | -7.39% |
| **Entry Rate** | 24.16% | 22.04% |
| **Exit Rate** | 16.72% | 29.42% |

*Notes: This table presents results on the entry of food establishments between 2015 and 2018 from Singapore Food Authority data. Entry rate is defined as the share of 2018 food establishments that did not exist in 2015. Exit rate is defined as the share of 2015 food establishments that do not exist in 2018. Downtown line subzones are subzones with a DTL station.*

Table A3: Spatial Differences-in-Differences Effect of the Downtown Line on Log Prices: Flats Within 1km vs 5km

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Pooled | 2 Years | 3 Years | 4 Years |
| Within 1km x Post | 0.0212*** | 0.0213*** | 0.0373*** | 0.0484*** |
| | (0.00260) | (0.00530) | (0.00597) | (0.00613) |
| Building - Flat Type FE | Yes | Yes | Yes | Yes |
| Observations | 114255 | 7706 | 7669 | 7622 |
| $R^2$ | 0.881 | 0.979 | 0.977 | 0.978 |
| Adjusted $R^2$ | 0.876 | 0.957 | 0.952 | 0.953 |

Standard errors in parentheses

$^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

*Notes: This table estimates Equation 1. Each observation is a residential unit - quarter pair. The dependent variable is log price. Within 1km is an indicator for all flats within 1km of a Phase 2 DTL station. Post is an indicator for quarters after the announcement of the alignment of the Phase 2 of the DTL in July 2008. We include fixed effects for the building interacted with the number of rooms. We include all quarters between 2007 and 2012 in Column 1. We include only the quarter before the announcement and the quarter 2 years from the announcement in Column 2. We include only the quarter before the announcement and the quarter 3 years from the announcement in Column 2. We include only the quarter before the announcement and the quarter 4 years from the announcement in Column 2. We use geo-coded Housing Development Board flat transaction data.*

# Proof of Proposition 1

We combine data on tradable vs non-tradable land quantities with the model structure to solve for the share of employment in each sector by location and type.

$$\frac{H_{jG}q_j}{1-\beta^G} = \frac{\tilde{N}_{jG\theta}w_{jG\theta}}{\beta^G\beta^{G\theta}} \implies$$

$$\frac{\tilde{N}_{jN\theta}w_{jN\theta}}{\tilde{N}_{jT\theta}w_{jT\theta}} = \frac{H_{jN}(1-\beta^T)\beta^N\beta^{N\theta}}{H_{jT}(1-\beta^N)\beta^T\beta^{T\theta}}$$

where

$$\tilde{N}_{iG\theta} = \sum_n R_n \lambda^{L\theta}_{ni|n}\lambda^{L\theta}_{niG|ni}\bar{a}^{\theta}_{niG}$$

where $\bar{a}^{\theta}_{inG}$ is the average productivity of type-$\theta$ workers who live in $n$ and decide to work in $i$ and sector $G$.

Using the properties of Frechet we have that

$$\bar{a}^{\theta}_{niG} = \mathbb{E}[a^{\theta}_{inG}|T^{L_I\theta}_i T^{L_G\theta}_G (w^{\theta}_{iG}\exp(-\kappa\tau_{ni}))^{\epsilon^{L\theta}} > T^{L_I\theta}_l T^{L_G\theta}_H (w^{\theta}_{lH}\exp(-\kappa\tau_{nl}))^{\epsilon^{L\theta}}, \quad \forall l, H]$$

$$= \Gamma(\epsilon^{L\theta})(T^{L_I\theta}_i T^{L_G\theta}_G/\lambda^{L\theta}_{niG|n})^{1/\epsilon^{L\theta}}$$

Together we have

$$\frac{\sum_n R_n \lambda^{L\theta}_{ni|n}\lambda^{L\theta}_{niN|ni}(T^{L_G\theta}_N \lambda^{L\theta}_{ni|n}\lambda^{L\theta}_{niN|ni})^{-1/\epsilon^{L\theta}} w_{jN\theta}}{\sum_n R_n \lambda^{L\theta}_{ni|n}\lambda^{L\theta}_{niT|ni}(T^{L_G\theta}_T \lambda^{L\theta}_{ni|n}\lambda^{L\theta}_{niT|ni})^{-1/\epsilon^{L\theta}} w_{jT\theta}} = \frac{H_{jN}(1-\beta^T)\beta^N\beta^{N\theta}}{H_{jT}(1-\beta^N)\beta^T\beta^{T\theta}}$$

Also, conditional on type $\theta$ working in location $i$

$$\lambda^{L\theta}_{niG|ni} = \lambda^{L\theta}_{iG|i} = \frac{T^{L_G\theta}_G {w^{\theta}_{iG}}^{\epsilon^{L\theta}}}{\sum_H T^{L_G\theta}_H {w^{\theta}_{iH}}^{\epsilon^{L\theta}}}$$

This implies that

$$\frac{\lambda^{L\theta}_{iN|i}}{\lambda^{L\theta}_{iT|i}} = \frac{\lambda^{L\theta}_{iN|i}(T^{L_G\theta}_N/\lambda^{L\theta}_{iN|i})^{1/\epsilon^{L\theta}}w_{jN\theta}}{\lambda^{L\theta}_{iT|i}(T^{L_G\theta}_T/\lambda^{L\theta}_{iT|i})^{1/\epsilon^{L\theta}}w_{jT\theta}} = \frac{H_{jN}(1-\beta^T)\beta^N\beta^{N\theta}}{H_{jT}(1-\beta^N)\beta^T\beta^{T\theta}}$$

Thus, we have

$$\lambda^{G\theta}_{iG|i} = \frac{H_{jG}(1-\beta^T)\beta^G\beta^{G\theta}}{\sum_H H_{jH}(1-\beta^H)\beta^H\beta^{H\theta}}$$

and

$$\lambda^{G\theta}_{niG|ni} = \lambda^{G\theta}_{ni|n}\lambda^{G\theta}_{iG|i}$$

for each $i \in \{1, ..., N\}, \theta \in \{-, +\}, G \in \{N, T\}$ as a function of observed data.

Last, we can solve for $H_n$ with the following system of linear equations

$$H^R_n Q_n = \sum_\theta \alpha^{H,\theta} R^\theta_n \mathbb{W}_{i\theta}$$

For all $n \in N$.