Online Appendix to "Urban Transit Infrastructure and Inequality"

by Kwok-Hao Lee and Brandon Joel Tan

A Data Description and Construction

Our primary source of data is public transit fare card data (EZ-Link) from the Land Transport Authority of Singapore. We observe all trips made by public transit (mass rail, light rail or bus) linked to an individual's "EZ-Link" card. Our data set covers one week every quarter between June 2015 and June 2018, and three full months between December 2015 and February 2016. The full three months of data captures the period directly before and after the opening of Phase 2 of the Downtown Line. In total, we observe over a billion trips. For each trip, we observe the origin, destination, and 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 on the Adult and Low-Income Worker categories. Low-Income Workers are those who earn a monthly salary below the 25th percentile (S$2,000 prior to 2020, against a median of S$4,534).

We use our fare card data to generate work and consumption travel probabilities conditional on residential subzone. In the core data set (Dec 2015 to Feb 2016) we observe 7,326,131 unique fare cards and 410,988,245 trips. First, we restrict our data set to fare cards with at least 20 trips over our panel.[66] 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. These restrictions leave 2,893,160 farecards (39.4%) and 246,661,091 trips (60.0%). Next, 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 (5am to 11 am) and origin during evening rush hour (3pm to 11pm). Modal morning destination/evening origin locations match 71% of the time. To defend

[66] We obtain similar travel probabilities when we selected different trip thresholds for fare cards to be included in our data set.

against some common trip chains (e.g., dropping off children at school before work), when these two locations do not match, we impute workplaces from the modal evening origin. At the end, 97% of farecards have a "home" and a "work" subzone.[67] Finally, we classify all remaining trips as consumption trips.[68]

We also use data from several other 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 (Department of Statistics 2015). 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 employ-

[67]These identified residences and workplaces match well their corresponding shares in administrative data. See Figure A15.

[68]Could some of these trips be work trips taken by individuals without a fixed employment location, or the self-employed? We argue that such individuals make up a minority of transit users. As of 2015, according to the World Bank Development Indicators Database, about 14% of labor force participants in Singapore are self-employed. In our conversations, we found that most self-employed tradespeople utilized work vans (taxed less punitively than personal vehicles) for work and leisure travel. Moreover, doctors in private practice would perform house calls in personal vehicles. Hence, we believe classifying the remaining trips as consumption travel is correct to first order.

A further note: we abstract away from trip chaining because we do not directly observe the "stays" each commuter makes and our data suggest that trip chaining is not of first order concern. In our main data set, 77% of weekday farecard-day observations are associated with two or fewer trips in that day. Incorporating trip chaining would require solving a complex combinatorial choice problem over trip itineraries (see, e.g., Arkolakis, Eckert, and Shi 2021). Given data use limitations, we would be cursed by dimensionality.

ment and wages by industry (Ministry of Manpower 2018). The Household Expenditure Survey 2018 provides detailed data on household and worker expenditures by income bracket (Department of Statistics 2019). 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 (Urban Redevelopment Authority 2022). 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 (Housing and Development Board 2019).

Finally, we also use spatial data on amenities from the government data portal (Open Government Products 2019). First of all, we have two cross-sectional data sets, dating to 2015 and 2018, covering the universe of licensed food establishments in Singapore from the Singapore Food Authority. Second, data on all wet markets and hawker centers in Singapore are from the National Environment Agency.[69] Third, the Ministry of Education provides data on the location of all schools. Fourth, the National Parks Board provides data on the location of all parks. Fifth, the People's Association provides data on the location of all community clubs. Fifth, the Singapore Ministry of Health provides data on the location of all clinics participating in the Community Health Assist Scheme. We geocode all data and link each address or coordinates to the subzones in which they are contained.

B Additional Descriptive Analysis

B.1 The Downtown Line and Housing Prices

We consider how housing prices evolved before and after July 15, 2008, comparing government apartments closer to and farther from Downtown Line stations. On that date, the precise station locations on Stage 2 of the Downtown Line (DTL2) were suddenly revealed.[70] Our analysis is

[69] We present a map of the data in Appendix Figure A9.

[70] Stage 2 of the DTL extends from Bukit Panjang, a suburb, to the neighborhood of Bugis in the city center.

an event study across space and time, leveraging said announcement as a policy shock.[71] Despite knowing the areas through which DTL2 passes, homebuyers could not anticipate precisely where within the area stations would have been constructed. Once the station locations were revealed, prices after the announcement would adjust to reflect the distance to confirmed station locations. Figure 2a plots the log difference between residential prices within 0 to 1 kilometer of a DTL station and prices between 1 to 5 kilometers of a DTL station over time.[72] The price difference trend was flat before the announcement. After the station locations were announced, prices within 1 kilometer on the DTL began increasing relative to those between 1 to 5 kilometers. Table A5 presents estimates corresponding to the following regression:

$$\log(Price_{it}) = \beta_0 + \beta_1 Post_t + \beta_2 Close_i + \beta_3 Post_t * Close_i + \alpha_i + \varepsilon_{it} \tag{34}$$

where i indexes postal code-apartment type pairs and t is the quarter in which the transaction was made.[73] $Price_{it}$ 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 kilometer of a DTL station and 0 or all flats between 1 and 5 kilometers of a DTL station. The fixed effect α_i is

[71]To interpret our subsequent results as causal (a difference-in-differences design), one would have to believe that prices of treated and control apartments would have evolved similarly in the absence of the DTL. Any confounding factor driving housing prices must differentially affect apartments close to and farther from a DTL station. While an inspection of Figure 2a suggests that the parallel trends assumption holds, a large section of the DTL was built to relieve congestion along Bukit Timah Road and Dunearn Road, along which the line runs. Moreover, control apartments may also be affected by existing stations on other lines receiving different amounts of traffic because of the opening of DTL Phase 2. See the arguments in Donaldson and Hornbeck (2016).

[72]We use public Housing and Development Board Data covering the universe of government apartment transactions on the secondary market ("resale flats").

[73]We observe apartment (building-specific) postal codes and the number of rooms in each apartment (indexing apartment sizes).

specific to an (location, apartment type) pair, and ε_{it} is an error term. We estimate that over 4 years, residential prices increased by 4.84%, all else equal.

We also test the robustness of our results to a smaller radius. In Table A5, we report results shrinking the outer limit from 5km (2-5km being the control group) to 4km (2-4km being the control group) and 3km (1-3km being the control group). We find that across specifications the results are similar in magnitude and remain significant.

In most locations, the line placement was not known and home buyers were only aware of the general area in which the DTL was expected to pass through. In these cases, assignment is random within the area. However, in some locations the line placement was indicated (such as along Bukit Timah Road) ahead of time. In these cases, assignment is random along the line. For robustness, we consider an alternative specification. We construct a measure of the *expected distance* to a Phase 2 DTL station before announcement for each address. This expected distance to a DTL2 station assumes the following data generating process: the same number of stations as in Phase 2 of the Downtown Line are placed along the final observed line, but their locations are chosen uniformly at random.[74] Instead of comparing flats within 1km vs 2-5km of a DTL station, we compare flats where the actual distance from DTL is less than the expected distance ("closer than expected") to flats where the actual distance is greater than the expected distance ("further than expected"). We find, consistent with our baseline specification, that flats that are "closer than expected" experienced an increase in prices relative to flats that are "further than expected" after the alignment and station position announcement in 2008. We report the results in Table A5 and Figure A14.

We also consider the relationship between prices and distance from the DTL over time. We run

[74]The intuition behind this expected distance measure is as follows. We can decompose the actual distance of a house to DTL2 as the expected distance plus a forecast error. The forecast error between expected and actual distances to DTL2 is random, relative to the announcement in 2008.

the following regression:

$$\log(Price_{it}) = \tau_t + \sum_w \gamma_w \times Distance_i \times \mathbb{1}\{t = w - T\} + \varepsilon_{it} \tag{35}$$

where $Distance_i$ is the distance in kilometers from the nearest DTL stop. We restrict the sample to include apartments within 5 kilometers of a DTL stop. The indicator $\mathbb{1}\{t = w - T\}$ isolates price observations w years relative to the announcement on quarter T. We plot the estimated regression coefficients γ in Figure 2b. Consistent with our previous result, the 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. Post-announcement, the closer the flats are to the DTL, the higher their prices, all else equal.

C Model Details

C.1 Factor Demand

We derive expressions for factor demand. Fix a location i and a sector j. Under perfect competition and free entry, firms make zero profits in equilibrium. Furthermore, the price of each variety is equal to its marginal cost:

$$p_i^j = \frac{1}{A_i^j}(W_i^j)^{\beta^j} q_i^{1-\beta^j} \tag{36}$$

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

$$W_i^j = w_i^j(+)^{\beta^j(+)} w_i^j(-)^{\beta^j(-)} \tag{37}$$

is the cost of labor for sector j in location i. In this definition, $w_i^j(\theta)$ is the wage per efficiency unit for a worker of type θ. 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, we find that the demand for labor and commercial floor space, as well as wages, can be written as

$$L_i^j = \beta^j \frac{p_i^j Y_i^j}{W_i^j} \tag{38}$$

$$H_i^j = (1-\beta^j)\frac{p_i^j Y_i^j}{q_i} \tag{39}$$

$$\tilde{N}_i^j(\theta) = \beta^j(\theta)\frac{L_i^j W_i^j}{w_i^j(\theta)} \tag{40}$$

C.2 Average Productivity

Using the properties of the Fréchet distribution, the average productivity of type θ workers living in n and working in sector j in location i is

$$\begin{aligned}\bar{a}_{ni}^j(\theta) &= \mathbb{E}\left[a_{ni}^j(\theta)|(i,j) = \arg\max_{\tilde{i}\in\mathbb{N},\tilde{j}\in\{0,1\}} T_{\tilde{i}}(\theta)T^{\tilde{j}}(\theta)(w_{\tilde{i}}^{\tilde{j}}(\theta)\exp(-\kappa(\theta)\tau_{n\tilde{i}}))^{\varepsilon(L,\theta)}\right], \\ &= \Gamma[\varepsilon(L,\theta)](T_i(\theta)T^j(\theta)/\lambda_{ni}^j(L,\theta))^{1/\varepsilon(L,\theta)}\end{aligned}$$

C.3 Land Market

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

$$q_i = \xi_i Q_i, \tag{41}$$

where ξ_i equals 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 perfectly competitive developers using land M_i and construction materials K_i, with constant returns to scale technology:

$$\mathbb{H}_i = K_i^{\varphi} M_i^{1-\varphi} \tag{42}$$

where $\mathbb{H}_i$ is total floor space and φ is the share of land in floor space production. Therefore, the corresponding dual cost function for floor space is

$$Q_i = \varphi^{-\varphi}(1-\varphi)^{-(1-\varphi)}(P^{(K)})^{\varphi}\mathbb{R}_i^{1-\varphi}, \tag{43}$$

where $P^{(K)}$ is the common price for construction materials across all neighborhoods, and $\mathbb{R}_i$ is the price for land in neighborhood i. Cost minimization implies that

$$Q_i = P^{(K)} M_i^{\frac{\varphi-1}{\varphi}} \mathbb{H}_i^{\frac{1-\varphi}{\varphi}} \varphi^{-1} \tag{44}$$

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

C.4 Comments Regarding Homophily of Residents by Income

In our model, instead of being a function of the total density of residents, endogenous amenities depend on the demographic composition of workers across income groups, similar to Tsivanidis (2019). In contrast to the literature on neighborhood spillovers in the United States (e.g., Guerrieri, Hartley, and Hurst 2013), in Singapore, there is qualitative evidence that low-income workers prefer to live in neighborhoods with others of similar socioeconomic status. A stark example comes from an ethnographic study of inequality (Teo 2022, 57, 63). Low-income apartments in Singapore "are not ghettoized spaces... [and] there are no slums. In terms of access to clean water, electricity, amenities, and transportation, the people [the author] meet[s] are not denied access to these things in any absolute or physical way." Yet "a single mother once told [the author] she is afraid to walk through a mall near her home... because her six-year-old daughter will ask her to 'buy this, buy that.' "

Workers are more willing to pay to live in neighborhoods that are high in amenities of their type. As workers locate in these neighborhoods, they increase these type-specific endogenous amenities even more, strengthening segregation.

C.5 Existence and Uniqueness of Equilibrium

In a version of the model without externalities, by similar arguments to those made in Ahlfeldt et al. (2015), the model's congestion forces — commuting costs, travel costs to consume non-traded goods, and an inelastic land supply — ensure that a unique equilibrium exists. In the full version of the model, by adapting arguments in Section 3.1 of Allen, Arkolakis, and Li (2022), uniqueness can be guaranteed when "agglomeration forces are small relative to congestion forces."

Characterizing formal sufficient conditions for uniqueness involves tedious algebra. We sketch the approach here. First, bound the absolute values of the elasticities of productivity spillovers with respect to worker population; and those of amenity spillovers with respect to residential population. Next, substitute the gravity and spillover equations into the market clearing equations in residential population, labor supply, and consumption travel. This substitution gives rise to bounds on the elasticities of populations in each workplace, residence, and consumption location by type. Collect these elasticity bounds in a matrix of "cross-elasticities." Then, by Remark 1 and part (i) of Theorem 1 of Allen, Arkolakis, and Li (2022), for equilibrium uniqueness, it is sufficient that the spectral density of the matrix of "cross-elasticities" be less than one.

D 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 of all, 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, exploiting the structure of our model, 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 (19). Thus it suffices to observe data on residential wages and population by type. The following proposition formalizes the intuition behind our model inversion process.

Proposition 1. ***Sector Employment by Type and Neighborhood*** *(i) Given data on residence and work travel by type and commercial land quantities by sector* $\{\lambda_{ni}(L,\theta),\lambda_n(R,\theta),R(\theta),H_i^j\}$ *in addition to model parameters, there exists a unique vector of neighborhood-sector employment* *shares conditional residence by type, $\{\lambda^{j}_{ni}(L,\theta)\}$, 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 $\{\lambda_n(R,\theta), R(\theta), \mathbb{W}_n(\theta), Q_n\}$ in addition to model parameters, there exists a unique vector of residential land quantities, $\{H_n(R)\}$, that rationalizes the observed data as an equilibrium of the model.*

D.1 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. Fix a location i. The ratio of expenditures on the labor input to tradable to non-tradable sectors can be written as

$$\frac{H^{j}_{i} q_{j}}{1-\beta^{j}} = \frac{\tilde{N}^{j}_{i}(\theta) w^{j}_{i}(\theta)}{\beta^{j}\beta^{j}(\theta)} \implies$$

$$\frac{\tilde{N}^{(0)}_{i}(\theta) w^{(0)}_{i}(\theta)}{\tilde{N}^{(1)}_{i}(\theta) w^{(1)}_{i}(\theta)} = \frac{H^{(0)}_{i}(1-\beta^{(1)})\beta^{(0)}\beta^{(0)}(\theta)}{H^{(1)}_{i}(1-\beta^{(0)})\beta^{(1)}\beta^{(1)}(\theta)}$$

where

$$\tilde{N}^{j}_{i}(\theta) = \sum_{n\in\mathbb{N}} R(\theta)\lambda_n(R,\theta)\lambda^{j}_{ni}(L,\theta)\bar{a}^{j}_{ni}(\theta).$$

Above, $\bar{a}^{j}_{ni}(\theta)$ is the average productivity of type-θ workers who live in n and decide to work in location i and sector j.

By derivations in Online Appendix Section C.2, write

$$\bar{a}^{j}_{ni}(\theta) = \Gamma[\varepsilon(L,\theta)](T_i(\theta)T^{j}(\theta)/\lambda^{j}_{ni}(L,\theta))^{1/\varepsilon(L,\theta)}.$$

Rewriting the ratio of labor expenditures on the LHS, we have

$$\frac{\sum_n R_n(\theta)\lambda^{(0)}_{ni}(L,\theta)(T^{(0)}(\theta)\lambda^{(0)}_{ni}(L,\theta))^{-1/\varepsilon(L,\theta)} w^{(0)}_{i}(\theta)}{\sum_n R_n(\theta)\lambda^{(1)}_{ni}(L,\theta)(T^{(1)}(\theta)\lambda^{(1)}_{ni}(L,\theta))^{-1/\varepsilon(L,\theta)} w^{(1)}_{i}(\theta)} = \frac{H^{(0)}_{i}(1-\beta^{(1)})\beta^{(0)}\beta^{(0)}(\theta)}{H^{(1)}_{i}(1-\beta^{(0)})\beta^{(1)}\beta^{(1)}(\theta)}$$

Also, conditional on type θ working in location i,

$$\lambda_{ni}^{j}(L,\theta)=\lambda_{i}^{j}(L,\theta)=\frac{T^{j}(\theta)w_{i}^{j}(\theta)^{\varepsilon(L,\theta)}}{\sum_{j'\in\{0,1\}}T^{j'}(\theta)w_{i}^{j'}(\theta)^{\varepsilon(L,\theta)}}$$

Then the ratio of labor expenditures on the LHS can be written

$$\frac{\lambda_{i}^{(0)}(L,\theta)}{\lambda_{i}^{(1)}(L,\theta)}=\frac{\lambda_{i}^{(0)}(L,\theta)(T^{(0)}(\theta)/\lambda_{i}^{(0)}(L,\theta))^{1/\varepsilon(L,\theta)}w_{i}^{(0)}(\theta)}{\lambda_{i}^{(1)}(L,\theta)(T^{(1)}(\theta)/\lambda_{i}^{(1)}(L,\theta))^{1/\varepsilon(L,\theta)}w_{i}^{(1)}(\theta)}=\frac{H_{i}^{(0)}(1-\beta^{(1)})\beta^{(0)}\beta^{(0)}(\theta)}{H_{i}^{(1)}(1-\beta^{(0)})\beta^{(1)}\beta^{(1)}(\theta)}$$

Thus, we have

$$\lambda_{i}^{j}(L,\theta)=\frac{H_{i}^{j}(1-\beta^{|1-j|})\beta^{j}\beta^{j}(\theta)}{\sum_{j'\in\{0,1\}}H_{i}^{j'}(1-\beta^{|1-j'|})\beta^{j'}\beta^{j'}(\theta)}$$

and

$$\lambda_{ni}^{j}(L,\theta)=\lambda_{ni}(L,\theta)\lambda_{i}^{j}(L,\theta)$$

for each $i\in\{1,...,N\},\theta\in\{-,+\}$, and $j\in\{0,1\}$, as a function of observed data.

Lastly, we can solve for H_n with the following system of linear equations:

$$H_{n}(R)Q_{n}=\sum_{\theta\in\{-,+\}}\gamma(\theta)R_{n}(\theta)\mathbb{W}_{n}(\theta)\quad\forall n\in\mathbb{N}.$$

E Exact-Hat Algebra

This section details the "exact-hat algebra" approach used to compute changes in commuting patterns from the pre-Downtown Line (DTL) equilibrium.

First, we rewrite the gravity equations (7), (4) and (10) in the main text to derive expressions for changes in commuting, consumption, and residential probabilities by worker type.

$$\hat{\lambda}_{ni}^{j}(L,\theta)=\frac{[\hat{w}_{i}^{j}(\theta)]^{\varepsilon(L,\theta)}\exp(-\kappa(\theta)\Delta\tau_{ni}\varepsilon(L,\theta))}{\sum_{i'\in\mathbb{N},j'\in\{0,1\}}\lambda_{ni'}^{j'}(L,\theta)[\hat{w}_{i'}^{j'}(\theta)]^{\varepsilon(L,\theta)}\exp(-\kappa(\theta)\Delta\tau_{ni'}\varepsilon(L,\theta))}\tag{45}$$

$$\hat{\lambda}_{nl}(C,\theta)=\frac{p_{l}^{-\alpha(\theta)\varepsilon(C,\theta)}\exp(-\kappa(\theta)\Delta\tau_{nl}\varepsilon(C,\theta))}{\sum_{l'\in\mathbb{N}}\lambda_{nl'}(C,\theta)p_{l'}^{-\alpha(\theta)\varepsilon(C,\theta)}\exp(-\kappa(\theta)\Delta\tau_{nl'}\varepsilon(C,\theta))}\tag{46}$$

$$\hat{\lambda}_n(R,\theta) = \frac{\hat{B}_n(\theta)\left(\hat{\mathbb{W}}_n(\theta)\hat{\mathbb{C}}_n(\theta)\hat{Q}_n^{-\gamma(\theta)}\right)^{\varepsilon(R,\theta)}}{\sum_{n'\in\mathbb{N}}\lambda_{n'}(R,\theta)\hat{B}_{n'}(\theta)\left(\hat{\mathbb{W}}_{n'}(\theta)\hat{\mathbb{C}}_{n'}(\theta)\hat{Q}_{n'}^{-\gamma(\theta)}\right)^{\varepsilon(R,\theta)}} \tag{47}$$

Next, we rewrite the land and labor market equations (19), (16), (20), and (21) to derive expressions for changes in land and labor quantities.

$$\hat{H}_n(R) = \frac{\sum_\theta \gamma(\theta)R(\theta)\lambda'_n(R,\theta)\mathbb{W}'_n(\theta)}{\hat{Q}_n\sum_\theta \gamma(\theta)R(\theta)\lambda_n(R,\theta)\mathbb{W}_n(\theta)} \tag{48}$$

$$\hat{H}_i^j = \frac{\hat{L}_i^j\hat{W}_i^j}{\hat{q}_i} \tag{49}$$

$$\hat{\mathbb{H}}_i = \frac{H_i(R)\hat{H}_i(R) + \sum_{j\in\{0,1\}}H_i^j\hat{H}_i^j}{H_i(R) + \sum_{j\in\{0,1\}}H_i^j} \tag{50}$$

$$\hat{\tilde{N}}_i^j(\theta) = \frac{\sum_n((\lambda_{ni}^j)'(L,\theta))^{1-\frac{1}{\varepsilon(L,\theta)}}\lambda'_n(R,\theta)}{\sum_n(\lambda_{ni}^j(L,\theta))^{1-\frac{1}{\varepsilon(L,\theta)}}\lambda_n(R,\theta)} \tag{51}$$

Then, we rewrite the price equations (23), (18), (41), and (15) to derive expressions for changes in rents and wages.

$$p_i = \frac{\sum_{n\in\mathbb{N},\theta\in\{+,-\}}\alpha(\theta)R(\theta)\mathbb{W}'_n(\theta)\lambda'_{ni}(C,\theta)\lambda'_n(R,\theta)}{\hat{A}_i^{(0)}(\hat{L}_i^{(0)})^{\beta^{(0)}}(\hat{H}_i^{(0)})^{1-\beta^{(0)}}\sum_{n\in\mathbb{N},\theta\in\{+,-\}}\alpha(\theta)R(\theta)\mathbb{W}_n(\theta)\lambda_{ni}(C,\theta)\lambda_n(R,\theta)} \tag{52}$$

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

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

$$\hat{w}_i^j(+) = \left(\frac{\hat{A}_i^j\hat{p}_i^j}{\hat{q}_i^{1-\beta^j}\hat{w}_i^j(-)^{\beta^j\beta^j(-)}}\right)^{\frac{1}{\beta^j\beta^j(+)}} \tag{55}$$

$$\hat{w}_i^j(-) = \frac{\hat{\tilde{N}}_i^j(+)\hat{w}_i^j(+)}{\tilde{N}_i^j(-)} \tag{56}$$

Thereafter, we rewrite the spillover equations (24) and (25) to derive expressions for changes in productivity and amenities by type:

$$\hat{A}_i = \left(\frac{\sum_{n\in\mathbb{N}, j\in\{0,1\}, \theta\in\{+,-\}} R(\theta)\lambda_n'(R,\theta)(\lambda_{ni}^j)'(L,\theta)}{\sum_{n\in\mathbb{N}, j\in\{N,T\}, \theta\in\{+,-\}} R(\theta)\lambda_n(R,\theta)\lambda_{ni}^j(L,\theta)} \right)^{\mu(A)}; \tag{57}$$

$$\hat{B}_n(\theta) = \left(\frac{\hat{\lambda}_n(R,\theta)}{\hat{\lambda}_n(R, \text{not } \theta)} \right)^{\mu(U,\theta)}. \tag{58}$$

Lastly, we rewrite the expected wage and consumption equations (5) and (8) to derive expressions for changes in expected income and utility from consumption by type. In any sector j,

$$\hat{\mathbb{W}}_n(\theta) = \left[\frac{\hat{w}_n^j(\theta)^{\varepsilon(L,\theta)} \exp(-\kappa(\theta)\Delta\tau_{nn}\varepsilon(L,\theta))}{\hat{\lambda}_{nn}^j(L,\theta)} \right]^{1/\varepsilon(L,\theta)} \tag{59}$$

$$\hat{\mathbb{C}}_n(\theta) = \left[\frac{\hat{p}_n^{-\varepsilon(C,\theta)} \exp(-\kappa(\theta)\Delta\tau_{nn}\varepsilon(C,\theta))}{\hat{\lambda}_{nn}(C,\theta)} \right]^{\alpha(\theta)/\varepsilon(C,\theta)} \tag{60}$$

We solve Equations (45) to (60) starting with an initial guess in each endogenous variable such that $\hat{x} = 1$, updating our guess till the algorithm has converged to an equilibrium. With the counterfactual changes in endogenous variables $\{\hat{B}_n(\theta), \hat{\mathbb{C}}_n(\theta), \hat{\mathbb{W}}_n(\theta), \hat{Q}_n\}$, the change in expected utility by type is:

$$\hat{U}(\theta) = \left(\sum_{n\in\mathbb{N}} \lambda_n(R,\theta)\hat{B}_n(\theta) \left(\hat{Q}_n^{-\gamma(\theta)} \hat{\mathbb{W}}_n(\theta)\hat{\mathbb{C}}_n(\theta) \right)^{\varepsilon(R,\theta)} \right)^{1/\varepsilon(R,\theta)} \tag{61}$$

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 A6 presents the distribution of the changes in travel time.

F Robustness Checks: Counterfactuals

F.1 Testing Model Predictions

To verify the veracity of our model, we test its predictions 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 A19a, we plot the reduced-form relationship between predicted changes in consumption travel according to the model, $\hat{\lambda}_{nl}(C,\theta)$, 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 A19b, we plot the reduced-form relationship between predicted changes in work location travel according to the model, $\hat{\lambda}^{j}_{ni}(L,\theta)$, 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 in travel patterns (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 A19c, we plot the reduced-form relationship between predicted changes in residential shares according to the model, $\hat{\lambda}_{n}(R,\theta)$, 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 in residence choices 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 A19d, we plot the reduced form relationship between predicted changes in non-tradable floor space according to the model, $\hat{H}^{(0)}_{i}$, and observed entry of food establishments from the data. We find a positive, statistically-significant correlation between these quantities. Thus, the model has predictive power over changes in commercial activity in response to the Downtown Line.

F.2 Impacts of Other Lines in Singapore

In our main results, we find that the poor experience almost no welfare gains from the DTL, driven by reductions in employment access. Do these results generalize to other train lines? We consider two auxiliary counterfactuals. First, we remove the North South Line (NSL), one of the two original train lines in the Singapore metro system. The NSL is a trunk line primarily serving the northern neighborhoods of Woodlands.[75] We feature this line because the suburban neighborhoods it serves are much more socioeconomically diverse than those of the DTL. Column 1 of Table A16 displays the associated results. When the NSL is removed, we predict that welfare will fall for both low and high types, by 4.23% and 1.96% respectively. Both types experience reductions in consumption and employment access. Low income workers experience a 2.2 p.p. greater loss in welfare compared to high income workers. Our takeaway is that line placement matters for understanding the impacts of transit expansion on welfare and equity.[76]

Second, in a stylized exercise, we consider what happens if the entire rail system never existed, maintaining the current level of bus service. Column 2 of Table A16 displays the associated results. We predict that welfare falls by about 13-14% for low and high types respectively, driven by reductions in consumption and employment access of about 4-5% and 6-7%. We view our results as suggestive evidence of the robustness of Singapore's (bus) transit system, though the additional funds saved from not building rail could be used to bolster the bus and road network. The welfare

[75] The NSL begins in the west-southwestern suburbs of Jurong, travels due north through Woodlands and then veers due south into downtown.

[76] Can the alignment of the DTL be improved? We discuss two alternative line placements, each with their own flaws. The DTL takes the most direct route from the Bukit Panjang suburbs to the downtown area, passing by the wealthy neighborhoods of Bukit Timah. A "northern" candidate would run underneath the Central Catchment Area and meet the Circle Line at Caldecott Station, incurring high construction costs from tunnel construction under a large water body. A "southern" candidate connects to the East West Line at Dover and/or the Circle Line at one-north before continuing downtown, duplicating the (wealthy) neighborhoods served by the latter two lines.

improvements from an optimal urban transit network in Singapore is an open question (see, e.g., Fajgelbaum and Schaal 2020).

F.3 No Spillovers

Ignoring spillovers, we find similar results to our baseline estimates. We eliminate agglomeration or residential externalities by setting $\mu(A)$ and $\mu(U,\theta)$ to zero, presenting the associated results in Table 2. We find that low-income workers are slightly worse off than under our baseline, 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 a smaller effect on segregation due to the lack of residential externalities.

F.4 Covariates-Based Approach

Since our spatial data are granular, we re-estimate our counterfactuals using the "covariates-based approach" as proposed by Dingel and Tintelnot (2020) to account for potential overfitting. Instead of using observed initial shares of travel flows across neighborhoods, $\{\lambda_{nl}(C,\theta),\lambda_{ni}(L,\theta)\}$, we use fitted shares from Poisson Pseudo Maximum Likelihood gravity regressions of Equations (4) and (7). We present these results in Table A14, finding that they are similar to our mainline estimates.

F.5 Only Reduction in Consumption Travel Time

Given changes in consumption travel patterns, what is the partial equilibrium effect of the induced change in labor demand for low-income workers? We re-estimate our counterfactual assuming that the only impact of the DTL on the economy is a reduction in consumption travel time. In this exercise, we assume commute times remain unchanged. We present these results in Table A17. We find that workers in the non-traded sector travel 1% longer because labor demand increases downtown. Then, welfare rises by 0.9% for high income workers and decreases by 0.3% for low income workers. The increase in welfare is smaller for high-income workers than in the baseline because they no longer benefit from shorter commutes on the DTL, only from improved consumption travel. Low-income workers are even more worse off because they have to travel farther to work downtown, but they cannot use the DTL to do so.

F.6 Local Consumption

To highlight the importance of factoring in consumption travel, we consider a version of our model where all consumption is local to the neighborhood of residence. Then the consumption location decision (l) and residential location decision (n) are made jointly (i.e., fixing $l = n$). First, the worker chooses a single location in which to live and work. Then, she chooses a workplace. Note that, in this specification, the travel time for consumption is always zero.

We present counterfactual results with this model in Table A17. We find similar results to that of the original comparison with only work travel, as exposited in Section 6.2.2. Without non-local consumption trips, the model understates the welfare effects of the DTL: it fails to capture the gains from faster consumption travel. Moreover, this model also misses the re-organization of non-tradable production across the city in response to changes in consumption travel.

F.7 Roundabout Production

We show that allowing for roundabout production in tradable and non-tradable sectors does not change our qualitative results.

We modify our production model following Caliendo and Parro (2015). In each location $i \in \mathbb{N}$, firms operate in tradable and non-tradable sectors $j \in \{0, 1\}$. Sectoral productivity differs across locations. Firms produce under perfect competition and with a constant returns to scale technology. Departing from our mainline specification, tradable goods ("materials") are also used as an input into production. Firm production is a Cobb-Douglas aggregate over labor L_i^j, commercial floor space H_i^j, and materials m_i^j. Output is thus specified as

$$Y_i^j = A_i^j (L_i^j)^{\beta^j(L)} (H_i^j)^{\beta^j(H)} (m_i^j)^{1-\beta^j(L)-\beta^j(H)}.$$

Both labor and floor space shares, $\beta^j(L)$ and $\beta^j(H)$, lie in $[0, 1]$. Labor input L_i^j remains a Cobb-Douglas aggregate over each worker group's effective units of labor, as in Equation (13).

In location i, let p_i^j be the price of variety j and W_i^j the cost of labor in sector j. Solving the firm's profit maximization problem, we find that the demand for labor, commercial floor space, and materials, as well as wages, can be written as

$$L_i^j = \beta^j(L)\frac{p_i^j Y_i^j}{W_i^j} \tag{62}$$

$$H_i^j = \beta^j(H)\frac{p_i^j Y_i^j}{q_i} \tag{63}$$

$$m_i^j = [1-\beta^j(L)-\beta^j(H)]\frac{p_i^j Y_i^j}{p_i^{(1)}} \tag{64}$$

$$\tilde{N}_i^j(\theta) = \beta^j(\theta)\frac{L_i^j W_i^j}{w_i^j(\theta)} \tag{65}$$

Market clearing in non-tradables now takes the form

$$\begin{aligned} &p_i^{(0)}A_i^{(0)}(L_i^{(0)})^{\beta^{(0)}(L)}(H_i^{(0)})^{\beta^{(0)}(H)}(m_i^{(0)})^{1-\beta^{(0)}(L)-\beta^{(0)}(H)} \\ =&\sum_{n\in\mathbb{N}}\sum_{\theta\in\{+,-\}}\alpha(\theta)\lambda_{ni}(C,\theta)\lambda_n(R,\theta)R(\theta)\mathbb{W}_n(\theta). \end{aligned} \tag{66}$$

In calibration, we use the 2015 input-output table for Singapore provided by the Asian Development Bank. All calibrated parameters remain identical except for two sets. In each sector, the aggregate labor share is now taken as an expenditure-weighted average of the share of value added at basic prices.[77] We maintain the same floor space share as in our main specification, with the residual expenditure accruing to materials. The new calibrated labor share is 0.450 for non-tradables and 0.300 for tradables, lower than in our main specification.

We perform hat algebra as in Section E. The only modification is to Equation (52) for the price of the non-tradable good. The new hat equation in output also requires relative changes in materials. Since the price of the tradable good is our numeraire, relative changes in materials used

[77] Industries c1 to c21, spanning "Agriculture, hunting, forestry, and fishing" and "Retail trade, except of motor vehicles and motorcycles; repair of household goods", are taken to be tradable sectors; and industries c22 to c34, spanning "Hotels and restaurants" and "Other community, social, and personal services", are taken to be non-tradable sectors.

in the non-tradable sector evaluates to[78]

$$\hat{m}_i^{(0)} = \hat{p}_i \hat{Y}_i^{(0)}.$$

Relative changes in non-tradable output can be written

$$\hat{Y}_i^{(0)} = \hat{A}_i^{(0)} (\hat{L}_i^{(0)})^{\beta^{(0)}(L)} (\hat{H}_i^{(0)})^{\beta^{(0)}(H)} (\hat{m}_i^{(0)})^{1-\beta^{(0)}(L)-\beta^{(0)}(H)}.$$

Finally, the relative changes in prices of non-tradables in each location i evaluates to

$$\hat{p}_i = \frac{\sum_{n\in\mathbb{N},\theta\in\{+,-\}} \alpha(\theta) R(\theta) \mathbb{W}_n'(\theta) \lambda_{ni}'(C,\theta) \lambda_n'(R,\theta)}{\hat{Y}_i^{(0)} \sum_{n\in\mathbb{N},\theta\in\{+,-\}} \alpha(\theta) R(\theta) \mathbb{W}_n(\theta) \lambda_{ni}(C,\theta) \lambda_n(R,\theta)}. \tag{67}$$

Results can be found in Table A17. We find that the results are qualitatively consistent with our baseline results. Welfare increases by 2.13% for high-income workers, while the welfare impact for low-income workers is very small at a 0.39% increase. Access to consumption improves for both workers, with high income workers benefiting by 1.4%, higher than low-income workers at 1.0%. However, while access to employment (expected wages net of commuting costs) increases for high-income workers by 0.8%, low-income workers experience a 0.6% decline. Overall, high income workers experience a 1.7 p.p. larger increase in welfare compared to low income workers.

F.8 Details of Model Decomposition

In this section, we conduct a decomposition exercise to assess how robust our inequality findings are to different model assumptions. Our baseline specification will be the version of our model with no spillovers, as in Column 1 of the bottom panel of Table 2. The main results are summarized in Table 3. First, low-income workers spend less on non-tradables than the high-income. Thus, they experience smaller gains in welfare from being better able to access consumption opportunities. Second, our calibration assumes low-income workers could have slightly lower travel costs than high-income workers, so they gain less from shorter travel times. Third, we assumed low-income workers have higher travel elasticities; hence, in equilibrium, they can better substi-

[78]Our choice of numeraire also means that Equation (55), the hat equation for high type workers, is unchanged.

tute to attractive work and consumption locations. Therefore, they may benefit less from lower travel times. Lastly, our calibration assumes low-income workers have relatively more-dispersed residential preferences; hence, they may be less able to move to take advantage of improved access elsewhere.

From our household expenditure data, we established that low-income workers have smaller expenditures on non-tradables than high-income workers. What happens if we had assumed low-income consumption shares were equal to their high-income counterparts? We set $(\alpha(-),\gamma(-))$ equal to $(\alpha(+),\gamma(+))$. We re-estimate the counterfactual, presenting our results in Row 2 of Table 3. We find slightly larger disparate impacts of the DTL across worker groups relative to the baseline. Overall the differential impact of the DTL is slightly exacerbated because increasing non-tradable consumption strengthens the impact of non-tradable jobs moving to less attractive locations. Therefore, access to employment is worsened. However, this effect is partially offset because consumption access is improved for low-income workers; thus, inequality in consumption is attenuated. In the previous baseline, high-income workers experienced 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. However, setting consumption shares to be equal only changes the disparity of the impact of the DTL slightly, widening the difference in the welfare impact across worker groups by 0.2 percentage points. The mechanisms from our main analysis apply.

Next, our calibration assumes that low-income workers have cheaper travel costs than high-income workers. To assess the importance of this assumption, we set low-income travel costs, $\kappa(-)$, to be equal to that of high-income workers, $\kappa(+)$. The counterfactual is re-estimated in Row 3 of Table 3.[79] We find the DTL has slightly smaller impacts on inequality relative to the baseline. Inequality effects are smaller across both access to consumption and employment for low-income workers. If their travel costs were higher, high-income workers could experience larger gains in welfare from improved access. However, setting travel costs to be equal only mitigates the

[79]One policy exercise corresponding to setting travel costs equal is if the government removed travel subsidies for low-income workers.

disparate effects of the DTL slightly, narrowing the difference in the welfare impact across worker groups by 0.3 percentage points. The mechanisms from our main analysis still apply.

Furthermore, we assumed that low-income workers have higher travel elasticities, or a smaller dispersion of idiosyncratic consumption amenities and productivities across locations, relative to high-income workers. To assess the robustness of our main findings, we set low-income dispersion parameters for work and consumption equal to those for high-income workers: $(\varepsilon(L,-),\varepsilon(C,-)) = (\varepsilon(L,+),\varepsilon(C,+))$. We then re-estimate the counterfactual, presenting results in Row 4 of Table 3. We find smaller impacts of the DTL on inequality relative to the 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. If they have lower travel elasticities, they are less able to substitute to more attractive work and consumption locations in equilibrium. Therefore, high-income workers would benefit more when travel time falls. However, setting travel elasticities to be equal only mitigates the disparate effects of the DTL partially, narrowing the difference in the welfare impact across worker groups by just 0.3 percentage points. These inequality results are still driven by the mechanisms highlighted in our main analysis.

Finally, our calibration assumes that low-income workers have greater dispersion in idiosyncratic residential amenities across locations than high-income workers do. To assess the robustness of our main findings, we set the low-income residential dispersion parameter, $\varepsilon(R,-)$, to be equal to the one for high-income workers, $\varepsilon(R,+)$. Then, we re-estimate the counterfactual in Row 5 of Table 3. We find the DTL has slightly smaller impacts on inequality relative to the baseline. If they have larger residential elasticities, 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 disparate effects of the DTL partially, narrowing the difference in the welfare impact across worker groups by just 0.1 percentage points. The main mechanisms from the previous paragraph still drive these results.

G Cost Effectiveness Analysis

In this subsection, we evaluate the cost effectiveness of the Downtown Line.

G.1 Framework

We calculate two measures of cost effectiveness. First, we compute Marginal Value of Public Funds (MVPF) for the construction of the Downtown Line. The MVPF measures the amount of welfare that can be delivered to policy beneficiaries per dollar of government spending on the policy. Equivalently, the MVPF measures the shadow price of raising revenue from the beneficiaries of the policy by reducing spending. Following Hendren and Sprung-Keyser (2020), we define MVPF as the aggregate willingness to pay for the DTL, WTP, divided by the net cost to the government:

$$\text{MVPF} = \frac{\text{WTP}}{\text{Net Cost}} = \frac{\int_i \text{WTP}_i}{\text{Net Cost}} \tag{68}$$

where WTP_i is the individual i's willingness to pay for the DTL out of his/her own income. Net cost is inclusive both of the cost of the DTL and all other impacts of behavioral responses on the government budget such that Net Cost = Cost − Gov. Revenue. Specifically, if the DTL increases wages, net costs should incorporate the impact of those increases in future tax receipts.

Next, we calculate a social benefit-cost ratio, BCR. Following Heckman et al. (2010), we compare the net social benefits of the DTL, inclusive of benefits that accrue back to the government, against the budgetary spending on the policy. We use the following formula:

$$\text{BCR} = \frac{\text{WTP} + (1+\phi)\text{Gov. Revenue}}{(1+\phi)\text{Cost}} \tag{69}$$

where $\phi = 0.5$ is the marginal dead weight loss of raising government revenue. Policies are then deemed to pass the cost-benefit test if BCR exceeds 1.

We assume a discount rate of 3% to compute the net present value of future cash flows.

G.2 Downtown Line Costs

We compute the total cost of the Downtown Line. According to the Land Transport Authority, the total cost of construction for the Downtown Line is 20.7 billion SGD (Straits Times 2012). The Ministry of Transport reports that the operational cost of the Downtown Line is 132 million SGD each year. However, a total of 82.2 million SGD is offset by revenue, of which 69 million SGD is from fare revenue and 13.2 million SGD is from non-fare revenue (advertising etc.).[80] We discount all future operating costs and revenues to the present. Together, the total net present value cost of the DTL (including construction, operation, and revenue), Cost, is 22.4 billion SGD (See Panel A in Table A18).

G.3 Government Revenue

Next, we account for the fiscal externalities resulting from the impact of the Downtown Line on worker earnings. Since low-income workers pay zero income taxes, we focus of high-income workers. In Section 6.2.1, we estimated that the DTL causes a 0.77% increase in the expected earnings of high income workers. According to the General Household Survey, the median high income worker earned 63.5 thousand SGD per year in 2015. Thus, the DTL increased earnings by 489 SGD per high income worker per year. At this income level, the marginal tax rate is 7%.[81] Aggregating over the high-income population and discounting future tax revenue to the present gives us that the net present value increase in government tax revenue, Gov. Revenue, is 2.3 billion SGD (See Panel B in Table A18).

G.4 Willingness to Pay

Last, we compute willingness to pay as the dollar amount such that workers are indifferent between the construction of the DTL and receiving an increase in wages, as consistent in our model. In Section 6.2.1, we estimated that the DTL causes a 1.82% increase in the expected utility of high income workers. Using Equation (11), this is equivalent to a 1,155.92 SGD increase in yearly income. Similarly, we estimated that the DTL causes a 0.01% increase in the expected utility of low income workers, which is equivalent to a 21.84 SGD increase in yearly income. Aggregating over the high- and low-income populations and discounting future earning gains to the present implies that the total net present value willingness to pay for the Downtown Line, WTP, is 100.9 billion SGD (See Table A18 Panel C).

[80] As reported in Straits Times (2019) and Ministry of Transport (2019).

[81] The marginal income tax rate between 40,000 to 80,000 SGD is 7%. See https://www.iras.gov.sg/.

G.5 Estimates of Cost Effectiveness

We find that the Downtown Line is highly cost-effective (See Table A18 Panel D). Using Equation (68), we estimate a Marginal Value of Public Funds of 5.02. This implies that the DTL provides more than five dollars of benefits per dollar of government spending. This is a larger "bang-for-buck" than for typical policies targeted at adults such as health insurance, housing vouchers or food stamps which Hendren and Sprung-Keyser (2020) estimate range between 0.5 to 2. However, the MVPF of the Downtown Line is smaller than that of many investments into low-income children's health and education with MVPFs above 10 (Hendren and Sprung-Keyser 2020).[82] Using Equation (69), we calculate that the Social Benefit-Cost Ratio is 3.1 with a marginal dead weight loss of raising government revenue, $\phi = 0.5$. At $\phi = 0$, the Social Benefit-Cost Ratio is 4.6.

References for the Online Appendix

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

Allen, Treb, Costas Arkolakis, and Xiangliang Li (2022). "On the equilibrium properties of network models with heterogeneous agents".

Arkolakis, Costas, Fabian Eckert, and Rowan Shi (2021). *Combinatorial discrete choice*. Working Paper.

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

Caliendo, Lorenzo and Fernando Parro (2015). "Estimates of the Trade and Welfare Effects of NAFTA". *The Review of Economic Studies* 82.1, 1–44.

Department of Statistics, Singapore (2015). *General Household Survey 2015*. Dataset.

[82] In our model estimation, we focused on low- and high-income workers. Our analysis omits the benefit of the DTL accruing to other transit users, including students, senior citizens and military conscripts. Therefore, our estimated Marginal Value of Public Funds likely understates its true value.

Department of Statistics, Singapore (2019). *Household Expenditure Survey 2017/18*. Dataset.

Department of Transportation (2017). *2017 National Household Travel Survey*. Survey.

Dingel, Jonathan and Felix Tintelnot (2020). “Spatial Economics for Granular Settings”.

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

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

Fajgelbaum, Pablo D and Edouard Schaal (2020). “Optimal transport networks in spatial equilibrium”. *Econometrica* 88.4, 1411–1452.

Guerrieri, Veronica, Daniel Hartley, and Erik Hurst (2013). “Endogenous gentrification and housing price dynamics”. *Journal of Public Economics* 100, 45–60.

Heckman, James J. et al. (Feb. 2010). “The rate of return to the HighScope Perry Preschool Program”. *Journal of Public Economics* 94.1-2, 114–128.

Hendren, Nathaniel and Ben Sprung-Keyser (Aug. 2020). *A Unified Welfare Analysis of Government Policies*.

Housing and Development Board (2019). *Resale Flat Prices*. Dataset.

Ministry of Manpower, Singapore (2018). *Labour Force in Singapore 2018*. Dataset.

Ministry of Transport (2019). *Oral Reply by Minister for Transport Khaw Boon Wan to Parliamentary Question on Government Subsidies for Operating Costs for the Thomson-East Coast Line*. Parliamentary Report.

Open Government Products, Singapore (2019). *Data.gov.sg*. Dataset.

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

Straits Times (2012). *Downtown Line Costs Soar By More Than 70%*. Newspaper.

— (2019). *Dollars and sense of a quality public transport system*. Newspaper.

Teo, You Yenn (2022). *This is what inequality looks like*. Ethos Books.

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

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

Urban Redevelopment Authority, Singapore (2022). *Real Estate Information System (REALIS)*. Dataset.

H Appendix Figures and Tables

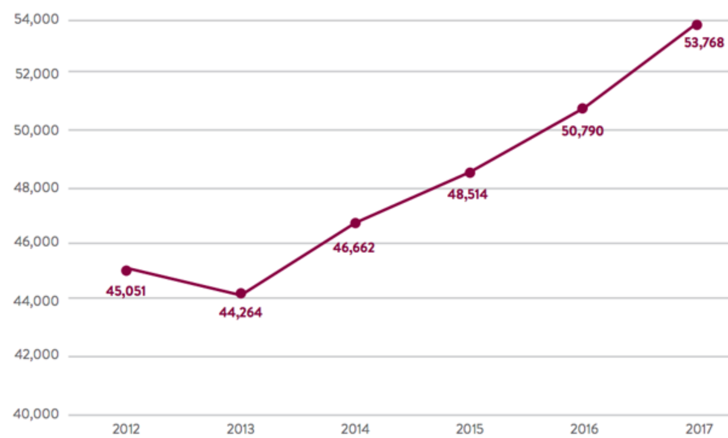


Figure A1: Mass-Transit Ridership over Time (in Millions)

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

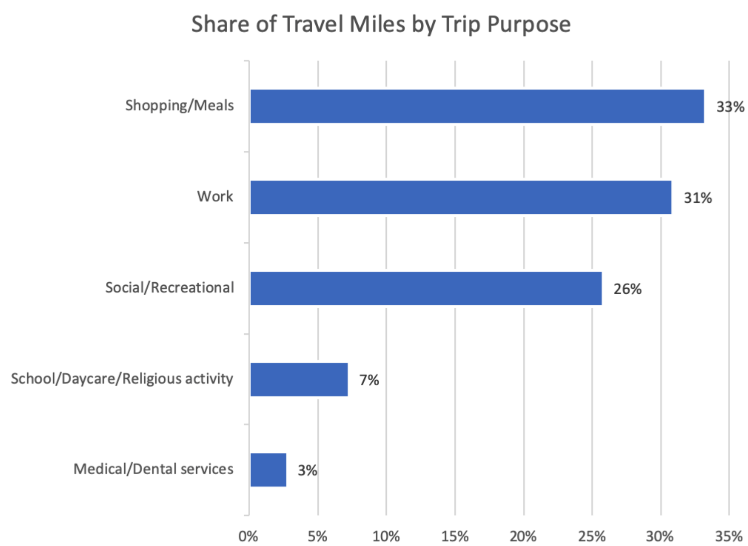
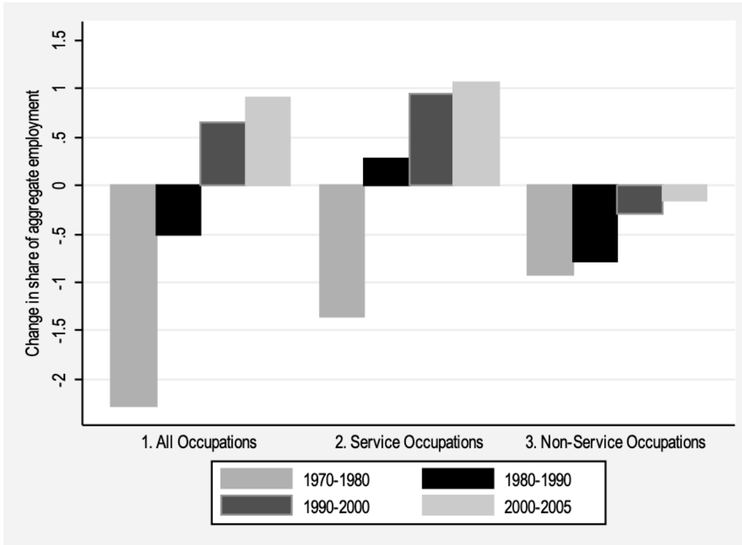


Figure A2: Purpose of Travel: National Household Travel Survey 2017

Notes: This figure presents the share of travel miles classified by purpose of trip in 2017.

Source: Department of Transportation (2017)



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

Figure A3: Rise in Low-Income Non-tradable Service Sector Jobs: Autor and Dorn (2013)

Notes: This figure displays 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)

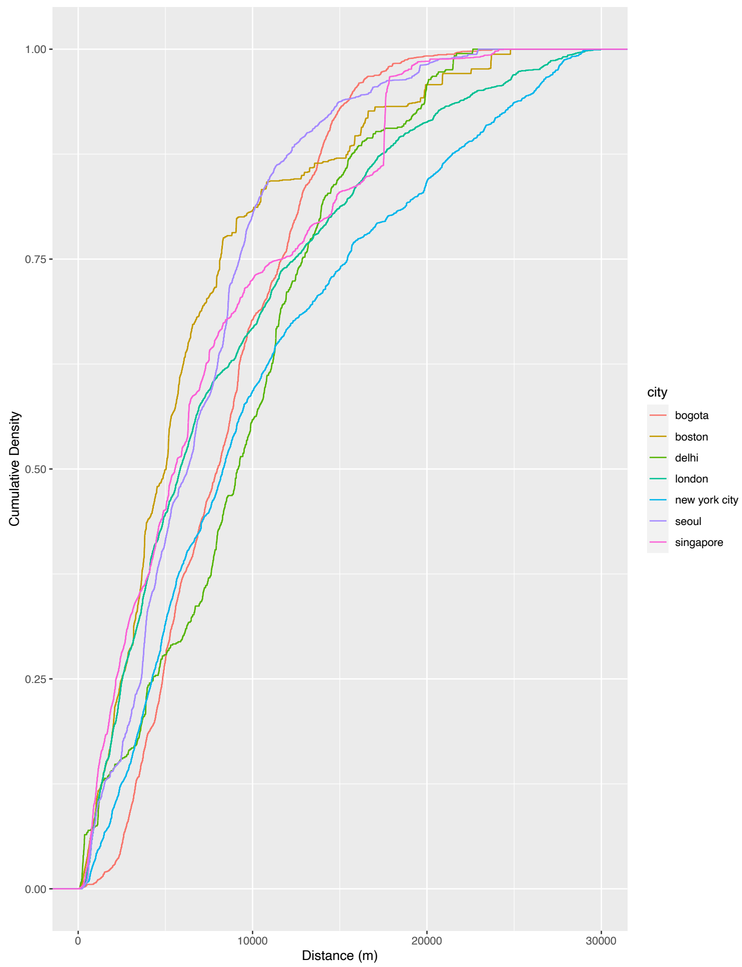
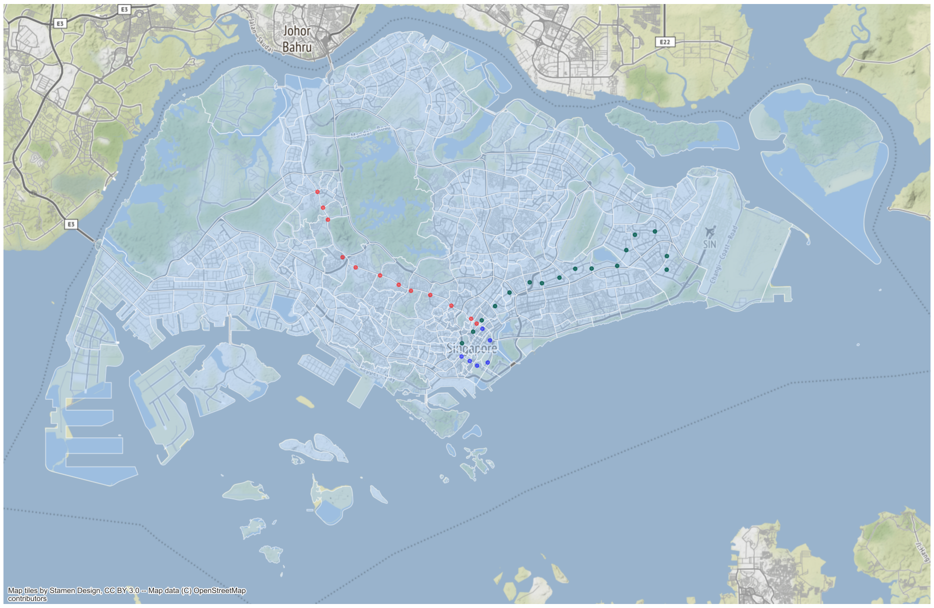


Figure A4: Distance between Restaurants and Central Business District: Cumulative Density Plot

Notes: This figure displays the cumulative density distribution of the distance between restaurants in various cities to the central business district. Data is taken from Open Street Map.



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.

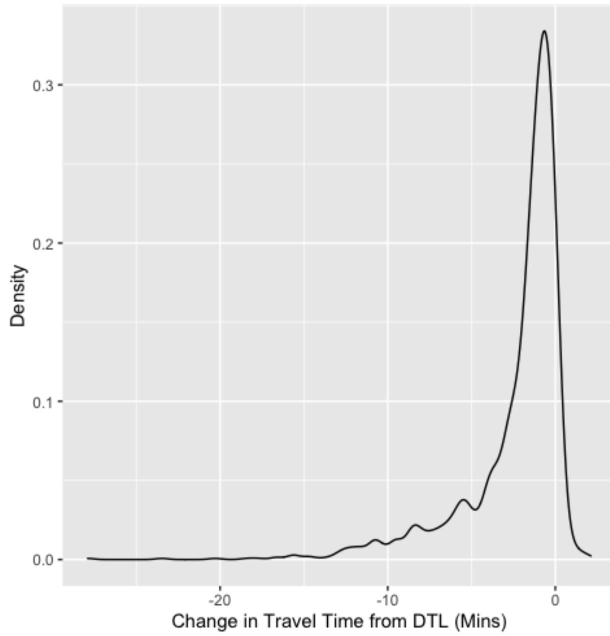
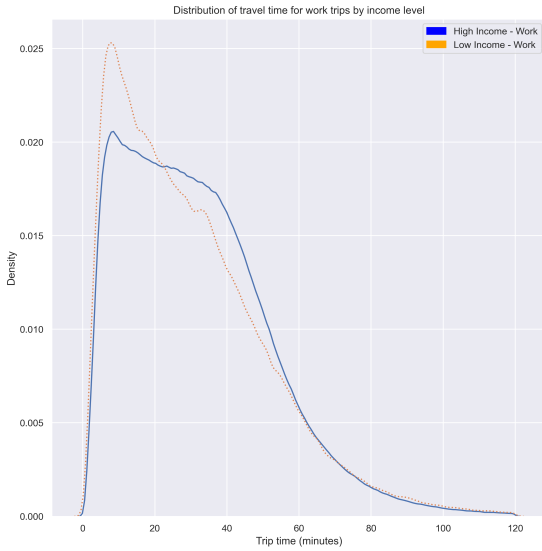


Figure A6: Changes in Bilateral Transit Time: Downtown Line

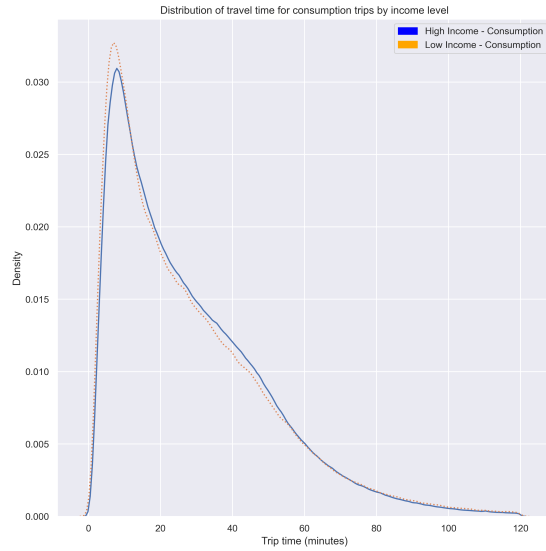
Notes: This figure displays the distribution of the changes in bilateral travel times in minutes across all neighborhoods. These farecard data, spanning 2015 to 2018, are from the Land Transport Authority of Singapore. Mean travel times are computed for subzone pairs before and after DTL2 opened.

Figure A7: Farecard Data Heatmap

Notes: This figure displays a heat map of log farecard data weekend trips by planning area. Farecard data are from the Land Transport Authority of Singapore.



(a) Work



(b) Consumption

Figure A8: Distribution of Travel Time

Notes: This figure presents the distribution of travel times for work in panel (a) and for consumption in panel (b) by high- and low-income workers respectively (above and below the 25th percentile). Travel time is computed at the trip level from farecard data, pooling across all years. These farecard data are from the Land Transport Authority of Singapore.

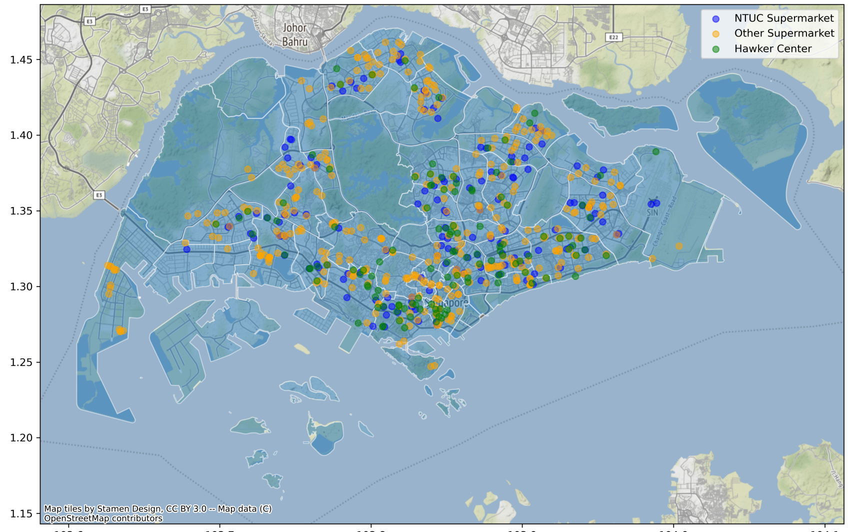


Figure A9: Amenities Data on Supermarkets and Hawker Centers

Notes: This figure plots all supermarkets and hawker centers in Singapore. Data are from the National Environment Agency of Singapore and the Singapore Food Agency.

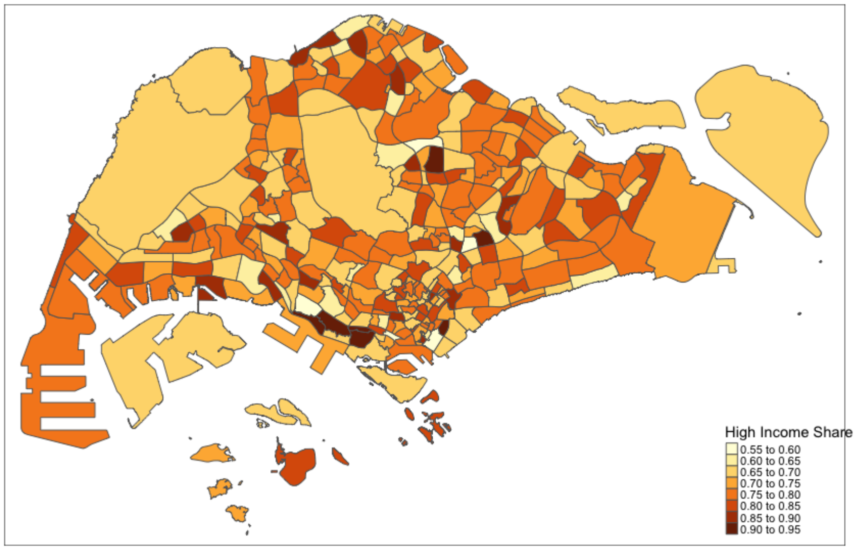


Figure A10: High-Income Employment Share

Notes: This figure plots the high-income employment share in each subzone/neighborhood. Shares are computed from farecard data from 2015 from the Land Transport Authority of Singapore. Workplaces are imputed as the modal morning destination and modal evening origin of each farecard.

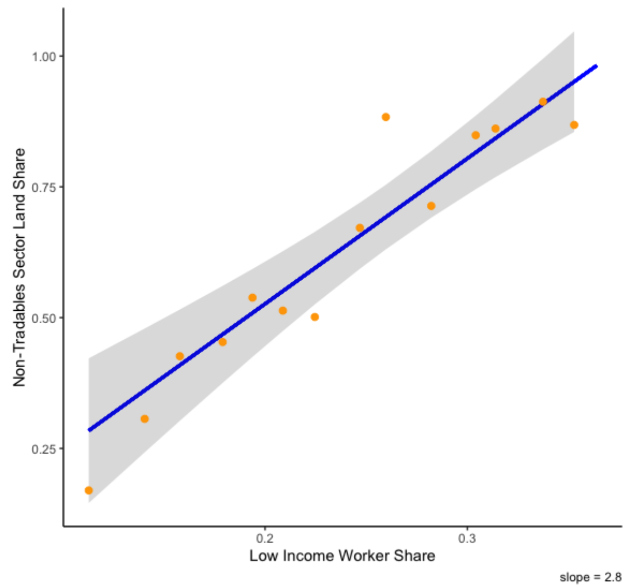


Figure A11: Relationship Between Low-Income Employment and Commercial Land Share for Non-Tradables

Notes: This figure displays 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. Low-income employment shares are computed from the modal morning destination and modal evening origin for Workfare farecards. The share of commercial land is taken from the REALIS data set from the Urban Redevelopment Authority of Singapore. Each point shows the mean of the variable with low-income employment share in the given bin.

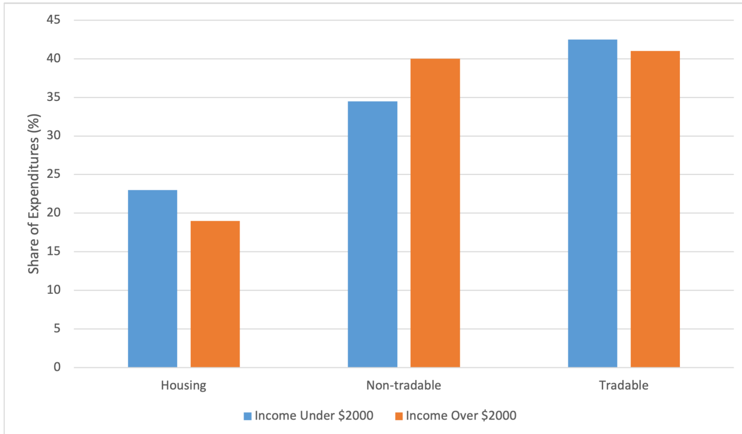


Figure A12: Expenditure Shares by High- and Low-Income Workers: Household Expenditure Survey 2018

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. Data are from the Household Expenditure Survey of Singapore.

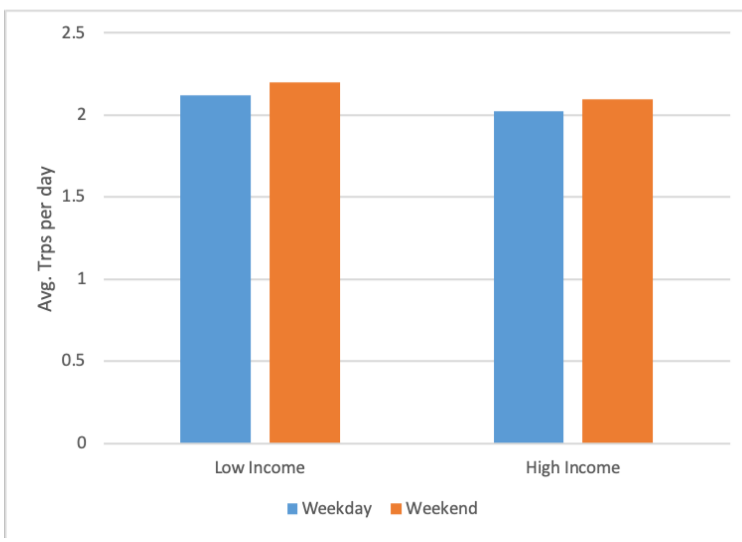


Figure A13: Daily Trips by High- and Low-Income Workers

Notes: This figure displays the average number of daily trips on weekdays and weekends by high- (Adult) and low-income (Workfare) workers, according to farecard data pooled over 2015 to 2018. Farecard data are from the Land Transport Authority of Singapore.

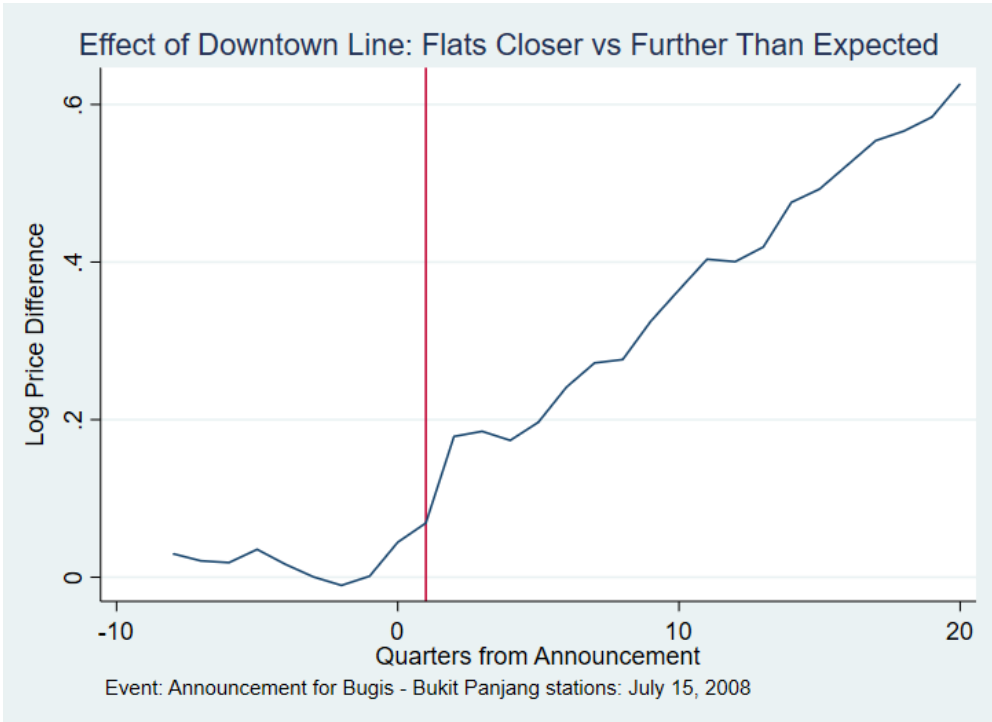
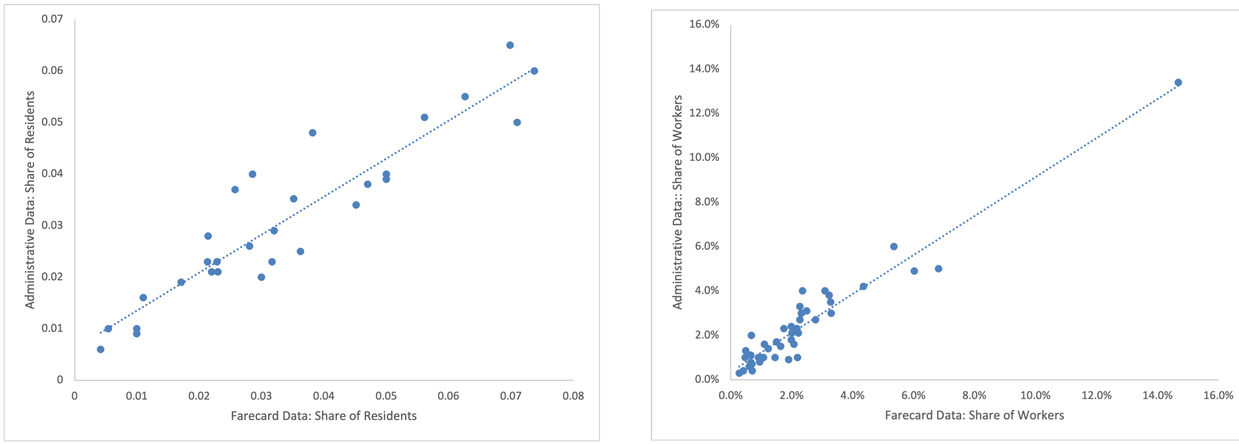


Figure A14: Effect of Downtown Line: Flats Closer vs Further Than Expected.

Notes: This figure plots the log difference in mean residential prices between government apartments where the distance to a Phase 2 DTL station is less than the expected distance to a station assuming that stations are random uniformly distributed along the DTL and apartments where the distance to a Phase 2 DTL station is more than the expected. We restrict our sample to addresses within 5km of a Phase 2 DTL station. Our data are a geocoded balanced panel of resale transactions for Housing and Development Board flats in Singapore.



(a) Residence

(b) Workplace

Figure A15: Identified Residence and Workplace: Fare card vs Administrative Data

Notes: The scatter plot in Panel A compares our imputed share of residents living in each neighborhood with shares from census data. are compared with residences reported in administrative data from the census. The scatter plot in Panel B compares our imputed share of workers employed in each neighborhood with workplaces reported in administrative data from the census. Farecard data are from the Land Transport Authority of Singapore. Residences are imputed from the modal morning origin and modal evening destination of each farecard, while workplaces are imputed from the modal morning destination and modal evening origin.

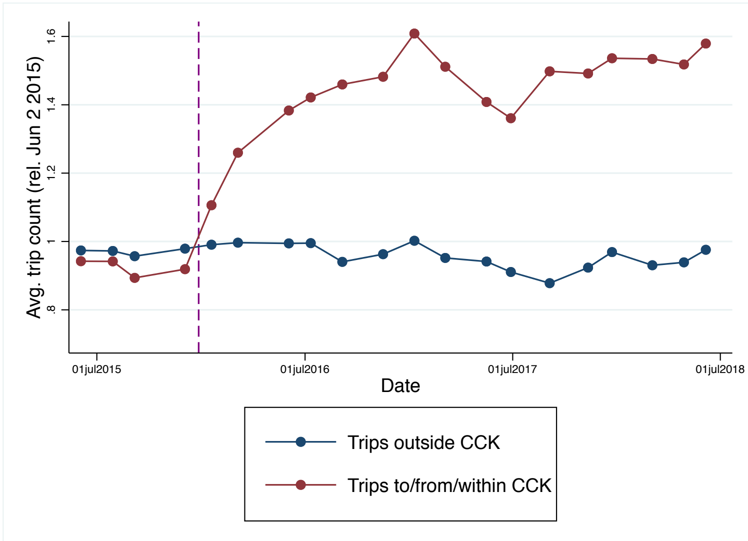
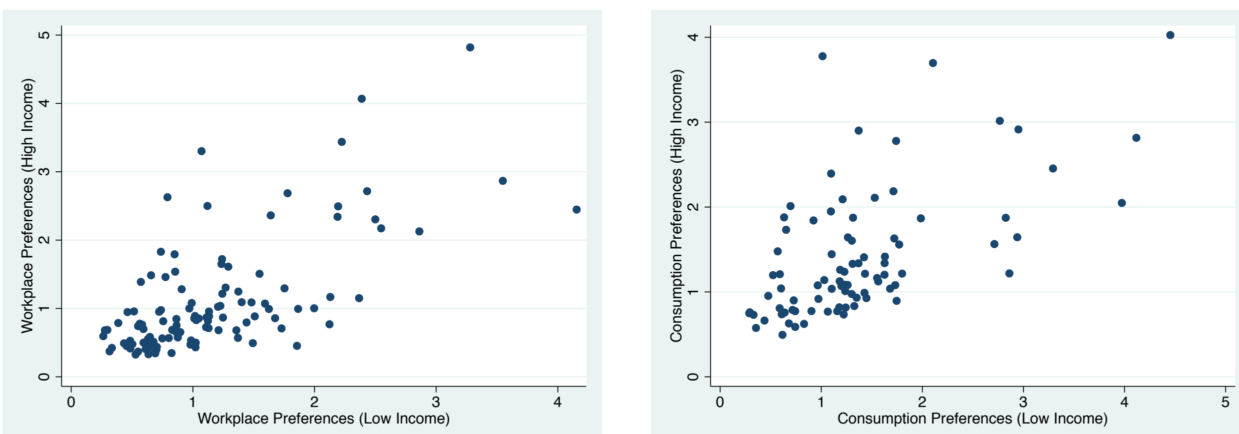


Figure A16: The Downtown Line and Long-Run Travel: Choa Chu Kang (CCK) Stations

Notes: This figure plots the quarterly volume of trips to and from Choa Chu Kang, compared to trips originating and ending at all other stations. The plot is made with farecard data from the Land Transport Authority between July 2015 and July 2018. The number of trips are shown relative to the number in June 2015.



(a) Work

(b) Consumption

Figure A17: 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). Location attractiveness measures arise as fixed effects from estimating gravity equations in Section 5. These measures differ between high- and low-income workers and at the subzone level.

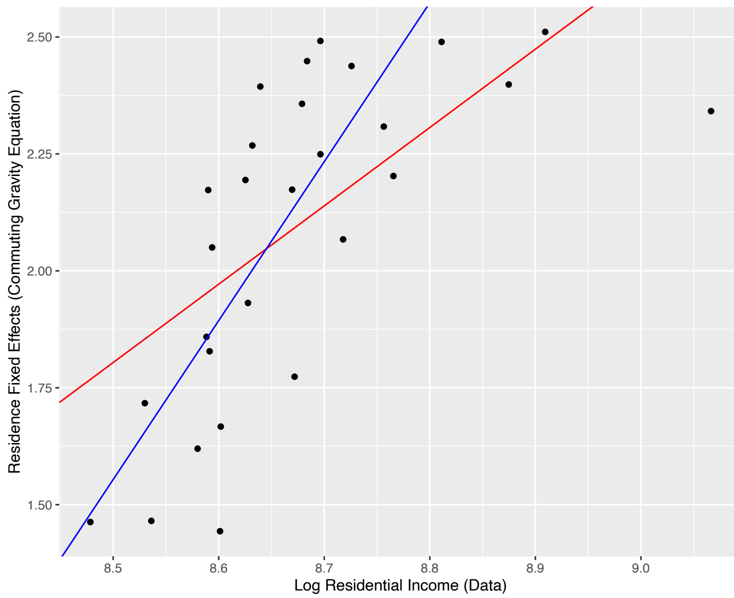
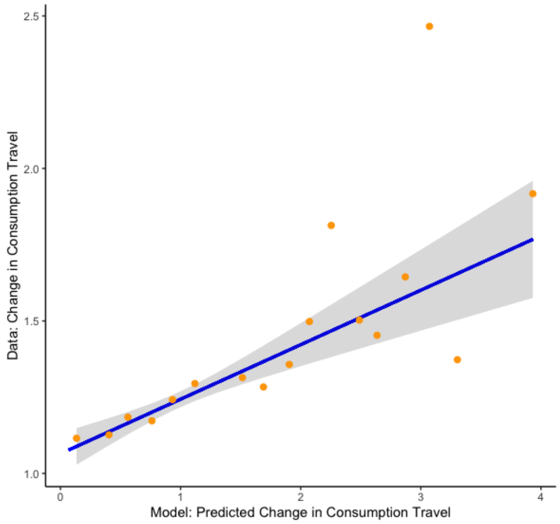
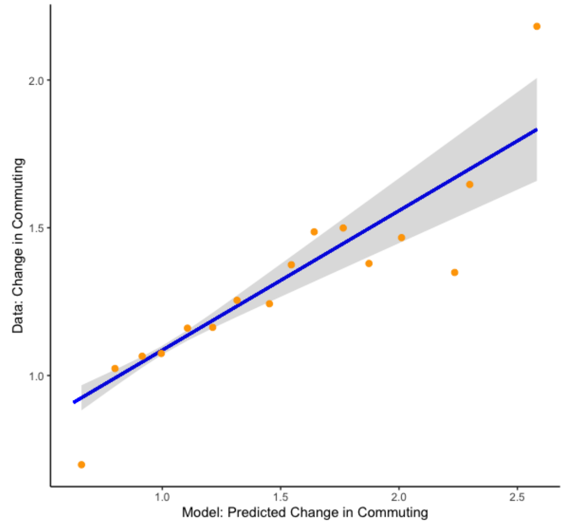


Figure A18: Expected Income: Data vs Model

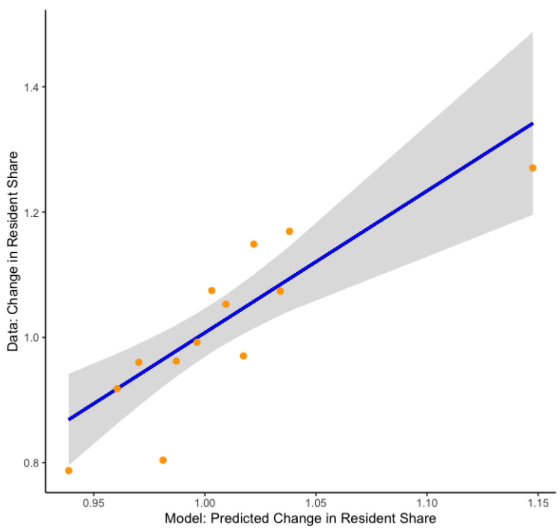
Notes: This figure displays a binned scatter plot of log model-estimated expected income by residence against log observed data on income by residence in 2015. Data are from the Singapore General Household Survey in that year. Each point shows the mean of the variable with observed data on income in the given bin. The red line is the linear fit across all neighborhoods, and the blue line is the linear fit dropping the 4 neighborhoods with the highest income.



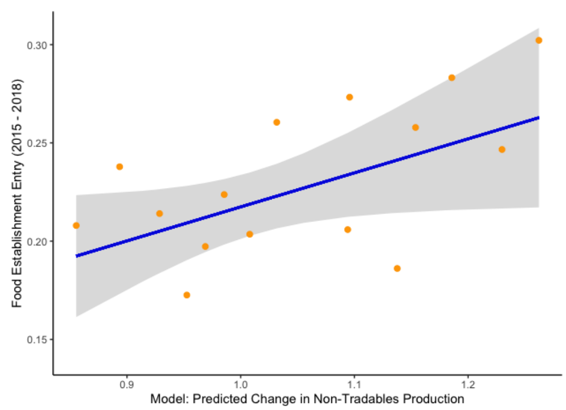
(a) Consumption Travel



(b) Work Travel



(c) Residential Shares



(d) Food Establishment Entry

Figure A19: 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 consumption 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 work 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, as well as the entry rate of food establishments between 2015 and 2018 by neighborhood. Data for panel (d) are from the Singapore Food Authority. Farecard data are from the Land Transport Authority of Singapore.

Table A1: Summary of Mass Rapid Transit Lines in Singapore

MRT line	Announced	Alignment announced	Open^
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
Thomson-East Coast	1/25/2008	8/29/2012 (Thomson) 8/15/2014 (East Coast)	~1/2020

- ^ date first section of line was opened
- * as Marina Line

Notes: This table displays announcement, alignment announcement and opening dates for selected mass-transit rail lines in Singapore.

Table A2: 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	**14%**	**86%**
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%

Notes: This table presents the share of employment, over high- and low-income workers, across sectors in Singapore. Data are from the 2018 Labour Force in Singapore Report.

Table A3: Share of Expenditures on Non-Tradable Goods and Services by High- and Low-Income Workers

(% of total expenditures)	**Income Under $2000**	**Income Over $2000**
Services		
Restaurants, cafes, and pubs	2.9	6.5
Fast food restaurants	2.0	1.4
Hawker centers, food courts, coffee shops,	11.5	9.4
canteens, kiosks and street vendors	11.5	9.4
Other food serving establishments	0.4	0.7
Personal care	7.9	7.4
Hospitality	0.7	1.3
Recreation and culture	4.1	6.5
Center-based education services	1.0	1.4
Retail Goods		
Clothing and footwear	1.7	2.6
Furnishings, carpets and other floor coverings	0.6	1.0
Household textiles	0.1	0.1
Household appliances	0.6	0.8
Glassware, tableware, and utensils	0.2	0.2
Other housewares and items	0.6	0.6
Total	34.4	40.0

Notes: This table presents the breakdown of expenditures (as a percentage of total expenditures) on non-tradable goods and services across high- and low-income groups in fiscal year 2017-2018 Data are from the Singapore Household Expenditure Survey of 2018.

Table A4: 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 displays statistics on the entry and exit of food establishments between 2015 and 2018. Data are from the Singapore Food Agency. The entry rate is defined as the share of 2018 food establishments that did not exist in 2015. The exit rate is defined as the share of 2015 food establishments that did not exist in 2018. Downtown Line Subzones are subzones with a DTL station.

Table A5: Effect of the Downtown Line on Housing Prices

	(1) Pooled	(2) 2 Years	(3) 3 Years	(4) 4 Years
Panel A. Flats Within 1km vs 5km				
Within 1km x Post	0.0180***	0.0190***	0.0346***	0.0450***
	(0.00257)	(0.00525)	(0.00591)	(0.00607)
Observations	139087	9383	9337	9276
Panel B. Flats Within 1km vs 4km				
Within 1km x Post	0.0212***	0.0213***	0.0373***	0.0484***
	(0.00260)	(0.00530)	(0.00597)	(0.00613)
Observations	114255	7706	7669	7622
Panel C. Flats Within 1km vs 3km				
Within 1km x Post	0.0185***	0.0142***	0.0278***	0.0386***
	(0.00268)	(0.00543)	(0.00611)	(0.00629)
Observations	76873	5186	5164	5134
Panel D. Robustness: Flats Closer vs Further Than Expected				
Closer than expected x Post	0.0238***	0.0111***	0.0326***	0.0410***
	(0.00173)	(0.00413)	(0.00484)	(0.00485)
Observations	139087	9383	9337	9276

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table displays regression estimates of Equation (34). Each observation is a residential unit - quarter pair. Panels A, B, and C include all flats within 5km, 4km and 3km of a Phase 2 DTL station respectively. 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 address of each apartment, interacted with its number of rooms. Standard errors are clustered at the address level. 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. Geocoded transaction data are from the Housing and Development Board. Panel D offers an additional specification for robustness. Closer than expected is an indicator for all flats where the distance to a Phase 2 DTL station is less than the expected distance to a station assuming that stations are random uniformly distributed along the DTL. All flats within 5km of a Phase 2 DTL station is included Panel D.*

Table A6: Gravity Estimation In Differences

	Change in Travel Share Conditional on Residence from 2015 to 2018			
	Commuting (High Income)	Commuting (Low Income)	Consumption (High Income)	Consumption (Low Income)
	(1)	(2)	(3)	(4)
Change in Travel Time (Minutes)	-0.045^{***}	-0.071^{***}	-0.048^{***}	-0.110^{***}
	(0.002)	(0.003)	(0.003)	(0.006)
Dest. FE	Yes	Yes	Yes	Yes
Origin FE	Yes	Yes	Yes	Yes
R^2	0.756	0.550	0.631	0.540
Adjusted R^2	0.749	0.545	0.623	0.534

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

Notes: This table displays results from estimating gravity equations in differences before and after DTL2 opened. Consumption equations (29) are in Columns 1 and 2, and workplace equations (28) are in Columns 3 and 4, for high- and low-income workers respectively. Destination and origin fixed effects are included. Each observation is a bilateral share of travel to a destination, conditional on residence, from 2015 and 2018 farecard data from the Land Transport Authority of Singapore.

Table A7: Workplace Attractiveness and Employment Density

	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 displays 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. Workplace attractiveness is measured in Section 5. Each observation is a neighborhood (subzone). Employment density is normalized and is computed by dividing total employment and total land area of the subzone (in square meters). Employment is imputed as modal morning destinations and modal evening origins in farecard data from the Land Transport Authority of Singapore. Subzone land areas are from the Singapore Land Authority.

Table A8: Employment Density and Idiosyncratic Productivity by Worker Type

	Parameter			
	$\varepsilon(L,+)$		$\varepsilon(L,-)$	
	(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 displays estimation results for Equation (30) *in Section* 5*. The dependent variable is* $\varepsilon(L,\theta)$ *multiplied by expected income by residence, as estimated in the model, and the explanatory variable is observed income by residence from the 2015 Singapore General Household Survey. We consider high-income workers (above the 25th percentile) in Columns 1 and 2, and consider low-income workers (below the 25th percentile) in Columns 3 and 4. In Columns 2 and 4, we drop the four neighborhoods with the highest income.*

Table A9: Consumption Location Attractiveness: Data vs Model

	Log Consumption 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 displays regression estimates of model-consistent measures of consumption attractiveness on various auxiliary 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 of food establishments from Singapore Food Agency Establishment data. Columns 3 and 4 only consider food establishments that are hawker stalls. Columns 5 and 6 only consider food establishments are food courts. Columns 7 and 8 consider the normalized total number of supermarkets from National Environment Agency data. Finally, columns 9 and 10 consider the normalized total number of clinics from Ministry of Health data.

Table A10: Estimates of Parameters Governing Travel Cost and the Dispersion of Idiosyncratic Consumption Amenities by High- and Low-Income

Parameter	**Estimate**
$\varepsilon(C,+)$	3.00
$\varepsilon(C,-)$	5.81
$\kappa(+)$	0.018
$\kappa(-)$	0.014

Notes: This table displays estimates of parameters governing travel costs, $\kappa(\cdot)$, and the dispersion of idiosyncratic consumption amenities, $\varepsilon(C,\cdot)$, by income type. (+) denotes high-income workers (above the 25th percentile) while (-) denotes low-income workers (below the 25th percentile).

Table A11: Residential Amenities and Externalities by High- and Low-Income

	Dependent variable:	
	$\Delta \ln \lambda_n(R,+)$	$\Delta \ln \lambda_n(R,-)$
	(1)	(2)
$\Delta \ln Q_n^{-\gamma(+)} \mathbb{W}_n(+) \mathbb{C}_n(+)$	1.481***	
	(0.407)	
$\ln(R_n(+)/R_n(-))$	0.332***	
	(0.085)	
$\Delta \ln Q_n^{-\gamma(-)} \mathbb{W}_n(-) \mathbb{C}_n(-)$		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 displays estimates of Equation (31). Column 1 considers high-income workers (above the 25th percentile), and Column 2 considers low-income workers (below the 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 neighborhood (subzone).

Table A12: 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 displays regression coefficients validating model amenities with auxiliary data. Model-consistent measures of residential amenities are regressed on, separately, the number of parks, community clubs, all schools and "neighborhood" schools in each neighborhood. We consider estimated high-income composite residential amenities in odd columns and low-income composite residential amenities in even columns. Model amenities are measured in Section 5. Each observation is a neighborhood (subzone). Columns 1 and 2 consider the normalized total number of 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 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 do not have gifted education programs.

Table A13: Calibrated parameters

Calibrated Parameter	Value	Source
$\alpha(-)$	0.34	2018 HES
$\alpha(+)$	0.40	2018 HES
$\gamma(-)$	0.23	2018 HES
$\gamma(+)$	0.19	2018 HES
$\beta^{(0)}$	0.605	2013 ESS
$\beta^{(1)}$	0.454	2013 ESS
$\beta^{(0)}(-)$	0.194	2018 LFS
$\beta^{(1)}(-)$	0.0596	2018 LFS
$\beta^{(0)}(+)$	0.806	$1-\beta(N,-)$
$\beta^{(1)}(+)$	0.940	$1-\beta(T,-)$
φ	0.25	Epple, Gordon, and Sieg (2010), etc.
$\mu(A)$	0.03	Rosenthal and Strange (2004), etc.

Notes: This table reports the values and sources of calibrated parameters. The 2018 HES is the 2018 Singapore Household Expenditure Survey. The 2013 ESS is the 2013 Economic Survey of Singapore. The 2018 LFS is the 2018 Singapore Labour Force in Singapore Report.

Table A14: Impact of the Downtown Line: Main Estimates using "Covariates-Based Approach"

	(1) Full Model		**(2) Only Commuting**	
% Change in	Low Type	High Type	Low Type	High Type
Welfare, $\bar{U}$	0.04	1.95	0.43	1.22
Access to Consumption, $\mathbb{C}$	1.15	1.53	0	0
Access to Employment, $\mathbb{W}$	-1.10	0.43	0.43	1.43
Gap in Welfare Impact	1.91		0.79	

Notes: This table displays results from the counterfactual estimation in Section 6 using the covariates-based approach from Dingel and Tintelnot (2020). Column 1 shows results for the full model. Column 2 displays results abstracting away from consumption trips. Welfare, access to consumption and access to employment are defined by the model in Section 4. Gap in Welfare Impact is the difference in percentage change in welfare across high-income and low-income workers. Segregation is measured by the dissimilarity index.

Table A15: Impact of the Downtown Line: Estimates with Perfectly Inelastic Housing Supply

	(1) Full Model		**(2) Only Commuting**	
% Change in	Low Type	High Type	Low Type	High Type
Welfare, $\bar{U}$	0.23	1.89	0.16	0.42
Access to Consumption, $\mathbb{C}$	1.03	1.41	0	0
Access to Employment, $\mathbb{W}$	-0.77	0.57	0.18	0.90
Gap in Welfare Impact	1.66		0.26	

Notes: This table displays robustness results corresponding to the counterfactual estimation in Section 6. In this specification, we assume perfectly inelastic housing supply. Column 1 displays results from the full model. Column 2 shows results abstracting away from consumption trips. Welfare, access to consumption and access to employment are defined by the model in Section 4. Gap in Welfare Impact is the difference in percentage changes in welfare across high- and low-income workers. Segregation is measured by the dissimilarity index.

Table A16: Impact of the North South Line and the Rail System in Singapore

	(1) Remove North South Line		**(2) Remove all train lines (Buses only)**	
% Change in	Low Type	High Type	Low Type	High Type
Welfare, $\bar{U}$	-4.23	-1.96	-14.32	-12.67
Access to Consumption, $\mathbb{C}$	-9.1	-0.59	-5.45	-4.23
Access to Employment, $\mathbb{W}$	-2.30	-0.98	-7.24	-6.56
Gap in Welfare Impact	2.27		1.65	

Notes: This table displays results from our auxiliary counterfactual analyses. Column 1 displays results from removing the North South Line. Column 2 shows the effect of removing all train lines. Welfare, access to consumption and access to employment are defined by the model in Section 4. Gap in Welfare Impact is the difference in percentage changes in welfare across high- and low-income workers.

Table A17: Impact of the Downtown Line: Additional Simulations

	(1) Only Consumption Shock		**(2) Local Consumption**		**(3) Roundabout Production**	
% Change in	Low Type	High Type	Low Type	High Type	Low Type	High Type
Welfare, $\bar{U}$	-0.32	0.93	0.38	1.66	0.39	2.13
Access to Consumption, $\mathbb{C}$	0.92	1.04	-0.17	0.33	1.01	1.39
Access to Employment, $\mathbb{W}$	-1.25	-0.18	0.60	1.40	-0.60	0.84
Gap in Welfare Impact	1.25		1.28		1.74	

Notes: This table displays results from the counterfactual estimation in Section 6. Column 1 displays results considering only the impact of DTL on reducing consumption travel time, while holding commuting time fixed. Column 2 shows results where households consume non-tradable goods and services in their neighborhood of residence. Column 3 shows results incorporating roundabout production as described in Section F.7. Welfare, access to consumption and access to employment are defined by the model in Section 4. Gap in Welfare Impact is the difference in percentage change in welfare across high-income and low-income workers.

Table A18: Cost Effectiveness of the Downtown Line

	Value	Source
Panel A: Costs		
Total cost of construction	\$2.07 bn.	Land Transport Authority
Yearly operating costs	\$132 mn.	Ministry of Transport
Yearly fare revenue	\$69 mn.	Ministry of Transport
Yearly non-fare revenue	\$13.2 mn.	Ministry of Transport
Net Present Value DTL Cost	\$22.36 bn.	
Panel B: Government Revenue		
Pre-DTL high-income earnings per year	\$63,512	General Household Survey 2015
Pre-DTL low-income earnings per year	\$18,187	General Household Survey 2015
High-income marginal tax rate	7%	Inland Revenue Authority of Singapore
Low-income marginal tax rate	0%	Inland Revenue Authority of Singapore
Change in high-income earnings	0.77%	Table 2
High-income worker population.	2.64 mn.	Ministry of Manpower
Low-income worker population.	0.88 mn.	Ministry of Manpower
Net Present Value Tax Revenue Change	\$2.26 bn.	
Panel C: Willingness to Pay		
Change in expected high-income utility	1.82	Table 2
Change in expected low-income utility	0.12	Table 2
WTP per year per high-income worker	\$1,155.92	Equation (11)
WTP per year per low-income worker	\$21.84	Equation (11)
Net Present Value WTP	\$100.9 bn.	
Panel D: Cost Effectiveness		
Marginal Value of Public Funds	5.02	Equation (68)
Social Benefits to Costs Ratio ($\phi = 0.5$)	3.11	Equation (69)
Social Benefits to Costs Ratio ($\phi = 0$)	4.61	Equation (69)

Notes: This table presents a cost effectiveness analysis of the Downtown Line following the calculations in Section G. Sources for each parameter are in Column 3. Figures are denominated in 2015 Singapore dollars. We assume a discount rate of 3% for all net present value calculations.