Online Appendix for "Consumption Access and Agglomeration: Evidence from Smartphone Data" (Not for Publication)

Yuhei Miyauchi[*]

Boston University

Kentaro Nakajima[†]

Hitotsubashi University

Stephen J. Redding[‡]

Princeton University, NBER and CEPR

February 2021

[*]Dept. Economics, 270 Bay State Road, Boston, MA 02215. Tel: 1-617-353-5682. Email: miyauchi@bu.edu.
[†]Institute of Innovation Research, 2-1 Naka, Kunitachi, Tokyo 186-8603, Japan. Tel: 81-42-580-8417. E-mail: nakajima.kentaro@gmail.com
[‡]Dept. Economics and SPIA, Julis Romo Rabinowitz Building, Princeton, NJ 08544. Tel: 1-609-258-4016. Email: reddings@princeton.edu.

A Data Appendix

This section of the appendix provides additional information about our smartphone data. In Section A.1 we provide further evidence on the representativeness of our smartphone data by comparing coverage by residence characteristics (income, age and distance to city center) and workplace characteristics (employment by industry and distance to city center). In Section A.2, we report further evidence validating our smartphone commuting data using census commuting data, supplementing the results reported in Section 2.3 of the paper.

In Section A.3, we present descriptive statistics on users with missing work locations, showing that they have more infrequent smartphone use, and demonstrating that the probability of assigning missing work locations is uncorrelated with the observable characteristics of users' municipality of residence. In Section A.4, we show that our findings from our smartphone data that non-commuting trips are more frequent than commuting trips are consistent with evidence from separate Japanese travel survey data that reports travel behavior during the working week.

In Section A.5, we provide further evidence on different types of non-commuting trips, supplementing the evidence for Fact 2 in Section 3 of the paper.

A.1 Coverage of Smartphone Data

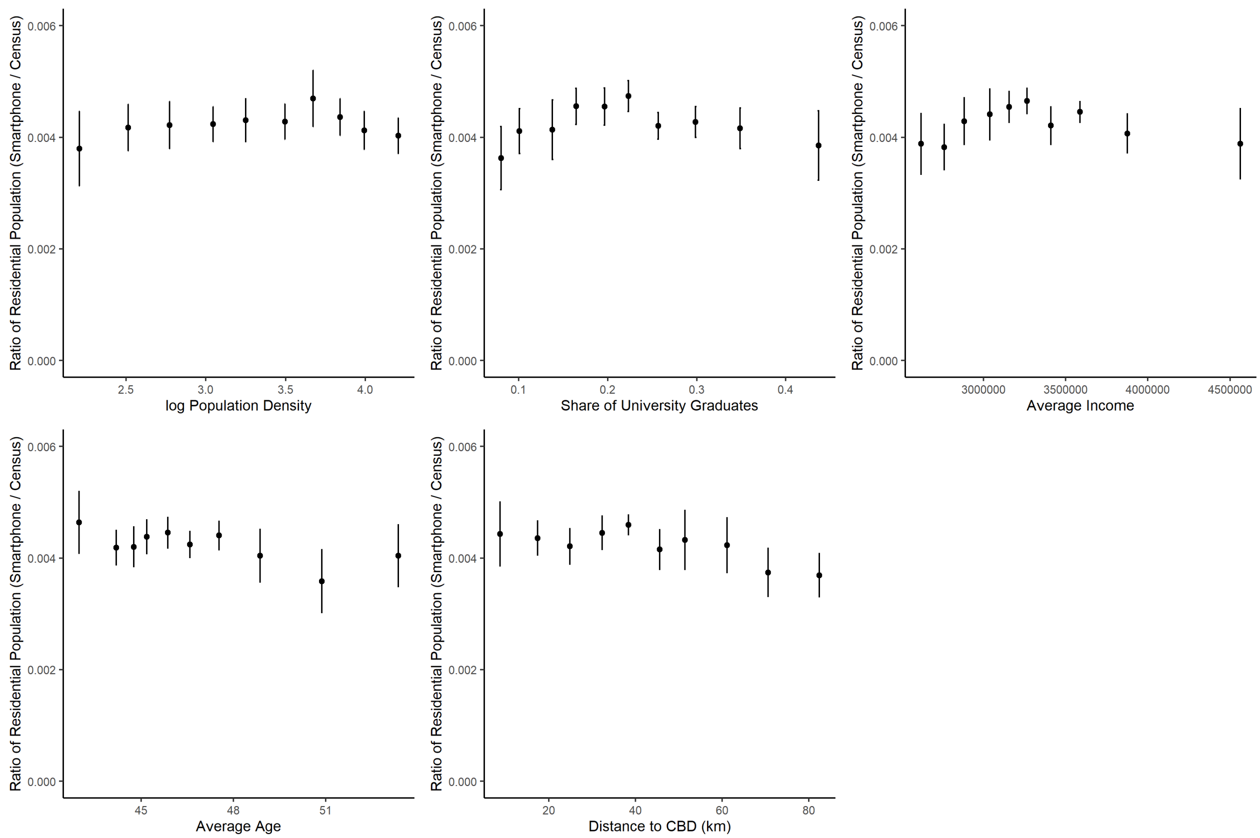
In this subsection of the appendix, we provide further evidence about the coverage of the samples of our smartphone GPS data. In Figure A.1.1, we plot the coverage rates of our smartphone data by the users' home municipality (defined by the number of users whose home location is in the municipality divided by the number of residents in the municipality in population census) against various characteristics of the municipalities, such as the population density, the share of university graduates, average income, average age and the distance to the central business district (Chiyoda Ward). The dots in each figure represents the average coverage rate for each decile of the characteristics described in the horizontal axis, and the line segments indicate the 95 percent confidence interval. We find no systematic patterns between the the coverage rates and municipality characteristics. The coverage rates have some inverse U-shaped pattern for the share of university graduates and the average income. However, these variations are not quantitatively large.

Similarly, in Figure A.1.2, we plot the coverage rates of our smartphone data by the users' work municipality (defined by the number of users whose work location is in the municipality divided by the number of employment in the municipality in population census) against various characteristics of the municipalities, such as the employment density, employment share of manufacturing, employment share of service, and the distance to the central business district (Chiyoda Ward). Similarly as the coverage rates by the residential locations, we find no systematic association between the coverage rates and these municipality characteristics.

A.2 Additional Validations for Commuting Flows from Smartphone Data

In Figure 4 of our main paper, we show that the commuting flows extracted from smartphone data has a similar spatial decay pattern as those in official census data. More specifically, we show that the gravity coefficients on bilateral distance between Tokyo municipalities controlling for residence and workplace fixed effects are close to each other when we use smartphone data and when we use official census data. In this section of the online appendix, we provide additional validation by comparing the residence and workplace fixed effects and the residual of the gravity

Figure A.1.1: Representativeness of Smartphone GPS Data by Residential Location



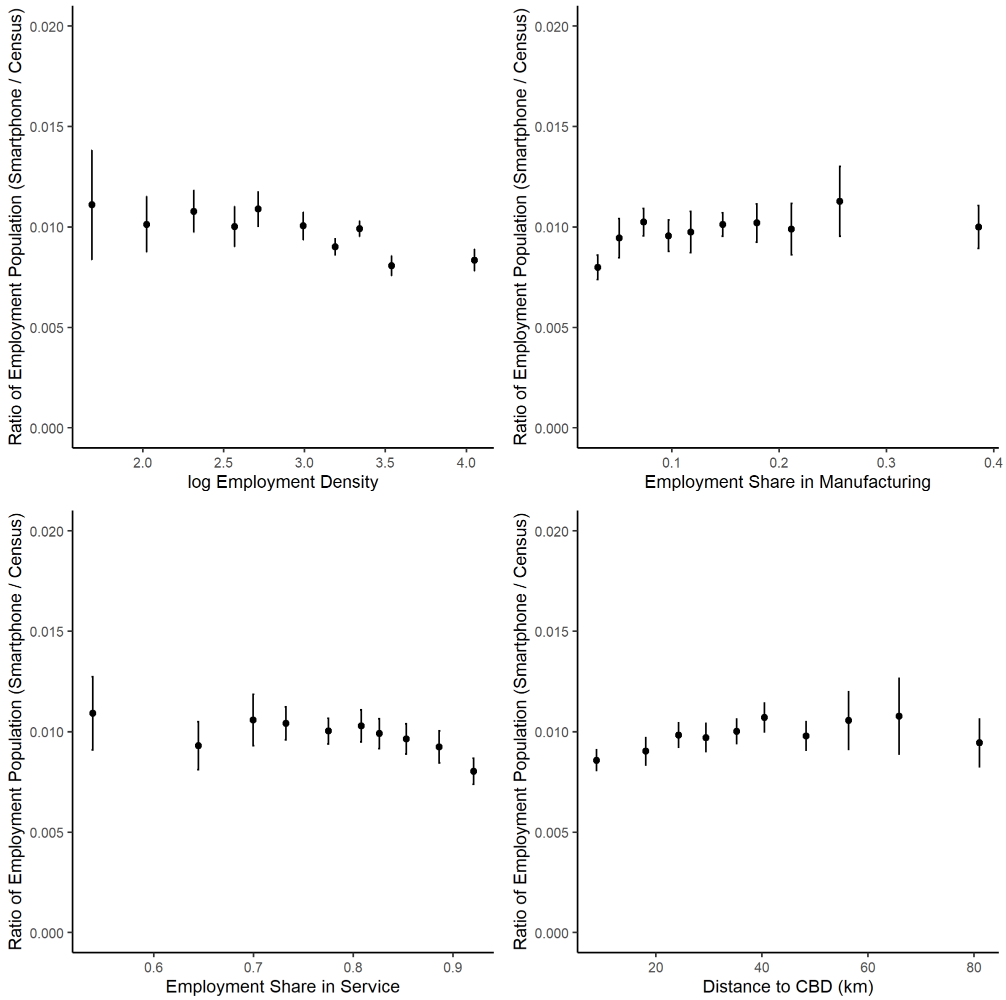
equations using these two separate data sets.

In Figure A.2.1, we compare the origin (residence) fixed effects (Panel A) and the destination (workplace) fixed effects (Panel B) from the gravity equations from the two data sources. In each panel, vertical axis corresponds to the estimates from the smartphone GPS data and horizontal axis corresponds to the estimates from the census data. Both fixed effects have approximately log-linear relationships with coefficients close to one and R-squared of above 0.9. Therefore, these results provide further supportive evidence that the attractiveness of residential locations taking out the effects of proximity to workplaces (origin fixed effects; Panel A), and the attractiveness of employment locations taking out the effects of proximity to residences (destination fixed effects; Panel B), are closely aligned between our smartphone data and official census data.

As another validation for our commuting flow data, in Figure A.2.2, we compare the residuals of the same gravity equations using the two data sets. The vertical axis represents the residuals of the gravity equations using smartphone data, and the horizontal axis takes the analogous objects using official census data. If these two residuals are closely aligned, it implies that the idiosyncratic shocks to commuting flows at the bilateral pair level that are not captured by the bilateral distances shows up in both data sets. We indeed find the regression coefficient of 0.976 that is close to one and R-squared of 0.482.

Together, these pieces of evidence support that our commuting flows constructed from smartphone data closely replicates the rate of spatial decay (gravity coefficients on distances), the attractiveness of employment and residences (residence and workplace fixed effects), and the residuals from gravity equations.

Figure A.1.2: Representativeness of Smartphone GPS Data by Employment Location



A.3 Patterns of Users without Workplace Assignment

As discussed in Section 2.1 of our main paper, "home" location and "work" locations are defined as the centroid of the first and second most frequent locations of geographically contiguous stays, respectively. To ensure that these two locations do not correspond to different parts of a single property, we also require that the "work" location is more than 600 meters away from "home" location. In particular, if the second most frequent location is within 600 meters of the "home" locations, we define the "work" location as the third most frequent location. To abstract from noise in geo-coordinate assignment, all stays within 500 meters of the home location are aggregated with the home location. Similarly, all stays within 500 meters of the work location are aggregated with the work location. We assign "work" location as missing if the user appears in that location for less than 5 days per month, which applies for about 30 percent of users in our baseline sample during April 2019. In this section of the online appendix, we characterize the pattern of users whose workplace is not assigned in this procedure.

We first provide suggestive evidence that a large fraction of cases with missing "work" locations is likely to be due to the inactive usage of smartphone devices themselves. Figure A.3.1 shows the distribution of the number of days that we observe any stays in April 2019, including stays at home. As discussed in Section 2.1, location information is

Figure A.2.1: Correlation of fixed effects in commuting gravity estimation

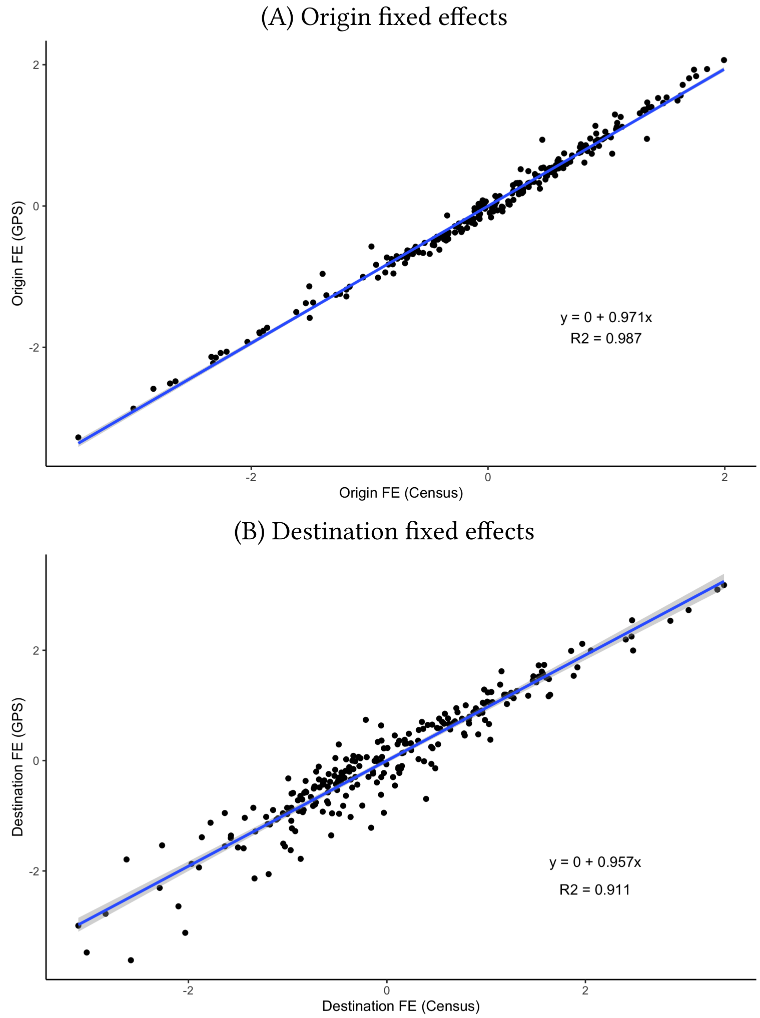
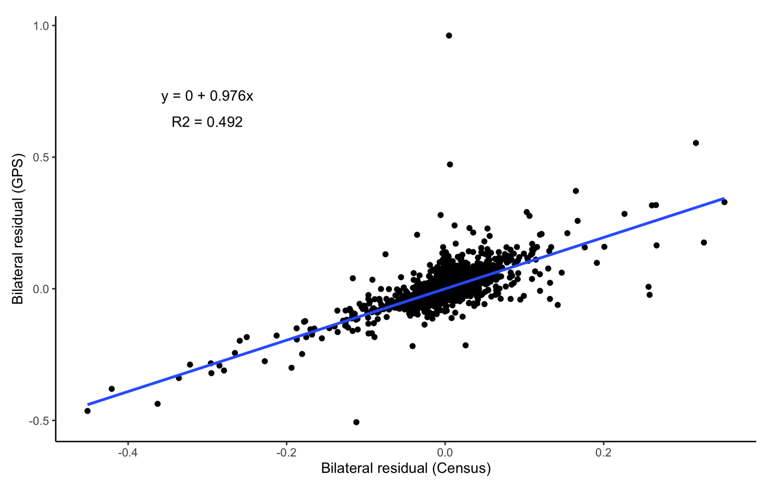
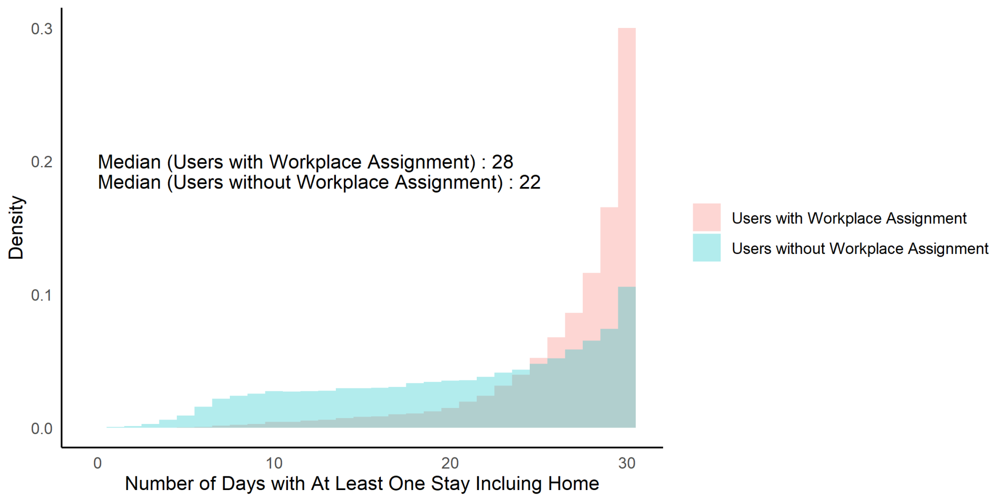


Figure A.2.2: Correlation of bilateral residuals in commuting gravity estimation



collected regardless of what application the user has open, as long as the device is turned on (upon users' consent). Therefore, the number of days with any stays is a proxy for how active the device is used (i.e., whether it is regularly turned on, or whether the user bring the device with them when they move out). We find that the median active days is 22 for users without workplace assignment, which is significantly smaller than 28 days for users with workplace assignment. Therefore, many devices we have “missing” workplaces are not actively used in April 2019.

Figure A.3.1: Number of Stays by Users with and without Workplace Assignment



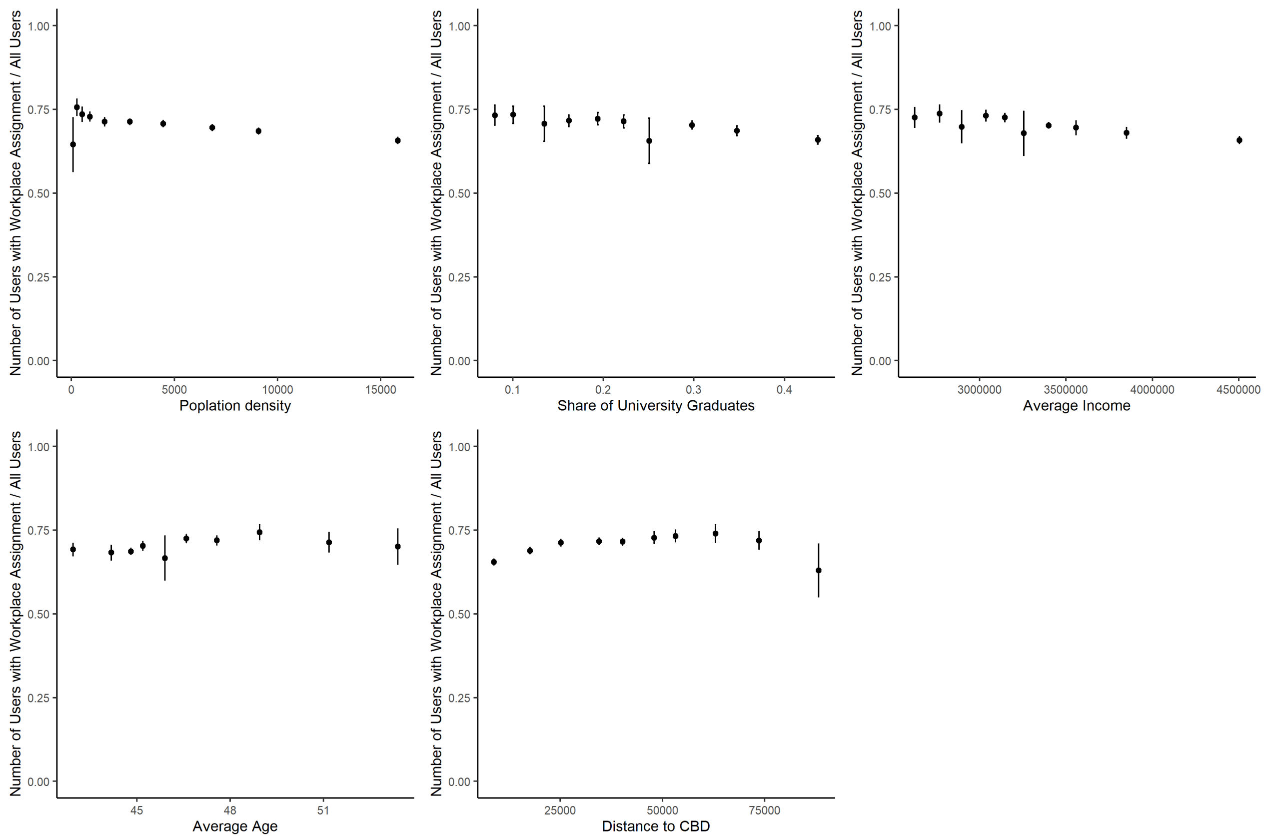
If these devices are not actively used, one may worry about the potential bias due to this attrition. To check this point, in Figure A.3.2, we display how the probability of missing “work” location is correlated with the residential characteristics. In each of the panel, we plot the fraction of users whose workplace is assigned (on the vertical axis) against the characteristics of their residential municipality, such as the population density, the share of university graduates, average income, average age and the distance to the central business district (Chiyoda Ward). The dots in each figure represents the average fraction of non-missing “work” for each decile of the characteristics described in the horizontal axis at the users’ “home” location, and the line segments indicate the 95 percent confidence interval. We find no strong associations between non-missing “work” with these municipality characteristics. There is a mild decreasing pattern in the share of university graduates and the average income, but the magnitudes are not quantitatively large. Therefore, along the dimensions of observable residential characteristics, we find limited evidence of the bias in the probability of missing “work” locations.

A.4 Comparison with Travel Survey

In this section of the online appendix, we validate our smartphone GPS data using a separate travel survey (person trip survey). We show that similar patterns of urban trips documented in our paper using our smartphone data also hold with the travel survey data.

In greater Tokyo Metropolitan Area, travel surveys (person trip surveys) are conducted at a decennial frequency. For this validation, we use the information from person trip survey data conducted in 2008. The respondents of this survey are members of households residing in Tokyo, Saitama, Chiba, and Kanagawa and a part of Ibaraki prefectures, which is mostly consistent with our definition of Greater Tokyo Metropolitan Area. The respondents are asked about their travel behavior on a specific weekday. The survey asks about a sequence of of travel she made during the day (a sequence of segments of movements from one place to another). For each trip segment, the survey asks where and

Figure A.3.2: Pattern of the Ratio of Users with Workplace Assignment



when the trip starts and ends, the purpose of the trip, and what transportation mode is used. This survey also asks basic demographic information, work status, and home address. We use all non-student samples for the following validation exercises.

Figure A.4.1 compares the number of stays per day per person in our smartphone data and in this person trip survey. The left panel depicts the number of stays at workplaces and the number of stays at other (non-work) stays separately for the two data sets. The right panel depicts the same information. Given that the travel survey only asks about the pattern on a specific weekday, the number of stays from the travel survey is missing in the right panel, showing frequency of stays on weekends.

Focusing on the pattern on weekdays (left panel), we find that the number of stays is greater for smartphone data than for the person trip survey. One possible reason for this difference is under-reporting of trips in the person trip survey. Recall that we define "stays" if a user is static for more than 15 minutes. In the person trip survey, if the actual stay length is as short as 15 minutes, one may not report the stay in the survey. Despite this difference, the relative number of work stays and other (non-work) stays is similar between the two types of data sets. In smartphone data, 58 percent ($= 1.6/(1.6+1.14)$) of all stays are related to other purposes. In person trip data, this number is 56 percent ($= 0.95/(0.95+0.73)$), which is approximately the same as the same number from smartphone data.

We next compare the types of non-commuting stays between our smartphone data and the person trip survey. In our main paper, we assign sectors of "Other" stays from our smartphone data using separate economic census data, and presented the decomposition of these non-commuting stays into different sector (Figure 7). As a comparison with

Figure A.4.1: Frequency of Stays at Work and Other Locations

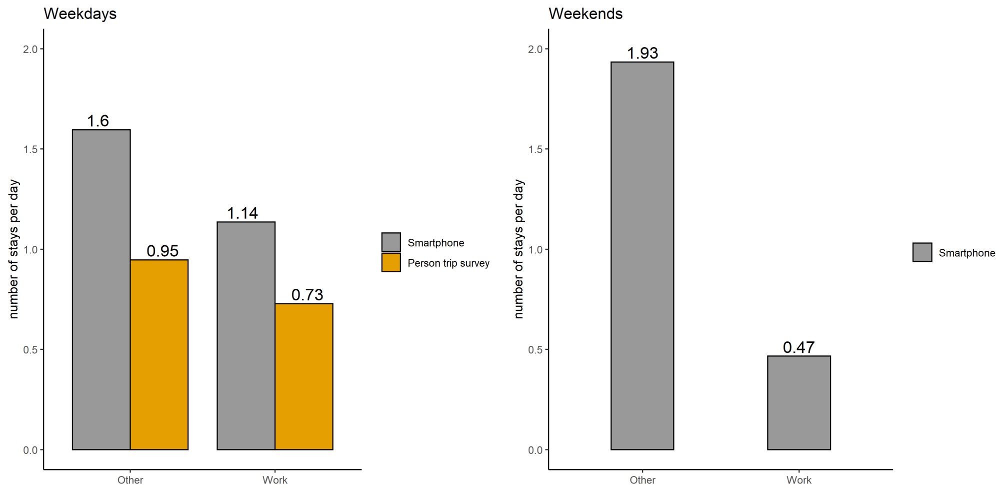
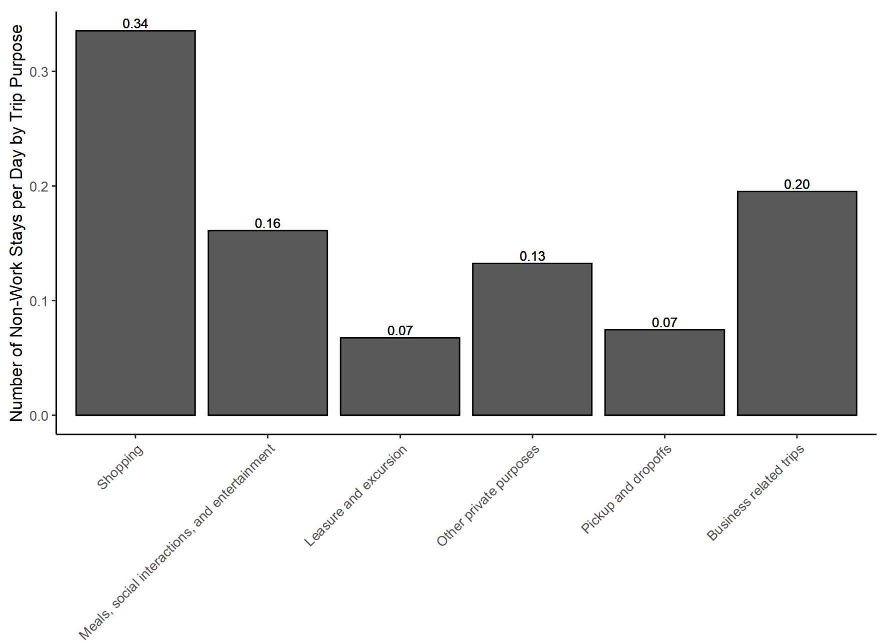


Figure 7, Figure A.4.2 displays the average number of stays per day for other locations in Person Trip Survey data by the stated purpose of the trip. While the precise comparison with our assignment of trip sectors from smartphone data is difficult due to different classifications, we find a similar overall pattern. In particular, we find that "Shopping" is the most frequent category (34 percent), consistent with out finding from our smartphone data that "Retail and Wholesales" sectors are the most frequent category of non-commuting stays with the share of 43 percent on weekdays (Figure 7). We also find a substantially smaller fraction of business-related trips (e.g., business meetings, procurement; 20 percent) compared to trips related to consumption.

Figure A.4.2: Frequency of Stays at Other Locations by purpose using trip survey data



As a final comparison between our smartphone data and the travel survey, Table A.4.1 displays the average distances to work and other stays in kilometers; from home location for smartphone GPS data (Column 1) and the person trip survey data (Column 2). In Column 1, we find that the destinations of other stays are more concentrated around home location than commuting trips from smartphone data (as documented by Fact 3 of Section 3 of our main paper). Consistent with this finding, we find a similar pattern using our travel survey data in Column 2. Moreover, the average distance to work locations form home locations are closely aligned between smartphone GPS data and the survey data (12.68 and 12.78 kilometers, respectively). For other (non-work) stays, the average distances are slightly more distant from home locations in our smartphone data (10.95 kilometer) than our travel survey data (7.39 kilometer). These differences potentially come from a noisier measure of distances in our travel survey data due to a coarser geographic aggregation, or due to underreporting of stays in the survey in far distances from home in travel survey data (for example, people may not report small errands along the commuting paths, or the lunch or coffee during the office hours).

Table A.4.1: Average Distances of Work Stays from Home Locations

	Smartphone	Person trip survey
Work	12.68	12.78
Other	10.95	7.39

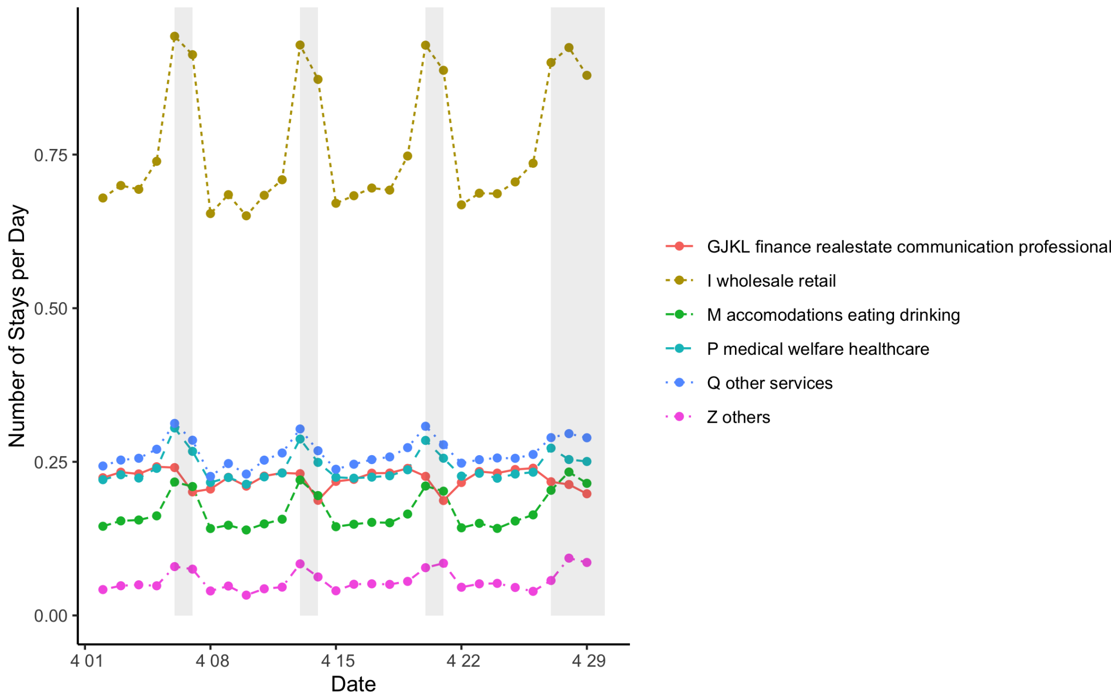
Note: Average distances of work stays and other (non-work) stays from home location (in kilometers); from smartphone data (first column) and from travel survey data (second column).

A.5 Additional Validation of Sector Assignment of Non-Commuting Stays

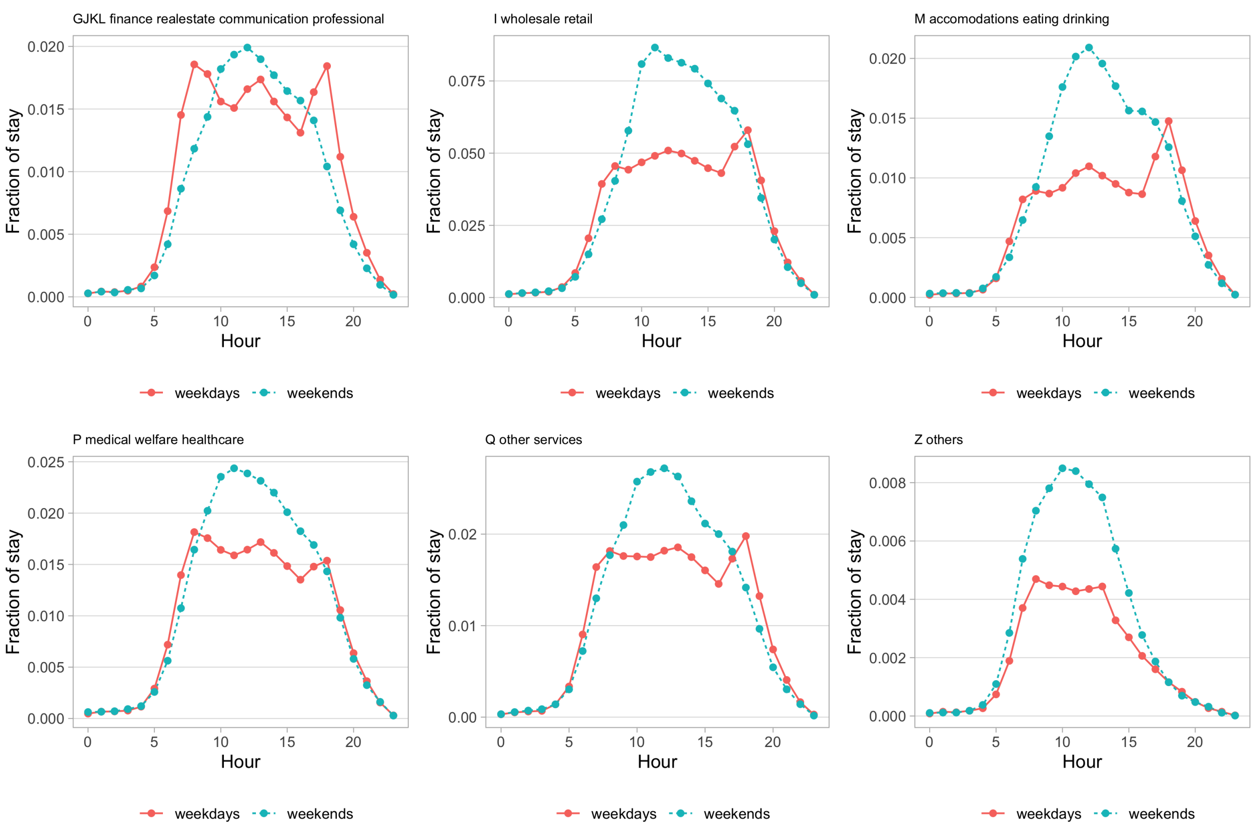
In this subsection of the appendix, we provide additional validation of our assignment of "other" stays (stays at neither home nor work locations) to sectors using separate economic census as defined in Figure 7. Figure A.5.1 displays the density of each type of other stay by starting hour and day, as a share of all stays for our baseline sample for the Tokyo metropolitan area in April 2019. We find that our probabilistic assignment captures the expected pattern of these different service-sector activities over the course of the week. First, we typically find a higher density of other stays during the middle of the day at weekends than during weekdays, which is in line with the fact that many of these services are consumed more intensively during leisure time. The one exception is "Finance, Real Estate, Communication, and Professional," which displays the opposite patter, consistent with the fact that establishments providing these services are often closed at the weekends in Japan. Second, we find that the peak densities of stays for "Wholesale and Retail" and "Accommodations, Eating, Drinking" occur at around 6pm on weekdays, corroborating the fact that these activities are typically concentrated after work during the week. For "Accommodations, Eating, Drinking," we find a smaller peak around noon on weekdays, as expected from the typical timing of lunch in Japan. Third, and finally, both of these activities are more concentrated in the middle of the day on weekends than during the week, which again is in line with workers having greater leisure time in the middle of day at weekends.

Figure A.5.1: Other Stays by Type, Day, and Hour

(A) Number of other stays by day



(B) Number of other stays by starting hour



B Theoretical Framework Appendix

In this section of the online appendix, we report the detailed derivations of the results in Section 4 of the paper.

We consider a city (Tokyo) that is embedded in a larger economy (Japan). We consider both a closed-city specification (in which total city population is exogenous) and an open-city specification (in which total city population is endogenously determined by population mobility with the wider economy that offers a reservation level of utility $\bar{U}$). The city consists of a discrete set of locations $i, j, n \in N$ that differ in productivity, amenities, supply of floor space and transport connections. Utility is defined over consumption of a single traded good, a number of different types of non-traded services (e.g. restaurants, coffee shops, stores), and residential floor space use. Both the traded good and the non-traded services are produced with labor and commercial floor space according to constant returns to scale under conditions of perfect competition. Floor space is supplied by a competitive construction sector using land and capital according to a constant returns to scale construction technology.

A continuous measure of workers ($\bar{L}$) choose a residence, a workplace and a set of locations to consume non-traded services in the city.[1] We assume the following timing or nesting structure for workers' location decisions. First, each worker observes her idiosyncratic preferences or amenities (b) for each location within the city, and chooses her residence n. Second, given a choice of residence, each worker observes her idiosyncratic productivities (a) for each workplace i and sector g, and chooses her sector and location of employment. Third, given a choice of residence and workplace, she observes idiosyncratic qualities (q) for each type of non-traded service k available in each location j, and chooses her consumption location for each type of non-traded service. Fourth, given a choice of residence, workplace, and the set of consumption locations, she observes idiosyncratic shocks (ν) of whether to visit the consumption location from their residence, returning to their residence after the visiting the consumption location, from their workplace, again a round trip, or on the way to work from home (which we call a "route"), and chooses the route for each consumption location r. We choose this nesting structure because it permits a transparent decomposition of residents and land prices into the contribution of travel accessibility and the residual of amenities. We also compare the predictions of our model with the special case abstracting from consumption trips, which corresponds to a conventional urban model, in which workers choose workplace and residence and consume only traded goods. In the open-city specification, population mobility ensures that the expected utility from living in the city equals the reservation utility in the wider economy.

B.1 Preferences

The indirect utility for worker ω who chooses residence n, works in location i and sector g, and consumes non-traded service k in location $j(k)$ with the route $r(k)$ (the choice of whether to visit consumption locations from home, from work, or in-between) is assumed to take the following Cobb-Douglas form:

$$U_{nig\{j(k)r(k)\}}(\omega) = \left\{ B_n b_n(\omega) \left(P_n^T\right)^{-\alpha^T} Q_n^{-\alpha^H} \right\} \left\{ a_{i,g}(\omega)\, w_{i,g} \right\}, \qquad \text{(B.1)}$$
$$\times \left\{ \prod_{k \in K^S} \left[P_{j(k)}^S / \left(q_{j(k)}(\omega) \right) \right]^{-\alpha_k^S} \right\} \left\{ d_{ni\{j(k)r(k)\}} \prod_{k \in K^S} \nu_{r(k)}(\omega) \right\}$$

[1] In our theoretical analysis, we assume for simplicity a continuous measure of workers in the model, which ensures that the expected values of variables equal their realized values. In our empirical analysis, we allow for granularity and a finite number of workers in both our estimation (using the PPML estimator) and our counterfactuals (using predicted shares) following Dingel and Tintelnot (2020).

$$0<\alpha^T,\alpha^H,\alpha_k^S<1, \qquad \alpha^T+\alpha^H+\sum_{k\in K^S}\alpha_k^S=1,$$

where we use the notation $j(k)$ to indicate that that non-traded service k is consumed in a single location j that is an implicit function of the type of non-traded service k; $r(k)\in\mathbb{R}\equiv\{HH,WW,HW,WH\}$ indicates the "route" choice of whether to visit consumption locations from home (HH), from work (WW), on the way from home to work (HW), or on the way from work to home (WH) for each non-traded service k; $K^S\subset K$ is the subset of sectors that are non-traded; the first term in brackets captures a residence component of utility; the second term in brackets corresponds to a workplace component of utility; the third term in brackets reflects a non-traded services component of utility; the fourth term in brackets reflects a travel cost component of utility.

The first, residence component, includes amenities (B_n) that are common for all workers in residence n; the idiosyncratic amenity draw for residence n for worker ω ($b_n(\omega)$); the price of the traded good (P_n^T); and the price of residential floor space (Q_n). We allow the common amenities (B_n) to be either exogenous or endogenous to the surrounding concentration of economic activity in the presence of agglomeration forces, as discussed further below. The second, workplace component, comprises the wage per efficiency unit in sector g in workplace i ($w_{i,g}$); the idiosyncratic draw for productivity or efficiency units of labor for worker ω in sector g in workplace i ($a_{i,g}(\omega)$).[2] The third, non-traded services component, depends on the price of the non-traded service k in the location $j(k)$ where it is supplied ($P_{j(k)}^S$ for $k\in K^S$); the idiosyncratic draw for quality for that service in that location ($q_{j(k)}(\omega)$ for $k\in K^S$). The fourth, travel cost component, includes iceberg form of travel cost for each combination of residence, workplace, the consumption locations and their routes ($d_{ni\{j(k),r(k)\}}$); and the idiosyncratic draw for route preference for each non-traded sector ($\nu_{r(k)}(\omega)$ for $k\in K^S$).

To make a precise mapping between the travel cost ($d_{ni\{j(k),r(k)\}}$) and the travel time between any bilateral locations, we parametrize $d_{ni\{j(k)r(k)\}}$ as follows:

$$d_{ni\{j(k)r(k)\}}=\exp(-\kappa^W\tau_{ni}^W)\prod_{k\in K^S}\exp(-\kappa^S\tau_{nij(k)r(k)}^S) \tag{B.2}$$

In this expression, the first term before the product sign captures the cost of commuting from residence n to workplace i without any detour to consume non-traded services, which depends on travel time (τ_{ni}^W) and the commuting cost parameter (κ^W), where overall commuting travel time is the sum of the travel time incurred in each direction:

$$\tau_{ni}^W=\tau_{ni}+\tau_{in}. \tag{B.3}$$

The second term in equation (2) captures the additional travel costs involved in consuming each type of non-traded service k in location $j(k)$ by the route $r(k)$, which depends on the additional travel time involved ($\tau_{nij(k)r(k)}^S$) and the consumption travel cost parameter (κ_k^S). This additional travel time depends on the route taken: whether the worker visits consumption location $j(k)$ from home ($r(k)=HH$), from work (WW), on the way from home to work (HW),

[2]Although we model the workplace idiosyncratic draw as a productivity draw, there is a closely-related formulation in which it is instead modelled as an amenity draw.

or on the way from work to home (WH):

$$\begin{aligned}
\tau^S_{nij(k)HH} &= \tau_{nj} + \tau_{jn}, \\
\tau^S_{nij(k)WW} &= \tau_{ij} + \tau_{ji}, \\
\tau^S_{nij(k)HW} &= \tau_{nj} + \tau_{ji} - \tau_{ni}, \\
\tau^S_{nij(k)WH} &= \tau_{ij} + \tau_{jn} - \tau_{in},
\end{aligned} \tag{B.4}$$

where the negative term at the end of the third and fourth lines above reflects the fact that the worker travels indirectly between residence n and workplace i via consumption location j on one leg of her journey between home and work, and hence does not incur the direct travel time between residence n and workplace i for that leg of the journey.

We make the conventional assumption in the location choice literature following McFadden (1974) that the idiosyncratic shocks are drawn from an extreme value distribution. In particular, idiosyncratic amenities (b), productivity (a), quality (q), route preferences (ν) for worker ω, residence n, workplace i and consumption location $j(k)$ for non-traded service k are drawn from the following independent Fréchet distributions:

$$\begin{aligned}
G^B_n(b) &= \exp\left(-T^B_n b^{-\theta^B}\right), && T^B_n > 0,\ \theta^B > 1, \\
G^W_{i,g}(a) &= \exp\left(-T^W_{i,g} a^{-\theta^W}\right), && T^W_{i,g} > 0,\ \theta^W > 1, \\
G^S_{j(k)}(q) &= \exp\left(-T^S_{j(k)} q^{-\theta^S_k}\right), && T^S_{j(k)} > 0,\ \theta^S_k > 1, k \in K^S, \\
G^R_{r(k)}(\nu) &= \exp\left(-T^R_{r(k)} \nu^{-\theta^R_k}\right), && T^R_{r(k)} > 0,\ \theta^R_k > 1, k \in K^S
\end{aligned} \tag{B.5}$$

where the scale parameters $\{T^B_n, T^W_{i,g}, T^S_{j(k)}, T^R_{r(k)}\}$ control the average draws and the shape parameters $\{\theta^B, \theta^W, \theta^S_k, \theta^R_k\}$ regulate the dispersion of amenities, productivity and quality respectively. The smaller these dispersion parameters, the greater the heterogeneity in idiosyncratic draws, and the less responsive worker decisions to economic variables.[3]

Using our assumption about the timing or nesting structure, the worker location choice problem is recursive and can be solved backwards. First, for given a choice of residence, workplace and sector, and the set of consumption locations, we characterize the probability that a worker chooses each route for each non-traded sector (whether to visit consumption locations from home, from work, or in-between). Second, for given a choice of residence, workplace and sector, we characterize the probability that a worker chooses each consumption location in each non-traded sector, taking into account the expected travel cost for consumption trips. Third, for given a choice of residence, we characterize the probability that a worker chooses each workplace and sector, taking into account the expected consumption access of that workplace and sector. Fourth, we characterize the probability that a worker chooses each residence, taking into account its expected travel accessibility for both commuting and consumption.

B.2 Route Choices

We begin with the worker's decision of the route choice for each non-traded service sector k. More specifically, conditional on residence, workplace, and consumption location, the worker chooses whether to visit consumption location $j(k)$ from home ($r(k) = HH$), from work (WW), on the way from home to work (HW), or on the way from

[3] Although we assume independent Fréchet distributions for amenities, productivity and quality, some locations can have high expected values for all three shocks if they have high values for T^B_n, T^W_{ig} and $T^S_{j(k)}$. Additionally, correlations between the shocks can be introduced using a multivariate Fréchet distribution, as in Hsieh, Hurst, Jones, and Klenow (2019).

work to home (WH). Given the indirect utility (B.1) and the specification of the travel cost (B.2), the indirect utility is rewritten as:

$$
\begin{aligned}
U_{nig\{j(k)r(k)\}}(\omega) = & \left\{B_n b_n(\omega)\left(P_n^T\right)^{-\alpha^T} Q_n^{-\alpha^H}\right\}\left\{a_{i,g}(\omega)\, w_{i,g}\right\}, \\
& \times\left\{\prod_{k\in K^S}\left[P_{j(k)}^S/\left(q_{j(k)}(\omega)\right)\right]^{-\alpha_k^S}\right\}\left\{\exp(-\kappa^W\tau_{ni}^W)\prod_{k\in K^S}\exp(-\kappa^S\tau_{nij(k)r(k)}^S)\nu_{r(k)}(\omega)\right\}
\end{aligned}
$$

The component of the utility that depends on the route $r(k)$ for non-traded service k is given by:

$$
\delta_{nij(k)r(k)}(\omega) = \exp(-\kappa^S\tau_{nij(k)r(k)}^S)\nu_{r(k)}(\omega). \tag{B.6}
$$

where the first component is the route-specific travel cost and the second component is the idiosyncratic route preference. Under our assumption of independent route-preference draws $\nu_{r(k)}(\omega)$ across each non-traded sector k, each worker chooses the route $r(k)$ that maximizes $\delta_{nij(k)r(k)}(\omega)$ independently for each sector k.

Using our independent extreme value assumption for idiosyncratic preference shocks for the route choice, the route choice probability is characterized by a logit form. In particular, the probability that a worker living in residence n and employed in workplace i consuming non-traded service k in location $j(k)$ chooses the route $r(k)$ ($\lambda_{r(k)|nij(k)}^R$) is derived as follows:

$$
\begin{aligned}
\lambda_{r(k)|nij(k)}^R &= \Pr\left[\delta_{nij(k)r(k)}(\omega) > \max\left\{\delta_{nij(k)\ell(k)}(\omega) : \ell \neq r\right\}\right], \\
&= \int_0^\infty \prod_{\ell\neq r} G_{\ell(k)}^R(\delta)\, g_{nr(k)}^R(\delta)\, d\delta, \\
&= \int_0^\infty \prod_{\ell\neq r} \exp\left(-\Phi_{nij(k)\ell(k)}^R \delta^{-\theta_k^R}\right)\theta_k^R \Phi_{nij(k)r(k)}^R \delta^{-\left(\theta_k^R+1\right)}\exp\left(-\Phi_{nij(k)r(k)}^R\delta^{-\theta_k^R}\right) d\delta, \\
&= \int_0^\infty \exp\left(-\Phi_{nij(k)}^R\delta^{-\theta_k^R}\right)\theta_k^R\Phi_{nij(k)r(k)}^R\delta^{-\left(\theta_k^R+1\right)} d\delta,
\end{aligned}
$$

where

$$
\Phi_{nij(k)}^R \equiv \sum_{r'(k)\in\mathbb{R}} \Phi_{nij(k)r'(k)}^R = \sum_{r'(k)\in\mathbb{R}} T_{r'(k)}^R \exp\left(-\theta_k^R\kappa^S\tau_{nij(k)r'(k)}^S\right).
$$

Note that

$$
\frac{d}{d\gamma}\left[\frac{1}{\Phi_{nij(k)}^R}\exp\left(-\Phi_{nij(k)}^R\delta^{-\theta_k^R}\right)\right] = \exp\left(-\Phi_{nij(k)}^R\delta^{-\theta_k^R}\right)\theta_k^R\delta^{-\left(\theta_k^R+1\right)}.
$$

Using this result to evaluate the integral above, we have:

$$
\lambda_{r(k)|nij(k)}^R = \Phi_{nij(k)}^R\left[\frac{1}{\Phi_{nij(k)}^R}\exp\left(-\Phi_{nij(k)}^R\delta^{-\theta_k^R}\right)\right]_0^\infty,
$$

which becomes:

$$
\lambda_{r(k)|nij(k)}^R = \frac{T_{r(k)}^R\exp(-\theta_k^R\kappa_k^S\tau_{nij(k)r(k)}^S)}{\sum_{r'\in\mathbb{R}} T_{r'(k)}^R\exp(-\theta_k^R\kappa_k^S\tau_{nij(k)r'(k)}^S)}. \tag{B.7}
$$

Using our independent extreme value assumption for idiosyncratic preference shocks for route choice, we can also compute the expected contribution to utility from the travel cost from consumption trips as follows. Using the property that the maximum of a sequence of Fréchet distributions is itself Fréchet distributed, the contribution to utility from the preferred route choice also has a Fréchet distribution:

$$
G_{nij(k)}^R(\delta) = \prod_{r'(k)\in\mathbb{R}} G_{nij(k)r'(k)}^R(\delta) = \prod_{r'\in\mathbb{R}}\exp\left(-\Phi_{nij(k)r'(k)}^R\delta^{-\theta_k^R}\right),
$$

$$G^R_{nij(k)}(\delta) = \exp\left(-\Phi^R_{nij(k)}\delta^{-\theta^R_k}\right), \qquad \Phi^R_{nij(k)} \equiv \sum_{r'(k)\in\mathbb{R}} \Phi^R_{nij(k)r'(k)}.$$

Given this Fréchet distribution for the contribution to utility from the preferred location, the expected contribution to utility from from the preferred route choice is:

$$\begin{aligned}
\mathbb{E}_{nij(k)}[\delta] &= \int_0^\infty \delta g^R_{nij(k)}(\delta)\, d\delta. \\
&= \int_0^\infty \theta^R_k \Phi^R_{nij(k)} \delta^{-\theta^R_k} \exp\left(-\Phi^R_{nij(k)}\delta^{-\theta^R_k}\right) d\delta
\end{aligned}$$

Now define the following change of variables:

$$y = \Phi^R_{nij(k)}\gamma^{-\theta^R_k}, \qquad dy = \left(\theta^R_k\right)\Phi^R_{nij(k)}\delta^{-\left(\theta^R_k+1\right)}d\delta.$$

$$\delta = \left(\frac{y}{\Phi^R_{nij(k)}}\right)^{-\frac{1}{\theta^R_k}}, \qquad d\delta = \frac{dy}{\left(\theta^R_k\right)\Phi^R_{nij(k)}\delta^{-\left(\theta^R_k+1\right)}}.$$

Using this change of variables, we can write the expected contribution to utility as:

$$\begin{aligned}
\mathbb{E}_{nij(k)}[\delta] &= \int_0^\infty \left(\theta^R_k\right)\Phi^R_{nij(k)} \frac{y}{\Phi^R_{nij(k)}} \exp(-y) \frac{dy}{\left(\theta^R_k\right)\Phi^R_{nij(k)}\left(\frac{y}{\Phi^R_{nij(k)}}\right)^{\frac{\theta^R_k+1}{\theta^R_k}}}, \\
&= \int_0^\infty y\exp(-y)\frac{dy}{\Phi^R_{nij(k)}\left(\frac{y}{\Phi^R_{nij(k)}}\right)^{\frac{\theta^R_k+1}{\theta^R_k}}}, \\
&= \int_0^\infty y^{-\frac{1}{\theta^R_k}}\exp(-y)\frac{dy}{\Phi^R_{nij(k)}\left(\Phi^R_{nij(k)}\right)^{-\frac{\theta^R_k+1}{\theta^R_k}}}, \\
&= \left(\Phi^R_{nij(k)}\right)^{\frac{1}{\theta^R_k}} \int_0^\infty y^{-\frac{1}{\theta^R_k}}\exp(-y)\,dy. \\
&= \vartheta^R_k\left(\Phi^R_{nij(k)}\right)^{\frac{1}{\theta^R_k}},
\end{aligned}$$

where

$$\vartheta^R_k \equiv \Gamma\left(\frac{\theta^R_k-1}{\theta^R_k}\right) = \int_0^\infty y^{-\frac{1}{\theta^R_k}}\exp(-y)\,dy,$$

and $\Gamma(\cdot)$ is the Gamma function. To summarize,

$$d^S_{nij(k)} = \mathbb{E}_{nij(k)}\left[\delta_{nij(k)r(k)}(\omega)\right] = \vartheta^R_k\left[\sum_{r'\in\mathbb{R}} T^R_{r'(k)}\exp(-\theta^R_k\kappa^S_k\tau^S_{nij(k)r'(k)})\right]^{\frac{1}{\theta^R_k}} \tag{B.8}$$

which corresponds to equation (8) in the paper.

B.3 Consumption Choices

We begin with the worker's decision of where to consume each type of non-traded service. Conditional on living in residence n, each worker chooses consumption location $j(k)$ for non-traded service k to to maximize the contribution to indirect utility (B.1) from consuming that non-traded service:

$$\gamma_{nij(k)}(\omega) = \left[P^S_{j(k)}/\left(q_{j(k)}(\omega)\right)\right]^{-\alpha^S_k} d^S_{nij(k)}, \qquad k\in K^S. \tag{B.9}$$

We thus have the following monotonic relationship between the contribution to utility from consuming nontradable services and quality:

$$q_{j(k)}(\omega) = \left(\gamma_{nij(k)}(\omega) / d^S_{nij(k)}\right)^{1/\alpha^S_k} \left(P^S_{j(k)}\right)$$

Therefore, using this relationship and the Fréchet distribution for idiosyncratic quality, we have:

$$\Pr\left[\gamma_{nij(k)} < \gamma\right] = G^S_{nij(k)}\left(\gamma^{1/\alpha^S_k} P^S_{j(k)} \left(d^S_{nij(k)}\right)^{-1/\alpha^S_k}\right),$$

$$G^S_{nij(k)}(\gamma) = e^{-\Phi^S_{nij(k)}\gamma^{-\theta^S_k/\alpha^S_k}}, \qquad \Phi^S_{nij(k)} \equiv T^S_{j(k)} \left(P^S_{j(k)}\right)^{-\theta^S_k} \left(d^S_{nij(k)}\right)^{\theta^S_k/\alpha^S_k}.$$

Using this distribution for the contribution of nontradable services to utility, the probability that a worker in residence n consumes nontradable service k in location $j(k)$ is:

$$\begin{aligned}
\lambda^S_{j(k)|ni} &= \Pr\left[\gamma_{nj(k),k} > \max\left\{\gamma_{n\ell(k),k} : \ell \neq j\right\}\right], \\
&= \int_0^\infty \prod_{\ell \neq j} G^S_{n\ell(k)}(\gamma)\, g^S_{nij(k)}(\gamma)\, d\gamma, \\
&= \int_0^\infty \prod_{\ell \neq j} \exp\left(-\Phi^S_{ni\ell(k)}\gamma^{-\theta^S_k/\alpha^S_k}\right)\left(\theta^S_k/\alpha^S_k\right)\Phi^S_{nij(k)}\gamma^{-\left(\left(\theta^S_k/\alpha^S_k\right)+1\right)} \exp\left(-\Phi^S_{nij(k)}\gamma^{-\theta^S_k/\alpha^S_k}\right) d\gamma, \\
&= \int_0^\infty \exp\left(-\Phi^S_{ni,k}\gamma^{-\theta^S_k/\alpha^S_k}\right)\left(\theta^S_k/\alpha^S_k\right)\Phi^S_{nij(k)}\gamma^{-\left(\left(\theta^S_k/\alpha^S_k\right)+1\right)} d\gamma,
\end{aligned}$$

where

$$\Phi^S_{ni,k} \equiv \sum_{\ell \in N} \Phi^S_{ni\ell(k)} = \sum_{\ell \in N} T^S_{\ell(k)} \left(P^S_{\ell(k)}\right)^{-\theta^S_k} \left(d^S_{nij(k)}\right)^{\theta^S_k/\alpha^S_k}.$$

Note that

$$\frac{d}{d\gamma}\left[\frac{1}{\Phi^S_{ni,k}} \exp\left(-\Phi^S_{ni,k}\gamma^{-\theta^S_k/\alpha^S_k}\right)\right] = \exp\left(-\Phi^S_{ni,k}\gamma^{-\theta^S_k/\alpha^S_k}\right)\left(\theta^S_k/\alpha^S_k\right)\gamma^{-\left(\left(\theta^S_k/\alpha^S_k\right)+1\right)}.$$

Using this result to evaluate the integral above, we have:

$$\lambda^S_{j(k)|ni} = \Phi^S_{nij(k)}\left[\frac{1}{\Phi^S_{ni,k}} e^{-\Phi^S_{ni,k}\gamma^{-\theta^S_k/\alpha^S_k}}\right]_0^\infty,$$

which becomes:

$$\lambda^S_{j(k)|ni} = \frac{\Phi^S_{ni\ell(k)}}{\Phi^S_{ni,k}} = \frac{T^S_{j(k)}\left(P^S_{j(k)}\right)^{-\theta^S_k}\left(d^S_{nij(k)}\right)^{\frac{\theta^S_k}{\alpha^S_k}}}{\sum_{\ell \in N} T^S_{\ell(k)}\left(P^S_{\ell(k)}\right)^{-\theta^S_k}\left(d^S_{ni\ell(k)}\right)^{\frac{\theta^S_k}{\alpha^S_k}}}. \tag{B.10}$$

and corresponds to equation (10) in the paper. We refer to this probability as the conditional consumption probability, since it is computed conditional on living in residence n.

Using the property that the maximum of a sequence of Fréchet distributions is itself Fréchet distributed, the contribution to utility from the preferred location for consuming nontradable services of type k from residence n also has a Fréchet distribution:

$$G^S_{ni,k}(\gamma) = \prod_{\ell \in N} G^S_{ni\ell(k)}(\gamma) = \prod_{\ell \in N} \exp\left(-\Phi^S_{ni\ell(k)}\gamma^{-\left(\theta^S_k/\alpha^S_k\right)}\right),$$

$$G^S_{ni,k}(\gamma) = \exp\left(-\Phi^S_{ni,k}\gamma^{-\left(\theta^S_k/\alpha^S_k\right)}\right), \qquad \Phi^S_{ni,k} \equiv \sum_{\ell \in N} \Phi^S_{ni\ell(k)}.$$

Given this Fréchet distribution for the contribution to utility from the preferred location, the expected contribution to utility from consuming nontradable services of type k from residence n is:

Now define the following change of variables:

$$y = \Phi_{ni,k}^{S} \gamma^{-\theta_k^S/\alpha_k^S}, \qquad dy = \left(\theta_k^S/\alpha_k^S\right) \Phi_{ni,k}^{S} \gamma^{-\left(\left(\theta_k^S/\alpha_k^S\right)+1\right)} d\gamma.$$

$$\gamma = \left(\frac{y}{\Phi_{ni,k}^{S}}\right)^{-\frac{1}{\theta_k^S/\alpha_k^S}}, \qquad d\gamma = \frac{dy}{\left(\theta_k^S/\alpha_k^S\right) \Phi_{ni,k}^{S} \gamma^{-\left(\left(\theta_k^S/\alpha_k^S\right)+1\right)}}.$$

Using this change of variables, we can write the expected contribution to utility as:

$$\begin{aligned}
\mathbb{E}_{nik}\left[\gamma\right] &= \int_0^\infty \left(\theta_k^S/\alpha_k^S\right) \Phi_{ni,k}^{S} \frac{y}{\Phi_{ni,k}^{S}} \exp\left(-y\right) \frac{dy}{\left(\theta_k^S/\alpha_k^S\right) \Phi_{ni,k}^{S} \left(\frac{y}{\Phi_{ni,k}^{S}}\right)^{\frac{\left(\theta_k^S/\alpha_k^S\right)+1}{\left(\theta_k^S/\alpha_k^S\right)}}}, \\
&= \int_0^\infty y \exp\left(-y\right) \frac{dy}{\Phi_{ni,k}^{S} \left(\frac{y}{\Phi_{ni,k}^{S}}\right)^{\frac{\left(\theta_k^S/\alpha_k^S\right)+1}{\left(\theta_k^S/\alpha_k^S\right)}}}, \\
&= \int_0^\infty y^{-\frac{1}{\theta_k^S/\alpha_k^S}} \exp\left(-y\right) \frac{dy}{\Phi_{ni,k}^{S} \left(\Phi_{ni,k}^{S}\right)^{-\frac{\left(\theta_k^S/\alpha_k^S\right)+1}{\left(\theta_k^S/\alpha_k^S\right)}}}, \\
&= \left(\Phi_{ni,k}^{S}\right)^{\frac{1}{\theta_k^S/\alpha_k^S}} \int_0^\infty y^{-\frac{1}{\left(\theta_k^S/\alpha_k^S\right)}} \exp\left(-y\right) dy. \\
&= \vartheta_k^S \left(\Phi_{ni,k}^{S}\right)^{\frac{1}{\left(\theta_k^S/\alpha_k^S\right)}},
\end{aligned}$$

where

$$\vartheta_k^S \equiv \Gamma\left(\frac{\left(\theta_k^S/\alpha_k^S\right)-1}{\left(\theta_k^S/\alpha_k^S\right)}\right) = \int_0^\infty y^{-\frac{1}{\left(\theta_k^S/\alpha_k^S\right)}} \exp\left(-y\right) dy,$$

and $\Gamma(\cdot)$ is the Gamma function. We thus obtain the following measure of residence n's consumption access for non-traded service k:

$$\mathbb{S}_{nik} \equiv \mathbb{E}_{nik}\left[\gamma_{nij(k)}\right] = \vartheta_k^S \left(\Phi_{ni,k}^{S}\right)^{\frac{\alpha_k^S}{\theta_k^S}} = \vartheta_k^S \left[\sum_{\ell\in N} T_{\ell(k),k}^{S} \left(P_{\ell(k)}^{S}\right)^{-\theta_k^S} \left(d_{ni\ell(k)}^{S}\right)^{\frac{\theta_k^S}{\alpha_k^S}}\right]^{\frac{\alpha_k^S}{\theta_k^S}}, \tag{B.11}$$

which corresponds to equation (11) in the paper.

Aggregating across sectors using the Cobb-Douglas function form for consumption of non-tradeable services, we arrive at the following expression for consumption access:

$$\mathbb{S}_{ni} = \prod_{k\in K^S} \mathbb{S}_{nik} = \prod_{k\in K^S} \vartheta_k^S \left[\sum_{\ell\in N} T_{\ell(k)}^{S} \left(P_{\ell(k)}\right)^{-\theta_k^S} \left(d_{ni\ell(k)}^{S}\right)^{\frac{\theta_k^S}{\alpha_k^S}}\right]^{\frac{\alpha_k^S}{\theta_k^S}},$$

which corresponds to equation (12) in the paper.

B.4 Workplace Choice

We next turn to the worker's choice of workplace. In making this choice, each worker takes into account access to surrounding consumption possibilities. In particular, conditional on living in residence n, each worker chooses the workplace i and sector g that offers the highest utility, taking into account the wage per efficiency unit ($w_{i,g}$), the idiosyncratic draw for productivity ($a_{i,g}(\omega)$), commuting costs (d_{ni}^W), and expected consumption access ($\mathbb{S}_{ni}$):

$$v_{ni,g}(\omega) = w_{i,g} a_{i,g}(\omega) \exp\left(-\kappa^W \tau_{ni}^W\right) \mathbb{S}_{ni}. \tag{B.12}$$

We thus have the following monotonic relationship between the contribution to utility from income and productivity:

$$a_{i,g}(\omega) = \frac{v_{ni,g}(\omega)}{w_{i,g} \exp\left(-\kappa^W \tau_{ni}^W\right) \mathbb{S}_{ni}}.$$

Therefore, using this relationship and the Fréchet distribution for idiosyncratic productivity, we have:

$$\Pr\left[v_{ni,g} < v\right] = G_{ni,g}^W\left(\frac{v}{w_{i,g} \exp\left(-\kappa^W \tau_{ni}^W\right) \mathbb{S}_{ni}}\right),$$

$$G_{ni}^W(v) = \exp\left(-\Phi_{ni,g}^W v^{-\theta^W}\right), \qquad \Phi_{ni,g}^W \equiv T_{i,g}^W w_{i,g}^{\theta^W} \exp\left(-\theta^W \kappa^W \tau_{ni}^W\right) \left(\mathbb{S}_{ni}\right)^{\theta^W}.$$

Using this distribution for the contribution to utility from income, the probability that a worker in residence n commute to workplace i in sector g is:

$$\begin{aligned}
\lambda_{ig|n}^W &= \Pr\left[v_{ni,g} \geq \max\left\{v_{n\ell,m}\right\}; \forall \ell, m\right], \\
&= \int_0^\infty \prod_{\ell \neq i} G_{n\ell,g}^W(v) \left[\prod_{\ell \in N} \prod_{m \neq g} G_{n\ell,m}(v)\right] g_{ni,g}^W(v)\, dv, \\
&= \int_0^\infty \prod_{\ell \neq i} \exp\left(-\Phi_{n\ell,g}^W v^{-\theta^W}\right) \left[\prod_{\ell \in N} \prod_{m \neq g} \exp\left(-\Phi_{n\ell,m}^W v^{-\theta^W}\right)\right] \theta^W \Phi_{ni,g}^W v^{-(\theta^W+1)} \exp\left(-\Phi_{ni,g}^W v^{-\theta^W}\right) dv, \\
&= \int_0^\infty \exp\left(-\Phi_n^W v^{-\theta^W}\right) \theta^W \Phi_{ni,g}^W v^{-(\theta^W+1)} dv,
\end{aligned}$$

where

$$\Phi_n^W \equiv \sum_{\ell \in N} \sum_{m \in G} \Phi_{n\ell,m}^W = \sum_{\ell \in N} \sum_{m \in G} T_{\ell,m}^W w_{\ell,m}^{\theta^W} \exp\left(-\theta^W \kappa^W \tau_{n\ell}^W\right).$$

Note that

$$\frac{d}{dv}\left[\frac{1}{\Phi_n^W} \exp\left(-\Phi_n^W v^{-\theta^W}\right)\right] = \exp\left(-\Phi_n^W v^{-\theta^W}\right) \left(\mathbb{S}_{ni}\right)^{\theta^W} \theta^W v^{-(\theta^W+1)}.$$

Using this result to evaluate the integral above, we have:

$$\lambda_{ig|n}^W = \Phi_{ni,g}^W \left[\frac{1}{\Phi_n^W} \exp\left(-\Phi_n^W v^{-\theta^W}\right)\right]_0^\infty,$$

which becomes:

$$\lambda_{ig|n}^W = \frac{T_{i,g}^W w_{i,g}^{\theta^W} \exp\left(-\theta^W \kappa^W \tau_{ni}^W\right) \left(\mathbb{S}_{ni}\right)^{\theta^W}}{\sum_{\ell \in N} \sum_{m \in K} T_{\ell,m}^W w_{\ell,m}^{\theta^W} \exp\left(-\theta^W \kappa^W \tau_{n\ell}^W\right) \left(\mathbb{S}_{n\ell}\right)^{\theta^W}}, \tag{B.13}$$

and corresponds to equation (14) in the paper. We refer to this expression as the conditional commuting probability, since again it is computed conditional on living in residence n. Aggregating across sectors, we also obtain the overall commuting probability between residence n and workplace i:

$$\lambda_{i|n}^W = \sum_{g \in K} \lambda_{ig|n}^W. \tag{B.14}$$

Using the property that the maximum of a sequence of Fréchet distributions is itself Fréchet distributed, the distribution of income for residence n across all workplaces ℓ and sectors m also has a Fréchet distribution:

$$G_n^W(v) = \prod_{\ell \in N} \prod_{m \in G} G_{n\ell,g}^W(v) = \prod_{\ell \in N} \prod_{m \in G} e^{-\Phi_{n\ell,m}^W v^{-\theta^W}},$$

$$G_n^W(v) = e^{-\Phi_n^W v^{-\theta^W}}, \qquad \Phi_n^W \equiv \sum_{\ell \in N} \sum_{m \in G} \Phi_{n\ell,m}^W.$$

Given this Fréchet distribution for the contribution to utility from income from the chosen workplace, the expected contribution to utility from income is:

$$\begin{aligned}
\mathbb{E}_n^W[v] &= \int_0^\infty v g_n^W(v)\, dv, \\
&= \int_0^\infty \theta^W \Phi_n^W v^{-\theta^W} \exp\left(-\Phi_n^W v^{-\theta^W}\right) dv.
\end{aligned}$$

Now define the following change of variables:

$$y = \Phi_n^W v^{-\theta^W}, \qquad dy = \theta^W \Phi_n^W v^{-(\theta^W+1)} dv. \tag{B.15}$$

$$v = \left(\frac{y}{\Phi_n^W}\right)^{-\frac{1}{\theta^W}}, \qquad dv = \frac{dy}{\theta^W \Phi_n^W v^{-(\theta^W+1)}}.$$

Using this change of variables, we can write the expected contribution to utility as:

$$\begin{aligned}
\mathbb{E}_n[v] &= \int_0^\infty \theta^W \Phi_n^W \frac{y}{\Phi_n^W} \exp(-y) \frac{dy}{\theta^W \Phi_n^W \left(\frac{y}{\Phi_n^W}\right)^{\frac{\theta^W+1}{\theta^W}}}, \\
&= \int_0^\infty y \exp(-y) \frac{dy}{\Phi_n^W \left(\frac{y}{\Phi_n^W}\right)^{\frac{\theta^W+1}{\theta^W}}}, \\
&= \int_0^\infty y^{-\frac{1}{\theta^W}} \exp(-y) \frac{dy}{\Phi_n^W \left(\Phi_n^W\right)^{-\frac{\theta^W+1}{\theta^W}}}, \\
&= \left(\Phi_n^W\right)^{\frac{1}{\theta^W}} \int_0^\infty y^{-\frac{1}{\theta^W}} \exp(-y)\, dy. \\
&= \vartheta^W \left(\Phi_n^W\right)^{\frac{1}{\theta^W}},
\end{aligned}$$

where

$$\vartheta^W \equiv \Gamma\left(\frac{\theta^W - 1}{\theta^W}\right) = \int_0^\infty y^{-\frac{1}{\theta^W}} \exp(-y)\, dy,$$

and $\Gamma(\cdot)$ is the Gamma function. We thus have the following expression for expected income conditional on living in residence n:

$$\mathbb{A}_n = \mathbb{E}_n[v_{ni,g}] = \vartheta^W \left[\sum_{\ell \in N} \sum_{m \in K} T_{\ell,m}^W w_{\ell,m}^{\theta^W} \exp\left(-\theta^W \kappa^W \tau_{n\ell}^W\right) \left(\mathbb{S}_{n\ell}\right)^{\theta^W}\right]^{\frac{1}{\theta^W}}, \tag{B.16}$$

which corresponds to equation (16) in the paper.

While expected utility is equalized across all workplaces conditional on residence, expected income is different because of the heterogeneity of consumption access $\mathbb{S}_{ni}$ for bilateral commuting pairs. Therefore, expected income for workers in residence n and workplace i is given by:

$$E_{ni} = \frac{\mathbb{A}_i}{\mathbb{S}_{ni}}, \tag{B.17}$$

and expected income by residents in n is given by:

$$E_n = \sum_{i \in N} E_{ni} \lambda_{i|n}^W. \tag{B.18}$$

B.5 Residence Choice

Having characterized a worker's consumption and workplace choices conditional on her residence, we now turn to her residence choice. Each worker chooses her residence after observing her idiosyncratic draws for amenities (b), but before observing her idiosyncratic draws for productivity (a) and the quality of non-traded services (q). Therefore, each worker ω chooses the residence n that offers her the highest utility given her idiosyncratic amenity draws ($b_n(\omega)$), expected travel accessibility ($\mathbb{A}_n$), and other residence characteristics (the price of floor space (Q_n), the price of the traded good (P_n^T) and common amenities (B_n)):

$$U_n(\omega) = B_n b_n(\omega) \left(P_n^T\right)^{-\alpha^T} Q_n^{-\alpha^H} \mathbb{A}_n,$$

We thus have the following monotonic relationship between idiosyncratic amenities and utility:

$$b_n(\omega) = U_n(\omega) B_n^{-1} \mathbb{A}_n^{-1} \left(P_n^T\right)^{\alpha^T} Q_n^{\alpha^H}$$

Therefore, using this relationship and the Fréchet distribution for idiosyncratic amenities, we have:

$$\Pr\left[U_n(\omega) < u\right] = G_n^B\left(u B_n^{-1} \mathbb{A}_n^{-1} \left(P_n^T\right)^{\alpha^T} Q_n^{\alpha^H}\right),$$

$$G_n^B(u) = \exp\left(-\Phi_n^B u^{-\theta^B}\right), \qquad \Phi_n^B \equiv T_n^B B_n^{\theta^B} \mathbb{A}_n^{\theta^B} \left(P_n^T\right)^{-\alpha^T\theta^B} Q_n^{-\alpha^H\theta^B}.$$

Using this distribution for utility, the probability that a worker chooses residence n is:

$$\begin{aligned}
\lambda_n^B &= \Pr\left[U_n > \max\{U_i : i \neq n\}\right], \\
&= \int_0^\infty \prod_{i \neq n} G_i^B(u)\, g_n^B(u)\, du, \\
&= \int_0^\infty \prod_{i \neq n} \exp\left(-\Phi_i^B u^{-\theta^B}\right) \theta^B \Phi_n^B u^{-\left(\theta^B+1\right)} \exp\left(-\Phi_n^B u^{-\theta^B}\right) du, \\
&= \int_0^\infty \exp\left(-\Phi^B u^{-\theta^B}\right) \theta^B \Phi_n^B u^{-\left(\theta^B+1\right)} du,
\end{aligned}$$

where

$$\Phi^B \equiv \sum_{i \in N} \Phi_i^B = \sum_{i \in N} T_i^B B_n^{\theta^B} \mathbb{A}_n^{\theta^B} \left(P_n^T\right)^{-\alpha^T\theta^B} Q_n^{-\alpha^H\theta^B}.$$

Note that

$$\frac{d}{du}\left[\frac{1}{\Phi^B} \exp\left(-\Phi^B u^{-\theta^B}\right)\right] = \exp\left(-\Phi^B u^{-\theta^B}\right) \theta^B u^{-\left(\theta^B+1\right)}.$$

Using this result to evaluate the integral above, we have:

$$\lambda_n^B = \Phi_n^B \left[\frac{1}{\Phi^B} \exp\left(-\Phi^B u^{-\theta^B}\right)\right]_0^\infty,$$

which implies that the probability that each worker chooses residence n (λ_n^B) is given by:

$$\lambda_n^B = \frac{T_n^B B_n^{\theta^B} \mathbb{A}_n^{\theta^B} \left(P_n^T\right)^{-\alpha^T\theta^B} Q_n^{-\alpha^H\theta^B}}{\sum_{\ell \in N} T_\ell^B B_\ell^{\theta^B} \mathbb{A}_\ell^{\theta^B} \left(P_\ell^T\right)^{-\alpha^T\theta^B} Q_\ell^{-\alpha^H\theta^B}}, \tag{B.19}$$

which corresponds to equation (17) in the paper.

Using the property that the maximum of a sequence of Fréchet distributions is itself Fréchet distributed, the distribution of utility across all locations is also Fréchet:

$$G^B(u) = \prod_{i \in N} G_i^B(u) = \prod_{i \in N} \exp\left(-\Phi_i^B u^{-\theta^B}\right),$$

$$G^B(u) = \exp\left(-\Phi^B u^{-\theta^B}\right), \qquad \Phi^B \equiv \sum_{i \in N} \Phi_i^B.$$

Given this Fréchet distribution for utility, expected utility is:

$$\begin{aligned}
\mathbb{E}^B[u] &= \int_0^\infty u g^B(u)\, du, \\
&= \int_0^\infty \theta^B \Phi^B u^{-\theta^B} \exp\left(-\Phi^B u^{-\theta^B}\right) du.
\end{aligned}$$

Now define the following change of variables:

$$y = \Phi^B u^{-\theta^B}, \qquad dy = \theta^B \Phi^B u^{-\left(\theta^B+1\right)} du.$$

$$u = \left(\frac{y}{\Phi^B}\right)^{-\frac{1}{\theta^B}}, \qquad du = \frac{dy}{\theta^B \Phi^B u^{-(\theta^B+1)}}.$$

Using this change of variables, we can write expected utility as:

$$\begin{aligned}
\mathbb{E}^B[u] &= \int_0^\infty \theta^B \Phi^B \frac{y}{\Phi^B} \exp(-y) \frac{dy}{\left(\theta^B\right) \Phi^B \left(\frac{u}{\Phi^B}\right)^{\frac{\theta^B+1}{\theta^B}}}, \\
&= \int_0^\infty y \exp(-y) \frac{dy}{\Phi^B \left(\frac{y}{\Phi^B}\right)^{\frac{\theta^B+1}{\theta^B}}}, \\
&= \int_0^\infty y^{-\frac{1}{\theta^B}} \exp(-y) \frac{dy}{\Phi^B \left(\Phi^B\right)^{-\frac{\theta^B+1}{\theta^B}}}, \\
&= \left(\Phi^B\right)^{\frac{1}{\theta^B}} \int_0^\infty y^{-\frac{1}{\theta^B}} \exp(-y)\, dy. \\
&= \vartheta_B \left(\Phi^B\right)^{\frac{1}{\theta^B}},
\end{aligned}$$

where

$$\vartheta_B \equiv \Gamma\left(\frac{\theta^B - 1}{\theta^B}\right) = \int_0^\infty y^{-\frac{1}{\theta^B}} \exp(-y)\, dy,$$

and $\Gamma(\cdot)$ is the Gamma function. We thus have the following expression for the expected utility from living in the city:

$$\mathbb{E}[u] = \vartheta^B \left[\sum_{\ell \in N} T_\ell^B B_\ell^{\theta^B} \mathbb{A}_\ell^{\theta^B} P_{\ell,T}^{-\alpha^T \theta^B} Q_\ell^{-\alpha^H \theta^B}\right]^{\frac{1}{\theta^B}}, \tag{B.20}$$

which corresponds to equation (18) in the paper.

Having characterized workplace and residence choices, we can also recover the demand for residential floor space in each location, using the implication of Cobb-Douglas utility that expenditure on residential floor space is a constant share of income:

$$H_{n,U} = \frac{\alpha^H E_n R_n}{Q_n}, \tag{B.21}$$

where $R_n = \lambda_n^B \bar{L}$ is the measure of residents in location n; recall that $\bar{L}$ is total city population; and E_n is expected income in residence n.

B.6 Production

When we undertake counterfactuals in our quantitative analysis below, we do need to take a stand on production technology and market structure, in which case consider a version of the canonical urban model. In particular, we assume that both the traded good and non-traded services are produced using labor and commercial floor space according a constant returns to scale technology. We assume for simplicity that this production technology is Cobb-Douglas and that production occurs under conditions of perfect competition. Together these assumptions imply that profits are zero in each location in which a tradable good and non-tradable service is produced:

$$\begin{aligned} P_i^T &= \frac{1}{A_{i,k}} w_{i,k}^{\beta^T} Q_i^{1-\beta^T}, \qquad 0 < \beta^T < 1, \qquad k \in K/K^S, \\ P_{i(k)}^S &= \frac{1}{A_{i,k}} w_{i,k}^{\beta^S} Q_i^{1-\beta^S}, \qquad 0 < \beta^S < 1, \qquad k \in K^S, \end{aligned} \tag{B.22}$$

where $A_{i,k}$ is productivity in location i in sector k. Using the first-order condition for profit maximization, we can obtain demand for commercial floor space in each sector and location ($H_{i,k}$) as a function of the goods or service price ($P_{i(k)}^S$), productivity ($A_{i,k}$), the price of floor space (Q_i) and labor input adjusted for effective units of labor ($\tilde{L}_{i,k}$):

$$H_{i,k} = \begin{cases} \frac{1-\beta^T}{\beta^T} \left(\frac{P_i^T A_{i,k}}{Q_i} \right)^{\frac{1}{\beta^T}} \tilde{L}_{i,k}, \qquad k \in K/K^S \\ \frac{1-\beta^S}{\beta^S} \left(\frac{P_{i(k)}^S A_{i,k}}{Q_i} \right)^{\frac{1}{\beta^S}} \tilde{L}_{i,k}, \qquad k \in K^S, \end{cases} \tag{B.23}$$

where $\tilde{L}_{i,g}$ denotes labor input adjusted for expected idiosyncratic worker productivity, i.e.,

$$\tilde{L}_{i,g} = \frac{1}{w_{i,g}} \sum_{n \in N} E_{ni} \lambda_{ig|n}^W R_n.$$

where E_{ni} is the labor income earned by workers who reside in n and work in i.

We allow productivity in equations (B.22) and (B.23) to be either exogenous or endogenous to the surrounding concentration of economic activity in the presence of agglomeration forces, as discussed further below. We assume no-arbitrage between residential and commercial floor space, and across the different sectors in which commercial floor space is used, such that there is a single price for floor space within each location (Q_i) in equation (B.22). In general, the wage per efficiency unit ($w_{i,k}$) differs across both sectors and locations in equation (B.22), because workers draw efficiency units for each combination of sector and location pair, and hence each sector and location pair faces an upward-sloping supply function for effective units of labor. Finally, we assume that the traded good is costlessly traded within the city and wider economy and choose it as our numeraire such that:

$$P_i^T = 1 \qquad \forall\, i \in N. \tag{B.24}$$

B.7 Market Clearing

The price for each type of non-traded service k in each location j ($P_{j(k)}^S$ for $k \in K^S$) is endogenously determined by market clearing, which requires that revenue equals expenditure for that non-traded service k and location j:

$$P_{j(k)}^S A_{j,k} \left(\frac{\tilde{L}_{j,k}}{\beta^S} \right)^{\beta^S} \left(\frac{H_{j,k}}{1-\beta^S} \right)^{1-\beta^S} = \alpha_k^S \sum_{n \in N} R_n \sum_{i \in N} \lambda_{j(k)|ni}^S \lambda_{i|n}^W E_{ni}, \qquad k \in K^S, \tag{B.25}$$

where expenditure on the right-hand side equals the sum across locations of workers travelling to consume non-traded service k in location j; R_n is the measure of residents in location n; recall that $\lambda^S_{j(k)|n}$ is the conditional consumption probability and E_{ni} is expected income by workers with residence n and workplace i.

Labor market clearing implies that the measure of workers employed in workplace j in sector k equals the total measure of workers from all residences n who commute to that workplace j in sector k:

$$L_{j,k} = \sum_{n \in N} \lambda^W_{jk|n} R_n, \qquad k \in K, \tag{B.26}$$

where we use $L_{j,k}$ without a tilde to denote the measure of workers without adjusting for effective units of labor; and recall that $\lambda^W_{jk|n}$ is the conditional commuting probability.

Land market clearing requires that the demand for residential floor space $(H_{i,U})$ plus the sum across sectors of the demand for commercial floor space in each sector $(H_{i,k})$ equals the total supply of floor space (H_i):

$$H_i = H_{i,U} + \sum_{k \in K} H_{i,k}. \tag{B.27}$$

B.8 General Equilibrium with Exogenous Location Characteristics

We begin by considering the case in which productivity $(A_{i,k})$, amenities (B_i) and the supply of floor space (H_i) are exogenously determined. The general equilibrium of the model is referenced by the price for floor space in each location (Q_i), the wage in each sector and location $(w_{i,k})$, the price of the non-traded good in each service sector and location $(P^S_{i(k)})$, the route choice probabilities $(\lambda^R_{r(k)|nij(k)})$, the conditional consumption probabilities $(\lambda^S_{j(k)|ni})$, the conditional commuting probabilities $(\lambda^W_{jk|n})$, the residence probabilities (λ^B_n), and the total measure of workers living in the city $(\bar{L})$, where we focus on the open-city specification, in which the total measure of workers is endogenously determined by population mobility with the wider economy. Given these seven equilibrium variables, we can solve for all other endogenous variables of the models. These equilibrium variables are determined by the system of seven equations given by the land market clearing condition for each location (B.27), the labor market clearing condition for each location (B.26), the non-traded goods market clearing condition for each location and service sector (B.25), the conditional consumption probabilities (B.10), the conditional commuting probabilities (B.13), the residence probabilities (B.19), and the population mobility condition that equates expected utility in the city (B.20) to the reservation level of utility in the wider economy $(\bar{U})$.

B.9 General Equilibrium with Agglomeration Forces and Endogenous Floor Space

We next extend the analysis to allow productivity and amenities to be endogenous to the surrounding concentration of economic activity through agglomeration forces and to allow for an endogenous supply of floor space.

Agglomeration in Production. In both the traded and non-traded sector, we allow productivity $(A_{i,k})$ to depend on production fundamentals and production externalities. Production fundamentals $(a_{i,k})$ capture features of physical geography that make a location more or less productive independently of neighboring economic activity (e.g. access to natural water). Production externalities capture productivity benefits from the density of employment across all sectors (L_i/K_i), where employment density is measured per unit of geographical land area.[4]

$$A_{i,k} = a_{i,k} \left(\frac{L_i}{K_i} \right)^{\eta^W} \tag{B.28}$$

where $L_i = \sum_{k \in K} L_{i,k}$ is the total employment in location i, and η^W parameters the strength of production externalities, which we assume to the same across all sectors.

Agglomeration in Residents. Similarly, we allow residential amenities (B_n) to depend on residential fundamentals and residential externalities. Residential fundamentals (b_n) capture features of physical geography that make a location a more or less attractive place to live independently of neighboring economic activity (e.g. green areas). Residential externalities capture the effects of the surrounding density of residents (L_i/K_i) and are modeled symmetrically to production externalities:[5]

$$B_n = b_n \left(\frac{R_n}{K_n} \right)^{\eta^B} \tag{B.29}$$

where η^B parameters the strength of residential externalities.

Floor Space Supply We follow the standard approach in the urban literature of assuming that floor space is supplied by a competitive construction sector that uses land K and capital M as inputs. Following Combes, Duranton, and Gobillon (2019) and Epple, Gordon, and Sieg (2010), we assume that floor space (H_i) is produced using geographical land (K_i) and building capital (M_i) according to the following constant return scale technology:

$$H_i = M_i^{\mu} K_i^{1-\mu}, \qquad 0 < \mu < 1. \tag{B.30}$$

Using cost minimization and zero profits, this Cobb-Douglas construction technology implies that payments for building capital are a constant share of overall payments for the use of floor space:

$$\mu Q_i H_i = \mathbb{P} M_i, \tag{B.31}$$

where $\mathbb{P}$ is the common user cost of building capital. Using the construction technology (B.30) to substitute for building captial (M_i) in equation (B.31) linking payments for floor space and building capital, we obtain a constant elasticity supply function for floor space as in Saiz (2010), with the inverse supply function given by:

$$Q_i = \psi_i H_i^{\frac{1-\mu}{\mu}} \tag{B.32}$$

where $\psi_i = \mathbb{P} K_i^{\frac{\mu-1}{\mu}} / \mu$ depends solely on geographical land area (K_i) and parameters.

Furthermore, the cost minimization and zero profit condition also implies that:

$$Q_i = \left(\frac{\mathbb{P}}{\mu} \right)^{\mu} \left(\frac{\tilde{Q}_i}{1-\mu} \right)^{1-\mu} \tag{B.33}$$

where $\tilde{Q}_i$ is the price of land per unit area.

[4] We assume for simplicity that production externalities depend solely on a location's own employment density, although it is straightforward to allow for spillovers of these production externalities across locations.

[5] As for production externalities above, we assume that residential externalities depend solely on a location's own residents density, but it is straightforward to allow for spillovers of these residential externalities across locations.

Given this specification of agglomeration forces and endogenous floor space, the determination of general equilibrium remains the same as above with exogenous location characteristics above, except that productivity (A_n), amenities (B_n) and the supply of floor space (H_n) are now endogenously determined by equations (B.28), (B.29) and (B.32).

C Model Extensions

In this section of the online appendix, we discuss a number of different extensions of our theoretical mode. In Section C.1, we generalize the model to incorporate different frequencies of trips across the non-traded sectors, and show that the resulting model is isomorphic to our baseline specification up to the interpretation of the parameter κ_k^S that captures the response of commuting costs to travel times. In Section C.2, we show that our specification of the supply-side of the model with competitive markets and external economies of scale (through agglomeration forces in production) is isomorphic to a model of monopolistic competition under free entry.

C.1 Incorporating Frequency of Consumption Trips

In Section 4 of the paper, we capture the relative importance of each non-traded sector using its expenditure share, assuming for simplicity that users make one trip for each type of non-traded service. More generally, the frequency of trips can also differ across the non-traded sectors, as shown in Figure 7 in the paper. In this section of the online appendix, we explicitly incorporate this additional type of heterogenegity and show that the model is isomorphic up to a reinterpretation of the parameters κ_k^S. Therefore, all of our counterfactual results are unaffected by this extension of the model except for the interpretation of the estimated κ_k^S.

Similarly to equation (2) in our main paper, we assume that the iceberg travel cost for each combination of residence n, workplace i, consumption location $j(k)$, and route $r(k)$ ($d_{ni\{j(k)r(k)\}}$) as follows:

$$d_{ni\{j(k)r(k)\}} = \exp(-\kappa^W \tau_{ni}^W) \prod_{k \in K^S} \exp(-\kappa_k^S \tau_{nij(k)r(k)}^S). \tag{C.1}$$

In this expression, the first term before the product sign captures the cost of commuting from residence n to workplace i without any detour to consume non-traded services, which depends on travel time (τ_{ni}^W) and the commuting cost parameter (κ^W), where overall commuting travel time is the sum of the travel time incurred in each direction:

$$\tau_{ni}^W = \tau_{ni} + \tau_{in}.$$

The second term in equation (C.1) captures the additional travel costs involved in consuming each type of non-traded service k in location $j(k)$ by the route $r(k)$. Unlike in our main text, here we assume that workers have to visit the location of consumption δ_k times to meet their needs, where δ_k is an exogenous parameter depending on the model. δ_k can be less than one, in which case workers do not have to make a trip every day, while δ_k can be greater than one, in which case workers have to visit the location multiple times. δ_k intuitively captures the relative frequency of travel for different sectors as documented in Figure 7. For example, relatively frequent trips are required for grocery shopping, while less frequent trips are required for visiting banks. Accommodating the differences of the total travel time driven by these frequencies of trips, the additional travel time for each route taken is given by:

$$\begin{aligned}
\tau_{nij(k)HH}^S &= \delta_k \left(\tau_{nj} + \tau_{jn}\right), \\
\tau_{nij(k)WW}^S &= \delta_k \left(\tau_{ij} + \tau_{ji}\right), \\
\tau_{nij(k)HW}^S &= \delta_k \left(\tau_{nj} + \tau_{ji} - \tau_{ni}\right), \\
\tau_{nij(k)WH}^S &= \delta_k \left(\tau_{ij} + \tau_{jn} - \tau_{in}\right).
\end{aligned}$$

Here, δ_k enters multiplicatively with consumption travel cost κ_k^S in equation (C.1). Therefore, this specification of the travel cost is isomorphic to our main specification by replacing κ_k^S in our main paper with $\kappa_k^S \delta_k$.

It is important to note that this specification does not affect any of our counterfactual simulation results. When undertaking counterfactual simulations, we only need the composite elasticity of travel time ($\phi_k^S (= \theta_k^S \kappa_k^S / \alpha_k^S)$), which we estimate following the procedure in Section 5.1. With this extension, we estimate the same values of ϕ_k^S, except that the interpretation of these parameters is different ($\phi_k^S = \theta_k^S \kappa_k^S \delta_k / \alpha_k^S$). Intuitively, given the observed spatial decay patterns ϕ_k^S, less frequency of making trips for nontradable sector k (a smaller δ_k) implies that travel cost for this sector k is higher (κ_k^S), but whether ϕ_k^S is driven by κ_k^S or δ_k do not matter for our counterfactual simulations. For this reason, our counterfactual simulation results are unaffected by this extension of our model.

C.2 Monopolistic Competition and Firm Entry

The model in our main paper assumes that firms are perfectly competitive in each location, and there are agglomeration spillovers for production. In this appendix, we instead assume that firms are monopolistically competitive, and firms enter in each location following the free-entry condition. We show that this alternative model leads to a similar expressions for the equilibrium conditions with a slight modification of the functional form of the agglomeration spillovers.

In this alternative model, we assume that there are potential entrepreneurs who consider setting up a store in each location in sector k. Below we discuss the case of nontradable sector $k \in K^S$, but the discussion is isomprphic to the tradable sector $k \in K/K^S$. Each entrepreneur has access to a distinct variety of goods and services. Entering in each location i requires a fixed cost payment $f_{i,k}^S$ in the unit of the Cobb-Douglas composite of labor and floor space with the labor share of β^S, such that an entry requires a lump-sum payment of $f_{i,k}^S w_{i,k}^{\beta^T} Q_i^{1-\beta^T}$. After the entry, she produces her good or service using the Cobb-Douglas production technology with labor and input. Therefore, the marginal cost is given by

$$c_{i(k)}^S = \frac{1}{a_{i,k}} w_{i,k}^{\beta^S} Q_i^{1-\beta^S}.$$

We assume that consumers have a constant elasticity of substitution (CES) utility over these differentiated varieties with elasticity of substitution σ. Firms are monopolistically competitive, and hence charge the prices with a constant markup:

$$p_{i(k)}^S = \frac{\sigma}{\sigma - 1} \frac{1}{a_{i,k}} w_{i,k}^{\beta^S} Q_i^{1-\beta^S}.$$

The measure of firms $M_{i(k)}^S$ is determined by the free entry condition. Under this condition, firm profit after entry is exactly offset by the fixed cost payment. Since firms earn $1/\sigma$ fraction of the revenue as profit, the free entry condition is given by

$$\begin{aligned} M_{i(k)}^S f_{i,k}^S w_{i,k}^{\beta^T} Q_i^{1-\beta^T} &= p_{i(k)}^S a_{i,k} \left(\frac{\tilde{L}_{i,k}}{\beta^S} \right)^{\beta^S} \left(\frac{H_{i,k}}{1-\beta^S} \right)^{1-\beta^S} \\ &= \frac{\sigma}{\sigma - 1} \left(\frac{w_{i,k} \tilde{L}_{i,k}}{\beta^S} \right)^{\beta^S} \left(\frac{Q_i H_{i,k}}{1-\beta^S} \right)^{1-\beta^S}. \end{aligned}$$

where $\tilde{L}_{i,k}$ and $H_{i,k}$ is the aggregate efficient unit of labor input and commercial floor space in location i and sector k. By reformulating this equation, we have

$$M_{i(k)}^{S} = \frac{1}{f_{i,k}^{S}} \frac{\sigma}{\sigma-1} \left(\frac{\tilde{L}_{i,k}}{\beta^{S}} \right)^{\beta^{S}} \left(\frac{H_{i,k}}{1-\beta^{S}} \right)^{1-\beta^{S}} .$$

Using the standard property of the CES utility function, the price index of the composite of goods offered in location in i in sector k is given by

$$\begin{aligned} P_{i(k)}^{S} &= p_{i(k)}^{S} \left(M_{i(k)} \right)^{\frac{1}{1-\sigma}} \\ &= \frac{\sigma}{\sigma-1} \frac{1}{a_{i,k}} w_{i,k}^{\beta^{S}} Q_{i}^{1-\beta^{S}} \left[\frac{1}{f_{i,k}^{S}} \frac{\sigma}{\sigma-1} \left(\frac{\tilde{L}_{i,k}}{\beta^{S}} \right)^{\beta^{S}} \left(\frac{H_{i,k}}{1-\beta^{S}} \right)^{1-\beta^{S}} \right]^{\frac{1}{1-\sigma}} \\ &= \frac{1}{\tilde{A}_{i,k}} w_{i,k}^{\beta^{S}} Q_{i}^{1-\beta^{S}}, \end{aligned} \tag{C.2}$$

where $\tilde{A}_{i,k}$ is defined by

$$\tilde{A}_{i,k} = \tilde{a}_{i,k} \left(\tilde{L}_{i,k} \right)^{\frac{\beta^{S}}{\sigma-1}} \left(H_{i,k} \right)^{\frac{1-\beta^{S}}{\sigma-1}} , \tag{C.3}$$

where $\tilde{a}_{i,k} \equiv \frac{\sigma-1}{\sigma} \left[\frac{\sigma}{\sigma-1} \left(\frac{1}{\beta^{S}} \right)^{\beta^{S}} \left(\frac{1}{1-\beta^{S}} \right)^{1-\beta^{S}} \right]^{\frac{1}{\sigma-1}} \left(\frac{1}{f_{i,k}^{S}} \right)^{\frac{1}{\sigma-1}} a_{i,k}$. Note that $\tilde{a}_{i,k}$ is exogenously determined by the parameters of the model.

Therefore, we have shown that the price index is a function of the cost of composite inputs ($w_{i,k}^{\beta^{S}} Q_{i}^{1-\beta^{S}}$), location fundamentals ($\tilde{a}_{i,k}$), and additional terms corresponding to the benefit of the love of variety in consumption ($\left(\tilde{L}_{i,k} \right)^{\frac{\beta^{S}}{\sigma-1}} \left(H_{i,k} \right)^{\frac{1-\beta^{S}}{\sigma-1}}$). Note that this expression of the price index is isomorphic to the model with perfect competition (equation 21), as long as we assume that external agglomeration spillover takes the form of equation C.3 (unlike in equation 27 of our main paper, where we assume $A_{i,k} = a_{i,k} \left(\frac{L_i}{K_i} \right)^{\eta^{W}}$).

D Additional Estimation Results and Model Validation

This section of the appendix provides additional estimation results and model validation, supplementing the results reported in Section 5 of our main paper. In Section D.1, we present a number of overidentification checks on the model's predictions using separate data not used its calibration, including data on residential income and prices. In Section D.2, we provide additional evidence on the fit of our extended gravity equations for consumption and commuting trips, supplementing the results reported in Section 5.1 of the paper.

In Section D.3, we demonstrate that we find a similar pattern of results whether we construct consumption access and travel access using destination or origin fixed effects, consistent with the predictions of our theoretical model. In Section D.4, we compare the estimates of our baseline theoretical model from the paper, which allows for endogenous route choice and trip chains, with those of an alternative specification in which all consumption trips are (falsely) assumed to originate from home.

D.1 Overidentification Test for Model Validation

In this subsection of the appendix, we provide additional model validation by comparing our model's predictions with separate data. In particular, we validate our model's prediction about residential income and the price index in each location, using separate data that we do not use for estimation (see Section 2.2 of our main paper for these other data sources.) Because we do not use these separate data in our model estimation procedure, these comparisons serve as an overidentification test for our model.

D.1.1 Residential Income

We first compare our model's prediction of residential income. Our model predicts that the aggregate residential income for workers in home location n is given by equations (B.17) and (20):

$$E_n^{Model} = \sum_{i \in N} \lambda_{i|n}^W E_{ni} = \sum_{i \in N} \lambda_{i|n}^W \frac{\mathbb{A}_n}{\mathbb{S}_{ni}}. \tag{D.1}$$

where $\mathbb{A}_n$ is our estimate of travel access, and $\mathbb{S}_{ni}$ is our estimate of consumption access under the parameter values of $\theta^W = 6$ and $\theta_k^S = 6$ for all nontradable sector k. We compare this model-predicted residential income with separate data of municipality income from tax-base information, E_n^{Data}.

Table D.1.1 presents the results of a linear regression of the log of income from the tax-base information on the log of model-predicted income. Our model is necessarily an abstraction and does not capture all of the idiosyncratic factors that can affect residential income in individual locations. For example, the tax base data includes non-labor income, whereas our model focuses on labor income. Nevertheless, we estimate a slope coefficient that is close to one and is not statistically significantly different from one: coefficient 0.987 (standard error 0.135). We find a regression R-squared of 0.213, which is in line with the values typically found in univariate regressions using cross-section micro data, and is consistent with the many idsiosycratic factors that affect residential income in individual locations.

This validation in particular supports our choice of the Fréchet dispersion parameter, particularly for θ^W. From equation (16), one can see that θ^W is directly related to the variation in travel access ($\mathbb{A}_n$), which affects residential income. Therefore, the choice of $\theta^W = 6$ is not only consistent with the estimates from previous literature (Ahlfeldt, Redding, Sturm, and Wolf 2015, Heblich, Redding, and Sturm 2020, Kreindler and Miyauchi 2019), but also consistent with external data on residential income for the Tokyo metropolitan area.

Table D.1.1: Model Validation of Price Index

	log income (data)
log income (model)	0.987^{***}
	(0.135)
Observations	198
R^2	0.213

Note: Results of the linear regression of the logarithm of the residential income from tax-base data on the logarithm of model-predicted residential income. Observations are weighted by the residential population. Heteroskedasticity robust standard errors in parentheses.

D.1.2 Price Index

In our last model validation exercise, we validate our model's prediction of price indices in each location. We construct model-predicted price index from our estimated destination fixed effects of our consumption gravity equations. We compare this model-predicted price index with the analogous objects constructed from external data. Because we do not directly observe price indices of each municipality, we construct the price index consistent with the constant elasticity of substitution (CES) utility function in each location, following our extended model with monopolistic competition and firm entry (Section C.2). While these auxiliary assumptions can potentially add substantial measurement errors for our price indices, we find that our model prediction of price index is closely aligned with the price index that we observe in the data.

For the purpose of this validation, we construct our model-predicted proxy for the price index using the destination fixed effects of our consumption location choice estimation. More specifically, the destination fixed effects of consumption location choice is given by equation (36) in our main paper, reproduced here:

$$\xi^S_{j(k)} = T^S_{j(k)} \left(P^S_{j(k)}\right)^{-\theta^S_k}, \tag{D.2}$$

where $P^S_{j(k)}$ is the price index, and $T^S_{j(k)}$ is the parameter that shifts the attractiveness of each consumption location other than from the factors related to price index $P^S_{j(k)}$. To recover the price index from equation (D.2), we impose an auxiliary assumption that $T^S_{j(k)}$ is uncorrelated with price index $P^S_{j(k)}$ such that $Var(T^S_{j(k)}, P^S_{j(k)}) = 0$. Under this assumption, we construct an mean-unbiased proxy for our model price index as

$$P^{S,Model}_{j(k)} \equiv \mathbb{E}\left[P^S_{j(k)}\right] = \mathbb{E}\left[T^S_{j(k)} \left(\xi^S_{j(k)}\right)^{\frac{1}{\theta^S_k}}\right] = \left(\xi^S_{j(k)}\right)^{\frac{1}{\theta^S_k}}, \tag{D.3}$$

where the last transformation used our auxiliary assumption of $Var(T^S_{j(k)}, P^S_{j(k)}) = 0$ and the normalization that $\mathbb{E}[T^S_{j(k)}] = 1$. We assume $\theta^S_k = 6$ consistent with our main calibration (Table 1).

We compare this model-predicted price index with the analogous object constructed from external data. Because the price index is not observed at the municipality level, we construct our proxy for price index under the constant elasticity of substitution (CES) demand system. More specifically, under CES demand system, the price index of location j in sector k is given by:

$$P^{S,Data}_{j(k)} = \left(M_{j(k)}\right)^{\frac{1}{\sigma-1}} p_{j(k)}, \tag{D.4}$$

where $M_{j(k)}$ is the number of varieties in location j in sector k; $p_{j(k)}$ is the price of each variety in location j in sector k; and σ is the elasticity of substitution. We proxy $M_{j(k)}$ by the number of establishments in each location j from economic census data. We proxy $p_{j(k)}$ from the separate retail survey data that provides the relative prices for broad categories of products at the level of prefecture level (4 prefectures covering 240 municipalities in Tokyo Metropolitan Area). Lastly, we set σ such that $\sigma = 5$ from the central estimates from the literature of this parameter in the context of retail products (Broda and Weinstein 2006).

Table D.1.2 presents the results of the OLS regression of the log of the price index constructed from the observed price data ($P_{j(k)}^{S,Data}$) on the log of our model-predicted price proxy ($P_{j(k)}^{S,Model}$). Again our model is necessarily an abstraction, but for all nontraded sectors, we find a strong and statistically significant correlation between our model's predictions and the price data. Furthermore, across the board, the slope coefficient ranges from 0.522-1.114, centered around one. As discussed above, the choice of θ_k^S crucially governs the dispersion of the model-predicted price index (equation (D.2)). Therefore, our choice of $\theta_k^S = 6$ is not only consistent with estimates from the existing empirical literature (Atkin, Faber, and Gonzalez-Navarro 2018, Couture, Gaubert, Handbury, and Hurst 2019), but also consistent with external data on price indices for the Tokyo metropolitan area.

Table D.1.2: Model Validation of Price Index

	Dependent Variable: log(Price Index (data))				
	Finance realestate communication professional	Wholesale retail	Accomoda-tions eating drinking	Medical welfare healthcare	Other services
	(1)	(2)	(3)	(4)	(5)
log(Price Index (model))	1.114***	0.677***	0.718***	0.874***	0.522***
	(0.103)	(0.108)	(0.108)	(0.106)	(0.103)
Observations	240	240	240	240	240
R^2	0.331	0.142	0.157	0.222	0.098

Note: Results of the linear regression of the logarithm of the price index constructed from external data on the logarithm of model-predicted price index. Heteroskedasticity robust standard errors in parentheses.

D.2 Fit of Gravity Equations

In this section of the appendix, we discuss the fit of our empirical models of consumption location choice probability and workplace choice probability (Table 3 and Table 4 of our main paper).

As a specification check, we estimate our workplace and consumption location choice probabilities by including the dummies of the bins of the travel time instead of the linear term. Namely, for the consumption location choice probabilities, we estimate:

$$\lambda_{j(k)|ni}^S = \frac{\xi_{j(k)}^S \exp\left(-\sum_b \tilde{\phi}_{k,b}^S 1\left[\log \tilde{d}_{nij(k)}^S = b\right]\right) \exp\left(u_{nij(k)}^S\right)}{\zeta_{ni,k}^S}, \tag{D.5}$$

where b indicates the deciles of the consumption travel cost $\tilde{d}_{nij(k)}^S$, and the difference from equation (35) is that we replace $\phi_k^S \log \tilde{d}_{nij(k)}^S$ with $\sum_b \tilde{\phi}_{k,b}^S 1\left[\log \tilde{d}_{nij(k)}^S = b\right]$. Therefore, if our specification of exponential travel cost is correct, we expect that $\tilde{\phi}^S_{k,b}$ has a linear relationship with b.

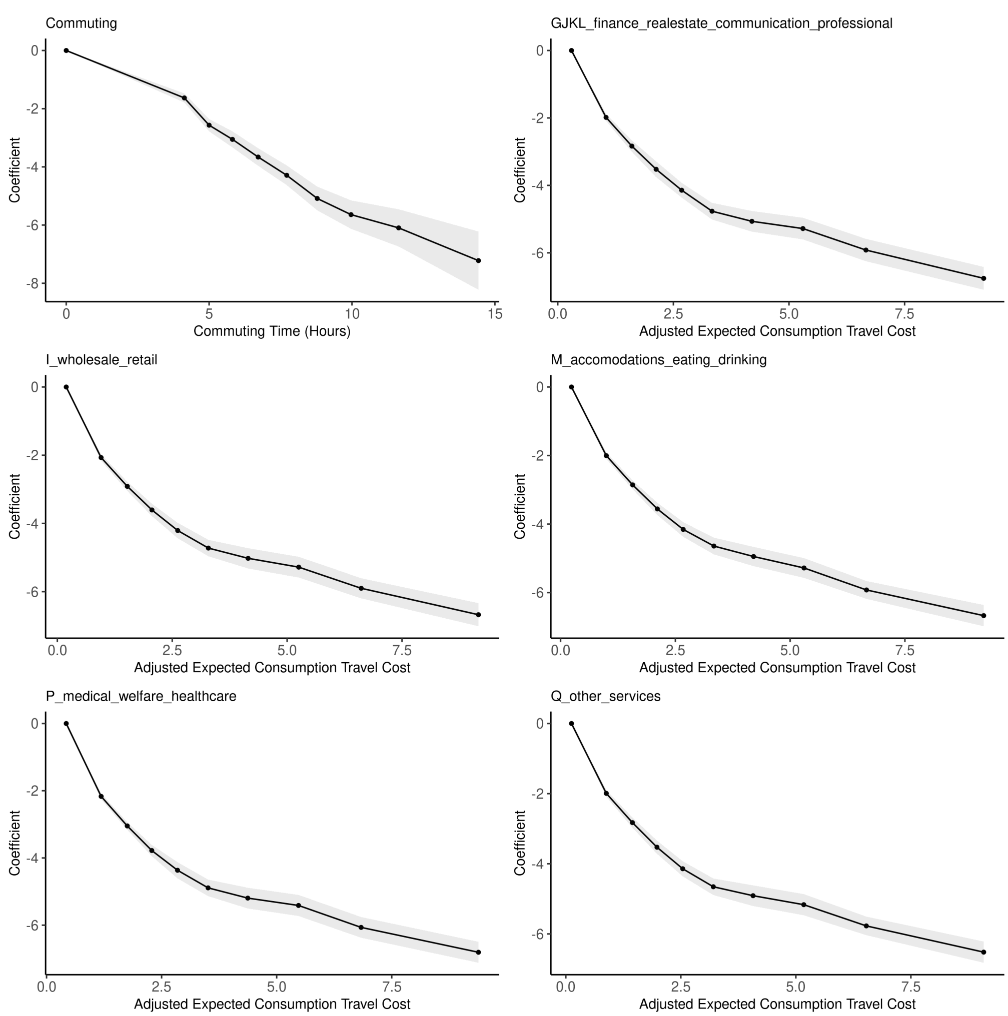
For the workplace choice probabilities, we estimate:

$$\lambda^W_{i|n} = \frac{\xi^W_i \exp\left(-\sum_b \tilde{\phi}^W_b 1\left[\tau^W_{ni} = b\right]\right)(\mathbb{S}_{ni})^{\theta^W} \exp\left(u^W_{ni}\right)}{\zeta^W_n}, \tag{D.6}$$

where b indicates the deciles of the commuting travel time τ^W_{ni}, and the difference from equation (39) is that we replace $\phi^W \tau^W_{ni}$ with $\sum_b \tilde{\phi}^W_b 1\left[\tau^W_{ni} = b\right]$. Therefore, if our specification of exponential travel cost is accurate, we expect that $\tilde{\phi}^W_b$ has a linear relationship with b.

Figure D.2.1 plots the estimated coefficients on these bins of travel time with their 95 percent confidence intervals. We plot the estimated coefficients $\tilde{\phi}^S_{k,b}$ and $\tilde{\phi}^W_b$ against the consumption travel cost and commuting travel time (b in the equations above). The coefficients exhibit an approximately linear relationship with travel time on the horizontal axis for the commuting choices. For consumption choices, there is some convex pattern, but the relationship is approximately linear for the first few deciles where most trips are concentrated. This pattern of results supports our specification of the workplace and consumption location choice, in which travel time enters exponentially in the iceberg travel cost.

Figure D.2.1: Estimation Results of Nonparametric Gravity Equations



D.3 Robustness of Access Measures by Using Origin Fixed Effects

In Section 5.1 of our main paper, we construct consumption access measure $\mathbb{S}_{ni}$ and travel access measure $\mathbb{A}_n$ as the distance-weighted sum of the destination effects of the consumption location choice and the commuting choice, respectively. As we discuss in our main paper, an alternative way of constructing these measures is to use the origin fixed effects. In this appendix, we show that our access measures are robust to these choices.

In Section 5.1.3 of our main paper, we estimate the consumption location choice probability using the following equation:

$$\lambda^S_{j(k)|ni} = \frac{\xi^S_{j(k)} \left(\tilde{d}^S_{nij(k)}\right)^{-\phi^S_k} \exp\left(u^S_{nij(k)}\right)}{\zeta^S_{ni,k}}, \tag{D.7}$$

and the consumption access measures can be constructed by:

$$\mathbb{S}_{nik} = \Gamma\left(\frac{\theta^S_k/\alpha^S_k - 1}{\theta^S_k/\alpha^S_k}\right)\left[\sum_{\ell \in N} \xi^S_{\ell(k)} \left(\tilde{d}^S_{ni\ell(k)}\right)^{-\phi^S_k}\right]^{\frac{\alpha^S_k}{\theta^S_k}}. \tag{D.8}$$

An alternative theory-consistent way of estimating the consumption access is to use residence-and-workplace fixed effects:

$$\mathbb{S}_{nik} = \left[\Gamma\left(\frac{\theta^S_k/\alpha^S_k - 1}{\theta^S_k/\alpha^S_k}\right)\right]\left(\zeta^S_{ni,k}\right)^{\frac{\alpha^S_k}{\theta^S_k}}. \tag{D.9}$$

Similarly, in Section 5.1.4 of our main paper, we estimate the workplace choice probability using the following equation:

$$\lambda^W_{i|n} = \frac{\xi^W_i \exp\left(-\phi^W \tau^W_{ni}\right) \left(\mathbb{S}_{ni}\right)^{\theta^W} \exp\left(u^W_{ni}\right)}{\zeta^W_n}, \tag{D.10}$$

and the travel access measures can be constructed by:

$$\mathbb{A}_n = \Gamma\left(\frac{\theta^W - 1}{\theta^W}\right)\left[\sum_{\ell \in N} \xi^W_\ell \exp\left(-\phi^W \tau^W_{n\ell}\right)\left(\mathbb{S}_{n\ell}\right)^{\theta^W}\right]^{\frac{1}{\theta^W}}. \tag{D.11}$$

An alternative theory-consistent way of estimating the consumption access is to use residence fixed effects:

$$\mathbb{A}_n = \Gamma\left(\frac{\theta^W - 1}{\theta^W}\right)\left(\zeta^W_n\right)^{\frac{1}{\theta^W}}. \tag{D.12}$$

In Table D.3.1, we present the results of the OLS regression of the origin fixed effects of commuting gravity equations (Column 1) and consumption gravity equations (Column 2-6) on the weighted sum of the destination fixed effects. We find that the regression slopes are essentially one, and the R-squared is also extremely close to one. Therefore, the two measures are essentially identical. In fact, Thibault (2015) show that these two measures are asymptotically equivalent. Even within a finite sample, we find that this approximation performs closely as the asymptotic theory predicts.

D.4 Estimation of Commuting and Consumption Location Choice without Trip Chains

In Section 5.1 of our main paper, we estimate our baseline model assuming that visits to consumption location may happen not only from home but also workplaces or on the way between home to work. In this section, we show how our estimation results compare if we instead assume that visits to consumption locations originate solely from home locations.

Table D.3.1: Robustness to Alternative Definitions of Consumption and Travel Access

	Dependent Variable: Origin fixed effects					
	Commuting	Finance realestate communication professional	Wholesale retail	Accomoda-tions eating drinking	Medical welfare healthcare	Other services
	(1)	(2)	(3)	(4)	(5)	(6)
Weighted sum of destination fixed effects	0.985***	1.000***	1.000***	1.000***	1.000***	1.000***
	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	242	12,322	12,330	12,327	12,310	12,329
R^2	0.997	1.000	1.000	1.000	1.000	1.000

Note: OLS regression of the origin fixed effects of commuting gravity equations (Column 1) and consumption gravity equations (Column 2-6) on the weighted sum of the destination fixed effects. Heteroskedasticity robust standard errors in parentheses.

More specifically, we consider a special case of our model where workers receive an infinitely negative preference shocks for the consumption route of ($r(k) = WW$), such that $T^R_{r(k)} = -\infty$ for $r(k) = WW, HW, WH$. In this special case, the expected travel cost for consumption trips (equation 8 of our main paper) simplifies to

$$d^S_{nij(k)} = \mathbb{E}_{nij(k)}\left[\delta_{nij(k)r(k)}(\omega)\right] = \vartheta^R_k \left[T^R_{HH}\right]^{\frac{1}{\theta^R_k}} \exp(-\kappa^S_k \tau^S_{nij(k)HH}) \tag{D.13}$$

where $\vartheta^R_k \equiv \Gamma\left(\frac{\theta^R_k - 1}{\theta^R_k}\right)$ and $\Gamma(\cdot)$ is the Gamma function, and $\tau^S_{nij(k)HH} = \tau_{nj} + \tau_{jn}$ is the travel time for visiting consumption location j from home location n. Therefore, our estimating equation for consumption location choice (equation 35 in our main paper) simplifies to:

$$\lambda^S_{j(k)|ni} = \frac{\xi^S_{j(k)} \exp\left(-\phi^S_k \tau^S_{nij(k)HH}\right) \exp\left(u^S_{nij(k)}\right)}{\zeta^S_{ni,k}}. \tag{D.14}$$

Therefore, the relevant consumption travel cost is simply a function of the bilateral travel time from home to consumption location. Furthermore, because the travel time for consumption trips $\tau^S_{nij(k)HH} = \tau_{nj} + \tau_{jn}$ does not depend on workplace i, the consumption location choice $\lambda^S_{j(k)|ni}$ does not depend on workplace i. Therefore, we can estimate the consumption location choice at the bilateral pair of home and consumption locations.

Furthermore, because the consumption location choice is independent of work location, the consumption access $\mathbb{S}_{ni}$ does not depend on workplace i. Therefore, this terms is now irrelevant for the commuting choice, and the commuting choice (equation 39 of our main paper) comes down to

$$\lambda^W_{i|n} = \frac{\xi^W_i \exp\left(-\phi^W \tau^W_{ni}\right) \exp\left(u^W_{ni}\right)}{\zeta^W_n}, \tag{D.15}$$

where the difference from equation (39) is again the omission of the consumption access term $(\mathbb{S}_{ni})^{\theta^W}$.

Table D.4.1 presents our estimation results of consumption location choice when we assume that all consumption trips originate from home ($T^R_{r(k)} = -\infty$ for $r(k) = WW, HW, WH$). This table corresponds to Table 3 of our main paper where we allow for trip chains for consumption trips. Because we can now estimate these equations at the bilateral location pairs of home location and consumption location (because work location is irrelevant to this choice), we report the results of the estimation at the unit of the bilateral location pairs of home location and consumption location in Panel (A), and the results of the estimation where we define the unit of the observations at the triplets of home, work and consumption locations in Panel (B). (The estimated coefficients are similar between the two panels, despite the large differences in sample size). We find that the travel time coefficient for consumption trips somewhat smaller in magnitude when we restrict consumption trips are only generated from home (ranging from -0.8 to -0.6, instead of -1.2 to -1.0 in Table 3). This is suggestive of the measurement error of travel time by assuming all trips originate from home location. Furthermore, we find a better model fit with the consumption gravity equation with route choice than this alternative specification, as evident from the smaller Akaike Information Criteria (AIC) or the Bayesian Information Criteria (BIC). This is the expected pattern of estimates if in reality consumption trips can originate from either work or home. Workers may frequently consume non-traded services that are close to work but far from home, precisely because they can easily access these non-traded services from work. In the model that falsely assumes that all consumption trips originate from home, the way the model tries to rationalize these consumption trips far from home is with artificially low semi-elasticities with respect to travel time.

Table D.4.1: Estimation of Consumption Location Choice without Trip Chains

(A) Estimate at the pair of home and consumption location

Dependent Variable:	Consumption Location Choice Probability				
	Finance realestate communication professional	Wholesale retail	Accomodations eating drinking	Medical welfare healthcare	Other services
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Consumption Travel Time (Hours)	-0.758***	-0.697***	-0.693***	-0.780***	-0.656***
	(0.040)	(0.038)	(0.039)	(0.041)	(0.037)
Fixed-effects					
Home Location	Yes	Yes	Yes	Yes	Yes
Consumption Location	Yes	Yes	Yes	Yes	Yes
Fit statistics					
AIC	1,374.5	1,805.7	1,409.3	1,412.7	1,527.3
BIC	5,719.8	6,151.0	5,754.6	5,758.0	5,872.6
Observations	58,564	58,564	58,564	58,564	58,564

(B) Estimate at the triplets of home, work and consumption location

Dependent Variable:	Consumption Location Choice Probability				
	Finance realestate communication professional	Wholesale retail	Accomodations eating drinking	Medical welfare healthcare	Other services
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Consumption Travel Time (Hours)	-0.690***	-0.699***	-0.698***	-0.737***	-0.665***
	(0.026)	(0.023)	(0.025)	(0.024)	(0.022)
Fixed-effects					
Home and Work Location	Yes	Yes	Yes	Yes	Yes
Consumption Location	Yes	Yes	Yes	Yes	Yes
Fit statistics					
AIC	140,263.4	141,209.4	141,197.0	139,977.4	142,431.2
BIC	302,440.5	303,498.0	303,443.7	301,987.4	304,705.8
Observations	2,981,924	2,983,860	2,983,134	2,979,020	2,983,618

Note: A version of Table 3 of our main paper where we assume that all consumption trips originate from home ($T^R_{r(k)} = -\infty$ for $r(k) = WW, HW, WH$). Results of the estimation of regression (D.14) by the Poisson Pseudo Maximum Likelihood (PPML) estimator. In Panel (A), we estimate the model with the observations at all bilateral pairs of municipalities in the Tokyo metropolitan area (residence n and consumption location $j(k)$). In Panel (B), we estimate the model with the triplets of municipalities in the Tokyo metropolitan area (residence n, workplace i, and consumption location $j(k)$). See the footnote of Table 3 for other comments.

Table D.4.2 presents our estimation results of commuting choices when we assume that all consumption trips originate from home ($T^R_{r(k)} = -\infty$ for $r(k) = WW, HW, WH$). This table corresponds to Table 4 of our main paper where we allow for trip chains for consumption trips. We find that the travel time coefficient for commuting trips somewhat larger when we restrict that all consumption trips are generated from home (-0.649 instead of -0.617 in Table 4). Furthermore, we find a better model fit with the consumption gravity equation with route choice than this alternative specification, as evident from the smaller Akaike Information Criteria (AIC) or the Bayesian Information Criteria (BIC). This pattern of results is consistent with the idea that workers willingness to commute longer bilateral distance may reflect not only higher wages or other characteristics of their workplace itself but also the greater access to consumption possibilities that this workplace provides. As for other large metropolitan areas such as London and New York, the downtown area of Tokyo to which workers commute long distances on average provide dense access to bars, restaurants and other non-traded consumption services.

Table D.4.2: Estimation of Commuting Choice on Residence without Trip Chains

Dependent Variable:	Commuting Choice Probability
Model:	(1)
Variables	
Commuting Time (Hours)	-0.649***
	(0.034)
Fixed-effects	
Home Location	Yes
dest_cityid	Yes
Fit statistics	
AIC	3,097.1
BIC	7,442.4
Observations	58,564

Note: A version of Table D.15 of our main paper where we assume that all consumption trips originate from home ($T^R_{r(k)} = -\infty$ for $r(k) = WW, HW, WH$). Results of the estimation of regression (D.15) by the Poisson Pseudo Maximum Likelihood (PPML) estimator. Observations are all bilateral pairs of municipalities in the Tokyo metropolitan area (residence n and work location i). See the footnote of Table 4 for other comments.

Lastly, we conduct the same decomposition exercise of the observed concentration of the attractiveness of the summary measure of the relative attractiveness of locations for residence. We follow the same procedure in Section 5.2 and present our regression-based decomposition with the following equations:

$$\ln \mathbb{A}_n = c_0^A + c_1^A \ln\left(\left(\lambda_n^A\right)^{1/\theta^B} Q_n^{\alpha^H}\right) + u_{nt}^A, \tag{D.16}$$
$$\ln \mathbb{B}_n = c_0^B + c_1^B \ln\left(\left(\lambda_n^B\right)^{1/\theta^B} Q_n^{\alpha^H}\right) + u_{nt}^B.$$

Table D.4.3 presents our results. Panel (A) presents the results of the regressions (45) from our baseline model, and Panel (B) presents the same regression results from our special case of omitting consumption trips ($\alpha_k^S = 0$ for all $k \in K^S$, $\alpha^T = 1 - \alpha^H$, $\lambda^S_{j(k)|ni} = 0$ and $\mathbb{S}_{ni} = 1$). Compared to the results which incorporate the possibility of trip chains, we find a greater coefficients for both panels compared to the case which include trip chains (0.56 and 0.35, respectively; Table 5). This difference primarily arise from the difference of the estimate of the travel cost elasticity (ϕ^W). As discussed above, ϕ^W is overestimated by omitting trip chains (Table D.4.2). A greater ϕ^W tend to give a greater variation of travel access $\mathbb{A}_n$. Despite this difference, our main conclusion stays the same: As evident from the comparison between Panel (A) and (B), omitting the consumption trips tend to underestimate the contribution of travel access ($\mathbb{A}_n$) and overestimate the contribution of $\mathbb{B}_n$ for the revealed attractiveness of residential location.[6]

Table D.4.3: Variance Decomposition of the Relative Attractiveness of Locations without Trip Chains

	$\log \mathbb{A}_n$	$\log \mathbb{B}_n$
	(1)	(2)
Panel A: Baseline Model		
$\log Q_n^{\alpha^H} \left(\lambda_n^B\right)^{1/\theta^B}$	0.845***	0.155**
	(0.067)	(0.067)
Observations	201	201
R^2	0.441	0.026
Panel B: No Consumption Trips		
$\log Q_n^{\alpha^H} \left(\lambda_n^B\right)^{1/\theta^B}$	0.479***	0.521***
	(0.038)	(0.038)
Observations	201	201
R^2	0.450	0.491

Note: A version of Table 5 of our main paper where we assume that all consumption trips originate from home ($T^R_{r(k)} = -\infty$ for $r(k) = WW, HW, WH$). Panel (A): Ordinary least squares (OLS) estimates of the regression-based variance decomposition in equation (45) when we commute travel access $\mathbb{A}_n$ with consumption trips; Panel (B) presents the results when we omit consumption access to construct $\mathbb{A}_n$ ($\alpha_k^S = 0$ for all $k \in K^S$, $\alpha^T = 1 - \alpha^H$, $\lambda^S_{j(k)|ni} = 0$ and $\mathbb{S}_{ni} = 1$). Note that Panel (B) is equivalent to a version of Panel (B) of Table 5 using the estimates of ϕ^W and ξ_ℓ^W from a conventional commuting gravity equation excluding consumption access. Observations are municipalities in the Tokyo metropolitan area. Heteroskedasticity robust standard errors in parentheses.

[6] Panel (B) of Table D.4.3 also corresponds to the exercise where we omit consumption trips and construct $\mathbb{A}_n^{nocons}$ using the estimates the commuting gravity equations by omitting consumption access terms $\log\left(\mathbb{S}_{n\ell}\right)^{\theta^W}$, as opposed to Panel (B) of Table 5 of our main paper where we construct $\mathbb{A}_n^{nocons}$ using the estimates of gravity equations including the term of $\log\left(\mathbb{S}_{n\ell}\right)^{\theta^W}$.

E Details of Calibration and Simulation Procedure

This section of the online appendix explains further details of the counterfactual simulation procedure of Section 6.1 of our main paper. In Section E.1, we report the derivation of the system of equations that we use to solve for a counterfactual equilibrium. In Section E.2, we provide further details on the calibration of the baseline variables in the initial equilibrium that are used in this system of equations to solve for a counterfactual equilibrium.

In Section F, we present additional empirical results for our counterfactuals for changes in travel costs in the Tokyo Metropolitan Area, as discussed in 6.2 of the paper. In Section G, we report further empirical results for our counterfactuals for the opening of the Tozai subway line in the city of Sendai, as discussed in Section 6.3 of the paper.

E.1 Mathematical Details for the System of Equations for Counterfactual Simulation

In this section of the online appendix, we derive the system of equations that we use to solve for a counterfactual equilibrium as discussed in Section 6.1 of our main paper. In our baseline specification, we consider the closed-city specification of the model, in which total population for the city as a whole ($\bar{L}$) is exogenous, and hence the change in travel costs affects worker welfare. We denote the value of a variable in the initial equilibrium by x, the value of this variable in the counterfactual equilibrium by x' (with a prime), the relative change in this variable by $\hat{x} = x'/x$ (with a hat). Given values for the model parameters (α^H, α^T, $\{\alpha_k^S\}$, $\{\theta_k^S\}$, θ^W, θ^B, κ^W, κ^S, η^B, η^W, β^S, β^T, μ), assumed bilateral changes in travel cost $\{\hat{d}_{ni}^W, \hat{d}_{nij(k)}^S\}$, and observed values for the endogenous variables in the initial equilibrium ($\{\lambda_{ig|n}^W, \lambda_{j(k)|ni}^S, \lambda_n^B\}$, $\{H_{j,k}, H_{n,U}\}$, $\{E_{ni}\}$), we solve for the counterfactual equilibrium by solving the following system of equations for the general equilibrium of the model.

(i) Changes in Commuting and consumption probabilities From equations (10) and (14), the counterfactual changes in conditional commuting probabilities ($\hat{\lambda}_{ig|n}^W$) and conditional consumption probabilities ($\hat{\lambda}_{jk|n}^S$) satisfy:

$$
\begin{aligned}
\hat{\lambda}_{ig|n}^W &= \frac{T_{i,g}^W w_{i,g}^{\prime\theta^W} \left(d_{ni}^{W\prime}\right)^{\theta^W} \left(\mathbb{S}_{ni}^{\prime}\right)^{\theta^W}}{T_{i,g}^W w_{i,g}^{\theta^W} \left(d_{ni}^W\right)^{\theta^W} \left(\mathbb{S}_{ni}\right)^{\theta^W}} \frac{\sum_{\ell \in N} \sum_{m \in K} T_{\ell,m}^W w_{\ell,m}^{\theta^W} \left(d_{n\ell}^W\right)^{\theta^W} \left(\mathbb{S}_{n\ell}\right)^{\theta^W}}{\sum_{\ell \in N} \sum_{m \in K} T_{\ell,m}^{W\prime} w_{\ell,m}^{\prime\theta^W} \left(d_{n\ell}^{W\prime}\right)^{\theta^W} \left(\mathbb{S}_{n\ell}^{\prime}\right)^{\theta^W}} \\
&= \frac{\hat{w}_{i,g}^{\theta^W} \left(\hat{d}_{ni}^W\right)^{\theta^W} \hat{\mathbb{S}}_{ni}^{\theta_W}}{\sum_{\ell \in N} \sum_{m \in K} \hat{w}_{\ell,m}^{\theta^W} \left(\hat{d}_{n\ell}^W\right)^{\theta^W} \hat{\mathbb{S}}_{n\ell}^{\theta_W} \lambda_{\ell m|n}^W},
\end{aligned}
\tag{E.1}
$$

$$
\begin{aligned}
\hat{\lambda}_{j(k)|ni}^S &= \frac{T_{j(k)}^S \left(P_{j(k)}^S\right)^{-\theta_k^S} \left(d_{nij(k)}^S\right)^{\frac{\theta_k^S}{\alpha_k^S}}}{T_{j(k)}^S \left(P_{j(k)}^{S\prime}\right)^{-\theta_k^S} \left(d_{nij(k)}^{S\prime}\right)^{\frac{\theta_k^S}{\alpha_k^S}}} \frac{\sum_{\ell \in N} T_{\ell(k)}^S \left(P_{\ell(k)}^S\right)^{-\theta_k^S} \left(d_{ni\ell(k)}^S\right)^{\frac{\theta_k^S}{\alpha_k^S}}}{\sum_{\ell \in N} T_{\ell(k)}^{S\prime} \left(P_{\ell(k)}^{S\prime}\right)^{-\theta_k^S} \left(d_{ni\ell(k)}^{S\prime}\right)^{\frac{\theta_k^S}{\alpha_k^S}}} \\
&= \frac{\left(\hat{P}_{j(k)}^S\right)^{-\theta_k^S} \left(\hat{d}_{nij(k)}^S\right)^{\theta_k^S}}{\sum_{\ell \in N} \left(\hat{P}_{\ell(k)}^S\right)^{-\theta_k^S} \left(\hat{d}_{ni\ell(k)}^S\right)^{\theta_k^S} \lambda_{\ell(k)|ni}^S}.
\end{aligned}
\tag{E.2}
$$

Using equations (10), (12), (14) and (16), the corresponding changes in travel access ($\hat{\mathbb{A}}_n$) and consumption access ($\hat{\mathbb{S}}_n$) can be written in terms of own commuting shares ($\hat{\lambda}^W_{nT|n}$) and own consumption shares ($\hat{\lambda}^S_{n(k)|ni}$):

$$\begin{aligned}\hat{\mathbb{A}}_n &= \left[\frac{\sum_{\ell\in N}\sum_{m\in K} T^{W'}_{\ell,m} w'^{\theta^W}_{\ell,m}\left(d^{W'}_{n\ell}\right)^{\theta^W}\left(\mathbb{S}'_{n\ell}\right)^{\theta^W}}{\sum_{\ell\in N}\sum_{m\in K} T^{W}_{\ell,m} w^{\theta^W}_{\ell,m}\left(d^{W}_{n\ell}\right)^{\theta^W}\left(\mathbb{S}_{n\ell}\right)^{\theta^W}}\right]^{\frac{1}{\theta^W}} \\ &= \left[\hat{w}^{\theta^W}_{n,g}\left(\hat{d}^W_{nn}\right)^{\theta^W}\hat{\mathbb{S}}^{\theta_W}_{nn}/\hat{\lambda}^W_{ng|n}\right]^{\frac{1}{\theta^W}}, \end{aligned} \tag{E.3}$$

for some $g \in K$, and

$$\begin{aligned}\hat{\mathbb{S}}_{ni} &= \prod_k \left[\frac{\sum_{\ell\in N} T^{S'}_{\ell(k)}\left(P^{S'}_{\ell(k)}\right)^{-\theta^S_k}\left(d^{S'}_{ni\ell(k)}\right)^{\frac{\theta^S_k}{\alpha^S_k}}}{\sum_{\ell\in N} T^{S}_{\ell(k)}\left(P^{S}_{\ell(k)}\right)^{-\theta^S_k}\left(d^{S}_{ni\ell(k)}\right)^{\frac{\theta^S_k}{\alpha^S_k}}}\right]^{\frac{\alpha^S_k}{\theta^S_k}} \\ &= \prod_k \left[\left(\hat{P}^S_{n,k}\right)^{-\theta^S_k}\left(\hat{d}^S_{nin(k)}\right)^{\theta^S_k}/\hat{\lambda}^S_{n(k)|ni}\right]^{\frac{\alpha^S_k}{\theta^S_k}}. \end{aligned} \tag{E.4}$$

(ii) Changes in residential location decision From equations (21) and (17), the counterfactual changes in residential probabilities ($\hat{\lambda}^B_n$) satisfy:

$$\begin{aligned}\hat{\lambda}^B_n &= \frac{T^B_n B'^{\theta^B}_n \mathbb{A}'^{\theta^B}_n \left(P^{T'}_n\right)^{-\alpha^T\theta^B} Q'^{-\alpha^H\theta^B}_n}{T^B_n B^{\theta^B}_n \mathbb{A}^{\theta^B}_n \left(P^{T}_n\right)^{-\alpha^T\theta^B} Q^{-\alpha^H\theta^B}_n} \frac{\sum_{\ell\in N} T^B_\ell B^{\theta^B}_\ell \mathbb{A}^{\theta^B}_\ell \left(P^{T}_\ell\right)^{-\alpha^T\theta^B} Q^{-\alpha^H\theta^B}_\ell}{\sum_{\ell\in N} T^{B'}_\ell B^{\theta^B}_\ell \mathbb{A}'^{\theta^B}_\ell \left(P^{T'}_\ell\right)^{-\alpha^T\theta^B} Q'^{-\alpha^H\theta^B}_\ell} \\ &= \frac{\hat{\mathbb{A}}^{\theta^B}_n \hat{Q}^{-\alpha^H\theta^B}_n \hat{B}^{\theta^B}_n}{\sum_{\ell\in N}\hat{\mathbb{A}}^{\theta^B}_\ell \hat{Q}^{-\alpha^H\theta^B}_\ell \hat{B}^{\theta^B}_n \lambda^B_\ell}. \end{aligned} \tag{E.5}$$

(iii) Changes in commercial and residential floor space demand From equation (22), the changes in commercial floor space in each sector ($\hat{H}_{i,g}$) are given by:

$$\hat{H}_{i,g} = \frac{\hat{w}_{i,g}\hat{\tilde{L}}_{i,g}}{\hat{Q}_i}, \tag{E.6}$$

where the change in labor input adjusted for effective units of labor ($\hat{\tilde{L}}_{i,g}$) can be derived from equation (B.6) as:

$$\hat{\tilde{L}}_{i,g} = \frac{1}{\hat{w}_{i,g}}\frac{\sum_{n\in N} E'_{ni}\lambda'^W_{ig|n}\lambda^{B'}_n}{\sum_{n\in N} E_{ni}\lambda^W_{ig|n}\lambda^B_n}. \tag{E.7}$$

From equation (19), the changes in residential floor space ($\hat{H}_{i,U}$) satisfy:

$$\hat{H}_{i,U} = \frac{\hat{E}_i\hat{\lambda}^B_i}{\hat{Q}_i}, \tag{E.8}$$

where the counterfactual residential income E'_n is given by equation (20):

$$E'_n = \sum_{i\in N} E'_{ni}\lambda'^W_{i|n}, \tag{E.9}$$

and the changes of commuting-pair specific income $\hat{E}_{ni}$ is derived from equation (B.17):

$$\hat{E}_{ni} = \frac{\hat{\mathbb{A}}_n}{\hat{\mathbb{S}}_{ni}}. \tag{E.10}$$

(iv) Changes in the price of floor space From equation (31), the change in the price of floor space ($\hat{Q}_i$) and the overall quantity of floor space ($\hat{H}_i$) are related as follows:

$$\hat{Q}_i = \hat{H}_i^{\frac{1-\mu}{\mu}}, \tag{E.11}$$

where the change in this overall quantity of floor space ($\hat{H}_i$) is a weighted average of the changes in the quantities of commercial floor space in each sector ($\hat{H}_{i,k}$) and the quantity of residential floor space ($\hat{H}_{i,U}$):

$$\hat{H}_i = \frac{H_{i,U}\hat{H}_{i,U} + \sum_{k\in K} H_{i,k}\hat{H}_{i,k}}{H_{i,U} + \sum_{k\in K} H_{i,k}}. \tag{E.12}$$

(v) Changes in endogenous productivities and amenities From equations (27) and (28), the changes in endogenous productivities ($\hat{A}_{i,k}$) and amenities ($\hat{B}_n$) as a result of agglomeration forces satisfy:

$$\hat{A}_{i,k} = \hat{L}_i^{\eta^W}, \tag{E.13}$$

$$\hat{B}_n = \hat{R}_n^{\eta^B}. \tag{E.14}$$

(vi) Changes in nontraded goods prices From equation (24), the changes in non-traded goods prices ($\hat{P}_{j,k}$) satisfy:

$$\hat{P}^S_{j,k} = \frac{1}{\hat{A}_{j,k}\hat{\hat{L}}_{j,k}^{\beta^S}\hat{H}_{j,k}^{1-\beta^S}} \frac{\sum_{n,i\in N} E'_{ni}\lambda'^S_{j(k)|ni}\lambda'^W_{i|n}\lambda'^B_n}{\sum_{n,i\in N} E_{ni}\lambda^S_{j(k)|ni}\lambda^W_{i|n}\lambda^B_n}. \tag{E.15}$$

(viii) Changes in Wages From the zero-profit condition (21), the changes in wages in each sector and location with positive production ($\hat{w}_{i,k}$) are given by:

$$\hat{w}_{i,k} = \left(\frac{\hat{A}_{i,k}\hat{P}^S_{i(k)}}{\hat{Q}_i^{1-\beta^S}}\right)^{1/\beta^S}. \tag{E.16}$$

We solve this system of equations (47)-(62), starting with an initial guess of the relative change in each endogenous variable ($\hat{x} = 1$), and updating this initial guess until the solution to this system converges to equilibrium. Using the resulting counterfactual changes in the endogenous variables of the model ($\hat{\lambda}^W_{ig|n}$, $\hat{\lambda}^S_{j(k)|ni}$, $\hat{\mathbb{A}}_n$, $\hat{\mathbb{S}}_n$, $\hat{\lambda}^B_n$, $\hat{H}_{i,g}$, $\hat{H}_{i,U}$, $\hat{\hat{L}}_{i,g}$, $\hat{Q}_i$, $\hat{A}_{i,k}$, $\hat{B}_n$, $\hat{P}^S_{i(k)}$, $\hat{w}_{i,k}$), together with equation (18), we can compute the implied change in expected utility ($\widehat{\mathbb{E}[u]}$) induced by the change in travel costs as follows:

$$\begin{aligned}\widehat{\mathbb{E}[u]} &= \left[\frac{\sum_{\ell\in N} T^B_\ell B'^{\theta^B}_\ell \mathbb{A}'^{\theta^B}_\ell \left(P^{T\prime}_\ell\right)^{-\alpha^T\theta^B} Q'^{-\alpha^H\theta^B}_\ell}{\sum_{\ell\in N} T^B_\ell B^{\theta^B}_\ell \mathbb{A}^{\theta^B}_\ell \left(P^T_\ell\right)^{-\alpha^T\theta^B} Q^{-\alpha^H\theta^B}_\ell}\right]^{\frac{1}{\theta^B}} \\ &= \left[\sum_{\ell\in N} \lambda^B_\ell \hat{B}^{\theta^B}_\ell \hat{\mathbb{A}}^{\theta^B}_\ell \hat{Q}^{-\alpha^H\theta^B}_\ell\right]^{\frac{1}{\theta^B}},\end{aligned} \tag{E.17}$$

where we have used our choice of numeraire ($P^T_\ell = 1$ for all $\ell \in N$).

E.2 Calibration of Baseline Variables

In this section, we discuss the calibration of the baseline variables in the initial equilibrium used in the system of equations from the previous section to solve for a counterfactual equilibrium. In general, the system of equations for a counterfactual equilibrium can be solved using either the observed initial travel shares (as in the conventional exact-hat algebra approach of Dekle, Eaton, and Kortum 2007) or using the initial travel shares predicted by the estimated model (as in the covariate-based approach of Dingel and Tintelnot 2020). In our baseline specification, we use this covariate-based approach to address concerns about granularity and the resulting potential for overfitting using the exact-hat algebra approach. In later sections of this online appendix, we report robustness tests, in which we use the observed initial travel shares following the conventional exact-hat algebra approach. In our empirical application, we find a relatively similar pattern of results using both approaches.

In addition to the parameters as presented in Table 1, we use a set of baseline variables ($\{\lambda^W_{ig|n}, \lambda^S_{j(k)|ni}, \lambda^B_n\}$, $\{\tilde{L}_{i,k}\}$, $\{H_{j,k}, H_{n,U}\}$, $\{E_{ni}\}$) and the changes of travel cost $\{\hat{d}^W_{ni}, \hat{d}^S_{nij(k)}\}$ for this counterfactual simulation. Below, we discuss how we construct these baseline variables using our smartphone data and separate data sources (see Section 2.2 of our main paper for these other data sources).

Spatial Units. We first need to take a stand on the spatial units at which we conduct counterfactual simulations. In the counterfactual of reducing travel cost in Tokyo Metropolitan Area (Section 6.2), the spatial unit is specified as 242 municipalities, which together form the Tokyo Metropolitan Area. In the counterfactual of the subway opening in Sendai City (Section 6.3), the spatial unit is specified as 370 Oazas, which together form the municipality of Sendai City.

Calibration of commuting probability $\{\lambda^W_{ig|n}\}$. We construct $\lambda^W_{ig|n}$ using our smartphone data and separate economic census data. From our smartphone data and its assignment of "home" and "work" locations for each device, we define the bilateral commuting flows $\lambda^W_{i|n}$ at specified bilateral spatial units. Our model assumes that the probability that each worker works in a certain industry ($\lambda^W_{ig|n}/\sum_{k\in K}\lambda^W_{ik|n}$) is independent of the worker's residential location n. Therefore, we use the separate economic census data to construct the employment share of each sector ($L_{i,g}/\sum_{k\in K} L_{i,k}$), and define the commuting probability for each bilateral location and sector by $\lambda^W_{ig|n} = \lambda^W_{ig|n} \times L_{i,g}/\sum_{k\in K} L_{i,k}$.

In measuring the commuting shares in the initial equilibrium, one potential concern is granularity (the difference between realized and expected values because of a small integer number of commuters on some bilateral routes). In particular, Dingel and Tintelnot (2020) show that counterfactuals using the observed initial commuting shares can be potentially biased, because of the model overfitting from using the observed commuting flows to calibrate the model. Instead, they recommend "covariate-based approach," in which one calibrates the commuting flows by the predicted flows using covariates such as travel time. Following their recommendations, we use the predicted commuting flows from our commuting choice probability estimation (equation (39)) to construct the initial commuting shares in the baseline specification for our counterfactuals. As a robustness test, we also report the results of counterfactuals based on the conventional exact-hat algebra approach using the observed initial commuting shares.

Calibration of consumption probability $\{\lambda^S_{j(k)|ni}\}$. We construct $\lambda^S_{j(k)|ni}$ using our smartphone data and separate economic census data. Using our smartphone data, we assign "Other" stays to each nontradable service sector as discussed in Section 3 of our main paper. Using this assignment, we define the travel for consumption trips for each location and sector given residential and work locations. In our baseline specification, we again follow the 'covariate-based approach," in which we use the predicted consumption flows from our consumption choice probability estimation (equation (35) to construct the initial travel shares in the baseline specification for our counterfactuals. As a robustness test, we also report the results of counterfactuals based on the conventional exact-hat algebra approach using the observed initial travel shares.

Calibration of residential share $\{\lambda_n^B\}$. We construct λ_n^B from our smartphone data using the devices' "Home" locations.

Calibration of floor space $\{H_{j,k}, H_{n,U}\}$. We construct $\{H_{j,k}, H_{n,U}\}$ using our building use data and economic census data. From the building data, we construct the floor space in each location separately for residential purposes ($H_{n,U}$) and commercial purposes ($\sum_{k \in K} H_{j,k}$). We allocate commercial floor space into each sector k proportionally to the employment share of the sector in each locations using economic census data.

Calibration of income $\{E_{ni}\}$. We construct $\{E_{ni}\}$ using equation (B.17) such that $E_{ni} = \mathbb{A}_n / \mathbb{S}_{ni}$, where $\mathbb{A}_n$ and $\mathbb{S}_{ni}$ are estimated following the procedures in Section 5.1 of our main paper.

Calibration of changes in travel cost $\{\hat{d}_{ni}^W, \hat{d}_{nij(k)}^S\}$. We construct the travel cost change $\{\hat{d}_{ni}^W, \hat{d}_{nij(k)}^S\}$ from the specified parameter changes of κ^W and κ_k^S (for the counterfactual in Tokyo Metropolitan Area in Section 6.2) and the travel time change (for the counterfactual in Sendai City in Section 6.3). We construct $\{\hat{d}_{ni}^W\}$ from equation (14) such that

$$\hat{d}_{ni}^W \equiv \exp(-\kappa^W \Delta \tau_{ni}^W) \tag{E.18}$$

where Δ indicates the difference operation. We construct $\{\hat{d}_{nij(k)}^S\}$ from our definition of d_{ni}^W in equation (34) such that

$$\hat{d}_{nij(k)}^S = \vartheta_k^R \left[\sum_{r' \in \mathbb{R}} \xi_{r'(k)}^R \exp(-\phi_k^R \Delta \tau_{nij(k)r'(k)}^S) \right]^{\frac{\alpha_k^S}{\theta_k^R}} . \tag{E.19}$$

F Additional Results on Counterfactual Simulations in Tokyo Metropolitan Area

In this section of the appendix, we present additional results of the counterfactual simulation to decrease travel cost within Tokyo Metropolitan Area in Section 6.2 of our main paper. In particular, we show the sensitivity of our results under different values of the elasticity of residential amenity spillover η^B and that of productivity spillover η^W. Overall, regardless of the choice of η^B and η^W, we find that a robust pattern that consumption access is quantitatively important for the spatial concentration of economic activity in urban areas relative to the workplace access.

We conduct three counterfactuals in section 6.2. In our first set of counterfactuals, we provide further evidence on the role of travel costs for commuting and consumption in shaping the spatial concentration of economic activity, by shutting down each of these sources of spatial frictions. In a first exercise, we halve travel costs for commuting trips ($\kappa^{W'} = 0.5 \cdot \kappa^{W'}$), maintain travel costs for consumption trips equal to their estimated value in the data ($\kappa_k^S > 0$), and solve for the counterfactual equilibrium distribution of economic activity. In a second exercise, we halve travel costs for consumption trips ($\kappa_k^{S'} = 0.5 \cdot \kappa_k^S$), maintain travel costs for commuting trips equal to their estimated value in the data ($\kappa^W > 0$), and solve for the counterfactual equilibrium. Finally, in a third exercise, we halve travel costs for both commuting and consumption trips ($\kappa^{W'} = 0.5 \cdot \kappa^{W'}$ and $\kappa_k^{S'} = 0.5 \cdot \kappa_k^S$), and solve for the counterfactual equilibrium.

Table F.0.1 presents the results of our three counterfactuals (reduce consumption travel cost, reduce commuting travel cost, and reduce both consumption and commuting travel cost). Each entry in the table corresponds to the regression slope of the counterfactual prediction of the employment by residence (Table (i)) and employment by workplace (Table (ii)) on those variables from actual data. Within each table, Panel A-C corresponds to the counterfactual simulations under different values of η^B and η^W, and each of the three columns indicates our three counterfactuals. If employment is unaffected by the change in travel costs, counterfactual employment is more decentralized than actual employment, because employment decreases in locations with higher actual employment, and increases in locations with lower actual employment.

By comparing the panels in Table (i) and (ii), we find that the counterfactual predictions are affected by the choice of η^B and η^W. In particular, when we increase η^B from 0 to 0.15, the counterfactual changes of regression slopes become larger (from Panel A to Panel B). Furthermore, we decrease the value of η^W from 0.08 to 0, the counterfactual changes become smaller (from Panel B to Panel C). Therefore, we find that both η^B and η^W amplify the changes in counterfactuals. Nonetheless, our main point in Section 6.2 holds under different choices of η^B and η^W: consumption access is quantitatively important for the spatial concentration of economic activity in urban areas relative to the workplace access (comparing Column 1 to 2).

Table F.0.1: Counterfactuals for Reducing Travel Costs for Commuting and Consumption Trips

(i) Employment by Residence

	Dependent variable: log(R)		
	Reduce travel costs for consumption trips	Reduce travel costs for commuting trips	Reduce travel costs for commuting and consumption trips
	(1)	(2)	(3)
Panel A: $\eta^B = 0$, $\eta^W = 0.08$			
log(R) (baseline)	0.943***	0.807***	0.778***
	(0.013)	(0.018)	(0.030)
Panel B: $\eta^B = 0.15$, $\eta^W = 0.08$			
log(R) (baseline)	0.771***	0.396***	0.209
	(0.069)	(0.117)	(0.263)
Panel C: $\eta^B = 0.15$, $\eta^W = 0$			
log(R) (baseline)	0.795***	0.481***	0.235
	(0.053)	(0.145)	(0.289)
Observations	242	242	242

(ii) Employment by Workplace

	Dependent variable: log(L)		
	Reduce travel costs for consumption trips	Reduce travel costs for commuting trips	Reduce travel costs for commuting and consumption trips
	(1)	(2)	(3)
Panel A: $\eta^B = 0$, $\eta^W = 0.08$			
log(L) (baseline)	0.976***	0.939***	0.919***
	(0.017)	(0.021)	(0.034)
Panel B: $\eta^B = 0.15$, $\eta^W = 0.08$			
log(L) (baseline)	0.921***	0.674***	0.783***
	(0.027)	(0.029)	(0.032)
Panel C: $\eta^B = 0.15$, $\eta^W = 0$			
log(L) (baseline)	0.948***	0.794***	0.855***
	(0.016)	(0.017)	(0.021)
Observations	242	242	242

Note: Each entry in the table corresponds to the regression slope of the counterfactual prediction of the employment by residence (Table (i)) and employment by workplace (Table (ii)) on those variables from actual data. Within each table, Panel A-C corresponds to the counterfactual simulations under different values of η^B and η^W, and each of the three columns indicates our three counterfactuals. Figure 10 of our main paper displays the graphical version of the relationships in Panel B ($\eta^B = 0.15; \eta^W = 0.08$).

G Additional Results on Sendai Subway Analysis

In this section of the appendix, we present additional results for our analysis of the opening of the new Tozai (East-West) Subway Line in the city of Sendai as discussed in Section 6.3 of the paper. In Subsection G.1, we discuss the extension of our baseline theoretical model to incorporate an endogenous transport mode choice between the railway and road transportation. In Subsection G.2, we discuss our estimation of travel access in the city of Sendai, following the same approach as developed in Section 5.1 of the paper.

In Subsection G.3, we report additional difference-in-differences estimates for the impact of the opening of the new Tozai (East-West) Subway Line in the city of Sendai, as discussed in Section 6.3 of the paper. In Subsection G.4, we present the results of our Placebo specification, in which we repeat our empirical analysis for the already-existing Nanbuko (North-South) Subway Line.

In Subsection G.5, we compare our baseline counterfactual results for the new Tozai (East-West) Subway Line using the covariate-based approach of Dingel and Tintelnot (2020) (to address concerns about granularity) to those using the conventional exact-hat algebra approach of Dekle, Eaton, and Kortum (2007). In our empirical application, we find a relatively similar pattern of results using both approaches.

G.1 Incorporating Transportation Mode Choice

In this subsection, we extend our baseline model (Section 4) to incorporate the mode choice between the railway and the road transportation. This extension is important for the correct assessment of subway opening, because the welfare estimates are greatly influenced by the intensity at which residents use railways as opposed to road traffic. In Section G.1.1, we set up our extended model and show that the extended model remains isomorphic to our main model by replacing the travel time with mode-adjusted travel time. In Section G.1.2, we estimate the mode choice and mode-adjusted travel time.

G.1.1 Model Extension to Incorporate Transportation Mode Choice

We introduce the mode choice as an additional choice of transportation made by each worker for each leg of travel (for each movement from one place to another). These decisions are made after workers decide their residence, workplace, the set of consumption locations, and the routes (as in our baseline model). After the worker chooses the set of routes for each consumption location (whether to visit her consumption location from home, from work, or in between), she observes an idiosyncratic preference shock for each transportation mode for each leg of her travel, and decides her optimal transportation mode for each leg of travel.

Utility. The indirect utility of the worker is given as follows:

$$\begin{aligned} U_{nig\{j(k)r(k)\}}(\omega) = & \left\{ B_n b_n(\omega) \left(P_n^T\right)^{-\alpha^T} Q_n^{-\alpha^H} \right\} \left\{ a_{i,g}(\omega) w_{i,g} \right\}, \\ & \times \left\{ \prod_{k \in K^S} \left[P_{j(k)}^S / \left(q_{j(k)}(\omega)\right) \right]^{-\alpha_k^S} \right\} \left\{ d_{ni\{j(k)r(k)\}}(\omega) \prod_{k \in K^S} \nu_{r(k)}(\omega) \right\} \\ & 0 < \alpha^T, \alpha^H, \alpha_k^S < 1, \qquad \alpha^T + \alpha^H + \sum_{k \in K^S} \alpha_k^S = 1, \end{aligned} \tag{G.1}$$

The only difference from our main specification (equation 1 in the main paper) is that the travel cost $d_{ni\{j(k)r(k)\}}(\omega)$ now depends on worker ω, which reflects the optimal mode choice specific to each individual. Similarly as in our baseline model, we specify that $d_{ni\{j(k)r(k)\}}(\omega)$ is decomposed into the component related to commuting and consumption trips. More specifically, we assume:

$$d_{ni\{j(k)r(k)\}}(\omega) = \exp(-\kappa^W \tau_{ni}^W(\omega)) \prod_{k \in K^S} \exp(-\kappa_k^S \tau_{nij(k)r(k)}^S(\omega)), \tag{G.2}$$

where $\tau_{ni}^W(\omega)$ is the travel cost required for commuting from residence n to workplace i without making any detour, and $\tau_{nij(k)r(k)}^S$ is the travel cost additionally required to visit consumption location $j(k)$ by the route $r(k)$ by deviating from the commuting path. Again, the only difference from the definition of travel cost from our baseline model (equation 2) is that $\tau_{ni}^W(\omega)$ and $\tau_{nij(k)r(k)}^S(\omega)$ now depend on worker ω through the optimal choice of transportation mode.

We proceed by specifying $\tau_{ni}^W(\omega)$. Noting that workers decide the transportation mode for each leg of travel (separately for traveling from home to work and from work to home), we have:[7]

$$\tau_{ni}^W(\omega) = \tau_{ni}(\omega, m_{ni}) + \tau_{in}(\omega, m_{in}),$$

where $\tau_{ni}(\omega, m_{ni})$ is the expected travel cost from n to i using mode $m_{ni} \in \mathbb{M}$ inclusive of the preference shock for each mode. Denoting these mode-specific preference shock by $\nu_{ni,m}^M$, we can express $\tau_{ni}(\omega, m)$ as:

$$\tau_{ni}(\omega, m) = \delta_{ni,m} + \log \nu_{ni,m}^M(\omega)$$

where $\delta_{ni,m}$ is the travel time from n to i using mode m, and $\nu_{ni,m}^M(\omega)$ is again the stochastic preference shock for mode m for the leg of travel from n to i.

Similarly, the consumption trip component of travel time ($\tau_{nij(k)r(k)}^S(\omega)$) in equation (G.2) is given by:

$$\begin{aligned}
\tau_{nij(k)HH}^S(\omega) &= \tau_{nj}(\omega, m_{nj}) + \tau_{jn}(\omega, m_{jn}), \\
\tau_{nij(k)WW}^S(\omega) &= \tau_{ij}(\omega, m_{ij}) + \tau_{ji}(\omega, m_{ji}), \\
\tau_{nij(k)HW}^S(\omega) &= \tau_{nj}(\omega, m_{nj}) + \tau_{ji}(\omega, m_{ji}) - \tau_{ni}(\omega, m_{ni}), \\
\tau_{nij(k)WH}^S(\omega) &= \tau_{ij}(\omega, m_{ij}) + \tau_{jn}(\omega, m_{ni}) - \tau_{in}(\omega, m_{ni}),
\end{aligned}$$

where each component is the the sum of the mode-specific travel time and the preference shock given by equation (G.1.1).

Similarly as the idiosyncratic shocks for residence choice, workplace choice, consumption location choice and the route choice, we assume that $\nu_{ni,m}^M(\omega)$ is drawn from the following independent Fréchet distributions:

$$G_m^M(\nu) = \exp\left(-T_m^M \nu^{-\theta^M}\right), \; T_m^M > 0, \; \theta^M > 1,$$

where the scale parameter $\{T_m^M\}$ control the average draws and the shape parameter θ^M regulates the dispersion of the shock.

[7] While we assume that workers can use different transportation modes for each direction of travel, it is straightforward to assume that the same transportation mode is used for both directions.

Mode choice. After making decisions about the residence, workplace, set of consumption locations, and the routes (as discussed in the main paper), and after observing the idiosyncratic shocks $\nu_{ni,m}^{M}(\omega)$, each worker chooses the optimal mode for each leg of travel (e.g., travel from home to work, work to consumption locations). The optimal choice of the mode for the leg of travel from n to i is given by:

$$m_{ni}(\omega) = \arg\max_{m \in \mathbb{M}} \left\{ \delta_{ni,m} + \log \nu_{ni,m}^{M}(\omega) \right\}.$$

where $\delta_{ni,m}$ is again the travel time from n to i using mode m. Using the same property of the Fréchet distributions as used in the main paper (see, for example, Section 4.2 for the route choice), the probability that workers use mode m for moving from n to i ($\lambda_{m|ni}^{M}$) is given by:

$$\lambda_{m|ni}^{M} = \frac{T_m^B \exp(-\theta^M \delta_{ni,m})}{\sum_{m' \in \mathbb{M}} T_{m'}^R \exp(-\theta^M \delta_{ni,m'})}. \tag{G.3}$$

As we assumed above, workers decide their residence, workplace, the set of consumption locations, and the routes (as discussed in the main paper) before observing the realizations of the idiosyncratic shock for each mode and travel leg ($\nu_{ni,m}^{M}(\omega)$). Therefore, workers make these decisions based on the expected travel cost for each leg of travel.

The expected commuting cost in the indirect utility function (equation G.2) is given by

$$\begin{aligned} \mathbb{E}\left[\exp(-\kappa^W \tau_{ni}^W(\omega))\right] &= \mathbb{E}\left[\exp(-\kappa^W \tau_{ni}(\omega, m_{ni})) \exp(-\kappa^W \tau_{in}(\omega, m_{in}))\right] \\ &= \mathbb{E}\left[\exp(-\kappa^W \tau_{ni}(\omega, m_{ni}))\right] \mathbb{E}\left[\exp(-\kappa^W \tau_{in}(\omega, m_{in}))\right], \end{aligned} \tag{G.4}$$

where the expectations are with respect to $\nu_{ni,m}^{M}(\omega)$, and the second transformation used the property that $\nu_{ni,m}^{M}(\omega)$ are independent for each segment of travel (for each movement of one location to another). The first term is transformed as:

$$\begin{aligned} \mathbb{E}\left[\exp(-\kappa^W \tau_{ni}(\omega, m_{ni}))\right] &= \mathbb{E}\left[\max_m \exp(-\kappa^W \left(\delta_{ni,m} + \log \nu_{ni,m}^{M}(\omega)\right))\right] \\ &= \vartheta^M \left(\sum_{m' \in \mathbb{M}} T_{m'}^R \exp(-\theta^M \delta_{ni,m'}) \right)^{\frac{\kappa^W}{\theta^M}}, \end{aligned}$$

where $\vartheta^M \equiv \Gamma\left(\frac{\theta^M/\kappa^M - 1}{\theta^M/\kappa^M}\right)$, and we use the same property of the Fréchet distribution (see, for example, Section 4.2 for the route choice).

Mode-adjusted travel cost. For notational convenience, we define "mode-adjusted travel time" $\tilde{\tau}_{ni}$ such that:

$$\tilde{\tau}_{ni} \equiv \log \vartheta^M \left(\sum_{m' \in \mathbb{M}} T_{m'}^R \exp(-\theta^M \delta_{ni,m'}) \right)^{\frac{1}{\theta^M}}, \tag{G.5}$$

where $\tilde{\tau}_{ni}$ is a summary statistic for the travel cost relevant for moving from n to i by anticipating the mode choice. The reason why we call this object by $\tilde{\tau}_{ni}$ will be immediately clear below. By combining this definition of $\tilde{\tau}_{ni}$ with equation (G.4), the expected commuting cost is given by:

$$\begin{aligned} \mathbb{E}\left[\exp(-\kappa^W \tau_{ni}^W(\omega))\right] &= \mathbb{E}\left[\exp(-\kappa^W \tau_{ni}(\omega, m_{ni}))\right] \mathbb{E}\left[\exp(-\kappa^W \tau_{in}(\omega, m_{in}))\right] \\ &= \exp\left(-\kappa^W \tilde{\tau}_{ni}\right) \exp\left(-\kappa^W \tilde{\tau}_{in}\right). \end{aligned}$$

Furthermore, we can derive the expected consumption travel cost similarly as in the derivation for the expected commuting cost.

Together, the expected travel cost over the mode choice (conditional on the choice of residence, workplace, consumption locations and the routes) is given by:

$$\begin{aligned}\mathbb{E}\left[d_{ni\{j(k)r(k)\}}(\omega)\right] &= \mathbb{E}\left[\exp(-\kappa^W \tau_{ni}^W(\omega))\right] \prod_{k \in K^S} \mathbb{E}\left[\exp(-\kappa^S \tau_{nij(k)r(k)}^S(\omega))\right] \\ &= \exp(-\kappa^W \tilde{\tau}_{ni}^W) \prod_{k \in K^S} \exp(-\kappa_k^S \tilde{\tau}_{nij(k)r(k)}^S)\end{aligned}$$

where

$$\tilde{\tau}_{ni}^W = \tilde{\tau}_{ni} + \tilde{\tau}_{in}$$

and

$$\begin{aligned}\tilde{\tau}_{nij(k)HH}^S &= \tilde{\tau}_{nj} + \tilde{\tau}_{jn}, \\ \tilde{\tau}_{nij(k)WW}^S &= \tilde{\tau}_{ij} + \tilde{\tau}_{ji}, \\ \tilde{\tau}_{nij(k)HW}^S &= \tilde{\tau}_{nj} + \tilde{\tau}_{ji} - \tilde{\tau}_{ni}, \\ \tilde{\tau}_{nij(k)WH}^S &= \tilde{\tau}_{ij} + \tilde{\tau}_{jn} - \tilde{\tau}_{in}.\end{aligned}$$

Note that this expression of the travel cost is isomorphic to the baseline model without the mode choice (equation 2) except that the travel time τ_{ni} is replaced by the "mode-adjusted travel time" $\tilde{\tau}_{ni}$. Therefore, the choices of residence, workplace, consumption locations, and the route choice are as discussed in the baseline model except that travel time τ_{ni} is replaced by mode-adjusted travel time $\tilde{\tau}_{ni}$.

G.1.2 Estimation of Mode Choice Probability and Estimation of Mode-Adjusted Travel Time $\tilde{\tau}_{ni}$

Now we discus the data that we use for estimating the mode choice probability and the mode-adjusted travel time $\tilde{\tau}_{ni}$. To construct $\tilde{\tau}_{ni}$ from expression (G.5), we need (1) the mode-specific travel time $\delta_{ni,m}$ for each bilateral pair of locations, and (2) the parameter θ^M and T_m^R. (Note that ϑ^M is a constant term common to all bilateral pair of locations and hence the value of ϑ^M does not affect any of our results.)

We construct the mode-specific travel time $\delta_{ni,m}$ for each bilateral pair of locations as follows. First, we assume that workers face two possible transportation modes: public transportation and cars. We extract the travel time after Tozai Subway Line has opened using public transportation by the web-based route choice service, Eki-spert API, as discussed in Section 2.2. Since Eki-spert API does not allows us to extract the travel time before Tozai Subway Line has opened, we construct the travel time with public transportation in the following procedure. We first compute the approximate travel time with and without the Tozai subway line for all pairs of Oazas in Sendai City using ArcGIS and the geocoded data of subways and bus lines under certain assumptions of the travel speed for each mode.[8] We then compute the ratio of the travel times without Tozai Subway Line relative to that with Tozai Subway Line. Finally, we multiply this ratio by the travel time from Eki-spert API to obtain our final estimates for the travel time with public transportation in the absence of Tozai Subway Line.

[8] We assume that 80 meter per minute for walk, 600 meter per minute for rail, and 150 meter per minute for buses. We obtain the geocoded railway and bus networks from the website of the Ministry of Land, Infrastructure, Transport and Tourism.

We construct car travel time by Open Source Routing Machine (OSRM), a routing service based on OpenStreetMap.[9] OSRM finds the shortest routes on public roads by car, bicycle and on foot between a pair of coordinates. We collected data on all the pairs of the centroids of Oazas in Sendai City in October 2020.

We estimate the mode choice probability (G.4) to obtain parameters θ^M and T^R_m. To do so, we use the travel travel survey data from Sendai City conducted in 2017. From a representative households, the survey collects information of the origin location, destination location and what travel mode is used. From this data, we construct the probability that railway (including other public transportation) or road transportation mode is used for each movement of the respondent.

Table G.1.1 presents our results of the mode choice estimation in equation (G.4) by Poisson Pseudo-Maximum Likelihood (PPML) with origin-destination fixed effects. Column (1) starts with the specification with the dummy of public transportation. The negative significant coefficient for this dummy indicates that people have a strong preference to use cars instead of using public transportation. In Column (2), we show our estimation results where we additionally control for travel time. The coefficient on travel time is negative and statistically significant, indicating that people have strong tendency to choose the transportation mode that provides shorter travel time. However, the dummy for public transportation is still negative and significant, indicating that people have strong preference to travel by cars even after controlling for the fact that travelling cars provides shorter travel time.

Table G.1.1: Mode Choice Estimation

Dependent Variable:	Mode Choice Probability	
Model:	(1)	(2)
Variables		
Dummy (Public Transportation Including Railways)	-1.50***	-0.953***
	(0.020)	(0.042)
Travel Time (Hours)		-0.318***
		(0.023)
Fixed-effects		
Origin and Destination Locations	Yes	Yes
Fit statistics		
Pseudo R^2	0.12210	0.12512
Log-Likelihood	-42,367.3	-42,221.4
AIC	121,474.7	121,184.8
BIC	290,555.3	290,274.6
Observations	73,436	73,436

Note: The results of the mode choice estimation in equation (G.4) by Poisson Pseudo-Maximum Likelihood (PPML). The unit of observations are all bilateral pairs of Oazas where there are positive flows. Heteroskedasticity robust standard errors in parentheses.

Using the estimates of these parameters and the mode-specific travel time, we construct the changes of the mode-adjusted travel time $\tilde{\tau}_{ni}$ using the expression (G.5). As discussed above, the counterfactual simulation procedure is unaffected from our baseline model except that we replace travel time τ_{ni} by the mode-adjusted travel time $\tilde{\tau}_{ni}$.

G.2 Estimation of Travel Access $\mathbb{A}_n$ in Sendai City

In Section 6.3, we study the impacts of Tozai Subway Line on the travel access $\mathbb{A}_n$. In this appendix, we discuss how we construct $\mathbb{A}_n$ in pre-period (before the subway opening) and post-period (after the subway opening) in more detail.

[9] See http://project-osrm.org/ for more detail.

We construct travel access $\mathbb{A}_n$ in the pre-period following the same procedure in Section 5.1, except that we use the mode-adjusted travel time as discussed in Appendix G.1. In this procedure, we also recover travel time elasticity (ϕ^W and ϕ_k^S). We then estimate the post-period consumption location choice and commuting choices given estimated parameters ϕ^W and ϕ_k^S. Using these estimates, we construct the post-period travel access $\mathbb{A}_n$ following the same expression.

Below, we present our estimation results of the route choice probability (Table G.2.1), consumption location choice probability (Table G.2.2), and commuting choice probability (Table G.2.3). Each of these tables corresponds to Table 2, Table 3, and Table 4 of the main paper estimated using data in Tokyo Metropolitan Area. The qualitative findings are broadly the same as in our estimates from Tokyo Metropolitan Area, except that the coefficients on travel time for route choice, consumption location choice and commuting choice are greater.

Table G.2.1: Estimation of Route Choice in Sendai City

Dependent Variable:	Route Choice Probability				
	Finance realestate communication professional	Wholesale retail	Accomodations eating drinking	Medical welfare healthcare	Other services
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Travel Time (Hours)	-0.394***	-0.387***	-0.332***	-0.376***	-0.355***
	(0.032)	(0.027)	(0.026)	(0.044)	(0.028)
Dummy (HW)	-1.37***	-1.49***	-1.57***	-1.44***	-1.44***
	(0.050)	(0.047)	(0.043)	(0.053)	(0.051)
Dummy (WH)	-0.784***	-0.911***	-0.809***	-0.964***	-0.816***
	(0.039)	(0.034)	(0.034)	(0.045)	(0.036)
Dummy (WW)	-0.917***	-1.01***	-1.16***	-1.03***	-1.00***
	(0.056)	(0.055)	(0.057)	(0.065)	(0.056)
Fixed-effects					
Home-Work-Consumption Location	Yes	Yes	Yes	Yes	Yes
Fit statistics					
AIC	74,821.5	185,222.9	69,038.2	83,774.2	97,380.3
BIC	96,834.8	207,420.2	90,941.0	105,604.8	119,501.4
Observations	98,385	102,227	96,137	94,331	100,975

Note: A version of Table 2 where we use pre-period data from Sendai City. See the footnote of Table 2 for other comments.

Table G.2.2: Estimation of Consumption Location Choice in Sendai City

Dependent Variable:	Consumption Location Choice Probability				
	Finance realestate communication professional	Wholesale retail	Accomodations eating drinking	Medical welfare healthcare	Other services
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
log $\tilde{d}^S_{nij(k)}$	-3.48***	-3.43***	-3.26***	-3.48***	-3.34***
	(0.117)	(0.100)	(0.104)	(0.099)	(0.110)
Fixed-effects					
Home and Work Location Pairs	Yes	Yes	Yes	Yes	Yes
Consumption Location	Yes	Yes	Yes	Yes	Yes
Fit statistics					
AIC	26,702.4	27,301.8	26,449.9	27,684.4	26,821.4
BIC	54,680.9	55,630.4	54,185.5	55,567.7	55,125.2
Observations	565,455	594,395	538,876	561,708	593,865

Note: A version of Table 3 where we use pre-period data from Sendai City. See the footnote of Table 3 for other comments.

Table G.2.3: Estimation of Commuting Choice in Sendai City

Dependent Variable:	Commuting Choice Probability
Model:	(1)
Variables	
Commuting Time (Hours)	-2.29***
	(0.105)
Fixed-effects	
Home Location	Yes
Work Location	Yes
Fit statistics	
AIC	1,289.9
BIC	7,334.8
Observations	100,650

Note: A version of Table 4 where we use pre-period data from Sendai City. See the footnote of Table 4 for other comments.

G.3 Additional Results of the Difference-in-Difference Estimates of Tozai Subway Line

In this appendix, we provide additional results of our difference-in-difference effects of Tozai Subway Line (Table 6 of our main paper). In Section G.3.1, we show how our results are biased when we omit the consumption trips to construct travel access $\mathbb{A}_n$. In Section G.3.2, we show the sensitivity of our results under different values of the residential spillover elasticity η^B and productivity spillover elasticity η^W. In Section G.3.3, we provide a procedure to estimate η^B using the difference-in-difference estimates and present our estimation results.

G.3.1 Bias of Omitting Consumption Trips

In this subsection, we show how we obtain biased conclusions about the reduced-form effects of Tozai Subway Line when we omit the consumption trips to construct travel access $\mathbb{A}_n$. In Table G.3.1, we compare our results by including and omitting consumption trips when constructing travel access $\mathbb{A}_n$, following the same difference-in-difference specification as in Panel (A) of Table 6 of our main paper. Panel (A) uses outcome variables constructed directly from observed data, and it is identical to Panel (A) of Table 6. In Panel (B), we construct travel access $\mathbb{A}_n$ and residential amenity $\mathbb{B}_n$ by omitting consumption trips ($\alpha^S_k = 0$ for all $k \in K^S$, $\alpha^T = 1 - \alpha^H$, $\lambda^S_{j(k)|ni} = 0$ and $\mathbb{S}_{nt} = 1$). Columns (1) and (2) are identical between the two panels because they are directly obtained from data and unaffected by the construction of travel access $\mathbb{A}_n$. In Column (3), we find an underestimation of the effects of Tozai Subway Line on travel access $\mathbb{A}_n$ (0.042 instead of 0.054), and an overestimation on residential amenity (0.017 instead of 0.004). Therefore, by omitting consumption trips, one would mis-attribute the changes of residential values from subways due to unobserved changes in residential amenity. Hence, we find that incorporating consumption trips is important for the quantitative success of the model's mechanism in explaining the observed data.

Table G.3.1: Bias of Difference-in-Difference Effects of Tozai Subway Line by Omitting Consumption Trips

	$\Delta \log Q_n$	$\Delta \log \lambda_n^B$	$\Delta \log \mathbb{A}_n$	$\Delta \log \mathbb{B}_n$
	(1)	(2)	(3)	(4)
Panel A: Data				
Dummy (Tozai Line Stations)	0.046***	0.311	0.054***	0.004
	(0.014)	(0.210)	(0.008)	(0.036)
Observations	368	305	305	305
R^2	0.030	0.007	0.123	0.0001
Panel B: Data (Travel Access Constructed by Omitting Consumption Trips)				
Dummy (Tozai Line Stations)	0.046***	0.311	0.042***	0.017
	(0.014)	(0.210)	(0.010)	(0.036)
Observations	368	305	305	305
R^2	0.030	0.007	0.057	0.001

Note: In Panel (A), we construct travel access $\mathbb{A}_n$ and residual amenity $\mathbb{B}_n$ by including consumption trips, and it is identical to Panel (A) of Table 6 of our main paper. In Panel (B), we construct travel access $\mathbb{A}_n$ and residential amenity $\mathbb{B}_n$ by omitting consumption trips ($\alpha_k^S = 0$ for all $k \in K^S$, $\alpha^T = 1 - \alpha^H$, $\lambda^S_{j(k)|ni} = 0$ and $\mathbb{S}_{nt} = 1$). By construction, columns (1) and (2) are identical between the two panels.

G.3.2 Robustness to Different Elasticities of Agglomeration Spillovers

In this section of the appendix, we show the sensitivity of our results of the difference-in-difference regression of Tozai Subway Line under different values of the elasticities of agglomeration spillovers.

Table G.3.2 presents the results. Panel (A) uses outcome variables constructed directly from observed data, and Panel (B) is those from model prediction under $\eta^B = 0$ and $\eta^W = 0.08$, and the two panels are identical to those of Table 6. As discussed in the main paper, the results between Panel (A) and Panel (B) closely align with each other, validating our model and our choice of the parameter values for η^B and η^W.

In Panel (C), we report the results of the difference-in-difference regression with outcome variables constructed from model prediction when we increase residential amenity spillover η^B from 0 to 0.15. The model substantially over-predicts the impacts on floor space price (Column 1) and the residential probability (Column 2). This overestimation primarily comes from the predicted increase of residential amenity (Column 4), which we do not find from our observed data (Panel A). Therefore, we find that the model is quantitatively able to explain the observed increase in floor space prices and residential population through its mechanism of an improvement in travel access, without requiring increases in the residual of residential amenities in these locations.

In Panel (D), we report the results of the difference-in-difference with outcome variables constructed from model prediction when we decrease the productivity spillover η^W from 0.08 to 0. By comparing with Panel (B), we find somewhat smaller changes of floor space price and residential population, while this difference is modest.

Table G.3.2: Robustness of Difference-in-Difference Effects of Tozai Subway Line under Different Elasticities of Agglomeration Spillovers

	$\Delta \log Q_n$	$\Delta \log \lambda_n^B$	$\Delta \log \mathbb{A}_n$	$\Delta \log \mathbb{B}_n$
	(1)	(2)	(3)	(4)
Panel A: Data				
Dummy (Tozai Line Stations)	0.046***	0.311	0.054***	0.004
	(0.014)	(0.210)	(0.008)	(0.036)
Observations	368	305	305	305
R^2	0.030	0.007	0.123	0.0001
Panel B: Model Prediction ($\eta^B = 0; \eta^W = 0.08$)				
Dummy (Tozai Line Stations)	0.091***	0.300***	0.073***	0.000
	(0.010)	(0.032)	(0.008)	(0.000)
Observations	370	370	370	370
R^2	0.197	0.191	0.199	
Panel C: Model Prediction ($\eta^B = 0.15; \eta^W = 0.08$)				
Dummy (Tozai Line Stations)	0.206***	1.182***	0.074***	0.175***
	(0.022)	(0.127)	(0.008)	(0.019)
Observations	370	370	370	370
R^2	0.198	0.189	0.196	0.190
Panel D: Model Prediction ($\eta^B = 0; \eta^W = 0$)				
Dummy (Tozai Line Stations)	0.077***	0.295***	0.068***	0.000
	(0.008)	(0.031)	(0.007)	(0.000)
Observations	370	370	370	370
R^2	0.203	0.197	0.203	

Note: The results of the difference-in-difference regression of Tozai Subway Line using the observed outcome variables (Panel A) and the model prediction (Panel B-D). Panels A and B are identical to Table 6 of our main paper. Panel C and D use our model prediction under different values of agglomeration spillovers (η^B and η^W) Standard errors are clustered at Oaza level. See the footnote of Table 6 for other comments.

G.3.3 Estimation of Residential Amenity Spillover Elasticity η^B

In this section of the appendix, we show how we can estimate the residential spillover elasticity η^B using the opening of Tozai Subway Line. To do so, we introduce an identification assumption that the changes unobserved residential amenity is uncorrelated with the proximity to the stations of Tozai Subway Line. More specifically, from equations (28) and (44), we have:

$$\begin{aligned}\mathbb{B}_n &= B_n \left(T_n^B\right)^{1/\theta^B} \left(P_n^T\right)^{-\alpha^T} \left(\bar{U}/\vartheta^B\right)^{-1} \\ &= b_n \left(\frac{L\lambda_n^B}{K_n}\right)^{\eta^B} \left(T_n^B\right)^{1/\theta^B} \left(P_n^T\right)^{-\alpha^T} \left(\bar{U}/\vartheta^B\right)^{-1}.\end{aligned}$$

By assuming the geographic area K_n and the shifter for idiosyncratic preference shocks T_n^B are time-invariant and tradable prices are equalized across locations as numeraire, we obtain our estimating equation for η^B by logdifferencing the above equation:

$$\Delta \log \mathbb{B}_n = c_0 + \eta^B \Delta \log \lambda_n^B + u_n \tag{G.6}$$

where $c_0 = \Delta \log \left(L^{\eta^B} \bar{U}^{-1} \right) + \mathbb{E}\Delta \log b_n$ and u_n is the mean-zero error term capturing the idiosyncratic shift of the exogenous component of the residential amenity ($u_n = \Delta \log b_n - \mathbb{E}\Delta \log b_n$). We assume that u_n is uncorrelated with whether Oaza n has a Tozai Subway Line station (T_n). Under this assumption, we obtain a consistent estimate of residential spillover elasticity η^B by the instrumental variable (IV) regression using T_n as an instrument for $\Delta \log \lambda_n^B$.

Table G.3.3 presents our estimates of η^B. In Column (1), we present our results when we include consumption trips for constructing $\Delta \log \mathbb{B}_n$. We find a small point estimate of $\eta^B = 0.015$. In Column (2), we present our results when we omit consumption trips for constructing $\Delta \log \mathbb{B}_n$ (use $\Delta \log \mathbb{A}_n^{\text{nocons}}$ instead of $\Delta \log \mathbb{A}_n$ as discussed in Section 5.2). Interestingly, when we omit the consumption trips to estimate η^B, we obtain a larger point estimate of $\eta^B = 0.054$. Therefore, by omitting consumption trips, one would overestimate the residential spillover elasticity.

Table G.3.3: Estimation of Residential Amenity Spillover Elasticity in Sendai City

	$\Delta \log \mathbb{B}_n$	
	(1)	(2)
$\Delta \log \lambda_n^B$	0.014	0.054
	(0.105)	(0.080)
First Stage F-Statistcs	1.1	1.1
Specification	Include Consumption Trips	Omit Consumption Trips
Observations	305	305
Adjusted R^2	0.155	0.509

Note: This table presents our estimation results of η^B. We estimate these parameters using the IV regression (G.6) where we instrument $\Delta \log \lambda_n^B$ by whether Oaza n has a Tozai Subway Line station (T_n).

G.4 Network Effects on Nanboku (North-South) Subway Line

As an additional specification check, we repeat the same difference-in-difference regression with Tozai Subway Line (Table 6), but use a dummy variable that takes the value one for Oazas that contain stations on the existing Nanboku (North-South) Subway Line (which opened in 1987) rather than stations on the new Tozai (East-West) Subway Line (which opened in 2015). If there are positive or negative network effects from the new Tozai Subway Line on locations with stations on the existing Nanboku Subway Line, we would expect to again detect statistically significant treatment effects.

In Panel (A) of Table G.4.1, we show that we find no evidence of statistically significant treatments effects on the price of floor space, residential population, travel access, and residential amenities for this existing Nanboku Subway Line. These results are consistent with a limited net impact of network effects on the existing subway line and suggest that our earlier estimates for the Tozai Subway Line are capturing effects specific to this new subway line. Consistent with these findings using the observed data, in Panel (B), we find no evidence of statistically significant treatment effects for existing Nanboku Subway Line using our counterfactual predictions of the model.

Table G.4.1: Network Effects of Nanboku Subway Line

	$\Delta \log Q_n$	$\Delta \log \lambda_n^B$	$\Delta \log \mathbb{A}_n$	$\Delta \log \mathbb{B}_n$
	(1)	(2)	(3)	(4)
Panel A: Data				
Dummy (Tozai Line Stations)	-0.002	0.086	0.004	0.009
	(0.011)	(0.160)	(0.006)	(0.027)
Observations	357	296	296	296
R^2	0.0001	0.001	0.002	0.0004
Panel B: Model Prediction ($\eta^B = 0; \eta^W = 0.08$)				
Dummy (Tozai Line Stations)	-0.003	-0.010	-0.003	0.000
	(0.008)	(0.027)	(0.006)	(0.000)
Observations	359	359	359	359
R^2	0.0005	0.0004	0.0005	

Note: The results of the difference-in-difference regression by defining treatment by the stations around Nanboku (North-South) Subway Line (which already opened in 1987) instead of Tozai (East-West) Subway Line (which actually opened in 2015) for the same time interval. We exclude the observations of Oazas containing Tozai Subway Line from this analysis. We use observed outcome variables in Panel A and the model prediction in Panel B. See the footnote of Table 6 for additional comments.

G.5 Implications of Granularity

As remarked in Section 6.1 of our main paper, one important issue of the counterfactual simulation of this type of model is the granularity of observed travel flows ($\{\lambda_{ig|n}^W, \lambda_{j(k)|ni}^S, \lambda_n^B\}$). In particular, Dingel and Tintelnot (2020) show that this type of counterfactual simulation may be potentially biased due to model overfitting when one uses actual travel flows to calibrate the model. Instead, they recommend "covariate-based approach," in which one calibrates the travel flows by the predicted flows using covariates such as travel time. In our main paper, we presented the results following Dingel and Tintelnot (2020) to use the predicted travel flows from our consumption location choice and workplace choice estimation to calibrate the model. In this section, we demonstrate how the two approaches provide different predictions for our second counterfactual of subway opening.

Table G.5.1 shows how our results of the difference-in-difference regression of Tozai Subway Line is affected by this choice of the calibration strategy. Identically to Table 6 of our main paper, Panel (A) corresponds to the estimates using the observed data. Panel (B) estimates the same reduced-form regressions using the model's counterfactual predictions when we calibrate the model using predicted travel flows (as in Panel (B) of Table 6 of our main paper). In Panel (C), we present the estimates based on the model prediction when we instead calibrate the model with actual travel flows. Comparing Panel (B) and (C), we find somewhat smaller impacts on the model prediction of floor space price, residential probability, and the travel access. This difference is intuitive. We observe zero travel flows in many pairs of locations in our data due to granularity. Some of these routes are directly affected by the new subway line. Therefore, if we use actual travel flows to calibrate the model, we implicitly assume no direct gains from the travel time reduction of these routes with zero travel flows in baseline. Therefore, the model predicts smaller effects from subway line. Lastly, we cannot conclude which of the two assumptions perform better (by comparing Panel B and C with Panel A) because of the relatively large standard errors in these estimates.

In Table G.5.2, we compare our estimates of the welfare gains from Tozai Subway Line under these two calibration strategies. Consistent with the observation above, we find a smaller welfare gain using the model calibrated with actual travel flows (2.74 percentage points as opposed to 2.55 percentage points).[10] While this difference is non-negligible, it is significantly less than the difference in the results arising from the omission of consumption trips (Table 7). Therefore, the main message of the paper about the importance of consumption trips remains robust to these calibration strategies.

Table G.5.1: Impacts of Tozai Subway Line: Granularity

	$\Delta \log Q_n$	$\Delta \log \lambda_n^B$	$\Delta \log \mathbb{A}_n$	$\Delta \log \mathbb{B}_n$
	(1)	(2)	(3)	(4)
Panel A: Data				
Dummy (Tozai Line Stations)	0.046***	0.311	0.054***	0.004
	(0.014)	(0.210)	(0.008)	(0.036)
Observations	368	305	305	305
R^2	0.030	0.007	0.123	0.0001
Panel B: Model Prediction with Smoothed Flows ($\eta^B = 0; \eta^W = 0.08$)				
Dummy (Tozai Line Stations)	0.091***	0.300***	0.073***	0.000
	(0.010)	(0.032)	(0.008)	(0.000)
Observations	370	370	370	370
R^2	0.197	0.191	0.199	
Panel C: Model Prediction with Actual Flows ($\eta^B = 0; \eta^W = 0.08$)				
Dummy (Tozai Line Stations)	0.062***	0.205***	0.050***	0.000
	(0.010)	(0.037)	(0.009)	(0.000)
Observations	370	370	370	370
R^2	0.092	0.079	0.083	

Note: The results of the difference-in-difference regression of Tozai Subway Line using the observed outcome variables (Panel A) and the model prediction (Panel B-C). Panels A and B are identical to Table 6 of our main paper. In Panel C, we calibrate the model using actual consumption travel flows and commuting flows ($\{\lambda^W_{ig|n}, \lambda^S_{j(k)|ni}, \lambda^B_n\}$) instead of the predicted flows as reported in Panel (B). See the footnote of Table 6 for other comments.

Table G.5.2: Counterfactual Increase in Expected Utility in Sendai from the new Tozai Subway Line: Granularity

	Percentage Point Increase in Residential Utility	Relative to Baseline (%)
(1) Use Smoothed Travel Flows in Baseline	2.74	1.00
(2) Use Actual Travel Flows in Baseline	2.55	0.93

Note: The second column reports model counterfactuals for the percentage point increase in expected utility as a result of the reduction in travel time from the opening of the new Tozai (East-West) subway line in the City of Sendai. The first row presents results for our baseline specification ($\eta^B = 0, \eta^W = 0.08$), and this corresponds to Table 7 of our main paper. The second row presents the same results by calibrating the model using actual consumption travel flows and commuting flows ($\{\lambda^W_{ig|n}, \lambda^S_{j(k)|ni}, \lambda^B_n\}$).

[10] In the granular model of Dingel and Tintelnot (2020), agents form travel decisions based on the beliefs that equilibrium wages follow that of the model with continuum agents. The expected utility reported here is consistent with this assumption of the belief.