Online Appendix for "The Economics of Spatial Mobility: Theory and Evidence Using Smartphone Data" (Not for Publication)

Yuhei Miyauchi*
Boston University

Kentaro Nakajima†
Hitotsubashi University

Stephen J. Redding‡
Princeton University, NBER and CEPR

May 15, 2025

Table of Contents

*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.

Structure of the Online Appendix

This Online Appendix contains additional supplementary material for the paper. Section A provides further information on our smartphone data and reports additional empirical results discussed in the paper (supplementing Sections 3 and 4.1 of the paper). Section B presents further evidence on the shift to Working from Home (WFH) in Japan following the COVID-19 pandemic (supplementing Section 4.2 of the paper). Section C provides presents further evidence on retail store closure (supplementing Section 4.3 of the paper).

Section D reports first-order comparative statics for changes in the prices of non-traded services and shows that locations can be either substitutes or complements (as discussed in Section 5.1 of the paper). Section E provides further details about the importance sampling method used to estimate the parameters of our travel itinerary model (supplementing Section 5.2 of the paper). Section F reports additional results for our estimation of the parameters of our travel itinerary model and provides additional evidence on model fit (supplementing Sections 5.4-5.6 of the paper).

Section G provides further details on the general equilibrium of the model and our counterfactual simulations (supplementing Sections 6 and 7 of the paper). Section H reports robustness tests for our counterfactual simulations (supplementing Section 7 of the paper).

A Data Appendix

This section of the Online Appendix provides further information on our smartphone data, reports a number of specification checks on our smartphone data, and presents additional empirical results that are discussed in the paper.

We begin with several validation exercises for our smartphone data. In Subsection A.1, we present evidence on the spatial resolution of the "stays" recorded in our smartphone data. In Subsection A.2 we provide further evidence on the representativeness of our smartphone data. We show that the coverage of our data is not systematically related to either residence characteristics (income, age and distance to city center) or workplace characteristics (employment by industry and distance to city center).

In Subsection A.3, we show that our measures of commuting based on our smartphone data display the same patterns as census commuting data for Tokyo, supplementing the results reported in Subsection 3.2 of the paper. We also show that our findings 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 work days.

We next provide further evidence on patterns of spatial mobility using our smartphone data. In Subsection A.4, we show that our data exhibit an intuitive pattern of the average number of work and non-work stays by day and hour of the week. In Section A.5, we combine our smartphone data with spatially-disaggregated economic census data to provide evidence on non-commuting stays for different types of nontraded sectors. In Subsection A.6, we show compare our findings for non-commuting trips to evidence from other data sources and empirical contexts, including the United States. In Subsection A.7, we provide additional spatial and temporal patterns of travel itineraries, supplementing the results of Section 4.1.

In Subsection A.8, we report a further specification check on our measures of work location. We find that users with missing work locations have more infrequent smartphone use. We show that the probability of assigning missing work locations is uncorrelated with the observable characteristics of users' municipality of residence.

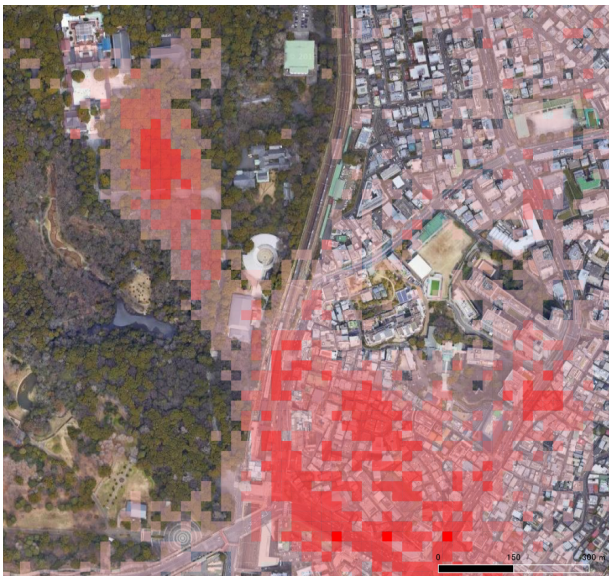
Finally, in Appendix A.9, we provide a detailed list of other data used in the main paper and this online appendix in addition to our smartphone data.

A.1 Spatial Resolution of our Smartphone data

In Figure A.1.1, we provide an example of the "stays" recorded in our smartphone data for a *Meiji Shrine* in the Shibuya municipality of Tokyo over the period from December 2017 to February 2018. Each red-shaded rectangle corresponds to a 25-meter by 25-meter grid cell. The darker the red shading of the grid cell, the larger the number of stays in that grid cell. We have

overlaid these grid cells on a satellite photograph. In this photograph, the building towards the top-left of the image, surrounded by trees, corresponds to the main building of the Meiji Shrine. Several features of our data are apparent from this image. First, we observe movement within the city at a high level of spatial resolution. Second, we find a sharp discontinuity in the density of stays at the road that separates the wooded area surrounding the shrine to the left from the developed area to the right, suggesting that the stays accurately capture the density of movement. Third, in the middle of this wooded area, the stays are concentrated tightly along the path that runs from the road to the main building of the shrine, again confirming the ability of our data to capture the main pathways of movement through the city.

Figure A.1.1: Example of Stays Around a Meiji Shrine in the Shibuya Municipality of Tokyo



Note: The map shows the geographic location of "stays" around a Meiji Shrine. Each red-shaded rectangle corresponds a grid cell of the size of approximately 25×25 meters. The darkness of the color represents the number of stays in each grid cell between December 2017 and February 2018. The building towards the top-left surrounded by trees is the main building of the shrine. The stays are concentrated tightly along the path that runs from the road to the main building of the shrine, consistent with them accurately capturing patterns of movement within the city.

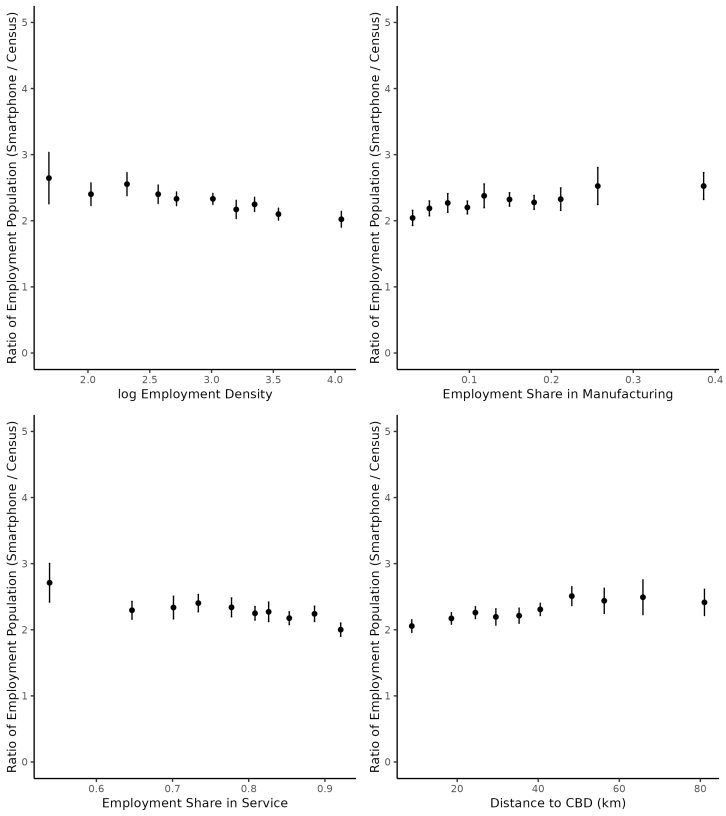
A.2 Representativeness of our Smartphone Data

In this subsection of the Online Appendix, we demonstrate the representativeness of our smartphone data by both workplace and residence characteristics.

Representativeness by Workplace We first compute a workplace coverage rate, defined as the number of users whose work location is in the municipality divided by employment in that

municipality in the population census. In Figure A.2.1, we plot this workplace coverage rate against municipality characteristics, including employment density, the employment share in manufacturing, the employment share in services, and distance to the Central Business District (the centroid of Chiyoda Ward). The dots in each figure represent the average coverage rate for each decile of municipality characteristics on the horizontal axis. The line segments indicate the 95 percent confidence intervals. We find little evidence of any systematic relationship between the workplace coverage rate and these municipality characteristics. There is a mild decreasing pattern for the employment density and increasing patterns for the distance to CBD, but the magnitudes are not quantitatively large. Therefore, we find little evidence of bias in the workplace coverage rates of our smartphone data along these dimensions of observable workplace characteristics.

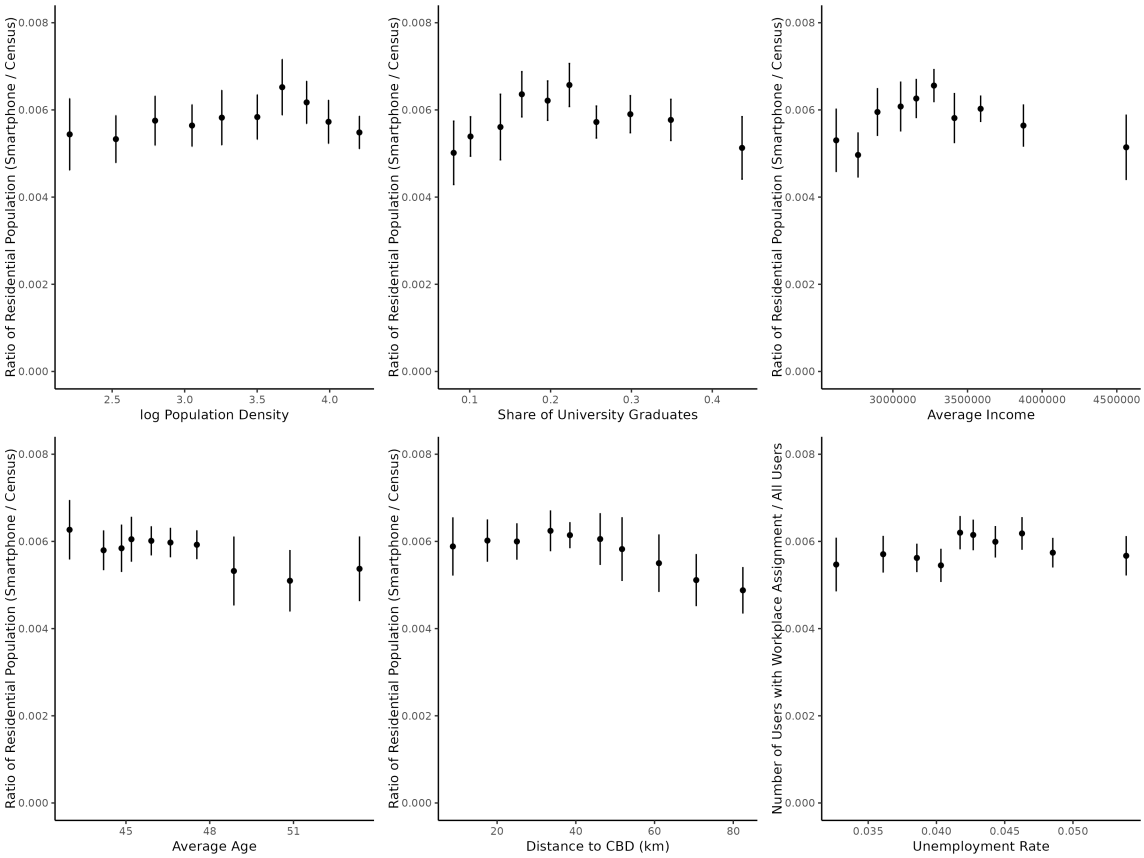
Figure A.2.1: Representativeness of our Smartphone Data by Workplace Characteristics



Notes: Vertical axis shows the workplace coverage rate, defined as the number of users in our smartphone data whose work location is in a municipality divided by employment in that municipality in the population census (242 municipalities in the Tokyo metropolitan area); horizontal axis shows workplace characteristics of municipalities; the dots represent the average coverage rate for each decile of municipality characteristics on the horizontal axis; the line segments indicate the 95 percent confidence intervals.

Representativeness by Residence We next compute a residence coverage rate for each municipality, defined as the number of users whose home location is in the municipality divided by the number of residents in that municipality in the population census. In Figure A.2.2, we plot this residence coverage rate against municipality characteristics, including population density, the share of university graduates, average income, average age and the distance to the Central Business District (the centroid of Chiyoda Ward). Again the dots in each figure represent the average coverage rate for each decile of municipality characteristics on the horizontal axis. The line segments indicate the 95 percent confidence intervals. We find little evidence of any relationship between the residential coverage rate and these municipality characteristics.

Figure A.2.2: Representativeness of our Smartphone Data by Residence Characteristics



Notes: Vertical axis shows the residential coverage rate, defined as the number of users in our smartphone data whose home location is in a municipality divided by the number of residents in that municipality in the population census (242 municipalities in the Tokyo metropolitan area); horizontal axis shows residential characteristics of municipalities; the dots represent the average coverage rate for each decile of municipality characteristics on the horizontal axis; the line segments indicate the 95 percent confidence intervals.

Taken together, these empirical findings provide support for the representativeness of our smartphone data, by both residence and workplace characteristics.

A.3 Validation of Smartphone Mobility Patterns

In this subsection of the Online Appendix, we provide further validation for our smartphone data using census commuting data and a travel survey from Tokyo.

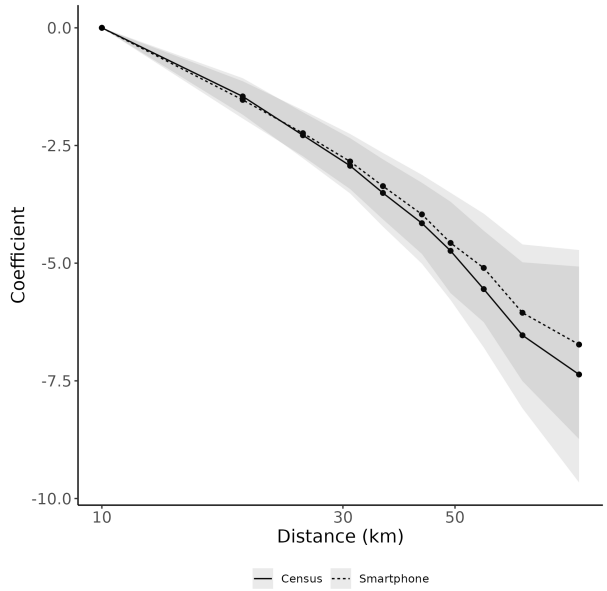
A.3.1 Validation using Census Commuting Data

In this subsection of the Online Appendix, we provide further validation for our smartphone data by showing that our measures of commuting based on home and work locations in our smartphone data display the same patterns as official census commuting data, supplementing the empirical results reported in Section 3.2 of the paper.

Spatial Decay with Distance In Figure A.3.1, we show that bilateral commuting flows from our smartphone data have a similar rate of spatial decay with geographic distance as the official census data. We estimate a gravity equation for bilateral commuting flows using both our smartphone commuting data and the census commuting data. We estimate this relationship for our baseline sample of 2019 using our smartphone data and for 2015 using the census commuting data. We estimate this gravity equation for the 242 municipalities of the Tokyo metropolitan area, because the census commuting data are only available at the municipality level. We use the Poisson Pseudo Maximum Likelihood (PPML) estimator and include indicators for deciles of bilateral distance and residence and workplace fixed effects. The figure displays the estimated coefficients on the decile indicators (black lines) and the 95 percent confidence intervals (gray shading). The solid black line and dark gray shading show results using the census commuting data. The dashed black line and light gray shading show results using our smartphone data. We find that these two sets of estimates lie close to one another, particularly for commutes of less than 50 kilometers, which account for the vast majority of observed bilateral commuting flows.

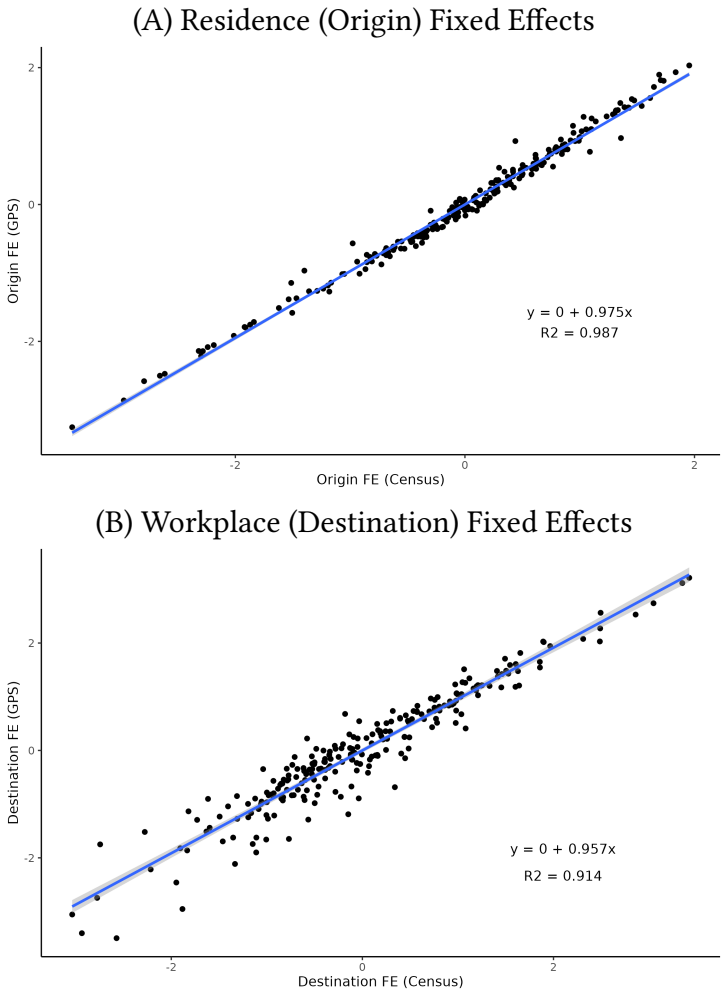
Residence and Workplace Fixed Effects In Figure A.3.2, we compare the estimated fixed effects using our smartphone data and the official census data from this gravity equation estimation. Panel A shows the estimated residence (origin) fixed effects. Panel B shows the estimated workplace (destination) fixed effects. In each panel, the vertical axis shows the estimated fixed effects using our smartphone GPS data, while the horizontal axis shows the estimated fixed effects using the census data. We find an approximately log linear relationship between the two sets of estimates with a slope coefficient close to one and an R-squared of above 0.9. Therefore, we not only find a similar relationship with bilateral distance, but also estimate similar measures of the attractiveness of residences and workplaces using our smartphone data as using the census data.

Figure A.3.1: Estimated Coefficients on Deciles of Geographic Distance from Gravity Equation Estimation using our Smartphone Data and Official Census Data



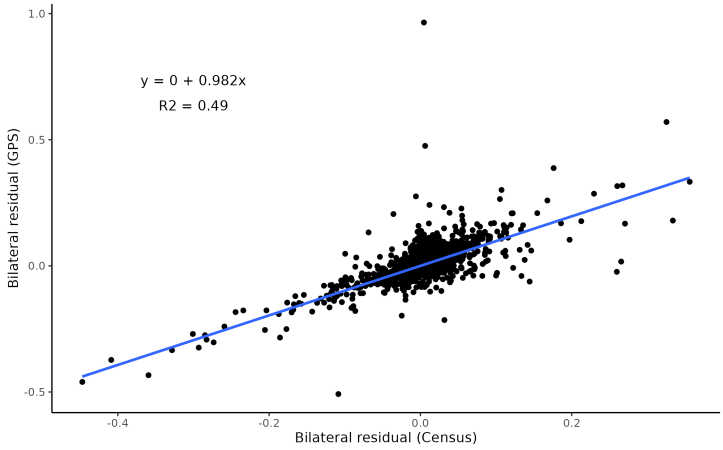
Notes: Gravity equation estimation including indicator variables for deciles of bilateral distance and residence and workplace fixed effects; gravity equation estimated using the Poisson Pseudo Maximum Likelihood (PPML) estimator; we estimate this relationship for our baseline sample of 2019 using our smartphone data and for 2015 using the census commuting data; the solid black line and dark gray shading show the estimated coefficients on the decile indicators and 95 percent confidence intervals using the official census data; the dashed black line and light gray shading show the estimated coefficients on the decile indicators and 95 percent confidence intervals using our smartphone data.

Figure A.3.2: Correlation of Gravity Equation Residence and Workplace Effects using our Smartphone Data and Official Census Data



Notes: Gravity equation estimation using the Poisson Pseudo Maximum Likelihood (PPML) estimator, including indicator variables for deciles of bilateral distance and residence and workplace fixed effects; we estimate this relationship for our baseline sample of 2019 using our smartphone data and for 2015 using the census commuting data; Panel (A) shows estimated residence (origin) fixed effects; Panel (B) shows estimated workplace (destination) fixed effects; vertical axis in each panel shows estimates using our smartphone GPS data; horizontal axis in each panel shows estimates using the census data; black circles correspond to the estimated fixed effects; blue solid line shows the linear regression fit.

Figure A.3.3: Correlation of Gravity Equation Bilateral Residuals Using our Smartphone Data and Official Census Data



Notes: Residuals from gravity equation estimation using the Poisson Pseudo Maximum Likelihood (PPML) estimator, including indicator variables for deciles of bilateral distance and residence and workplace fixed effects; we estimate this relationship for our baseline sample of 2019 using our smartphone data and for 2015 using the census commuting data; vertical axis shows estimates using our smartphone GPS data; horizontal axis shows estimates using the census data; black circles correspond to the estimated residuals; blue solid line shows the linear regression fit.

Residual Variation As a final specification check, Figure A.3.3 compares the estimated gravity equation residuals using the two different datasets. The vertical axis shows the estimated residuals using our smartphone data. The horizontal axis shows the estimated residuals using the census data. We find a close relationship between the two estimated residuals, with a regression slope coefficient of 0.982 that is close to one and a regression R-squared of 0.49. Therefore, idiosyncratic shocks to commuting flows that are not captured by bilateral distance or the fixed effects show up in both data sets, again confirming the ability of our smartphone data to accurately capture the commuting patterns in the official census data.

Taken together, we find strong evidence that our smartphone data closely replicate patterns of commuting in official census data, in terms of the rate of spatial decay with distance (coefficients on bilateral distance in the gravity equation), the attractiveness of residences and workplaces (as captured by residence and workplace fixed effects), and idiosyncratic variation in commuting flows (the residuals from the gravity equation).

A.3.2 Comparison with Travel Survey Data for Tokyo

In this section of the Online Appendix, we provide further validation of our smartphone data using travel survey data for Tokyo. These travel survey data are only available for weekdays (not weekends) and have a much lower level of temporal and spatial resolution than our smart-

phone data. Nevertheless, when we measure non-commuting trips at levels of aggregation that can be compared in both datasets, we find the same pattern of spatial mobility in the travel survey data as in our smartphone data.

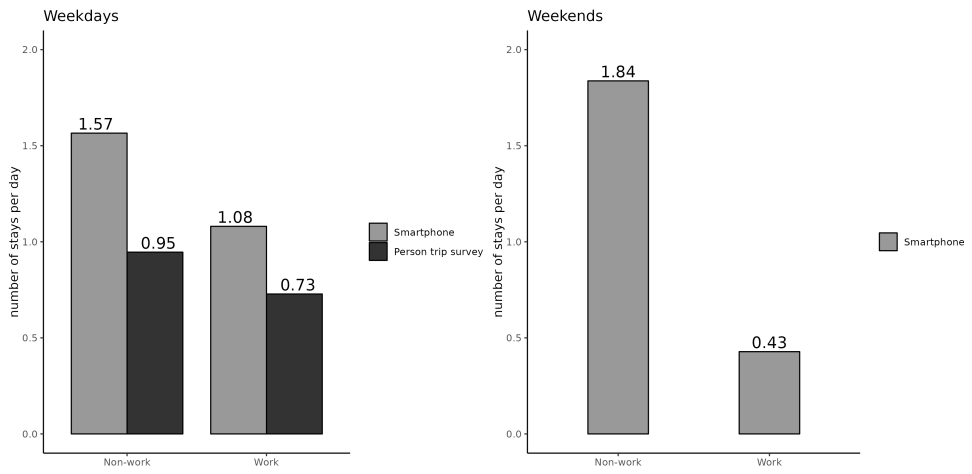
Travel Survey Data for Tokyo In the Greater Tokyo Metropolitan Area, travel surveys (person trip surveys) are conducted at a decennial frequency. We use the data from the travel survey data conducted in 2008. The respondents of this survey are members of households residing in the Tokyo, Saitama, Chiba, and Kanagawa and a part of Ibaraki prefectures, which is mostly consistent with our definition of the boundaries of the Greater Tokyo Metropolitan Area. The respondents are asked about their travel behavior on a specific weekday. The survey asks about the sequence of travel made during the day (a sequence of segments of movements from one place to another). For each trip segment, the survey asks where and when the trip starts and ends, the purpose of the trip, and what transportation mode is used. The survey also asks basic demographic information, work status, and home address. We use the data for all non-student respondents.

Frequencies of Work and Non-work Stays We begin by comparing the number of stays per day per person in our smartphone data and in the travel survey. In Figure A.3.4, the left panel depicts the number of work stays and non-work stays (at locations other than home or work) during weekdays for the two data sets. The right panel shows the same information for weekends, which is only available in our smartphone data (and hence the statistics for the travel survey are missing in the right panel).

Focusing on weekdays (left panel), we find that the number of stays is greater in our smartphone data than in the travel survey data. One possible reason for this difference is the under-reporting of trips in the travel survey data. Recall that our smartphone data measures a "stay" if a user is static for more than 15 minutes. In the travel survey, respondents may not think a stop as short as 15 minutes as worthy of recording. Despite this difference, the relative number of work stays and non-work stays (i.e., stays at locations other than home or work) is similar in the two data sets. In our smartphone data, 59 percent ($= 1.57/(1.57 + 1.08)$) of all stays outside the home are non-work stays. In the travel survey data, the corresponding figure is similar at 56 percent ($= 0.95/(0.95 + 0.73)$).

Non-work Stays by Type We next compare the types of non-work stays in our smartphone data and the travel survey. As discussed in Section A.5 of this Online Appendix, we assign non-work stays in our smartphone data to non-traded sectors using separate economic census data. In Figure A.5.2, we decompose non-work stays into these non-traded sectors in

Figure A.3.4: Frequencies of Work and Non-work Stays in our Smartphone Data and Travel Survey Data for Tokyo

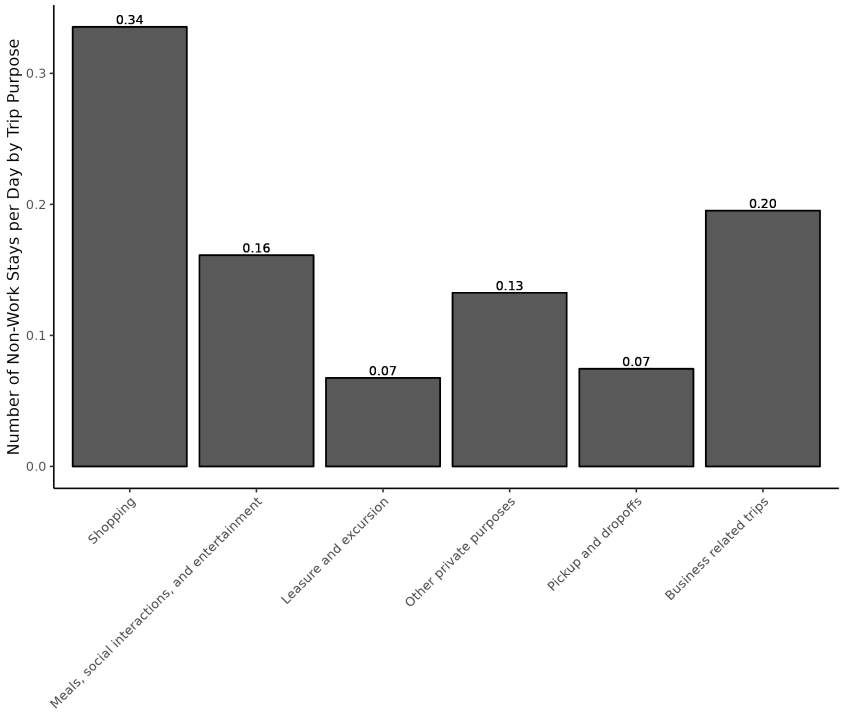


Notes: Frequencies of work stays and non-work stays (at locations other than home and work) in our smartphone data for our baseline sample in 2019 and travel survey data for the Greater Tokyo metropolitan area in 2008; left panel shows results for weekdays; right panel shows results for weekends; the travel survey data are only available for weekdays (and hence the statistics for the travel survey data are missing in the right panel).

our smartphone data. In Figure A.3.5, we display the average number of non-work stays by their stated purpose in the travel survey data. While a precise comparison between the two sets of data is difficult, because of the different classifications, we find a similar overall pattern. In particular, we find that "Shopping" is the most frequent category in the travel survey data (34 percent of all trips), which is consistent with our finding that the "Retail and Wholesales" sectors are the most frequent category of non-work stay in our smartphone data (43 percent of all non-work stays on weekdays). We also find that business trips (20 percent of all trips) are substantially less frequent than consumption trips (57 percent of all trips are shopping, entertainment, and leisure) in the travel survey data, which is consistent with our interpretation of non-work stays in our smartphone data as consumption trips.

Average Distances Travelled As a final comparison between our smartphone data and the travel survey, Table A.3.1 examines the average distances of work stays and non-work stays from home. The first and second columns report these average distances in kilometers for our smartphone data and the travel survey data, respectively. In both cases, we find that non-work stays are concentrated closer to home than work stays. For work stays, the average distances travelled are closely aligned between the two data sets (12.58 kilometers in our smartphone data compared to 12.78 kilometers in the travel survey). For non-work stays, the average distances travelled are larger in our smartphone data (10.72 kilometer) than in the travel survey

Figure A.3.5: Share of Non-work Stays by Type in the Travel Survey Data for Tokyo



Notes: Shares of non-work stays by the stated purpose reported by respondents in the travel survey data for the Greater Tokyo metropolitan area in 2008; shares sum to one; the travel survey data are reported for weekdays only (not weekends).

data (7.39 kilometer). This pattern of results could be explained by a systematic underreporting of non-work stays further from home in the travel survey data. As discussed above, our smartphone data measures a "stay" if a user is static for more than 15 minutes, whereas respondents to the travel survey may not think a stop as short as 15 minutes as worthy of recording. If many of these unreported short stops happen relatively far from home along the way to work (e.g., short errands or grabbing lunch or coffee during the working day), this would generate lower average reported distances in the travel survey data than in our smartphone data.

Table A.3.1: Average Distances of Work Stays from Home Locations in our Smartphone Data and Travel Survey Data for Tokyo

	Smartphone	Person trip survey
Work	12.58	12.78
Non-work	10.72	7.39

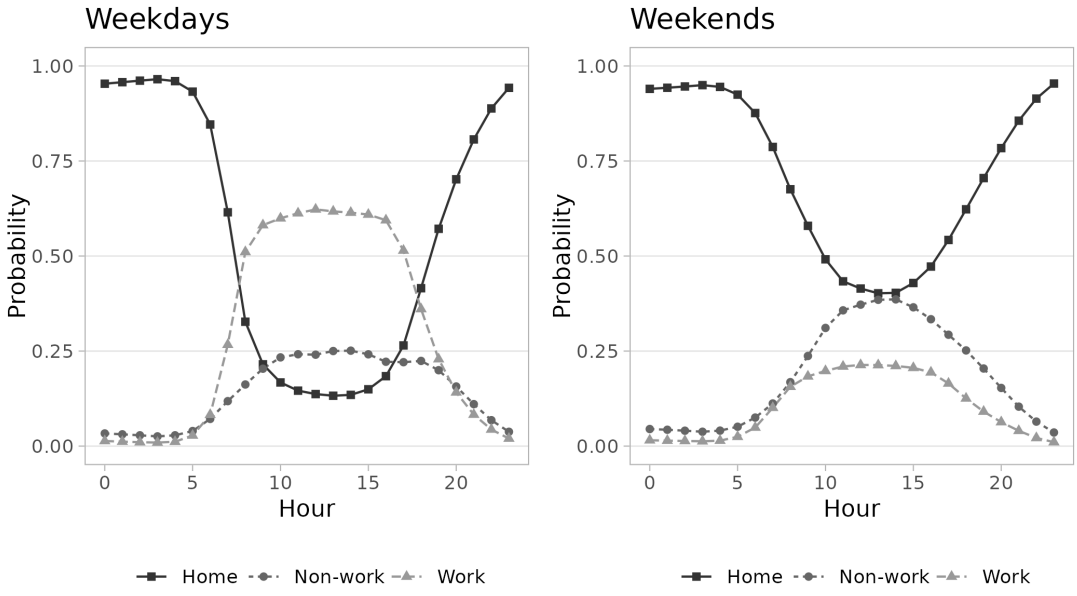
Notes: Average distances in kilometers of work stays and non-work stays from home; the first column reports these average distances in our smartphone data for our baseline sample in 2019; the second column reports these average distances for the travel survey data for 2008.

A.4 Work and Non-Work Stays by Hour

In this section of the Online Appendix, we show that our smartphone data exhibit an intuitive pattern of the average number of work and non-work stays by day and hour of the week, which provides further support for the idea that they accurately capture patterns of spatial mobility within Tokyo.

In Figure A.4.1, we show the average probability that a user stays at home, work or non-work locations by hour, based on the stay closest to the beginning of the hour for our baseline sample in 2019. The three probabilities sum to one, since home, work, and non-work stays are mutually exclusive of one another, and together they sum to the total number of stays. We find that non-work stays are frequent relative to work stays during both weekdays and weekends. Comparing across hours of the day, we find the expected pattern that home stays fall during the daytime (from around 6am-9pm), whereas work and non-work stays both rise. During weekdays, the probability of a stay rises more rapidly during the waking hours for work stays than for non-work stays. During weekends, we find the opposite pattern, with the probability of a stay rising more rapidly during the waking hours for non-work stays than for work stays.

Figure A.4.1: Home, Work and Non-work Stays in our Smartphone Data by Hour



Notes: This figure shows the probability that each smartphone user stays at home, work or non-work locations by each hour of the day for our baseline sample in 2019, where these three probabilities sum to one. To construct this figure, for each user and for each hour of the clock for each day (e.g., at 11am), we measure the user's location as the stay location that has started most recently. We then compute the probability of each type of stay by averaging across days, separately for weekdays and weekends, and for each hour. See Section 3 of the paper for the definitions of home, work and non-work stays.

A.5 Non-commuting Trips by Nontradable Service Sector

We next combine our smartphone data with spatially-disaggregated economic census data to provide evidence on non-commuting stays for different non-traded sectors.

We stochastically assign non-work stays (stays at neither home nor work locations) to different types based on the local economic activity undertaken at each geographical location, as captured by the share of service sectors in employment. For each 500×500 meter grid cell in the Tokyo metropolitan area, we compute the employment share of each service sector in total service sector employment. We disaggregate service-sector employment into the following five categories: "Finance, Real Estate, Communication, and Professional", "Wholesale and Retail", "Accommodations, Eating, Drinking", "Medical and Health Care", and "Other Services".[1] For each non-work stay in a given grid cell, we allocate that stay to these five categories probabilistically using their shares of service-sector employment. If no service-sector employment is observed in the grid cell, we allocate that non-work stay to the category "Z Others." Therefore, if non-commuting trips are unrelated to the availability of nontraded services, our algorithm assigns these stays to "Z Others."

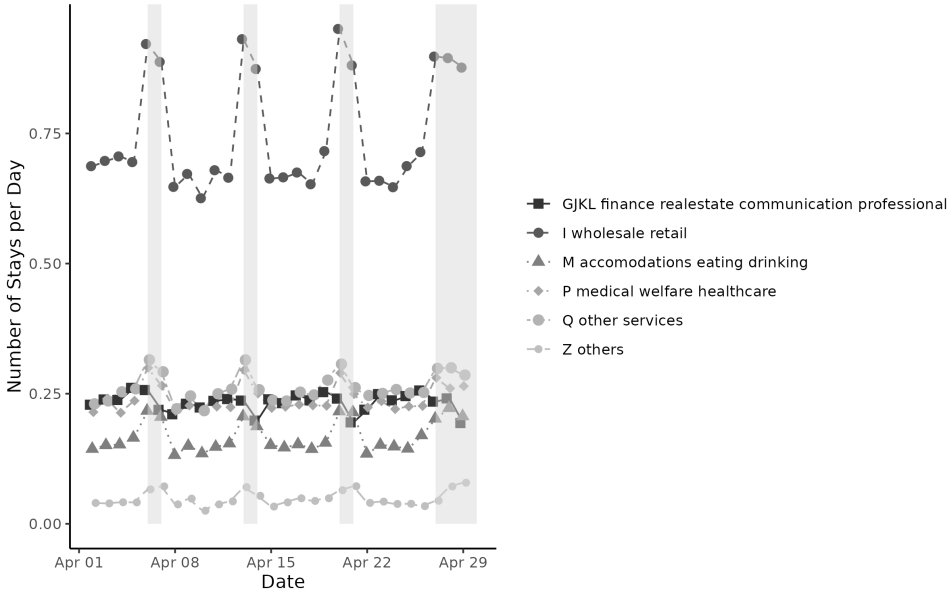
As a check on this probabilistic assignment, Figure A.5.1 displays the density of each type of non-work stay by hour and day, as a share of all stays for the month of 2019. Panel (A) shows results by day of the month. Panel (B) shows results by hour of the day for weekdays and weekends separately. We 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 non-work 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 pattern, consistent with the fact that establishments providing these services are often closed on 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 (18:00) 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 on weekends.

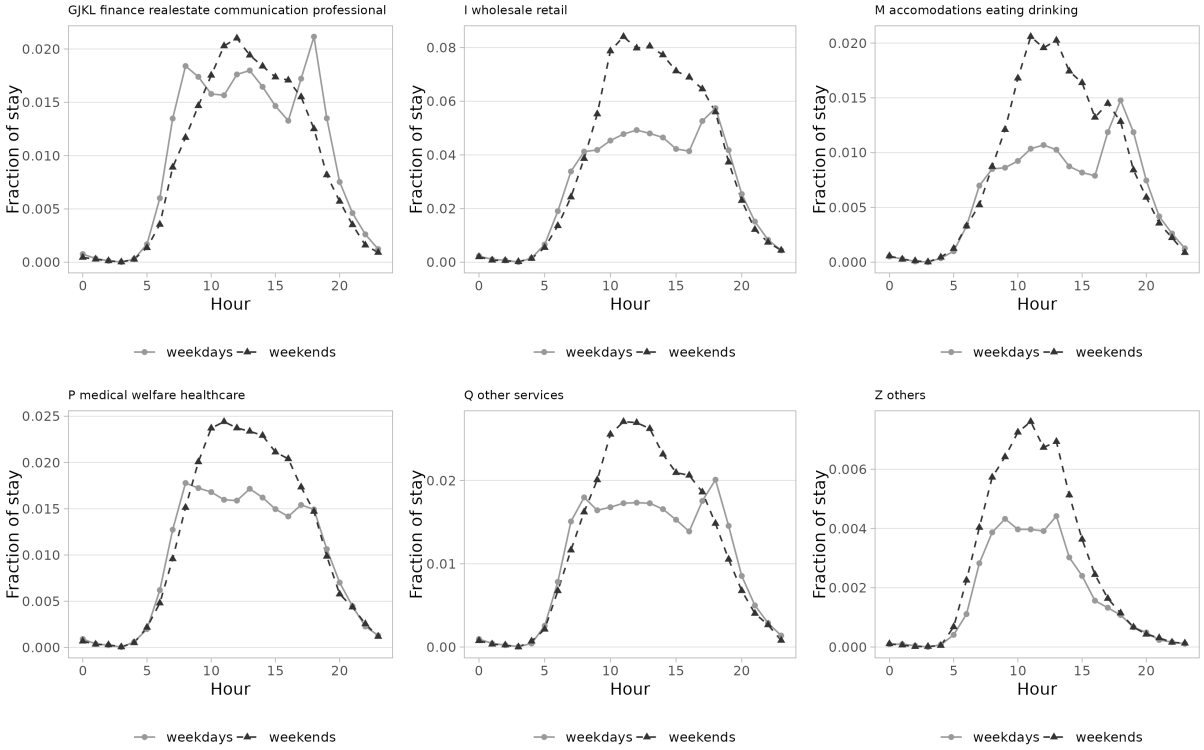
[1] This categorization of service sectors follows the one-digit classification of the Japan Standard Industrial Classification (JSIC), for which we have data available by 500×500 meter grid cells. "Finance, Real Estate, Communication, and Professional" correspond to sectors G, J, K, L; "Wholesale and Retail" corresponds to I, "Accommodations, Eating, Drinking" corresponds to M, "Medical and Health Care" corresponds to P, and "Other Services" corresponds to Q.

Figure A.5.1: Non-work Stays in our Smartphone Data by Type, Day, and Hour

(A) Number of Non-work Stays in our Smartphone Data by Day



(B) Number of Non-work Stays in our Smartphone Data by Starting Hour



Notes: For each 500×500 meter grid cell in the Tokyo metropolitan area, we compute the employment share of each service sector in total service sector employment. For each non-work stay in a given grid cell in 2019, we allocate that stay to these five categories probabilistically using their shares of service-sector employment. If no service-sector employment is observed in the grid cell, we allocate that non-work stay to the category "Z Others." Therefore, if non-commuting trips are unrelated to the availability of nontraded services, our algorithm assigns these stays to "Z Others."

Panels (A) and (B) of Figure A.5.2 report average numbers of these different types of non-work stays for weekdays and weekends separately. We find that "Wholesale and Retail" stays are by far the most frequent, with an average of 0.68 per day on weekdays and 0.90 per day on weekends. As a point of comparison, Panel (B) also reports the share of each individual service sector in overall service-sector employment for the Tokyo metropolitan area as a whole (penultimate column) and the average share of each individual service sector in overall service-sector employment across the 500×500 meter grid cells (final column).

Comparing the two panels, we find that "Wholesale and Retail" stays are substantially more frequent than would be implied by their shares of overall service-sector employment, accounting for 42.9 percent of weekday stays and 45.9 percent of weekend stays, compared to an aggregate employment share of 32.0 percent and an average employment share of 28.7 percent. This pattern of results implies that non-work stays are disproportionately targeted towards locations with relatively high shares of the "Wholesale and Retail" sector in employment, which is consistent with these non-work stays capturing consumption trips.

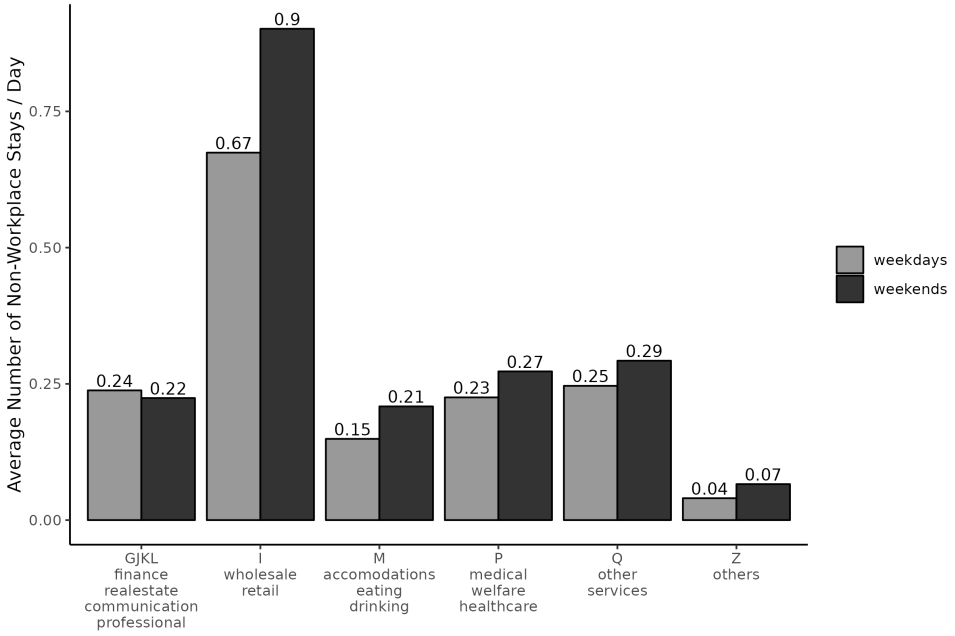
Although "Wholesale and Retail" stays are by far the most frequent, there is considerable variation in the composition of service-sector employment across the locations visited by users, with "Accommodations, Eating and Drinking," "Finance, Real Estate, Communication, and Professional," and "Medical and Health Care" all accounting for around 10 percent or more of the total number of stays. Lastly, "Other" stays are infrequent, consistent with the idea that non-commuting trips often capture the availability of nontraded services.[2]

Overall, our findings in this section for the temporal variation in non-work stays based on the local composition of economic activity in each destination provide support for the idea that non-work stays are closely related to the consumption of non-traded services.

[2]Some of these non-commuting trips could include business-related trips rather than consumption trips (e.g., business meetings, procurement). In Figure A.3.5 of this Online Appendix, we show that business-related trips are a minor fraction (20 percent) of all non-commuting weekday trips using separate travel survey data.

Figure A.5.2: Non-work Stays by Type on Weekdays and Weekends

(A) Average Number of Non-Work Stays by Type



(B) Average Number of Non-work Stays by Type and Service Sector Employment Shares

Industry	Weekdays		Weekends		Employment Share in Service (%)	
	Stays / Day	Share (%)	Stays / Day	Share (%)	Total	Average (500m Grids)
GJKL finance realestate communication professional	0.24	15.1	0.22	11.4	11.9	23.2
I wholesale retail	0.67	42.9	0.90	45.9	32.0	28.7
M accomodations eating drinking	0.15	9.5	0.21	10.6	13.2	13.2
P medical welfare healthcare	0.23	14.3	0.27	13.9	18.7	15.2
Q other services	0.25	15.7	0.29	14.9	24.3	19.8
Z others	0.04	2.6	0.07	3.4		

Notes: Panel (A): Average number of each type of non-work stay per day for weekdays and weekends (excluding stays at home locations) for our baseline sample of users in the metropolitan area of Tokyo in 2019. Non-work stays are allocated probabilistically to each of these five categories using the shares of these service sectors in total service-sector employment, as discussed in the main text. Panel (B) reports the same information in table form, together with the share of each type of stay in the total number of non-work stays, the share of each service sector in total service-sector employment for the Tokyo metropolitan area, and the average share of each service sector in total service-sector employment across the 500×500 meter grid cells. See Section 3 of the paper for the definitions of home, work and non-work stays.

A.6 Stylized Facts about Non-commuting Trips

In this section of the online appendix, we show that patterns of non-commuting trips in our smartphone data for Japan are consistent with existing evidence for non-commuting trips from other empirical settings, as discussed in Section 2 of the paper.

We establish five stylized facts about non-commuting trips in our smartphone data, and show that they are consistent with existing findings from other empirical contexts. In Subsection A.6.1, we corroborate existing evidence that non-commuting trips are pervasive (Fact 1). In Subsection A.6.2, we verify that non-commuting trips are closely related to the availability of non-traded services (Fact 2). In Subsection A.6.3, we establish that non-commuting trips are closer to home than commuting trips (Fact 3). In Subsection A.6.4, we ascertain that non-commuting trips frequently occur as part of travel itineraries (Fact 4). In Subsection A.6.5, we check that non-commuting and commuting trips became less frequent and shorter during the COVID-19 pandemic (Fact 5). We provide further evidence on the characteristics of travel itineraries in Section 4.1 of the paper.

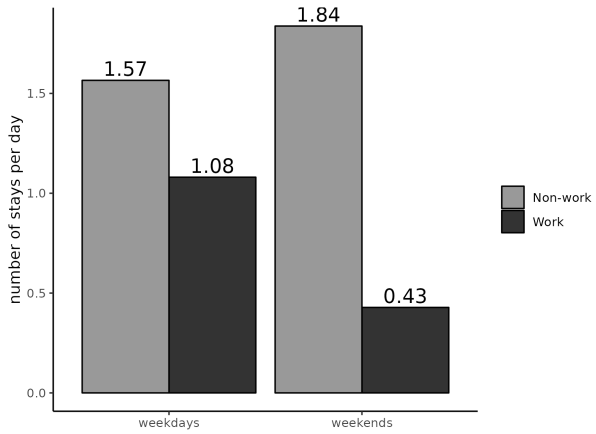
A.6.1 Non-commuting Trips are Pervasive (Fact 1)

Our first stylized fact is that non-commuting trips are pervasive relative to commuting trips. In Figure A.6.1, we display the average number of stays per day for work and non-work locations (excluding home locations) using our baseline sample for the Tokyo Metropolitan Area in 2019. In this figure, the average number of work stays can be greater than one during weekdays, because workers can leave their workplace during the day and return there later the same day (e.g., after lunch elsewhere). Similarly, the average number of work stays can be greater than zero on the weekend, because some workers can be employed during the weekend (e.g., in restaurants and stores). As shown in the figure, even during weekdays, we find that non-commuting trips are more frequent than commuting trips, with an average of 1.57 non-work stays compared to 1.08 work stays per day. This pattern is magnified at weekends, with an average of 1.84 non-work stays compared to 0.43 work stays per day.

These findings from our smartphone data are consistent with those from existing research using travel survey data, including Transportation Research Board (2006) and Agarwal, Jensen and Monte (2020). For example, page 2 of Transportation Research Board (2006) states "The journey to work is only one of a large number of purposes that generate daily travel activity. In 1956, the landmark metropolitan transportation study that ushered in the modern era of transportation studies, the Chicago Area Transportation Study (CATS 1956), identified about two trips per day per capita, of which approximately 40 percent were work trips." In the 2017 National Household Transportation Survey (NHTS) for the United States, more than 67

percent of all trips by privately-owned vehicles were taken for purposes like shopping and errands and recreation. In Online Appendix A.3.2, we show that our finding using our Japanese smartphone data of more frequent non-commuting stays than commuting stays is confirmed in travel survey data for the Tokyo Metropolitan Area, though these travel survey data for Tokyo are only available for weekdays for more aggregated spatial units than those in our smartphone data.

Figure A.6.1: Frequencies of Work and Non-work Stays (Excluding Home Stays)



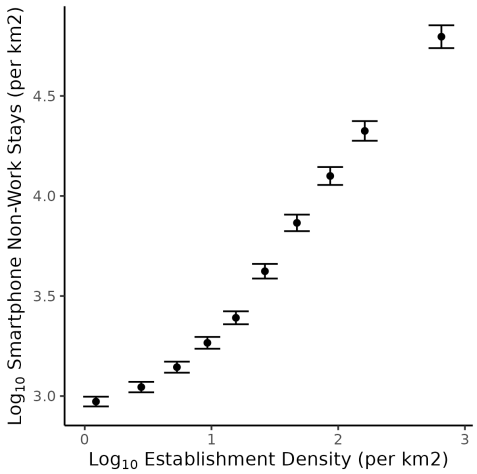
Notes: Average number of work and non-work stays per day for weekdays and weekends (excluding home stays) for our baseline sample users in the metropolitan area of Tokyo in 2019. See Section 3 of the paper for the definitions of home, work and non-work stays.

A.6.2 Non-commuting Trips are Closely Related to Nontradable Services (Fact 2)

Our second stylized fact is that non-commuting trips are closely related to the availability of non-traded services. In Figure A.6.2, we display the logarithm of the number of non-work stays for each 1×1 kilometer grid cell (vertical axis) against the number of establishments in nontraded service sectors in the cell (horizontal axis) using our baseline sample for the Tokyo Metropolitan Area in 2019. The nontraded service sector corresponds to “Finance, Real Estate, Communication, and Professional" and “Wholesale and Retail" (categories G, J, K, L of the Japanese Standard Industrial Classification (JSIC)); “Accommodations, Eating, Drinking" (category I); “Medical and Health Care" (category P); and “Other Services" (category Q). We find a strong, positive and statistically significant relationship. Therefore, locations that are more frequently visited for non-work stays have systematically larger numbers of non-traded services establishments. In Appendix A.5, we use economic census data on employment by sector for 500×500 meter grid cells to probabilistically assign non-work stays to service sectors. We find that the “Retail and Wholesale” sectors are the most frequent category of non-work stay (43 percent of all non-work stays).

These findings are consistent with existing evidence that local foot traffic plays a key role in shaping the performance of non-traded service-sector firms. A strand of empirical research documents that firm bankruptcies and the closure of national retail chains have significant negative externalities on nearby firms, leading to declines in employment and business closures, including Shoag and Veuger (2018); Bernstein et al. (2019); Benmelech et al. (2019); Pollman (2020); and Knight (2022). Using U.S. cellphone data, Qian et al. (2024) show that grocery-store openings lead to significant growth in foot traffic for businesses within 0.1 miles. Using U.S. Visa debit and credit card transactions data, Einav et al. (2021) show that the extensive margin of the number of customers accounts for most sales variation across merchants, across stores within retail chains, and over time for individual merchants and stores. The number of unique customers, as opposed to transactions per customer or dollar sales per transaction, consistently accounts for about 80 percent of sales variation.

Figure A.6.2: Non-work Stays and the Density of Retail Establishments



Notes: The figure plots the logarithm of the number of non-work stays for each 1 kilometer by 1 kilometer grid cell (on the y-axis) against the number of establishments in nontraded service sector in the cell (on the x-axis). The nontraded service sector is defined to include "Finance, Real Estate, Communication, and Professional" (category G, J, K, L of the Japanese Standard Industrial Classification (JSIC)), "Wholesale and Retail" (category I), "Accommodations, Eating, Drinking" (category M), "Medical and Health Care" (category P), and "Other Services" (category Q). See Section 3 above for the definition of non-work stays. Results for our baseline sample of users in the Tokyo Metropolitan Area in 2019.

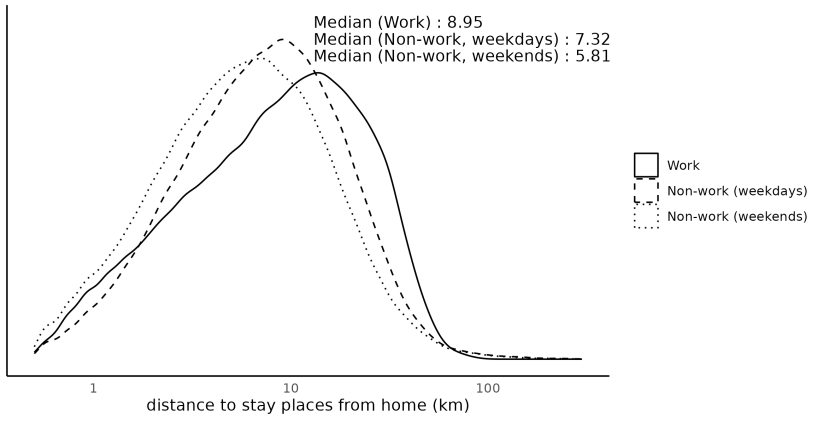
A.6.3 Non-commuting Trips are Closer to Home (Fact 3)

Our third stylized fact is that non-commuting trips are concentrated closer to home than commuting trips. Figure A.6.3 displays the distribution of distances from home for work and non-work stays using our baseline sample for the Tokyo Metropolitan Area in 2019. On weekdays, we find that non-work stays have an average distance from home of 7.32 kilometers compared to a median distance from home of 8.95 kilometers for work stays. This difference is magnified

at the weekend, with a median distance from home of only 5.81 kilometers for non-work stays, which is in line with the idea that users remain closer to home at the weekend.[3]

This pattern of results is consistent with a range of existing evidence of the role of geographical distance in shaping non-traded services consumption. Using travel survey and Google business data, Couture (2016) finds that the median distance travelled to a restaurant is 3 miles and takes 10 minutes, with an average distance travelled of 6 miles and time taken of 14.5 minutes. Using smartphone data for Korea, Oh and Seo (2022) finds that a stronger spatial decay with distance for non-commuting trips than for commuting trips. Using credit card transactions for the United States, Agarwal et al. (2020) find that card holders transact in only a few of the geographically proximate locations, with the median transaction occurring at a distance of about 9 kilometers from a card holder's home location.

Figure A.6.3: Distances of Work and Non-work Stays from Home



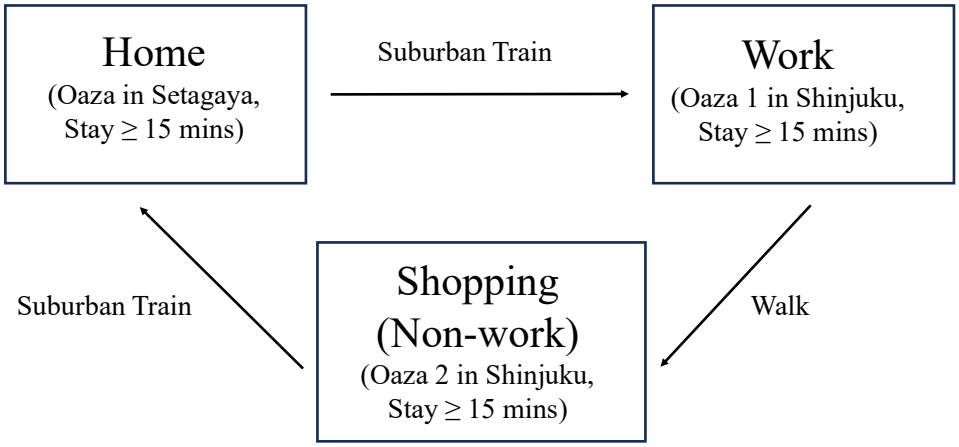
Notes: Distributions of distance in kilometers of work and non-work stays from the home location during weekdays and weekends. See Section 3 above for the definition of home, work and non-work stays. Results for our baseline sample of users in the Tokyo Metropolitan Area in 2019.

[3] In Section 5.7 of the paper, we show that our smartphone data are characterized by the property of extended gravity, such that the probability of travelling from an origin to a destination depends not only on the travel time of the destination from the origin, but also on the travel time of the destination from home and work.

A.6.4 Non-commuting Trips Frequently occur as a part of Travel Itineraries (Fact 4)

Our fourth stylized fact is that non-commuting trips frequently occur as part of travel itineraries, which we define as a journey starting and ending at home that can include more than one intermediate stay on a given day (see Figure A.6.4 for an example).[4]

Figure A.6.4: Example of a Travel Itinerary

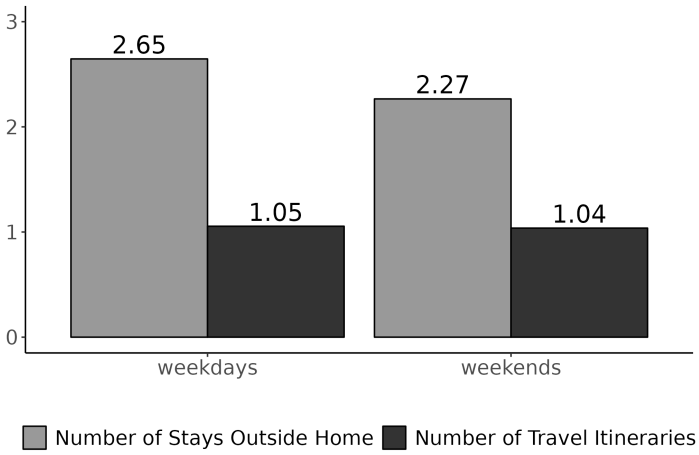


In Figure A.6.5, we display the average number of stays outside home and the average number of travel itineraries per day. On weekdays (weekends), we find that users make 2.65 (2.27) stays outside the home per day, and undertake 1.05 (1.04) travel itineraries. This pattern of results suggests that the typical user's travel behavior involves one travel itinerary per day, stopping at intermediate destinations along the way. In Section 4.1, we provide further evidence on the properties of travel itineraries beyond their prevalence, including the frequency distributions of the number of stays and the sequence of different types of stays.

Our findings for the empirical relevance of travel itineraries are consistent with a range of existing evidence in the transportation and urban economics literatures. Early studies of individual travel behavior documented the prevalence of multi-stop travel, sometimes referred to as travel itineraries or trip chains, including O'Kelly and Miller (1984) and Oster (1973), as reviewed in Thill and Thomas (1987). More recent research has examined the determinants of trip chaining and its implications of the location decisions of non-traded services firms, including Anas (2007), Primerano et al. (2008), and Oh and Seo (2023). More broadly, travel itineraries and trip chains can be viewed as a form of the classical travelling salesman problem

[4]This definition is closest to "Travel Day Trip" rather than "Trip Chain" in the National Household Travel Survey (NHTS) for the United States, which restricts the use of the word "Trip Chains" to journeys between an origin and destination that have an intervening stop of 30 minutes or less.

Figure A.6.5: Frequency of Travel Itineraries



Notes: Number of stays per day and number of travel itineraries per day per user, disaggregated by weekdays and weekends. Travel itineraries are defined as a journey starting from and ending at home that can include more than one intermediate stay on a given day. See Section 3 above for the definition of home, work and non-work stays. Results for our baseline sample of users in the Tokyo Metropolitan Area in 2019.

of finding the shortest possible route between a set of locations, which involves visiting each location once and returning to the origin location.

A.6.5 Both Non-commuting and Commuting Trips Became Less Frequent and Shorter During the early period of COVID-19 Pandemic (Fact 5)

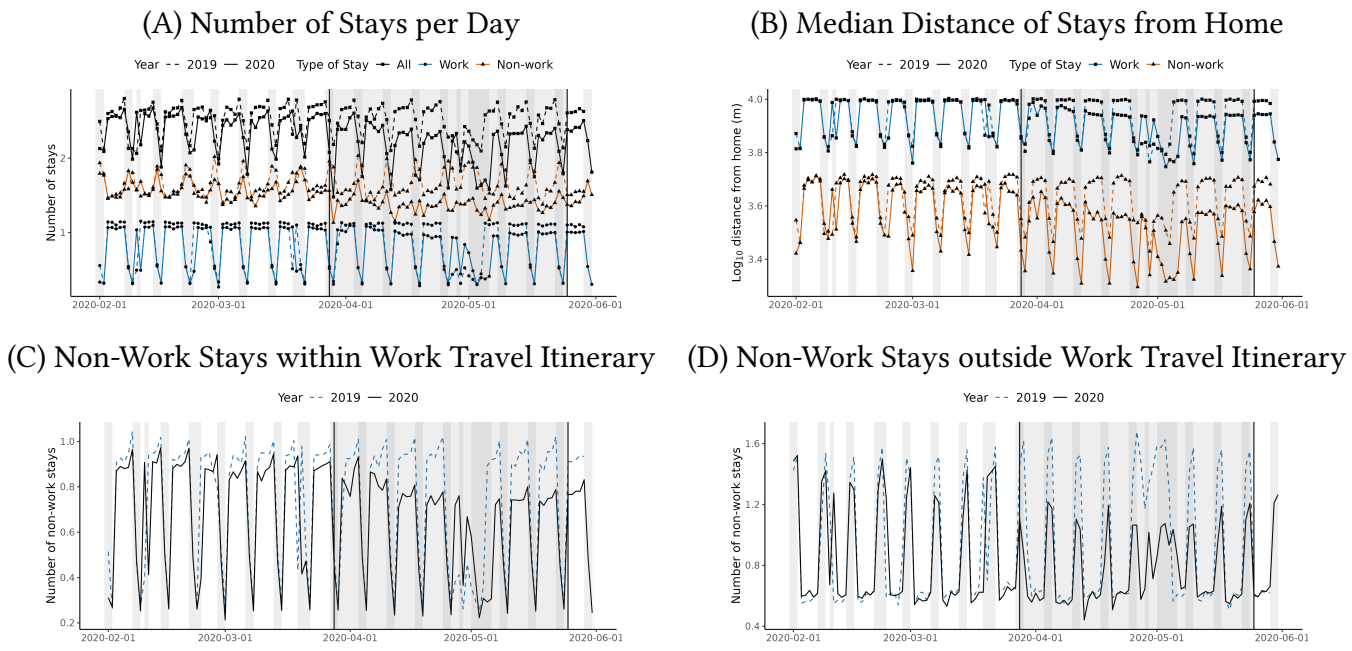
The outbreak of the COVID-19 pandemic, starting in early 2020 in Japan, had a dramatic impact on spatial mobility in cities around the world. In an effort to prevent a surge of cases, national and city governments in many places imposed lockdowns and mobility restrictions. Even in places where there were no explicit mobility restrictions, people chose on their own initiative to make fewer trips to reduce the risk of infection. Our fifth stylized fact about patterns of spatial mobility in our smartphone data is a decline in the frequency and distance of non-commuting and commuting trips during the COVID-19 pandemic.

COVID-19 in Japan Compared to other developed countries, Japan experienced relatively low infection rates. Partly because of this, Japanese cities did not impose a strict form of lockdown. However, governments issued multiple waves of "emergency orders" that asked residents to stay home unless travel was "absolutely necessary." In response to the spread of COVID-19 to Tokyo prefecture, the first emergency order in Tokyo was announced on March 28, 2020, and was temporarily lifted on May 25, 2020. Therefore, we define the period from March 28, 2020 to May 25, 2020 as the period when Tokyo residents were asked to stay home.

Frequency and Length of Travel Figure A.6.6 shows the frequency of trips and length of travel from February-May 2020 (including the period of the emergency order) compared to the same months in 2019 (before the COVID-19 pandemic). Panel A shows the number of stays outside the home location per day; Panel B gives the median distance of these stays from the home location; Panel C reports the number of non-work stays away from home that occur within travel itineraries including work; Panel D displays the number of non-work stays away from home that take place outside travel itineraries including work. In all figures, the dashed lines correspond to 2019, and the solid lines represent 2020. In Panels A and B, the red lines denote work stays, whereas the blue lines indicate non-work stays. In Panel A, the green lines represent all stays away from home (work plus non-work).

In Panel A, we find a reduction in the total number of stays away from home during the period of the emergency order, which is driven by a fall in both commuting and non-commuting trips. In Panel B, we observe a decline in the distance travelled for both work and non-work stays, which is somewhat larger for non-work stays. In Panels C and D, we show that this decline in non-work stays occurs both within and outside travel itineraries including work.

Figure A.6.6: Commuting and Non-Commuting Trips from February-May in 2020 (including the period of the Emergency Order) and 2019 (before the COVID-19 Pandemic)



Notes: Each figure shows the average number of stays away from home for different types of stays (Panels A, C and D) and the median distance from home (Panel B) for each day from February-May in 2020 (including the period of the emergency order) and in 2019 (before the COVID-19 pandemic). For the 2019 series, we start from February 2nd (instead of February 1st) to align the day of the week between 2019 and 2020. The dark shaded area indicates the period when there was an emergency order that discouraged people from travelling within the Tokyo Metropolitan Area (March 28, 2020 to May 25, 2020). Lighter shaded days are weekends and holidays. See Section 3 above for the definition of home, work and non-work stays.

Evidence from Other Empirical Settings Our findings for Tokyo are consistent with evidence of changes in spatial mobility during the COVID-19 pandemic from other empirical settings. Using U.S. smartphone data, Couture et al. (2022) construct a location exposure index (LEX) that describes county-to-county movements and a device exposure index (DEX) that quantifies the exposure of devices to each other within venues. Following the onset of the pandemic, there is sharp decline in geographical mobility, with distances travelled between states according to the LEX measure declining by around 40 percent by April 2019 relative to their pre-pandemic values.

Using pre-pandemic questionnaire data from the O*NET database, Dingel and Neiman (2021) develop a measure of the feasibility of working from home (WFH) for each occupation. Using this measure of the feasibility of working from home, Althoff et al. (2022) show that the centers of large cities were most exposed to the shift to remote working during and after the COVID-19 pandemic. First, these large cities have the biggest concentrations of business services jobs that can be done remotely. Second, business service workers left their big-city residences during the COVID-19 pandemic to work from elsewhere, as remote work became more prevalent. Third, this change in business service workers' location decisions led to a large decline in visits to establishments in the non-traded services sector.

A.7 Additional Spatial and Temporal Patterns of Travel Itineraries

In this section, we provide evidence on the spatial and temporal patterns of travel itineraries in our smartphone data. We examine the robustness of our findings for the properties of travel itineraries in Section 4.1 of the paper to the time threshold used to measure stays and the level of spatial aggregation used to measure stays.

Table A.7.1 reports the frequency distribution of the number of stays (including both work and non-work stays) outside the home across travel itineraries. Panel (A) defines stays using our baseline time threshold of 15 minutes or more and 250m $\times$ 250m grid cells, which corresponds to Panel (B) of Table 1. Panel (B) defines stays using a time threshold of 20 minutes or more, Panel (C) defines stays using a time threshold of 30 minutes or more, and Panel (D) aggregates stays at the Oaza level, such that multiple non-work stays within an Oaza are counted as a single non-work stay, using our baseline time threshold of 15 minutes or more. While the number of stays mechanically goes down in Panels (B)-(D), the main patterns of the frequency distribution remain the same as in Panel (A).

In Figure A.7.1, we present evidence on the spatial patterns of non-work stays relative to the home location, work location, and previous stays within the same travel itinerary. Specifically, we show the median straight-line distances between non-work stays and home locations, work locations, and previous stay locations, by hour of the day, and by weekdays (left panel)

Table A.7.1: Number of Stays within a Travel Itinerary (Robustness to Stay Duration and Spatial Unit)

(A) Baseline (15 minutes cutoff; 250m × 250m mesh)

Number of Stays	All (%)	Weekdays (%)	Weekends (%)
1	42	40	47
2	25	25	25
3	14	14	13
4	7	8	6
5+	12	12	8

(B) 20 minutes cutoff

Number of Stays	All (%)	Weekdays (%)	Weekends (%)
1	45	44	50
2	25	25	25
3	13	13	12
4	7	7	6
5+	10	11	7

(C) 30 minutes cutoff

Number of Stays	All (%)	Weekdays (%)	Weekends (%)
1	50	48	55
2	25	25	25
3	12	13	11
4	6	7	5
5+	4	7	4

(D) Aggregate stays at Oaza

Number of Stays	All (%)	Weekdays (%)	Weekends (%)
1	47	45	51
2	29	29	27
3	13	14	12
4	6	6	5
5+	5	6	5

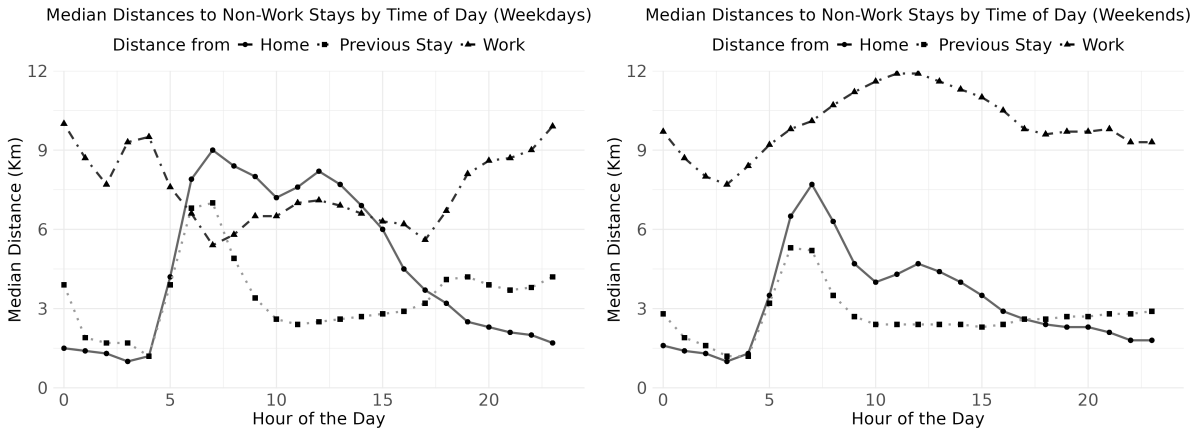
Note: Frequency distribution of the number of stays (including both work and non-work stays) outside the home location across travel itineraries; Panel (A) defines stays using our baseline time threshold of 15 minutes or more and 250m × 250m grid cells, which corresponds to Panel (B) of Table 1; Panel (B) defines stays using a time threshold of 20 minutes or more; Panel (C) defines stays using a time threshold of 30 minutes or more; and Panel (D) aggregates stays at the Oaza level, such that multiple non-work stays within an Oaza are counted as a single non-work stay, using our baseline time threshold of 15 minutes or more.

and weekends (right panel).

We find an intuitive pattern that is consistent with the predictions of our model of travel itineraries and the spatial clustering. On weekdays (left panel), the distances to home locations

are low at night and increase during work hours. In contrast, the distances to work locations are high at night and decrease during work hours. These patterns are consistent with the idea that during weekdays people tend to concentrate non-work stays around work. Distances to the previous stay within the itinerary are generally shorter during the daytime and are elevated at night, which is consistent with the idea that people finish travel itineraries at the end of the day, travelling back home. On weekends (right panel), non-work stays tend to be closer to home and farther from work on average than on weekdays. These patterns are consistent with the idea that people tend not to go to work on weekends and instead concentrate non-work stays around home.

Figure A.7.1: Median Distance of Non-work Stays from Home, Work, and Previous Stays in the Same Travel Itinerary, by Hour and Day Type

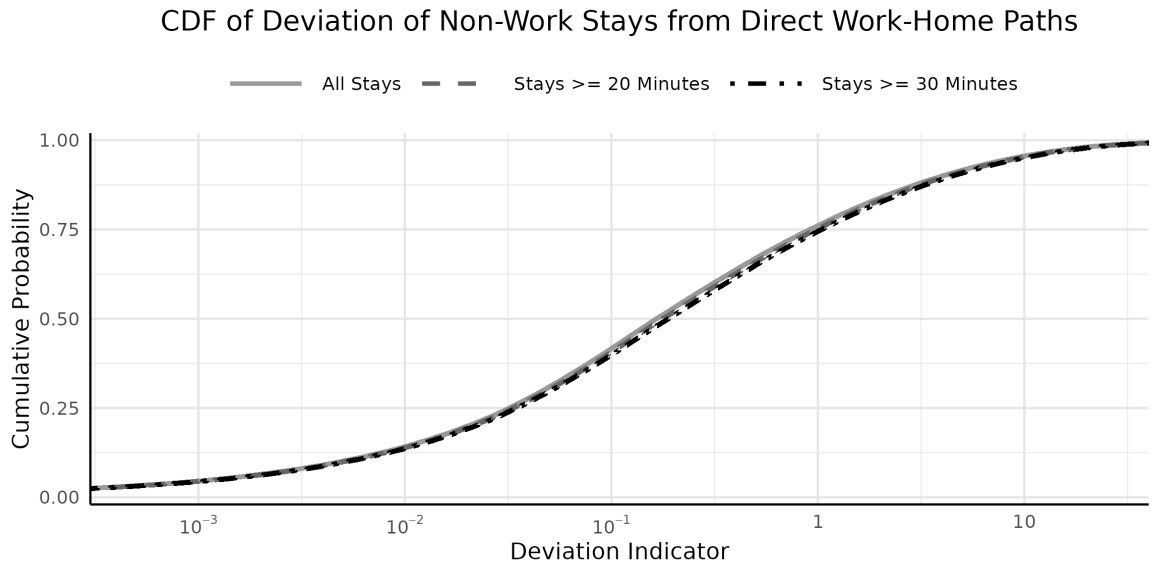


Note: Median straight-line distances of non-work stays from home locations, work locations, and previous stays within the same travel itinerary, by hour of the day, and by weekdays (left panel) and weekends (right panel).

In Figure A.7.2, we show the spatial patterns of non-work stay locations relative to the straight-line path between the home and work locations, using our baseline time threshold for measuring stays of 15 minutes or more and 250m $\times$ 250m meter grid cells. Specifically, for all non-work stays that occur as a part of the travel itinerary involving the work location, we compute the deviation from a straight-line path, defined as follows. For each non-work stay, we compute the straight-line distance between (a) the non-work stay and home, (b) non-work stay and work, and (c) home and work. We construct the deviation from a straight-line path as $\frac{(a)+(b)-(c)}{(c)}$. By the Pythagorean theorem, this ratio takes a value weakly greater than zero, and it is equal to zero if the non-work stay is exactly on the straight-line path between home and work. We compute this ratio using straight-line distances, instead of travel time, because it is not computationally feasible to compute the bilateral travel time matrix between the millions of bilateral pairs of 250 meter grid cells observed in our data.

We find that the distribution of these deviations from straight-line paths is relatively dispersed. Around 30 percent of this distribution has deviations of 0.05 or less, indicating that many travel itineraries have only minor deviations from the straight-line path. This finding is in line with our extended gravity regression results in Table 3 in the paper, which shows that non-work stays gravitate towards both home and work. At the same time, we find that the 70th percentile of this distribution has a deviation of 0.61, and the 90th percentile has a deviation of 7.4, indicating that some non-work stays occur at substantial distances from the straight-line path. Our model of travel itineraries captures this substantial heterogeneity in the spatial pattern of non-work stays through the idiosyncratic preference shocks. Finally, we find a relatively similar distribution of these deviations from straight-line paths across alternative time thresholds for measuring stays of 15 minutes (solid line), 20 minutes (dashed line) and 30 minutes (dashed-dotted line).

Figure A.7.2: Deviation of Non-work Stays from the Straight-line Path Between Home and Work Locations



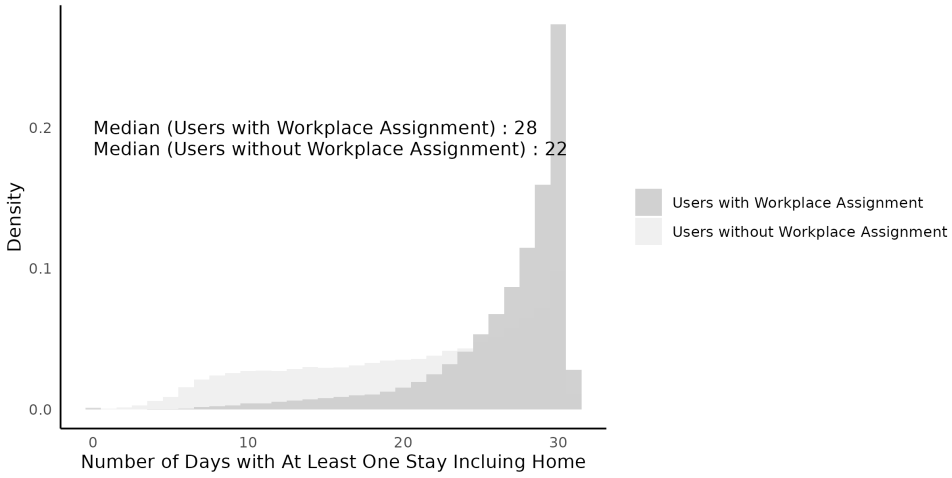
Note: For all non-work stays that occur as a part of a travel itinerary including work, we compute the deviation from a straight-line path between home and work, defined as $\frac{(a)+(b)-(c)}{(c)}$, where (a)-(c) correspond to the straight-line distance between (a) the non-work stay and home, (b) non-work stay and work, and (c) home and work. The figure shows cumulative distribution functions of these deviations using time thresholds for measuring stays of 15 (baseline), 20, and 30 minutes.

A.8 Patterns of Users without Workplace Assignment

In this section of the Online Appendix, we provide evidence on users whose workplace is not assigned according to the proprietary algorithm discussed in Section 3.1 of the paper.

We first show that smartphones without workplace assignment have significantly fewer numbers of active days (even at home locations). This pattern of results is consistent with the idea that one reason why the algorithm is not able to assign a workplace for these devices is that they are infrequently used. As discussed in Section 3.1 of the paper, after a user has downloaded the mapping application and given consent, location information is collected regardless of what application the user has open, as long as the device is turned on. Therefore, the number of days with any stays is a proxy for how actively the device is used (i.e., whether it is regularly turned on, or whether the user brings the device with them when they travel). In Figure A.8.1, we show the distribution of the number of days that we observe any stays in April 2019, including stays at home. We show this distribution separately for users with and without workplace assignment. We find a median number of active days of 22 for users without workplace assignment, which is substantially smaller than the median of 28 for users with workplace assignment. Therefore, smartphones without workplace assignment are substantially more infrequently used.

Figure A.8.1: Number of Days per Month with at Least One Stay for Smartphone Users With and without Workplace Assignment



Notes: Distribution of the number of days that we observe any stays in April 2019, including stays at home; lighter shading shows this distribution for smartphone users without workplace assignment; darker shading show this distribution for smartphone users with workplace assignment; the maximum possible number of days is 31 (instead of the number of calendar days in April of 30) because we count travel days to start at 4.00am and end at 3:59am.

We next show that the probability of workplace not being assigned is uncorrelated with the characteristics of smartphone user's municipality of residence. This pattern of results suggests that dropping these users without workplace assignment does not introduce any bias into our measures of residential population shares, which is also consistent with our findings above that our smartphone data closely replicate commuting patterns in official census data.

In Figure A.8.2, we display the share of users with missing workplace assignment for each municipality of residence against municipality characteristics, including population density, the share of university graduates, average income, average age, the distance to the Central Business District (centroid of Chiyoda Ward), and unemployment rates. The dots in each figure represent the average shares of users with missing workplace assignment for each decile of the characteristics on horizontal axis. The line segments indicate the 95 percent confidence intervals. We find little evidence of any relationship being the probability of missing workplace assignment and these municipality characteristics. There is a mild decreasing pattern for the share of university graduates and average income, but the magnitudes are not quantitatively large. Therefore, we find little evidence of any bias in the probability of missing workplace assignment along these dimensions of observable residential characteristics.

A.9 Additional Data Sources

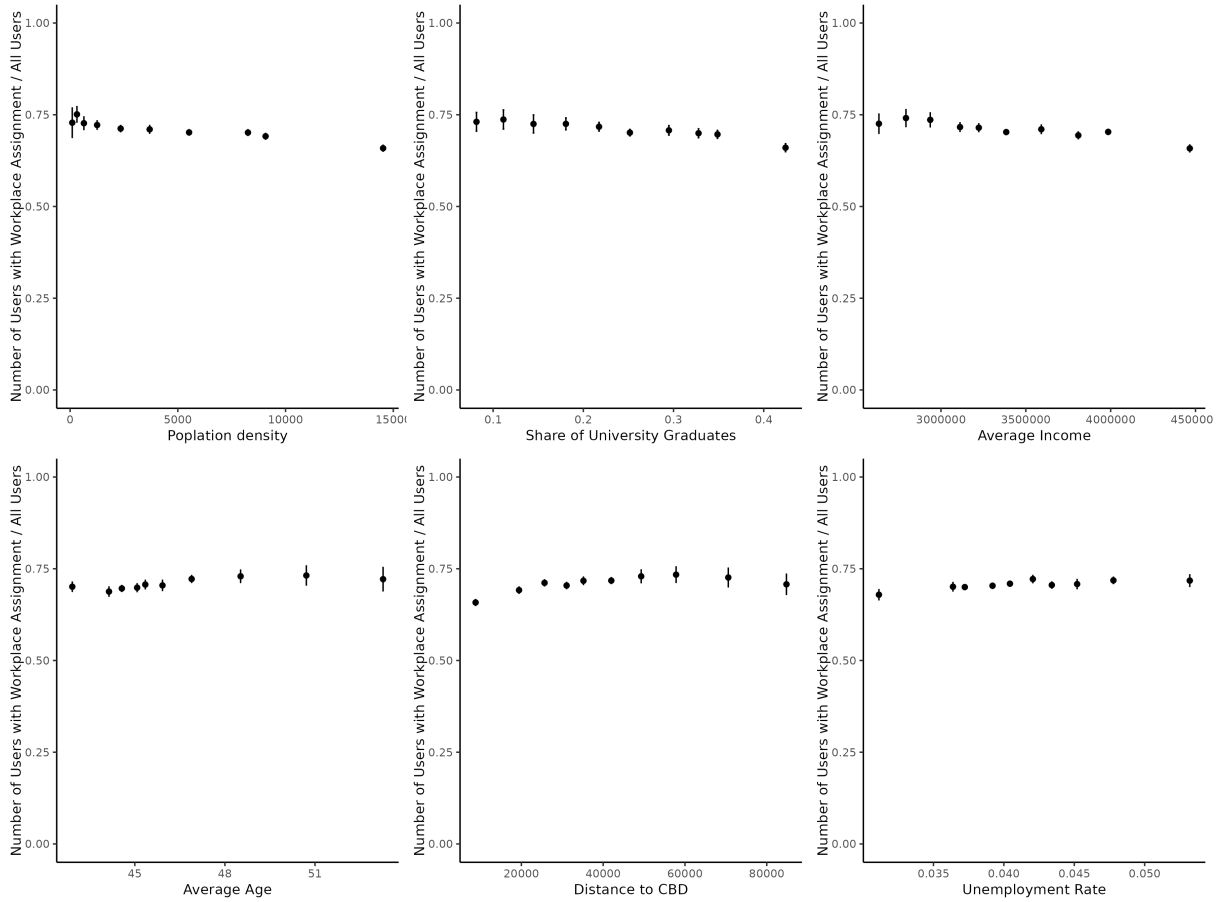
Population Census We measure residential population, employment by workplace and bilateral commuting flows using the 2015 population census, which is conducted by the Statistics Bureau, Ministry of Internal Affairs and Communications every five years. Residential population and total employment by workplace are available for 250-meter grid cells in the Tokyo metropolitan area. Bilateral commuting flows are reported between pairs of municipalities. Data on the residential population by age cohort is also available at the municipality level. We also use the 2010 population census to measure the residential population by the highest completed educational attainment. These data are available on e-Stat, the portal site for Japanese government statistics provided by the Statistics Bureau of Japan.

Economic Census We use data from the Economic Census on total employment by workplace and the number of establishments by two-digit industry for each 500-meter grid cell in the Tokyo Metropolitan Area in 2009, 2012, 2016, and 2021. The data is provided by the Center for Spatial Information Science, the University of Tokyo. We also use data on total revenue and factor inputs that are available at the municipality level in 2016 from e-Stat.

Retail Price Survey We use data from the Retail Price Survey in 2019 to proxy price indexes for each municipality. The data are available on e-Stat.

Household Expenditure Survey We use the Household Expenditure Survey to measure the expenditure share for each spending category. This data covers all of Japan and measures the per-household expenditure for households with two or more members across disaggregated spending categories. The data are available on e-Stat.

Figure A.8.2: Shares of Smartphone Users with Missing Workplace Assignment for each Municipality of Residence Against Observable Characteristics of these Municipalities of Residence



Notes: Vertical axis shows the share of smartphone users with missing workplace assignment by municipality of residence (242 municipalities in the Tokyo metropolitan area) in April 2019; horizontal axis shows observed characteristics of these municipalities of residence; the dots represent the average share of users with missing workplace assignment for each decile of the characteristic of the municipality of residence on the horizontal axis; the line segments indicate the 95 percent confidence intervals.

Basic Population Registration We use the Basic Resident Registration to measure the number of inmigrations and outmigrations. Data are available for each municipality in the Tokyo metropolitan area for the period from 2017-2023. The data provides the number of residents in January and the number of people who moved into/from the municipality during the 12 months from January to December for each year. The data is available on e-Stat.

Geographical Information Data on railway lines, railway stations, bus stops, and bus routes from the National Land Numerical Information are used to compute least-cost path distances and travel times between bilateral pairs of locations. We use geographical information systems (GIS) data on railway lines, railway stations, bus stops, and bus routes, provided

in shapefile format. Additionally, for each station and route, information on the establishment date is available. We also use the boundary data at the municipality level in shapefile format to calculate the area of the municipality.

Bilateral Travel Time We compute the least-cost path distances and travel times between the centroids of each of our locations using the ArcGIS Network Analyst. We assume the following speeds for each mode of transport: 80 meters per minute for walking, 150 meters per minute for a bus, and 600 meters per minute for rail. When changing between rail and bus or between rail lines, we assume a slower walking speed of 40 meters per minute to allow for climbing stairs and navigating busy intersections.

Travel Survey Data We use Person Trip Surveys conducted in 2008 for the detailed travel patterns of the residents in the Greater Tokyo Metropolitan Area (Tokyo, Saitama, Chiba, Kanagawa, and a part of Ibaraki prefectures). We use the data on where and when the trip starts and ends, and the purpose of the trip for each trip segment. We also use the work and home address information for each respondent. The data is provided by the Center for Spatial Information Science, the University of Tokyo.

Office/Commercial Building Data The office and commercial building market data for Central Tokyo (the five central wards: Chuo, Chiyoda, Minato, Shinjuku, and Shibuya) are available from Miki Shoji Co., Ltd. The data is available on the company's website.

We also use office market data from Sanko Estate Company Ltd., which is available for both central Tokyo and its suburbs. Data are available by month from January 2019 to September 2023. Locations are aggregations of Oaza in Central Tokyo and municipalities in the suburbs of Tokyo. Data are reported on the average office rent (measured as the average rent for current office listings) and the average office vacancy rate (measured as the ratio of the currently vacant floor area to the total floor area available for leasing for office use). The data is provided by the company.

Phone Directory We use the telephone directory named Telepoint provided by Zenrin Co., Ltd. The directory contains information on each establishment, including the business name, telephone number, and address. The dataset is available for the years 2018, 2019, 2020, 2021, and 2023. The data is provided by the Center for Spatial Information Science, the University of Tokyo.

Large Retail Store Listing We use the 2019 Edition of the Large Retail Store Data published by Toyo Keizai Shimposha, which provides information on the locations and closing dates of

retail stores with a sales floor area of 1,000 meters squared or more that were present in the Tokyo metropolitan area from 2015 to 2019. The data is provided by the Center for Spatial Information Science, the University of Tokyo.

B Appendix for Working from Home (WFH)

In this section of the Online Appendix, we provide further evidence on the shift to WFH in Japan, supplementing the material in Section 4.2 of the paper. Although the size of the shift to WFH is smaller in Japan than in the United States, it remains substantial and exhibits similar patterns as observed in the United States. In the paper, we show that our estimated model of travel itineraries is quantitatively successful in matching this change in patterns of spatial mobility in Tokyo following the shift to WFH. In contrast, a conventional urban model that abstracts from travel itineraries is unsuccessful in capturing this change in patterns of spatial mobility. In particular, the conventional urban model is unsuccessful in capturing the decline in non-work stays in Central Tokyo, because it assumes that all consumption occurs through direct trips from home, regardless of where work takes place.

B.1 WFH in a Cross-Country Perspective

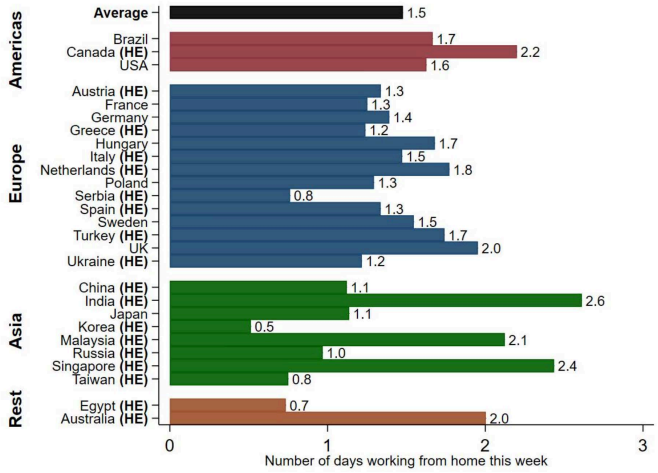
We begin by briefly summarizing existing evidence on the shift to WFH in the United States and other countries in the aftermath of the COVID-19 pandemic. In the year of 1965, the share of full days worked from home (WFH) in the United States was less than 0.5 percent of all paid workdays. In the wake of innovations in communications technology, such as email, the internet and videoconferencing, this share had risen to around 7 percent immediately before the COVID-19 pandemic. During the pandemic itself, social-distancing and the fear of disease transmission led to a large-scale shift to WFH, which has persisted in the aftermath of the pandemic, even after the development of widespread effective vaccines and the end of social distancing requirements. In June 2023, the share of full days worked from home in the U.S. was 28 percent of all paid workdays in the United States, namely around four times larger than immediately before the COVID-19 pandemic in 2019.[5]

Aksoy et al. (2022) report the results of a large-scale survey of WFH in 2022 across 27 countries, including Japan. Figure B.1.1 reproduced from that paper summarizes survey responses to the question "How many full paid days are you working from home this week?" Responses range from 0 to 5+ days per week. The figure shows conditional mean responses, which are

[5]For a review of the evidence on WFH in the United States, see Barrero et al. (2023). For estimates of the feasibility of occupations working from home in the United States, see Dingel and Neiman (2020). For evidence of the impact of this shift to WFH in US cities, see Ramani and Bloom (2021), Althoff et al. (2022), Gupta et al. (2022), and Monte et al. (2023).

obtained from the estimated coefficients on country dummies in an OLS regression, where the U.S. is the excluded category. The regression includes controls for gender, age groups, education groups, industry sectors, and survey wave. On average across the 27 countries, around 1.5 days are worked from home each week in 2022. We find that Japan's experience is broadly in line with that of many other countries, with an average of 1.1 days worked from home each week in 2022.

Figure B.1.1: Cross-Country Survey Evidence on WFH



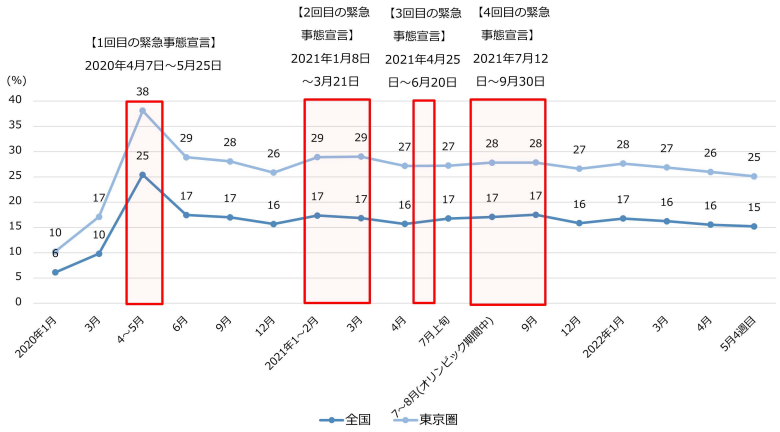
Notes: The figure is from Aksoy et al. (2022). The figure shows responses to the question "How many full paid days are you working from home this week?" Response options range from 0 to 5+ days per week. The figure reports conditional mean responses, which are obtained from the coefficients on country-level dummies in an OLS regression, treating the raw U.S. mean as the baseline. The regression controls for gender, age groups (20-29, 30-39, 40-49, 50-59), education groups (Secondary, Tertiary, Graduate), 18 industry sectors, and survey wave. The survey for Japan was conducted from January 27 to February 4, 2022.

B.2 WFH in Japan

We next provide further evidence on the evolution of WFH in Japan over time. Figure B.2.1 displays a time-series for the percentage of days worked from home in Japan, using survey data from Okubo and NIRA (2022). This survey is conducted on monitors affiliated with Nikkei Research and includes 10,000 employed people aged 15 years and older in Japan. Respondents are selected to be representative in terms of gender, age (6 categories), and region (5 categories).

Consistent with the idea that WFH is more prevalent in the traded-services occupations that are concentrated in large cities (e.g., Dingel and Neiman 2020, Althoff et al. 2022), the percentage of days worked from home is higher for Tokyo (Sky Blue) than for Japan as a whole (dark blue). During the COVID-19 pandemic, WFH in Tokyo increased sharply from 10 percent immediately beforehand to a peak of 38 percent. In the aftermath of the pandemic, with the development of widespread effective vaccines and the ending of stay-at-home orders,

Figure B.2.1: Work from Home in Japan



Notes: This figure is from Okubo and NIRA (2022). The sky blue line shows the average for Tokyo. The dark blue line shows the average for Japan as a whole. The red shaded area indicates the periods when there was an emergency order that discouraged people from travelling within the Tokyo Metropolitan Area.

the share of days worked from home falls from this peak. Nevertheless, it remains persistently higher than before the pandemic at around 25-28 percent in 2022. Therefore, this time-series evidence for Tokyo suggests a substantial change in patterns of spatial mobility following the shift to WFH. In our counterfactual simulation of WFH in Section 7.1, we reduce the fraction of commuting days (ξ) by 20 percent ($= 1-(1-0.28)/(1-0.10)$) following the survey evidence reported in Okubo (2022).

B.3 Commercial Activity in Central Tokyo

We next show that the shift to WFH has led to a large-scale reduction in commercial activity in Central Tokyo. We use data published by an office real estate agent company: Miki Shoji Company Ltd. The sample frame is major rental office buildings with a floor area of $33,000$ meters squared in Tokyo's central business district (the five central wards of Tokyo). Data are reported on the total floor area, the vacancy rate and the average rent for offices in these five central wards.

In our theoretical model, the shift to WFH leads to a reduction in work trips to the central city for workers employed in traded sectors, which reduces the demand for commercial floor space in the central city. Additionally, this reduction in work trips for workers employed in traded services leads to a collapse in demand for non-traded services that are consumed as part of travel itineraries on the way to and from work. This collapse in demand for non-traded services further reduces the demand for commercial floor space in the central city. Finally, the decline in the availability of non-traded services in the central city reduces its attractiveness to workers in both traded and non-traded sectors, further depressing the demand for commercial

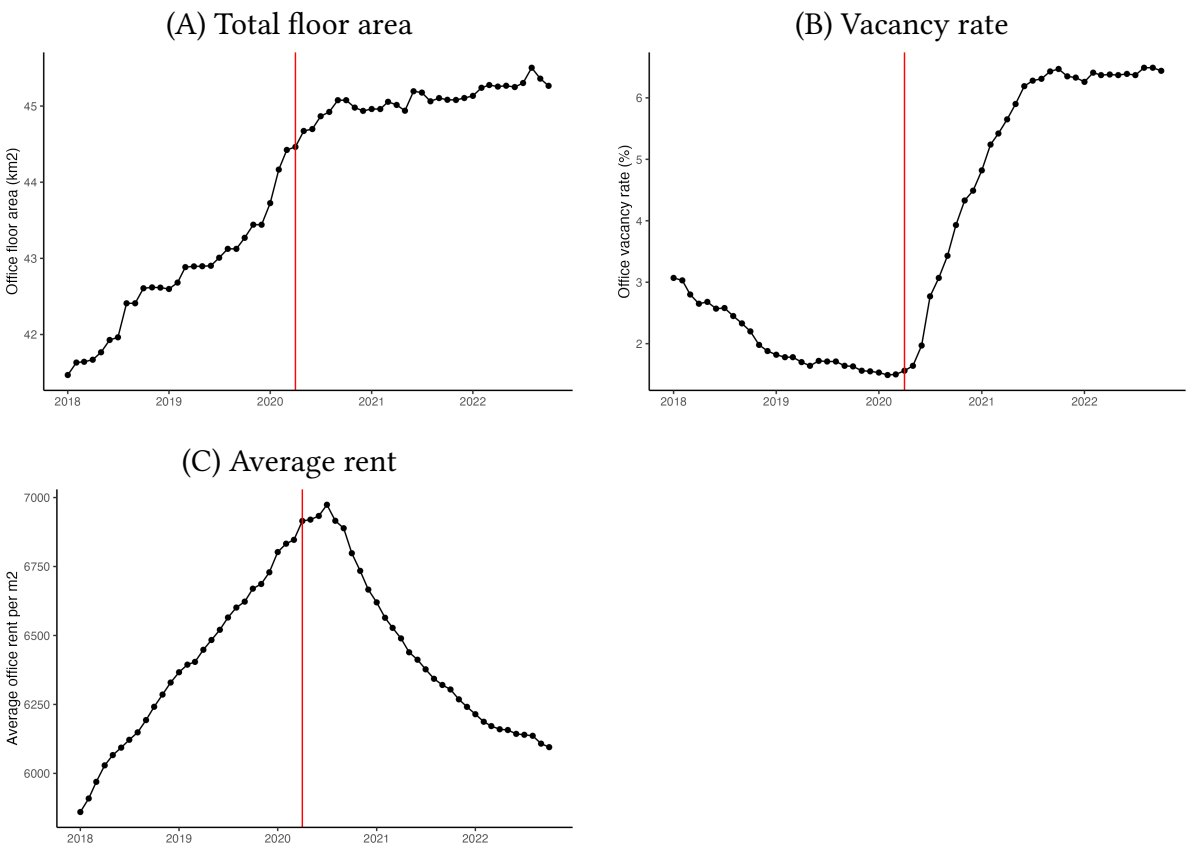
floor space. The supply of existing floor space is inelastic in the short run. Therefore, in the face of this decline in demand, we expect an increase in vacancy rates and/or rents for existing commercial floor space, and a slowdown in new construction.

In Figure B.3.1, we show the total floor area (Panel (A)), the vacancy rate (Panel (B)) and the average rent (Panel (C)) for offices in the five central wards of Tokyo over time (Chiyoda, Chuo, Minato, Shibuya, and Shinjuku Wards). The red vertical line denotes the first emergency order discouraging people from travelling within Tokyo. As shown in Panel (A), we find a marked slowing of the trend growth in total office floor area in central Tokyo following this first emergency order. This slowdown has persisted in the aftermath of the COVID-19 pandemic. This pattern of results is consistent with the economic mechanism in our model, in which there is a persistent reduction in the demand for commercial floor space in central areas following the shift to WFH.

As shown in Panel (B), we find a substantial increase in the office vacancy rate in central Tokyo following the first emergency order. The office vacancy rate continues to grow even after the end of the COVID-19 pandemic, before levelling off and remaining persistently higher in 2022 than in the years immediately before the pandemic. As shown in Panel (C), we find that average office rents in Central Tokyo begin to decline shortly after the first emergency order. Again, this decline continues even after the end of the COVID-19 pandemic, and begins to slow in 2022, consistent with adjustment towards a new equilibrium. Both patterns are again consistent with the persistent reduction in the demand for commercial floor space in central areas following the shift to WFH in our theoretical model.

Taken together, these findings for commercial office space in central Tokyo reinforce our findings using our cellphone data in the paper of a collapse of both work and non-work trips in central Tokyo following the shift to WFH. These results for Tokyo are also consistent with evidence for major metropolitan areas in the United States of a large-scale decline in the demand for commercial office space in central cities (the so-called commercial office apocalypse) following the shift to WFH (see in particular Gupta et al. 2022).

Figure B.3.1: Data on the Office/Commercial Building Market in Central Tokyo



Notes: Office market data is from Miki Shoji Company Ltd. The red vertical line shows April 2020, the date of the first emergency order that discouraged people from travelling within the Tokyo metropolitan area during COVID-19. The figures show the total floor area, vacancy rate, and average rent per square meter for five most central wards in Tokyo Metropolitan Area (Chiyoda, Chuo, Minato, Shibuya, and Shinjuku Wards).

B.4 Event-Study Evidence

In the previous two subsections, we have provided time-series evidence on the impact of WFH for the Tokyo metropolitan area. In this subsection, we use an event-study specification to examine the statistical significance of these changes in patterns of spatial mobility, and to examine other economic outcomes.

B.4.1 Econometric Specification

To assess the statistical significance of these changes in the spatial organization of economic activity over time, we estimate the following event-study specification:

$$\log Y_{it} = \sum_{t=-T}^{t=T} \beta_t \left(\mathbb{C}_i \times d_t \right) + \eta_i + d_t + u_{it} \tag{B.1}$$

where i indexes locations and t denotes time periods; Y_{it} denotes an economic outcome; $\mathbb{C}_i$ is an indicator that equals one if the location is in Central Tokyo and zero otherwise; η_i is a location fixed effect; d_t is a time period fixed effect; and u_{it} is a stochastic error.

The coefficients of interest are β_t, which capture the differential changes over time in an economic outcome in Central Tokyo relative to the other parts of Tokyo. The identifying assumption is parallel trends between the city center and other parts of the metropolitan area in the absence of the COVID-19 pandemic and the subsequent shift to WFH. As a specification check on this identifying assumption, we can examine whether these city-center coefficients (β_t) are flat in the periods leading up to the COVID-19 pandemic. If the shift to WFH leads to a persistent decentralization of economic activity away from the city center, we expect these city-center coefficients (β_t) to decline following the onset of the COVID-19 pandemic and remain persistently low thereafter. We use the two-way fixed effects specification, and cluster the standard errors by location to allow for serial correlation in the error term over time. Since this specification relies on a single timing of the treatment, the two-way fixed estimator is valid and not subject to a concern raised by the recent econometrics literature on differences-in-differences design (e.g., Borusyak et al. (2024)).

Our baseline definition of locations is Oaza (10,170 in the Tokyo metropolitan area), although in a few cases we report results for municipalities (242 in the Tokyo metropolitan area), when data are only available at this higher level of spatial aggregation. Our baseline definition of time periods is months, although we also report results for years, when data are only available at this higher level of temporal aggregation.

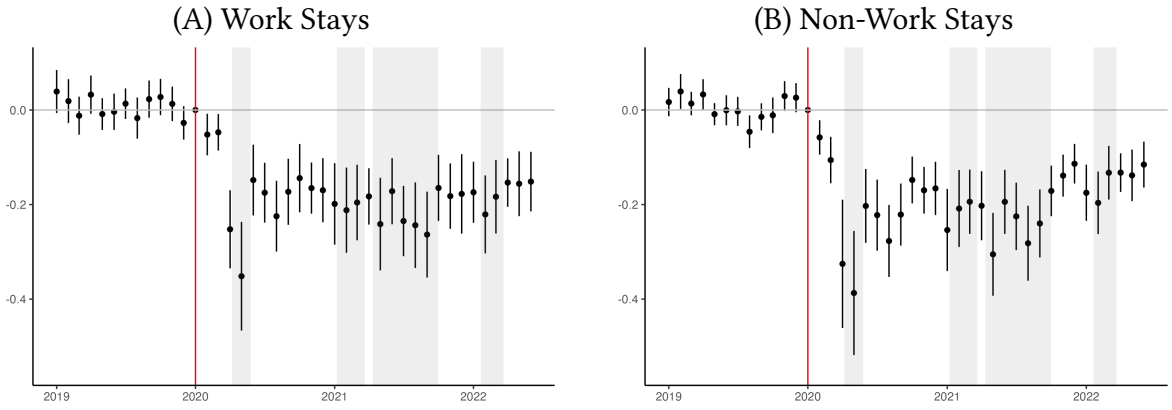
In our baseline specification, we define Central Tokyo as all locations with centroids within 10 km of the centroid of Chiyoda Ward, which is the historical center of Tokyo (containing the Imperial Palace, Tokyo Station, and the largest concentration of office space). This 10-km

central-city definition includes the Yamanote Railway Line (the urban circular transportation link as studied in Section 7.3 of our paper) and encloses an area of approximately the same size as Manhattan. The wider metropolitan area of Tokyo encompasses some locations more than 50 km from this central point. In the next subsection, we provide further evidence using continuous measures of distance from the center.

B.4.2 Empirical Results

In Figure B.4.1, we report the results of estimating our event-study specification (B.1) using work and non-work stays as measured using our smartphone data. We find that the city-center coefficients are relatively flat in the months leading up to the COVID-19 pandemic, implying similar employment trends in Central Tokyo and the other parts of the city before the pandemic. We find that the city-center coefficients drop sharply following the onset of the pandemic and remain persistently low thereafter, which is consistent with a shift to either fully remote or hybrid WFH, such that fewer workers commute into workplaces in Central Tokyo each day. This pattern is accompanied by the reduction of non-work stays in Central Tokyo, consistent with Figure 2 of the paper.

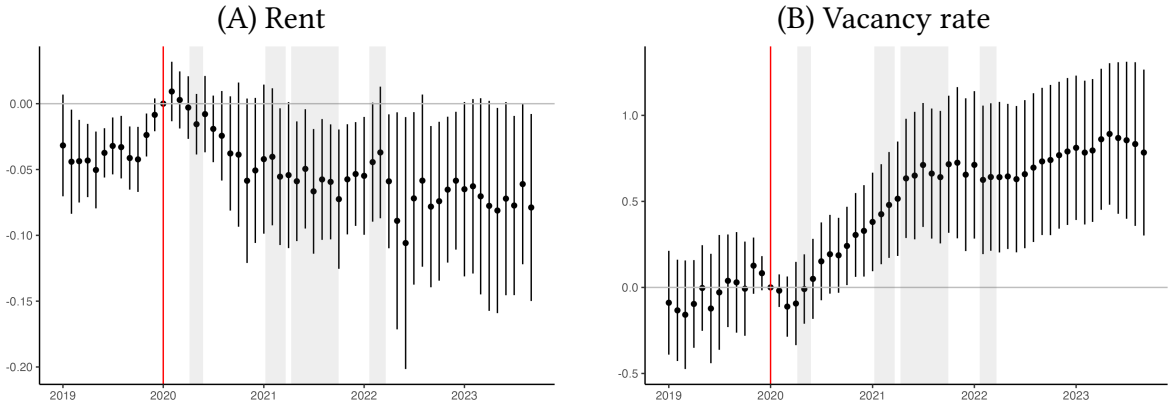
Figure B.4.1: Event-study Estimates for Work and Non-work Stays



Notes: Estimated event-study coefficients (β_t) on interaction terms between an indicator for locations in Central Tokyo (within 10 kilometers of the most central point) and month indicators from estimating equation (B.1) using two-way fixed effects; outcomes are work and non-work stays as measured using our smartphone data; observations are Oaza and month pairs; excluded month is January 2020.

In Figure B.4.2, we estimate the same event-study specification (B.1) using our office market data on average office rents and vacancy rates. In our model, the shift to working from home directly leads to a direct reduction in the demand for commercial floor space in the main concentrations of workplaces in Central Tokyo. Additionally, this decline in work trips is

Figure B.4.2: Event-study Estimates for Average Office/Commercial Building Rents and Vacancy Rates

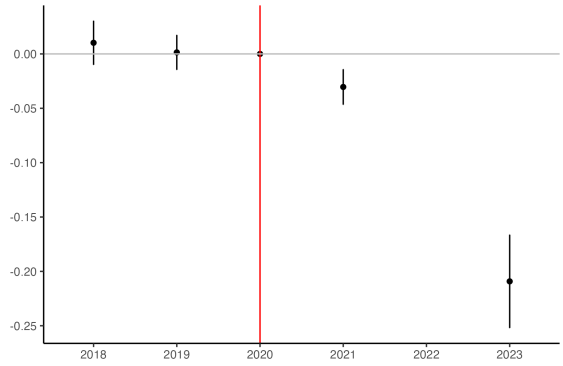


Notes: Estimated event-study coefficients (β_t) on interaction terms between an indicator for locations in Central Tokyo (within 10 kilometers of the most central point) and month indicators from estimating equation (B.1) using two-way fixed effects; outcomes are average office rents and vacancy rates; observations are aggregations of Oaza and month pairs; excluded month is January 2020.

accompanied by a decline in non-work trips to the central city, as fewer workers stop off to consume non-traded services along the way, thereby leading to a further indirect reduction in the demand for commercial floor space. Consistent with these predictions, we find a persistent decline in average office rents and increase in average office vacancy rates following the shift to WFH.

Figure B.4.3 estimates the event-study specification using the number of establishments in the non-traded services sector (retail) obtained from the telephone directory. We find a significant decline in the number of establishments in the non-traded service sector in Central Tokyo than elsewhere. Therefore, we find direct empirical support for the prediction of our model of travel itineraries of a collapse in demand for non-traded services in Central Tokyo.

Figure B.4.3: Event-study Estimates for the Number of Non-traded Services Establishments



Notes: Estimated event-study coefficients (β_t) on interaction terms between an indicator for locations in Central Tokyo (within 10 kilometers of the most central point) and year indicators from estimating equation (B.1) using two-way fixed effects; outcome is the number of establishments in the non-traded services sector (retail).

B.5 Event-Study Estimates by Distance

The event-study specifications in the previous section have provided evidence on the impact of the shift to WFH on economic activity in Central Tokyo relative to other parts of Tokyo, using a definition of Central Tokyo based on a 10 km distance threshold. In this section of the Online Appendix, we demonstrate that these findings are not sensitive to the precise choice of distance threshold, and provide further evidence on the heterogeneous impacts of the shift to WFH by distance from the center of Tokyo.

B.5.1 Econometric Specification

To explore the heterogeneous impact if WFH, we construct a series of distance rings $dr \in DR$, which are indexed by distance from the center of Tokyo, as measured by the centroid of Chiyoda Ward. For areas within 10 km of the center, we define these rings every 2 km. For areas over 10 km from the center, we define these rings every 10 km. The excluded category is locations 50 km or more from the center. Using these distance rings, we estimate the following event-study specification:

$$\log Y_{it} = \sum_{dr \in DR} \beta_{dr} \left(\mathbb{C}_i^{dr} \times T_t \right) + \eta_i + d_t + u_{it} \tag{B.2}$$

where again i indexes locations and t denotes time periods; Y_{it} denotes an economic outcome; $\mathbb{C}_i^{dr}$ is an indicator that equals one if the location lies within distance ring dr and zero otherwise; T_t is an indicator that equals one if t is after the first COVID-19 emergency order (March 2020 and after for monthly data, or 2020 and after for yearly data); η_i is a location fixed effect; d_t is a time period fixed effect; and u_{it} is a stochastic error. We cluster the standard errors by location to allow for serial correlation over time.

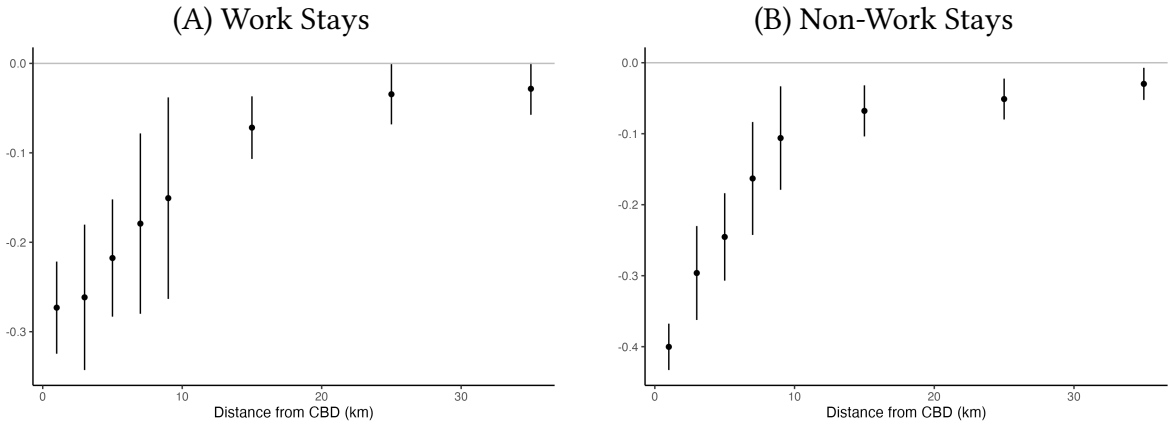
The coefficients of interest (β_{dr}) capture the differential change in an economic outcome before and after March 2020 for distance ring dr relative to the excluded category. The identifying assumption is parallel trends for distance ring dr and the excluded category in the absence of the COVID-19 pandemic and the subsequent shift to WFH. If the shift to WFH leads to a decentralization of economic activity away from the city center, we expect the estimated coefficients (β_{dr}) to be negative and significant for distance rings close to the city center.

Our baseline definition of locations is again oaza (10,170 in the Tokyo metropolitan area), although in a few cases we report results for municipalities (242 in the Tokyo metropolitan area), when data are only available at this higher level of spatial aggregation. Our baseline definition of time periods is months, although we also report results using years, when data are only available at this higher level of temporal aggregation.

B.5.2 Empirical Results

In Figure B.5.1, we report the results of estimating this event-study specification (B.2) using work and non-work stays, as measured using our smartphone data. We find negative and statistically significant coefficients (β_{dr}) for distance rings close to the city center. Therefore, the shift to WFH led to a reduction in both work and non-work stays in locations close to the city center relative to those further away from the city center. This pattern of results is in line with the predictions of our model of travel itineraries, in which work and non-work stays are jointly and endogenously determined.

Figure B.5.1: Distance Event-study Estimates for Work and Non-Work Stays

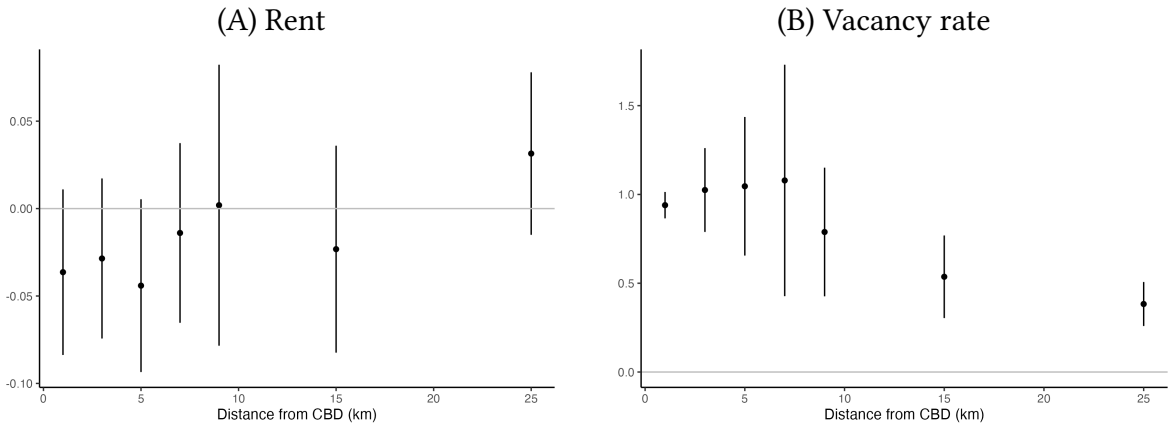


Notes: Estimated distance event-study coefficients (β_{dr}) on interaction terms between an indicator ($\mathbb{C}_i^{dr}$) that is one for distance ring $dr \in DR$ and zero otherwise and an indicator (T_t) that is one from March 2020 onwards and zero otherwise; coefficients estimated from equation (B.2) using the two-way fixed effects estimator; the excluded distance category is locations more than 50 km from the center of Tokyo (centroid of Chiyoda ward) outcome; outcomes are work and non-work stays as measured using our smartphone data; observations are oaza and month pairs.

In Figure B.5.2, we estimate this event-study specification for average office rents and vacancy rates. Consistent with the shift to WFH reducing the demand for commercial office space in Central Tokyo, we find larger declines in office rents and larger increases in office vacancy rates for distance rings closest to the city center.

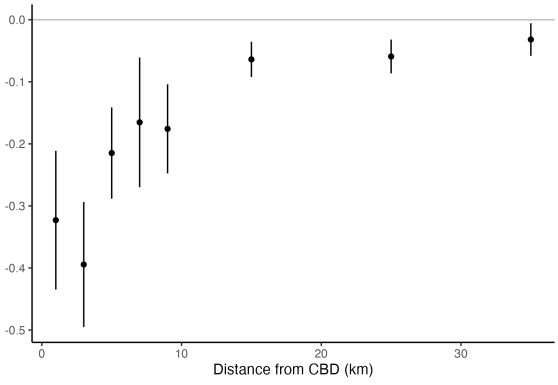
In Figure B.5.3, we estimate this event-study specification using the number of establishments in the non-traded services sector (the retail sector). We find the largest drops in non-traded services activity in the most central parts of Tokyo. This pattern of results is again consistent with the predictions our travel itinerary model. As workers commute into the central city less often (reducing work stays), this diminishes the number of intermediate stays at which workers consume non-traded services along the way to and from work (thereby leading to an accompanying reduction in non-work stays).

Figure B.5.2: Distance Event-study Estimates for Office/Commercial Building Rents and Vacancy Rates



Notes: Estimated distance event-study coefficients (β_{dr}) on interaction terms between an indicator ($\mathbb{C}_i^{dr}$) that is one for distance ring $dr \in DR$ and zero otherwise and an indicator (T_t) that is one from March 2020 onwards and zero otherwise; coefficients estimated from equation (B.2) using the two-way fixed effects estimator; the excluded distance category is locations more than 50 km from the center of Tokyo (the centroid of Chiyoda ward); outcomes are average office rents and vacancy rates; observations are aggregations of oaza and month pairs.

Figure B.5.3: Distance Event-study Estimates for Establishments in the Non-Traded Services Sector)



Notes: Estimated distance event-study coefficients (β_{dr}) on interaction terms between an indicator ($\mathbb{C}_i^{dr}$) that is one for distance ring $dr \in DR$ and zero otherwise and an indicator (T_t) that is one from March 2020 onwards and zero otherwise; coefficients estimated from equation (B.2) using the two-way fixed effects estimator; the excluded distance category is locations more than 50 km from the center of Tokyo (the centroid of Chiyoda ward); outcome is number of establishments in the non-traded services sector (the retail sector); observations are oaza and year pairs.

C Appendix for Retail Store Closures

In this section of the Online Appendix, we provide further detail on our quasi-experimental estimates of consumption externalities from large retail store closure, supplementing the material from Section 4.3 of the paper.

C.1 Data

We use the 2019 Edition of the Large Retail Store Data published by *Toyo Keizai Shimposha*, which provides information on the locations and closing dates of retail stores with a sales floor area of 1,000 m^2 or more. We use data on all large-scale retail stores with a sales floor area of 5,000 m^2 or more that were present in the Tokyo metropolitan area from 2015 to 2019. Among these large retail stores, we classify those that had closed by 2019 as the treatment group, while the remaining stores correspond to the control group.

The changes in foot traffic induced by large retail store closure can be quite local. Therefore, we use our smartphone data disaggregated by 250 $\times$ 250 meter geographical grid cells. We define a grid cell that contains a large retail store as the directly-affected grid cell. We construct a series of concentric distance rings around each directly-affected grid cell to capture spillover effects on neighboring grid cells. We construct these concentric rings at 300-meter intervals, based on the distance between the centroids of grid cells, with the outermost ring including grid cells with centroids 1,800-2,100 meters from the centroid of the directly-affected grid cell.

Especially in the central areas of Tokyo, where there is a high concentration of large retailers, there are many cases where multiple large retailers are located in the same grid cell or in nearby grid cells. Therefore, the concentric rings of different large retailers can overlap, or the directly-affected grid cell for one large retailer can lie in the concentric grid ring of another large retailer. To avoid spurious findings of spillovers driven by the directly-affected grid cell of one large retailer lying in the concentric ring of another larger retailer, our baseline specification drops all grid cells in which there is overlap between either the directly-affected and/or neighboring grid cells of multiple large retailers.

C.2 Empirical Specification

We estimate the impact of the closure of large retail stores on both the directly-affected and neighboring grid cells using an event-study regression specification. Our sample includes the directly-affected and neighboring grid cells for closed large retail stores (treatment group) and surviving large retail stores (control group). We index large retailers by r and the grid cells containing them by $i(r)$. Observations correspond to grid cells ($i(r)$) and months (t). We

denote distance rings by $d \in DR$, where DR is the set of distance rings. We define distance ring 0 as the directly-affected grid cell. The remaining distance rings are of 300-meter width and contain grid cells with centroids up to 2,100 meters from a directly-affected grid cell.

Before estimating our event-study specification, we consider the following more parsimonious differences-in-differences regression specification:

$$\log Y_{i(r)t} = \sum_{d \in DR} \beta_d \left(\mathbb{D}^d_{i(r)} \times \mathbb{C}_{i(r)t} \right) + \eta_{i(r)} + \xi_{dt} + u_{i(r)t} \tag{C.1}$$

where $Y_{i(r)t}$ is the number of non-work stays in grid cell $i(r)$ and month t; β_d are the treatment coefficients of interest on interaction terms between dummy variables for a grid cell lying within distance ring d ($\mathbb{D}^d_{i(r)}$) and a dummy variable for the period after the closure of a retail store ($\mathbb{C}_{i(r)t}$); $\eta_{i(r)}$ are grid cell fixed effects; ξ_{dt} are distance ring $\times$ month fixed effects; and $u_{i(r)t}$ is a stochastic error. We weight observations by the number of non-work stays in each grid cell in 2014 before the beginning of our sample period. We cluster the standard errors by grid cell to allow for serial correlation in the error term over time.

One challenge in estimating the impact of retail store closure is that the decision to close a store is unlikely to be random. Focusing on large stores that are typically part of national chains partially helps to alleviate this concern, because the decision to close a store is often influenced by national considerations, such as the trajectory of sales across the entire national chain, rather than simply the trajectory of sales for the closed store. Our differences-in-differences specification allows time-invariant unobserved heterogeneity between treatment and control grid cells (through the location fixed effect). Our identifying assumption is parallel trends between the treatment and control grid cells without the closure of the retail store. As a check on this identifying assumption, we estimate the following event-study specification:

$$\log Y_{i(r)t} = \sum_{x=-T}^{T} \sum_{d \in DR} \beta_{dx} \left(\mathbb{D}^d_{i(r)} \times \mathbb{C}_{i(r)x} \right) + \eta_{i(r)} + \xi_{dt} + u_{i(r)t} \tag{C.2}$$

where x denotes relevant to treatment (before or after the closure of a retail store); the excluded category is the month in which the closure occurs ($x = 0$); we consider windows of 24 months before and after closure ($T = 24$); β_{dx} are the treatment coefficients of interest on interaction terms between dummy variables for a grid cell lying within distance ring d ($\mathbb{D}^d_{i(r)}$) and dummy variables for period x before or after the closure of a retail store ($\mathbb{C}_{i(r)x}$); and the definition of the other variables remains as above. Again we weight observations by the number of non-work stays in each grid cell in 2014 before the beginning of our sample period. We cluster the standard errors by grid cell to allow for serial correlation in the error term over time.

A recent empirical literature has highlighted that the interpretation of the two-way fixed effects estimator in event-study specifications can be problematic in the presence of a variable

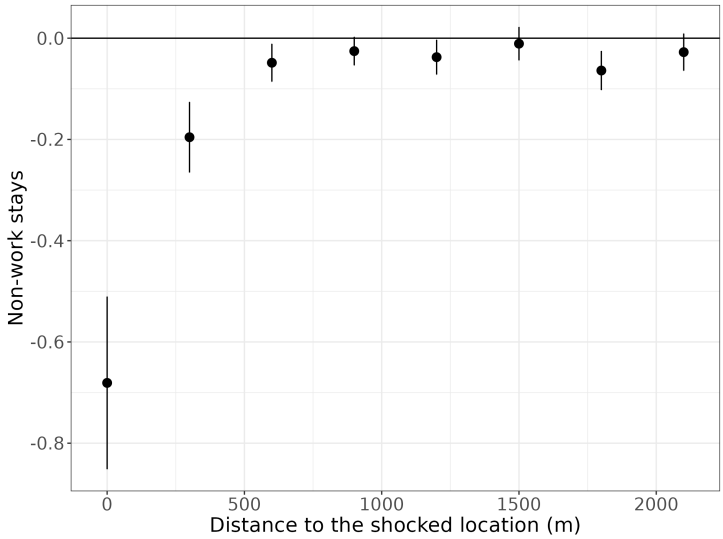
timing of the treatment (as in this specification) and treatment heterogeneity. Therefore, we use the estimator of Borusyak *et al.* (2024) for our baseline specification, which is robust to this concern. In our empirical application, we find a relatively similar pattern of estimated treatment effects using this estimator as using the two-way fixed effects estimator or other alternative event-study estimators, including Callaway and Sant'Anna (2020), Sun and Abraham (2021), and Gardner (2021), and Roth and Sant'Anna (2020), as surveyed in Roth et al. (2023).

C.3 Empirical Results

Figure C.3.1 shows the results of estimating our more parsimonious difference-in-differences specification (C.1) using the estimator of Borusyak *et al.* (2024). We display estimated coefficients (β_d) for all distance rings up to 2,100 meters. Following the closure of a large retail store, we find a large and statistically significant drop in non-work stays in the directly-affected grid cell by around 60 percent. We also find negative and statistically significant spillover effects on neighboring grid cells, with a drop of non-work stays in the 300-meter distance ring of around 20 percent. These spillover effects decline sharply with distance from the directly-affected grid cell and fall to close to zero by 1,000 meters. These findings of statistically significant and localized consumption externalities using our Japanese smartphone data are consistent with a range of existence quasi-experimental evidence of consumption externalities using the opening and closings of large retail stores and online shopping, including Shoag and Veuger (2018), Benmelech et al. (2019), Koster et al. (2019), Relihan (2022) and Qian et al. (2024).

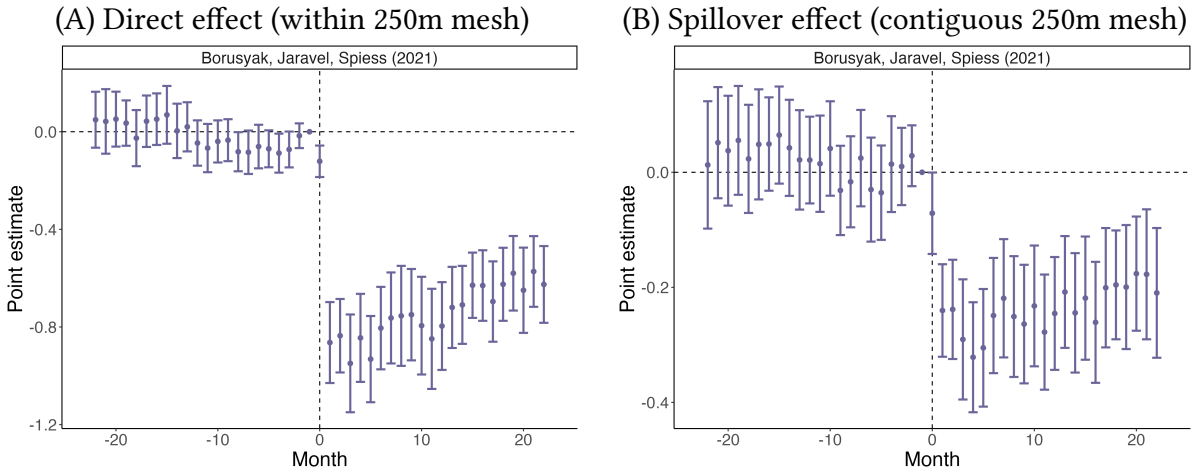
Figure C.3.2 shows the results of estimating our event-study specification (C.2) using the estimator of Borusyak *et al.* (2024). Given our findings immediately above that consumption externalities are highly-localized, we focus on the estimated treatment interactions for the directly-affected grid cell (Panel A) and the first 300-meter distance ring (Panel B). In line with the results from our more parsimonious difference-in-difference specification above, we find a large and statistically significant drop in non-work stays in the directly-affected grid cell of around 60 percent following the closure of a large retail store. We again find negative and statistically significant spillover effects on immediately neighboring grid cells, with a drop of non-work stays in the 300-meter distance ring of around 20 percent. This decline persists for at least 24 months, with some slight attenuation over time, consistent with new uses being found for the vacated floor space. Reassuringly, we find no evidence of pre-trends before the closure of a large retail store, for either the directly-affected grid cell or the 300-meter distance ring. Again these findings using our smartphone data are consistent with existing quasi-experimental evidence of consumption externalities using the opening and closings of large retail stores and online shopping.

Figure C.3.1: Difference-in-Differences Estimates of Large Retail Store Closure



Notes: Estimated treatment coefficients (β_d) from equation (C.1) on interaction terms between indicators for 300-meter distance rings around a large retail store and indicators for retail store closure; dependent variable is the log number of non-work stays; treatment coefficients estimated using the Borusyak et al. (2024) estimator; circles denote the point estimates; vertical bars indicate the 95 percent confidence intervals based on standard errors clustered by 250 × 250 meter grid cell; estimation with the full sample is infeasible because of the large control group and hence we report results using a 5 percent random sample of the control group.

Figure C.3.2: Event-study Estimates of Large Retail Store Closure



Notes: Estimated treatment coefficients (β_{dx}) from equation (C.2) on interaction terms between indicators for 300-meter distance rings around a large retail store and indicators for months before and after retail store closure; Panel (A) shows estimates for the directly-affected grid cell; Panel (B) shows estimates for the 300-meter distance ring; dependent variable is the log number of non-work stays (neither home nor work); treatment coefficients estimated using the Borusyak et al. (2024) estimator; treatment months before -24 and after +24 are binned into the first and last categories; circles denote the point estimates; vertical bars indicate the 95 percent confidence intervals based on standard errors clustered by 250 × 250 meter grid cell; estimation with the full sample is infeasible because of the large control group and hence we report results using a 5 percent random sample of the control group.

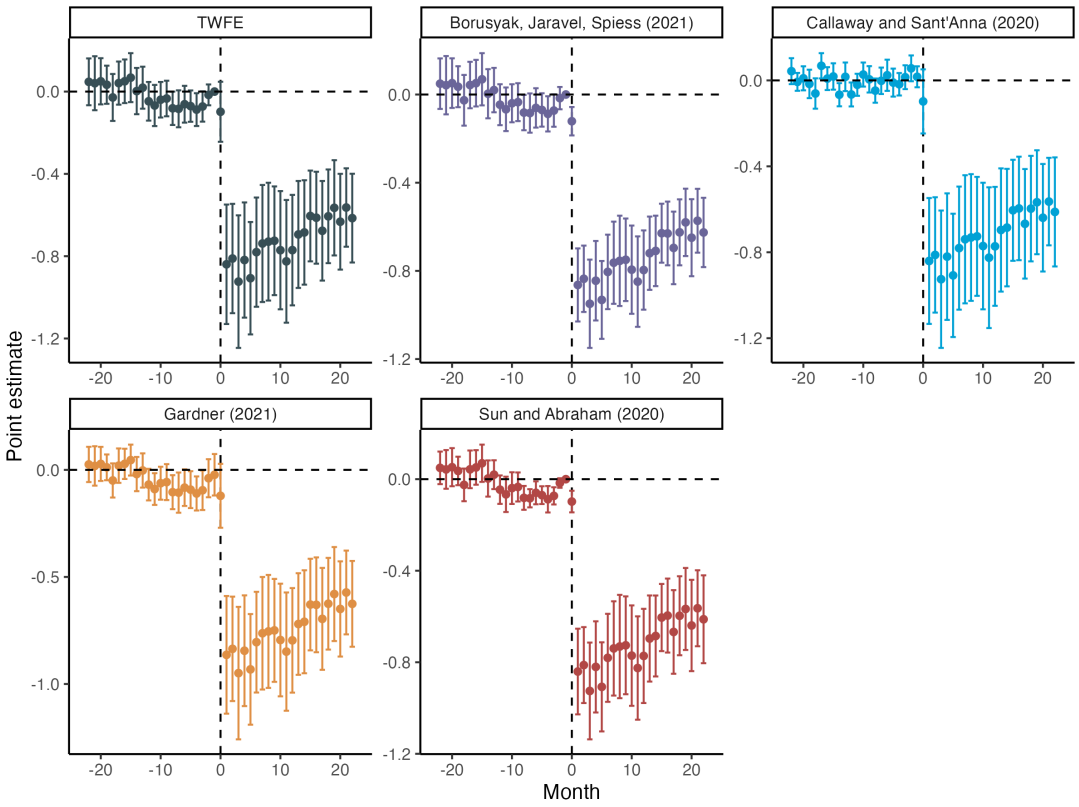
C.4 Robustness

As discussed above, we use the event-study estimator of Borusyak et al. (2024) in our baseline specification to address potential concerns about treatment heterogeneity and a variable timing of the treatment. In this section of the Online Appendix, we show that we find a similar pattern of estimated coefficients using the two-way fixed effects estimator, and a range of alternative event-study estimators.

Figure C.4.1 compares the estimated treatment effects of retail store closure on the directly-affected location (corresponding to Panel A of Figure 3 in the paper) using the two-way fixed effects estimator and the estimators of Borusyak et al. (2024), Callaway and Sant'Anna (2020), Gardner (2021), and Sun and Abraham (2020). Across all of these different specifications, we find a large and statistically significant drop in non-work stays in the directly-affected grid cell of around 60 percent following the closure of a large retail store. Again this decline persists for at least 24 months, with some slight attenuation over time, consistent with new uses being found for the vacated floor space. Therefore, our findings of a large decline in foot traffic in the directly-affected location following the closure of a large reatail store are not sensitive to the precise event-study estimator used.

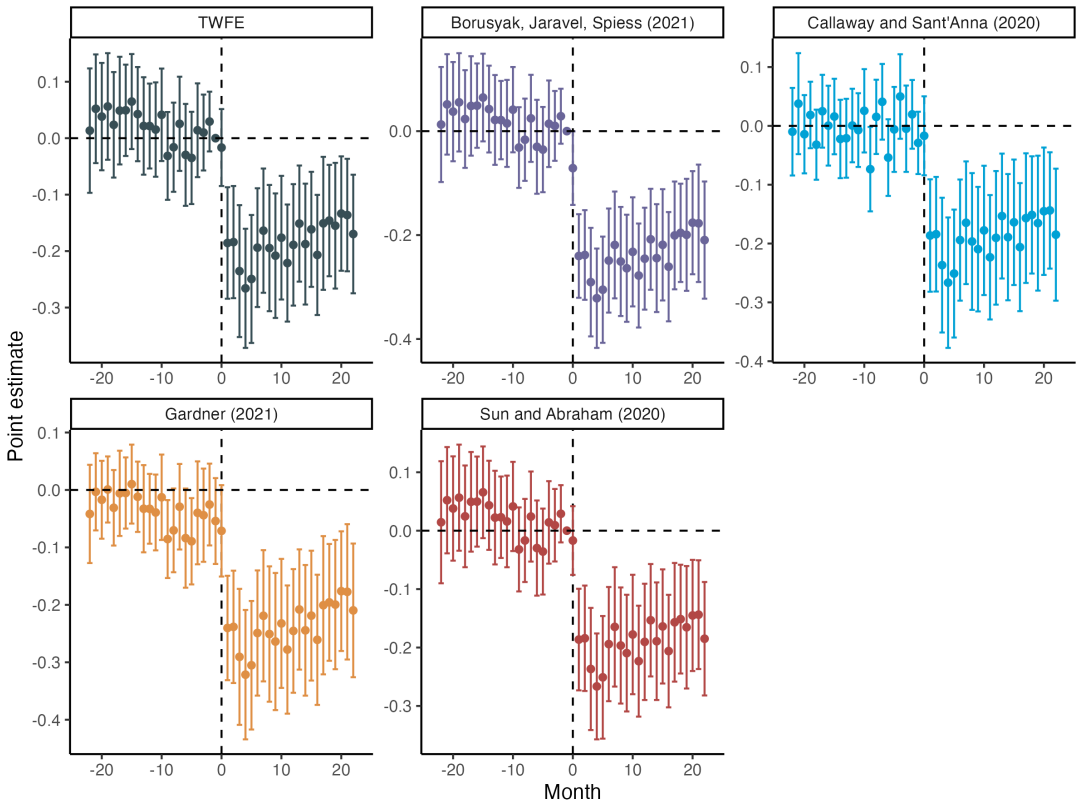
Figure C.4.2 compares the estimated treatment effects of retail store closure on grid cells in the 300-meter distance ring (corresponding to Panel B of Figure 3 in the paper) using the two-way fixed effects estimator and the estimators of Borusyak et al. (2024), Callaway and Sant'Anna (2020), Gardner (2021), Roth and Sant'Anna (2021), and Sun and Abraham (2020). Across all of these different specifications, we find negative and statistically significant spillover effects on grids cells in the 300-meter distance ring of around 20 percent. As for the directly-affected grid cell, this decline persists for at least 24 months, with some slight attenuation over time, consistent with new uses being found for the vacated floor space. Therefore, our findings of negative consumption externalities following the closure of a large retail store are not sensitive to the precise event-study estimator used.

Figure C.4.1: Event-study Estimates of Large Retail Store Closure for the Directly-affected Location (Robustness)



Notes: Estimated treatment coefficients (β_{dx}) from equation (C.2) on interaction terms between indicators for 300-meter distance rings around a large retail store and indicators for months before and after retail store closure; dependent variable is the log number of non-work stays (neither home nor work); figure shows results for the directly-affected grid cell using the two-way fixed effects estimator and the estimators of Borusyak et al. (2024), Callaway and Sant'Anna (2020), Gardner (2021), and Sun and Abraham (2020); treatment months before -24 and after +24 are binned into the first and last categories; circles denote the point estimates; vertical bars indicate the 95 percent confidence intervals based on standard errors clustered by 250 $\times$ 250 meter grid cell; estimation with the full sample is infeasible because of the large control group and hence we report results using a 5 percent random sample of the control group.

Figure C.4.2: Event-study Estimates of Large Retail Store Closure for the 300-meter Distance Ring (Robustness)



Notes: Estimated treatment coefficients (β_{dx}) from equation (C.2) on interaction terms between indicators for 300-meter distance rings around a large retail store and indicators for months before and after retail store closure; dependent variable is the log number of non-work stays (neither home nor work); figure shows results for the 300-meter distance ring using the two-way fixed effects estimator and the estimators of Borusyak et al. (2024), Callaway and Sant'Anna (2020), Gardner (2021), Roth and Sant-Anna (2021) and Sun and Abraham (2020); treatment months before -24 and after +24 are binned into the first and last categories; circles denote the point estimates; vertical bars indicate the 95 percent confidence intervals based on standard errors clustered by 250 × 250 meter grid cell; estimation with the full sample is infeasible because of the large control group and hence we report results using a 5 percent random sample of the control group.

D First-Order Effects of Localized Price Shocks

In this section of the Online Appendix, we analyze how a change in the price index for non-traded services in one location i affects non-work stays in each location n using the partial equilibrium model of travel itineraries from Section 5.1 of the paper.

We define total non-work stays in each location n as the sum of the non-work stays of agents with each home h and workplace j:

$$S_n = \sum_{h,j} \Omega_{hj} \sum_{I \in \mathcal{I}_{hj}: n \in I} \Lambda_{I|hj}, \tag{D.1}$$

where Ω_{hj} is the fraction of population with residence h and workplace j, and $\Lambda_{I|hj}$ is the probability of choosing travel itinerary I given home h and work j from equation (6):

$$\Lambda_{I|hj} = \frac{\mathbb{P}_{I|hj}^{-\theta}}{\sum_{\ell \in \mathcal{I}_{hj}} \mathbb{P}_{\ell|hj}^{-\theta}}, \qquad \mathbb{P}_{I|hj} = g_{I|hj}\left(\{P_{nS}\}_{n \in C(I)}\right), \tag{D.2}$$

where we suppress the travel cost $\tau_{I|hj}$ from the cost function $g_{I|hj}(\cdot)$ for notational convenience. We analyze the effects of a shock to the price index for non-traded services in location i (P_{iS}) on non-work stays in each location n (S_n). We take the fraction of agents with each home h and workplace j (Ω_{hj}) and the shock to the price index for non-traded services in location i (P_{iS}) as given here. We solve for these variables in terms of primitives when we embed our model of travel itineraries in general equilibrium in Section 6 of the paper. Taking derivatives in equation (D.1), we have:

$$\frac{\partial S_n}{\partial \log P_{iS}} = \sum_{h,j} \Omega_{hj} \sum_{I \in \mathcal{I}_{hj}: n \in C(I)} \Lambda_{I|hj} \frac{\partial \log \Lambda_{I|hj}}{\partial \log P_{iS}}, \tag{D.3}$$

where

$$\begin{aligned} \frac{\partial \log \Lambda_{I|hj}}{\partial \log P_{iS}} &= \frac{\partial}{\partial \log P_{iS}} \left(\log \mathbb{P}_{I|hj}^{-\theta} - \log \sum_{\ell \in \mathcal{I}_{hj}} \mathbb{P}_{\ell|hj}^{-\theta} \right) \\ &= -\theta \frac{\frac{\partial}{\partial \log P_{iS}} \mathbb{P}_{I|hj}}{\mathbb{P}_{I|hj}} + \theta \frac{\sum_{\ell \in \mathcal{I}_{hj}: i \in C(\ell)} \mathbb{P}_{\ell|hj}^{-\theta-1} \frac{\partial}{\partial \log P_{iS}} \mathbb{P}_{\ell|hj}}{\sum_{\ell \in \mathcal{I}_{hj}} \mathbb{P}_{\ell|hj}^{-\theta}}. \end{aligned} \tag{D.4}$$

From Shephard's lemma applied to the households' optimal consumption decision, we have:

$$\frac{\partial}{\partial \log P_{iS}} \mathbb{P}_{I|hj} = \mathbb{P}_{I|hj} \Psi_{i|I}, \tag{D.5}$$

where $\Psi_{i|I}$ is the share of expenditure on non-traded services from location i among consumers taking itinerary I. Combining (D.4) and (D.5), we have:

$$\begin{aligned}\frac{\partial \log \Lambda_{I|hj}}{\partial \log P_{iS}} &= -\theta\Psi_{i|I} + \theta\frac{\sum_{\ell\in\mathcal{I}_{hj}:i\in C(\ell)}\mathbb{P}_{\ell|hj}^{-\theta}\Psi_{i|\ell}}{\sum_{\ell\in\mathcal{I}_{hj}}\mathbb{P}_{\ell|hj}^{-\theta}} \\ &= -\theta\Psi_{i|I} + \theta\sum_{\ell\in\mathcal{I}_{hj}:i\in C(\ell)}\Lambda_{\ell|hj}\Psi_{i|\ell}.\end{aligned} \tag{D.6}$$

Using (D.6) in (D.3), we have:

$$\begin{aligned}\frac{\partial S_n}{\partial \log P_{iS}} &= \theta\sum_{h,j}\Omega_{hj}\sum_{I\in\mathcal{I}_{hj}:n\in C(I)}\Lambda_{I|hj}\left(-\Psi_{i|I} + \sum_{\ell\in\mathcal{I}_{hj}:i\in C(\ell)}\Lambda_{\ell|hj}\Psi_{i|\ell}\right) \\ &= \theta\sum_{h,j}\Omega_{hj}\left[\underbrace{-\sum_{I\in\mathcal{I}_{hj}:n\in C(I)}\Lambda_{I|hj}\Psi_{i|I}}_{\text{complementarity}} + \underbrace{\left(\sum_{I\in\mathcal{I}_{hj}:n\in C(I)}\Lambda_{I|hj}\right)\left(\sum_{\ell\in\mathcal{I}_{hj}:i\in C(\ell)}\Lambda_{\ell|hj}\Psi_{i|\ell}\right)}_{\text{substitution}}\right],\end{aligned} \tag{D.7}$$

Equation (D.7) shows that shocks to the price index for non-traded services in one location can induce either complementarity or substitution effects in other locations. The first term captures complementarity effects, reflecting the fact that an increase in the price index in location i discourages the choice of an itinerary that involves both i and n. The second term captures substitution effects, reflecting the fact that an increase in the price index in location i induces substitution from itineraries that involve i to those that involve n.

It is also straightforward to see that the special case of a conventional urban model, in which all consumption of non-traded services occurs through direct trips from home, cannot capture such complementarity. Given our finding that agents typically make a single travel itinerary each day in the data (from Table 1), we assume in this special case that agents make one direct trip to consume non-traded services (such that I includes only one non-work stay) and another direct trip to commute into work. In this special case, equation (D.7) reduces to

$$\frac{\partial S_n}{\partial \log P_{iS}} = \theta\sum_{h,j}\Omega_{hj}\left[-\Lambda_{\{hih\}|hj}1\left[n=i\right] + \underbrace{\Lambda_{\{hnh\}|hj}\Lambda_{\{hih\}|hj}}_{\text{substitution}}\right] \tag{D.8}$$

Therefore, in this special case of a conventional urban model, the cross-price elasticity for non-work stays for $n \neq i$ necessarily has the opposite sign to the own-price elasticity (the case of substitutes), such that an increase in the price index for non-traded services in location i necessarily increases non-work stays in location $n \neq i$.

An implication of these results is that our model of travel itineraries can rationalize our findings of positive spillovers to neighboring locations from the exit of large retail stores. When neighboring locations are complementary, the exit of a large retail store that reduces non-work stays in a directly-affected location also reduces non-work stays in neighboring locations, consistent with our empirical findings from our event-study regressions in Section 4.3 of the paper.

In contrast, in the special case of a conventional urban model with only direct consumption trips, neighboring locations are necessarily substitutes. Therefore, in this special case, the exit of a large retail store that reduces non-work stays in a directly-affected location necessarily increases non-work stays in neighboring locations, which is the opposite of our empirical findings from our event-study regressions in Section 4.3 of the paper.

E Appendix for Importance Sampling Method

In this section of the Online Appendix, we provide additional details about our importance sampling method, as described in Algorithm 1 in the paper. Section E.1 describes our practical choice of auxiliary distribution. Section E.2 shows that, given our choice of the number of importance samples R, the sampling errors are negligible for our quantitative results.

E.1 Choice of Auxiliary Distribution for Importance Sampling

In this section, we describe our choice of the auxiliary distribution $F_{hj}(\cdot)$ that we use to implement Algorithm 1 in the paper. As discussed in Section 5.2 of the paper, an advantage of this algorithm is that the choice of the auxiliary distribution $F_{hj}(\cdot)$ does not affect the results asymptotically as $R \to \infty$. At the same time, under finite R, the precision of the approximation depends on how close $F_{hj}(\cdot)$ is to the original distribution $\Lambda_{I|hj}$, as discussed by Kloek and van Dijk (1978) and Ackerberg (2009).

In our application, the following choice of $F_{hj}(\cdot)$ performs well in practice. For each home h and work j (including non-workday $j \neq \emptyset$), we first sample the number of stays and when the work stays occur during the itinerary following the empirical distribution in the data. We then simulate that agents choose a location to visit in sequence myopically, without taking into account the continuation value of visiting a destination, which implies that the probability of choosing each leg of the travel itinerary depends only on the origin and destination (and not the entire itinerary). Formally, the procedure is given as follows:

Algorithm 2 (Auxiliary Distribution for Importance Sampling) *The proposed travel itinerary I is simulated in the following order:*

1. *For each home h and work j, randomly generate the total number of non-work stays for a daily travel itinerary I, $|C(I)|$. On work days ($j \neq \emptyset$), randomly generate the number ι, such that the ι-th stay is a work stay.*

2. *Set the first location $i = 1$ as home h, and set the last location as home h, such that each daily travel itinerary starts and ends at home.*

3. *Determine the i-th location n_i starting from $i = 2$. If the i-th location of the day is at work (i.e., $i = \iota$), set the stay location as $n_i = j$. If the i-th location of the day is not at work (i.e., $i \neq \iota$), assume that agents choose the location myopically following some probability. Namely, denoting the itinerary up to $(i-1)$-th stay by I_{i-1}, we sample n_i following some probability $\Pi_{n_i|n_{i-1}}$ depending on the previous location n_{i-1}.*

4. *Repeat the previous step for all other stays included in the itinerary.*

This distribution $F_{hj}(I)$ has full support (with $\Pi_{n|\ell} > 0$ for all n, ℓ), which guarantees the consistency of the importance sampling procedure in Algorithm 1 as $R \to \infty$. Furthermore, this proposed auxiliary distribution is easy to simulate. We can also compute the likelihood using the following expression:

$$F_{hj}(I) = K_{hj}(|C(I)|, \iota) \prod_{i=2,\ldots,|I|-1; i \neq \iota} \Pi_{n_i|n_{i-1}},$$

where $K_{hj}(|C(I)|, \iota)$ is the empirical probability of $|C(I)|$ and ι as used in Step 1.

We construct $\Pi_{n_i|n_{i-1}}$ in the following manner. To ensure that this probability has full support over all possible itineraries, we use smoothed bilateral origin-destination flows for nonwork trips using our smartphone data for this auxiliary probability. Namely, we set

$$\Pi_{n|\ell} = \frac{T_{n\ell}^{\varrho} \xi_\ell}{\sum_{\ell'} T_{n\ell'}^{\varrho} \xi_{\ell'}},$$

where $T_{n\ell}$ is travel time from n to ℓ. We estimate the parameters ϱ and $\{\xi_\ell\}$ using the observed origin-destination flows for non-work trips from our smartphone data (collapsing all travel itineraries) using the Pseudo-Poisson maximum likelihood (PPML) estimator. Notice that this auxiliary probability is different from the true likelihood, because our model predicts extended gravity as discussed in Section 5.7. But this procedure provides a valid auxiliary distribution for our importance sampling. We adjust the sampling rate from this auxiliary distribution based on the likelihood ratio between the true distribution and the auxiliary distribution, which allows our importance sampling to capture extended gravity.

E.2 Stability of the Importance Sampling Simulations

In Table E.2.1, we show that our counterfactual simulation results in Section 7 of our main paper are not sensitive to the particular draw of importance samples, given our choice of auxiliary distribution and the number of importance draws per work-home pair of $R = 200$. We undertake our counterfactuals in Section 7 twenty times, and report the mean and the standard deviation of the model's counterfactual predictions across these simulations.

In the first row, we report the mean log change in non-work stays in the CBD in our WFH counterfactual using our baseline travel itinerary model from Section 7.1 of the paper. We find a mean log change in non-work stays in the CBD of -0.16, which compares with a value of -0.18 for our baseline importance sample in Table 5 of the paper. In the second and third rows, we report the mean welfare gains from our transportation infrastructure counterfactuals using our baseline travel itinerary model from Section 7.3 of the paper. We find mean welfare gains of 0.44 and 0.70, respectively, which compares with values of 0.44 and 0.71 for our baseline importance sample in Table 6 of the paper.

Therefore, the mean values across the importance samples in Table E.2.1 closely approximate those using our baseline importance sample in the paper. Furthermore, across the twenty sets of simulations, the standard deviations of these statistics are less than or equal to 0.01. These results confirm that the sampling errors are negligible for our application, given our choice of the number of importance samples $R = 200$ for our quantitative results.

Table E.2.1: Stability of the Importance Sampling Simulations

	Mean	SD
Changes of non-work stays in CBD by WfH	-0.16	0.01
Welfare gains from Yamanote Railway Line	0.44	0.00
Welfare gains from Chuo Railway Line	0.70	0.01

Note: Stability of the importance sampling simulations for our counterfactuals. We undertake our counterfactuals in Section 7 of the paper twenty times, and report the mean and the standard deviation (SD) of the model's counterfactual predictions across these simulations. In the first row, we report the mean log change in non-work stays in the CBD in our WFH counterfactual using our baseline travel itinerary model from Section 7.1 of the paper (compare with -0.18 in Table 5 in the paper). In the second and third rows, we report the mean welfare gains from our transportation infrastructure counterfactuals using our baseline travel itinerary model from Section 7.3 of the paper (compare with the values of 0.44 and 0.71 in Table 6 in the paper).

F Appendix for Parametrization and Model Estimation

In this section of the Online Appendix, we report additional details and results for our parametrized travel itinerary model and its estimation in Section 5 of the paper.

F.1 Geographic Area

In the theoretical model in Section 5 of the paper, we assume that the idiosyncratic preference draws have the same mean for each itinerary (equation (4)). When we take the model to the data, we have to aggregate our cellphone data to spatial units with administrative boundaries of different sizes.

To incorporate these features of the data, we extend our model so that the probability of a travel itinerary scales with the geographic size of the spatial units. We assume that $\epsilon_{\omega I}$ is drawn from the following distribution:

$$F_{I|hj}(\epsilon) = e^{-Z_{I|hj}\epsilon^{-\theta}}, \qquad \theta > 1, \tag{F.1}$$

where $Z_{I|hj} = \prod_{i \in C(I)} Z_i$ is the total geographical area of the spatial units included in the itinerary; and Z_i is the geographic area of each of these individual spatial units.

This specification (F.1) can be microfounded by the assumption that the idiosyncratic preference draws are drawn at the level of disaggregated geographical grid cells of equal size, instead of at the level of our aggregated spatial units (see Kreindler and Miyauchi 2023). The corresponding itinerary choice probability becomes:

$$\Lambda_{I|hj} = \frac{Z_{I|hj}\mathbb{P}_{I|hj}^{-\theta}}{\sum_{\ell \in \mathcal{I}_{hj}} Z_{\ell|hj}\mathbb{P}_{\ell|hj}^{-\theta}}. \tag{F.2}$$

F.2 Additional Evidence for Model Fit

In this section of the Online Appendix, we provide additional evidence on model fit, supplementing the results reported in Section 5.6 of the paper.

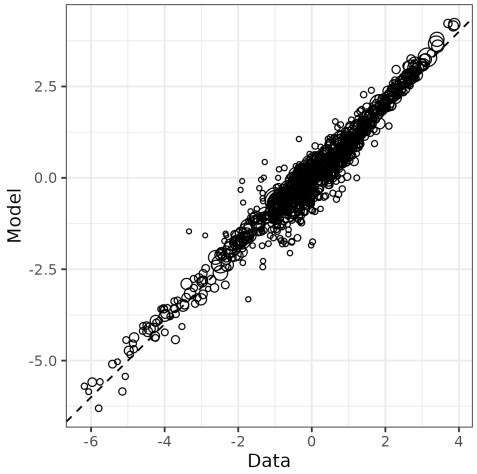
In Figure 4 of the paper, we show model fit for travel itineraries with one work stay (Panel (A), targeted) and more than one work stay (Panel (B), untargeted). We show that our model provides a good fit to the untargeted moments. Therefore, although our model is necessarily an abstraction, we are able to use its structure to successfully extrapolate from the subset of travel itineraries with one non-work stay to those with more than one non-work stay.

In Figure F.2.1 below, we provide further evidence on model fit by reporting the number of travel itineraries with a single non-work stay for workdays and non-workdays separately. The size of each circle is proportional to the overall number of non-work stays in a location.

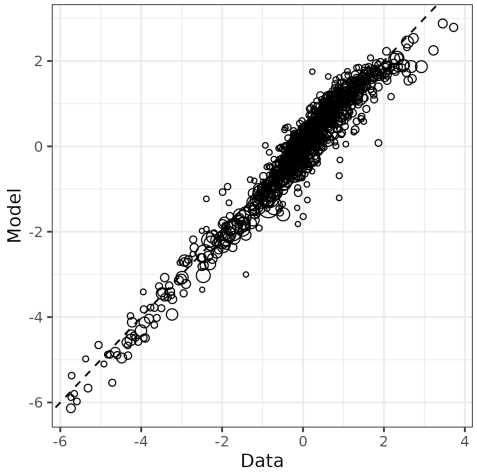
Although the moment targeted in our estimation pools both sets of observations, we find that the model provides a good fit to the data for these two types of days separately.

Figure F.2.1: Model Fit for the Number of Non-work Stays in Each Location for Travel Itineraries with a Single Non-work Stay (Work and Non-Work Days Separately)

(A) Density of Stays at Destination (Single Non-work Stays; Workdays)



(B) Density of Stays at Destination (Single Non-work Stays; Non-workdays)



Note: Panel A shows the log number of non-work stays per unit area in each location for travel itineraries with a single non-work stay in addition to home and work on workdays; Panel B shows the log number of non-work stays per unit area in each location for travel itineraries with a single non-work stay in addition to home and work on non-workdays; in both panels, vertical axis shows model predictions, and horizontal axis shows the empirical moment in our smartphone data; each circle corresponds to a location; the size of each circle is proportional to the overall number of non-work stays in a location in our smartphone data.

In Table F.2.1 below, we report the model fit for each of the targeted moments used in our parameter estimation separately. Although we are overidentified with more moments than parameters, we find that the model is relatively successful in capturing each of the targeted moments.

Table F.2.1: Targeted Moments for our Baseline Travel Itinerary Model

Moments	Data	Model
Frac 2 Non-Work Stays in Itinerary (Workdays)	0.16	0.21
Frac >2 Non-Work Stays in Itinerary (Workdays)	0.06	0.05
Frac 2 Non-Work Stays in Itinerary (Non-Workdays)	0.25	0.21
Frac >2 Non-Work Stays in Itinerary (Non-Workdays)	0.15	0.17
log 10-pctile Orig-Dest Travel Time (Workdays)	2.28	2.23
log 50-pctile Orig-Dest Travel Time (Workdays)	2.97	2.97
log 90-pctile Orig-Dest Travel Time (Workdays)	3.68	3.73
log 10-pctile Orig-Dest Travel Time (Non-Workdays)	1.70	1.63
log 50-pctile Orig-Dest Travel Time (Non-Workdays)	2.83	2.79
log 90-pctile Orig-Dest Travel Time (Non-Workdays)	3.74	3.97

Note: Targeted moments for the estimation of the parameters of our baseline travel itinerary model ($\Theta \equiv \{\theta, \rho_W, \rho_N, \eta_W, \eta_N\}$). First column reports the values of the targeted moments in the data. Second column reports the values of the targeted moments in the model simulation.

Table F.2.2: Targeted Moments for the Special Case of a Conventional Urban Model (All Consumption of Non-traded Services Through Direct Trips from Home)

Moments	Data	Model
log 10-pctile Orig-Dest Travel Time (Workdays)	-1.16	1.61
log 50-pctile Orig-Dest Travel Time (Workdays)	2.78	2.83
log 90-pctile Orig-Dest Travel Time (Workdays)	3.83	3.98
log 10-pctile Orig-Dest Travel Time (Non-Workdays)	1.68	1.64
log 50-pctile Orig-Dest Travel Time (Non-Workdays)	2.81	2.90
log 90-pctile Orig-Dest Travel Time (Non-Workdays)	3.70	4.01

Note: Targeted moments for the estimation of the special case of a conventional urban model, in which all consumption of non-traded services occurs through direct trips from home. In this specification, the multiplicative parameters (η_W, η_N) for the cost of each non-work stay cancel from the numerator and denominator of the travel itinerary choice probabilities, because all travel itineraries have a single non-work stay, and hence these parameters can be set equal to one without loss of generality. We estimate the remaining parameters (θ, ρ_W, ρ_N) using the targeted moments listed in the table. We omit the moments for the fraction of multi-stay travel itineraries because this specification mechanically predicts no variation. First column reports the values of the targeted moments in the data. Second column reports the values of the targeted moments in the model simulation.

G Appendix for General Equilibrium and Counterfactual Simulations

G.1 Market Clearing

We require that demand equals supply in each location in the markets for floor space, commuters and non-traded goods.

Floor Space Market Clearing We assume segmented markets for commercial and residential floor space, consistent with differential land use regulation in Tokyo between these two alternative uses of land. The price of commercial floor space in location j (q_j) is determined by the requirement that income for the owners of commercial floor space is equal to payments for its use in the two sectors:

$$q_j H_j^Y = \sum_{k \in \{T,S\}} (1 - \beta_k) P_{jk} A_{jk} \left(\frac{L_{jk}}{\beta_k} \right)^{\beta_k} \left(\frac{H_{jk}}{1 - \beta_k} \right)^{1-\beta_k}, \tag{G.1}$$

where H_j^Y is the supply of commercial floor space, as determined by land use regulations.[6]

The price of residential floor space in location j (Q_j) is determined by an analogous condition that equates the income of owners of residential floor space with payments for its use:

$$Q_j H_j^R = \alpha_H \overline{w}_j R_j, \tag{G.2}$$

where H_j^R is the supply of residential floor space, as again determined by land use regulations. The average income in each home h, $\overline{w}_h$, is a weighted average of the wage in each workplace j and sector k, weighted by the share of residents commuting to that workplace and sector:

$$\overline{w}_h = \sum_{j=1}^{N} \sum_{k \in \{T,S\}} \frac{\Omega_{hjk}}{\sum_{j'} \sum_{k' \in \{T,S\}} \Omega_{hj'k'}} w_{jk}. \tag{G.3}$$

Commuter Market Clearing Wages in each sector $k \in \{T, S\}$ and workplace j are determined by the requirement that the demand for labor equals the measure of workers choosing to commute to that sector and workplace:

$$\frac{\beta_k}{w_{jk}} P_{jk} A_{jk} \left(\frac{L_{jk}}{\beta_k} \right)^{\beta_k} \left(\frac{H_{jk}}{1 - \beta_k} \right)^{1-\beta_k} = \sum_{h=1}^{\mathcal{N}} \Omega_{hjk}, \tag{G.4}$$

where the left-hand side equals total payments for labor ($w_{jk} L_{jk}$) divided by the wage and we used our normalization of a unit measure of agents.

[6] While we assume fixed supplies of commercial and residential floor space, it is straightforward to introduce an upward-sloping supply function for space supply following Saiz (2010).

Nontraded Services Market Clearing Market clearing for non-traded services requires that the revenue from supplying non-traded services in each location n equals the expenditure of workers from all locations traveling to consume non-traded services in that location:

$$P_{nS}A_{nS}\left(\frac{L_{nS}}{\beta_S}\right)^{\beta_S}\left(\frac{H_{nS}}{1-\beta_S}\right)^{1-\beta_S} = \alpha_S \sum_{h,j,k}\sum_{I} w_{jk}\Omega_{hjk}\tilde{\Lambda}_{I|hj}\Psi_{n|I}, \tag{G.5}$$

where the left-hand side corresponds to the total revenue of nontraded service firms. The right-hand side represents total expenditure on nontraded services by consumers traveling to location j, where $\tilde{\Lambda}_{I|hj} = \xi\Lambda_{I|hj} + (1-\xi)\,\Lambda_{I|h\emptyset}$ is the weighted probability of choosing itinerary I across workdays and non-workdays, and $\Psi_{j|I} = P_{jS}^{1-\sigma}/\left(\sum_{j'\in C(I)} P_{j'S}^{1-\sigma}\right)$ is the expenditure share on location j conditional on choosing itinerary I.

G.2 System of Equations for Counterfactual Simulation

To undertake counterfactuals, we follow an exact-hat algebra approach (Dekle et al. 2007), which uses only the initial values of endogenous variables and model parameters. We denote the values of variables in a counterfactual equilibrium by a prime (e.g, x'), the values of variables in the initial equilibrium without a prime (e.g., x), and the relative changes in variables between the counterfactual and initial equilibria by a hat (e.g., $\hat{x} = x'/x$). Given an assumed change in the exogenous variables of the model (e.g., travel times, τ_{in}), the counterfactual equilibrium can be computed using the values of the structural parameters $\{\sigma, \xi, \theta, \rho_W, \rho_N, \eta_W, \eta_N, \phi, \alpha_T, \alpha_S, \alpha_H, \beta_S, \beta_T, \gamma_B, \gamma_W\}$ and the initial values of nontraded price indexes and wages $\{P_{jS}, w_{jk}\}$, and commuting probabilities, residents and employment $\{\Omega_{hjk}, R_h, L_{jk}\}$.

(i) Changes in travel itineraries and consumption access Using equations (10) and (13) in the paper,

$$\Lambda'_{I|hj} = \frac{Z_{I|hj}\left[\left(\sum_{n\in C(I)} P'^{1-\sigma}_{nS}\right)^{-\frac{1}{1-\sigma}}/\tau'_{I|hj}\right]^{\theta}}{\sum_{\ell\in\mathcal{I}_{hj}} Z_{\ell|hj}\left[\left(\sum_{n\in C(\ell)} P'^{1-\sigma}_{nS}\right)^{-\frac{1}{1-\sigma}}/\tau'_{\ell|hj}\right]^{\theta}},$$

$$\mathbb{A}'_{hj} = \varrho\sum_{\ell\in\mathcal{I}_{hj}} Z_{\ell|hj}\left[\left(\sum_{n\in C(\ell)} P'^{1-\sigma}_{nS}\right)^{-\frac{1}{1-\sigma}}/\tau'_{\ell|hj}\right]^{\theta},$$

$$\hat{\hat{\mathbb{A}}}_{hj} = \hat{\mathbb{A}}'^{\xi}_{hj}\hat{\mathbb{A}}^{(1-\xi)}_{h\emptyset}.$$

(ii) Changes in workplace and residential probabilities Using equation (19) in the paper,

$$\hat{\Omega}_{hjk} = \frac{\left(\hat{B}_h \hat{w}_{jk} \hat{\tilde{\mathbb{A}}}_{hj}^{\alpha_S}\right)^{\phi} \left(\hat{\kappa}_{hj} \hat{Q}_h^{\alpha_H}\right)^{-\phi}}{\sum_{h'=1}^{\mathcal{N}} \sum_{j'=1}^{\mathcal{N}} \sum_{k' \in \{T,S\}} \left(\hat{B}_{h'} \hat{w}_{j'k'} \hat{\tilde{\mathbb{A}}}_{h'j'}^{\alpha_S}\right)^{\phi} \left(\hat{\kappa}_{h'j'} \hat{Q}_{h'}^{\alpha_H}\right)^{-\phi} \Omega_{h'j'k'}}.$$

Together with Ω_{hjk} in the initial equilibrium, we can also compute the changes of the residential and employment population $\hat{R}_h$ and $\hat{L}_{ik}$ using equation (20) in the paper.

(iii) Changes in floor space prices By noticing that floor space market clears separately for production and for residence, and $\frac{1-\beta_k}{\beta_k} w_{ik} L_{ik}$ corresponds to the expenditure for floor space by sector k,

$$\hat{q}_i = \frac{1}{\sum_{k \in \{S,T\}} \frac{1-\beta_k}{\beta_k} w_{ik} L_{ik}} \sum_{k \in \{S,T\}} \hat{w}_{ik} \hat{L}_{ik} \frac{1-\beta_k}{\beta_k} w_{ik} L_{ik}, \tag{G.6}$$

$$\hat{Q}_i = \hat{\overline{w}}_i \hat{R}_i. \tag{G.7}$$

(iv) Changes in productivity and amenity From equations (23), (25) and (18) in the paper,

$$\hat{A}_{iS} = \left(\hat{L}_{iS}\right)^{\frac{\beta_S}{\sigma-1}} \left(\hat{H}_{iS}\right)^{\frac{1-\beta_S}{\sigma-1}}, \tag{G.8}$$

$$\hat{A}_{iT} = \hat{L}_{iT}^{\gamma_A}, \tag{G.9}$$

$$\hat{B}_n = \hat{R}_n^{\gamma_B}. \tag{G.10}$$

(v) Changes in nontraded service prices From equation (G.5), the changes in non-traded goods prices ($\hat{P}_{nS}$) satisfy:

$$\hat{P}_{nS} = \frac{1}{\hat{A}_{nS} \hat{L}_{nS}^{\beta_S} \hat{H}_{nS}^{1-\beta_S}} \frac{\sum_{h,j,k} \sum_I w'_{jk} \Omega'_{hjk} \tilde{\Lambda}'_{I|jh} \Psi'_{n|I}}{\sum_{h,j,k} \sum_I w_{jk} \Omega_{hjk} \tilde{\Lambda}_{I|jh} \Psi_{n|I}}$$

where $\tilde{\Lambda}_{I|hj} = \xi \Lambda_{I|hj} + (1-\xi)\, \Lambda_{I|h\emptyset}$.

(vii) Changes in wages From the zero-profit condition (22), the changes in wages in each sector and location with positive production ($\hat{w}_{ik}$) for $k \in \{S,T\}$ are given by:

$$\hat{w}_{ik} = \left(\frac{\hat{A}_{ik} \hat{P}_{ik}}{\hat{q}_{ik}^{1-\beta_k}}\right)^{1/\beta_k}. \tag{G.11}$$

G.3 Calibration of Wages

We calibrate baseline wages to ensure they are consistent with the market clearing assumptions. In particular, using equations (G.4) and (G.5),

$$w_{nS} = \frac{\alpha_S \beta_S}{L_{nS}} \sum_{h,j,k} \sum_{I} w_{jk} \Omega_{hjk} \tilde{\Lambda}_{I|hj} \Psi_{n|I} \tag{G.12}$$

Second, from the commuting choice probabilities in equation (19), we set

$$w_{nT} = \left(\frac{L_{nT}}{L_{nS}} \right)^{1/\phi} w_{nS}. \tag{G.13}$$

These two equations above determine the wages $\{w_{nS}, w_{nT}\}$ up to scale.

H Appendix for Counterfactual Simulations

In this section of the Online Appendix, we report additional sensitivity and robustness results for our counterfactual simulations in Section 7 of the paper.

H.1 Working from Home (WFH)

In this subsection, we report additional sensitivity analysis for our counterfactual simulations for WFH in Section 7.1 of the paper.

In Table H.1.1, we report changes in non-work stays (Panel A), commercial floor space prices (Panel B), and the number of non-traded varieties (Panel C) in the data and for different model counterfactuals. "(1) Data" corresponds to the observed data; "(2) Model: Baseline" is our estimated model of travel itineraries; "(3) Model: Only direct consumption trips" is the special case of a conventional urban model, in which all consumption trips occur directly from home. These first three specifications are identical to those reported in Table 5 in the paper. As discussed in the paper, our baseline travel itinerary model (row 2) is successful in explaining the observed drop of non-work stays, commercial floor space price, and nontraded varieties in the CBD. In contrast, the conventional urban model assuming direct consumption trips from home (row 3) is unsuccessful in explaining these patterns.

Rows (4) and (5) report results for alternative model specifications, in which we allow for travel itineraries, but restrict the number of possible intermediate stays relative to our baseline specification, which allows for up to four intermediate stays. Row (4) "At most two stays" restricts travel itineraries to include at most 2 stays during a workday (including work) and 2 stays during a non-workday. Row (5) "At most one nonwork stay" restricts travel itineraries to include at most 2 stays during a workday (including work) and 1 stay during a non-workday. We find broadly similar patterns to our baseline specification (row 2). Therefore, to explain the observed impact of the shift to WFH in the data, it is critical to model the complementarity between work and non-work stays through travel itineraries. In contrast, the complementarity between multiple non-work stays plays a more modest role for the shift to WFH.

In row "(6) Model: Baseline (Change parameters)," on top of changing the number of days each week that people commute into work (ξ) and residential amenities in each location ($\overline{B}_n$) in our counterfactual simulation, we also change the travel itinerary parameters. Specifically, we estimate travel itinerary parameters using our smartphone data in 2023 (instead of 2019 in our baseline analysis) following the same procedure described in Section 5.4, as reported in Table H.1.3. We find similar estimated parameter values as for our baseline sample of 2019, indicating that the travel itinerary parameters remain relatively stable following the shift to WFH. Consistent with this observation, we find that the counterfactual simulation results for

this specification in Table H.1.1 are broadly similar to our baseline specification (row 2).

In row "(7) Model: Baseline (Smoothed commuting)," instead of calibrating our model using the observed bilateral commuting flows Ω_{hjk}, we use smoothed commuting flows predicted by the bilateral commuting gravity equations.[7] We study this sensitivity analysis to address the potential randomness from granularity in commuting data, as suggested by Dingel and Tintelnot (2023). We find that the patterns are broadly similar to our baseline specification (row 2), with a slightly smaller drop of non-work stay and nontraded varieties, but still substantially larger than the model with only direct consumption trips (row 3).

In row "(8) Model: Baseline ($\gamma_A = 0.1$)," we consider an alternative value for the productivity spillover parameter of $\gamma_A = 0.1$, instead of our baseline value of $\gamma_A = 0.26$. We find a broadly similar pattern to our baseline specification (row 2), with a slightly smaller drop for commercial floor space prices in the CBD. Therefore, the agglomeration externality in the traded sector makes a modest contribution to the drop in commercial floor space prices in the CBD in response to the shift to WFH.

Overall, regardless of the model specifications considered here, we find that modeling travel itineraries is key to explaining the observed drop of non-work stays, commercial floor space price, and nontraded varieties in the CBD following the shift to WFH.

In Table H.1.2, we also report the patterns of employment by workplace in our data and counterfactual simulations. In row "(1) Data," we report the observed patterns of users' work locations based on our smartphone data in 2019 and 2023. All remaining rows report the same set of model counterfactuals as reported in Table H.1.1. Similar to the patterns for non-work stays, commercial floor space prices, and nontraded varieties, we find that our baseline model (row 2) is successful in explaining the observed drop of employment in the CBD, whereas the conventional urban model assuming direct consumption trips from home (row 3) is unsuccessful in explaining these patterns. This conclusion holds robustly across the same set of alternative specifications considered in Table H.1.1.

[7] Specifically, we estimate by PPML the extended gravity equation of bilateral commuting flows with a proxy of origin fixed effects, destination fixed effects, log bilateral travel time, and our consumption access proxy $\mathbb{A}_{hj}$, corresponding to equation (19) of our paper. We use the predictor of the commuting flows to calibrate our model, instead of the directly observed commuting flows from our smartphone data.

Table H.1.1: Actual and Counterfactual Changes in Non-work Stays, Commercial Floor Space Prices and the Number of Non-traded Varieties Following the Shift to WFH (Additional Model Specifications)

		CBD	High	Medium	Low
(A) Non-Work Stays					
(1)	Data	-0.18	-0.11	0.04	0.03
(2)	Model: Baseline	-0.15	-0.11	0.02	0.04
(3)	Model: Only direct consumption trips	0.00	0.00	0.00	0.00
(4)	Model: At most two stays	-0.16	-0.11	0.02	0.03
(5)	Model: At most one nonwork stay	-0.17	-0.10	0.03	0.03
(6)	Model: Baseline (Change parameters)	-0.17	-0.12	0.01	0.05
(7)	Model: Baseline (Smoothed commuting)	-0.13	-0.11	0.02	0.04
(8)	Model: Baseline ($\gamma_A = 0.1$)	-0.15	-0.11	0.01	0.04
(B) Commercial floor space price (normalized)					
(1)	Data	-0.12	-0.11	-0.02	[0.00]
(2)	Model: Baseline	-0.13	-0.09	-0.01	[0.00]
(3)	Model: Only direct consumption trips	0.00	0.00	0.00	[0.00]
(4)	Model: At most two stays	-0.12	-0.09	-0.01	[0.00]
(5)	Model: At most one nonwork stay	-0.12	-0.08	-0.01	[0.00]
(6)	Model: Baseline (Change parameters)	-0.13	-0.09	-0.01	[0.00]
(7)	Model: Baseline (Smoothed commuting)	-0.12	-0.08	0.01	[0.00]
(8)	Model: Baseline ($\gamma_A = 0.1$)	-0.10	-0.07	-0.01	[0.00]
(C) Nontraded service varieties (normalized)					
(1)	Data	-0.10	-0.04	0.06	[0.00]
(2)	Model: Baseline	-0.11	-0.08	-0.02	[0.00]
(3)	Model: Only direct consumption trips	0.00	0.00	0.00	[0.00]
(4)	Model: At most two stays	-0.11	-0.08	-0.01	[0.00]
(5)	Model: At most one nonwork stay	-0.12	-0.09	-0.01	[0.00]
(6)	Model: Baseline (Change parameters)	-0.11	-0.08	-0.01	[0.00]
(7)	Model: Baseline (Smoothed commuting)	-0.10	-0.08	-0.01	[0.00]
(8)	Model: Baseline ($\gamma_A = 0.1$)	-0.10	-0.08	-0.02	[0.00]

Note: "(1) Data" corresponds to the observed patterns in our data (see Online Appendix A.9 and B for the data sources); "(2) Model: Baseline" is our estimated model of travel itineraries; "(3) Model: Only direct consumption trips" is the special case of a conventional urban model, in which all consumption trips occur directly from home; "(4) Model: At most two stays" restricts travel itineraries to include at most 2 stays during a workday (including work) and 2 stays during a non-workday. "(5) At most one nonwork stay" restricts travel itineraries to include at most 2 stays during a workday (including work) and 1 stay during a non-workday. "(6) Model: Baseline (Change parameters)" uses our estimated baseline model of travel itineraries as in (1), but with the travel itinerary parameters estimated using the post-WFH data (2023) reported in Table H.1.3; "(7) Model: Baseline (Smoothed commuting)" uses smoothed commuting flows from estimating a bilateral commuting equation using the Poisson Pseudo Maximum Likelihood (PPML) estimator to address randomness due to granularity in commuting data following Dingel and Tintelnot (2023); "(8) Model: Baseline ($\gamma_A = 0.1$)" uses an alternative value for the productivity spillover parameter γ_A. Panel (B) reports mean log changes in commercial floor space prices; Panel (C) reports mean log changes in the number of non-traded varieties; in Panels (B) and (C), commercial floor space prices and the number of non-traded varieties normalized such that the mean for low-density locations is zero (as indicated by $[0.00]$); central business district (CBD) corresponds to locations within 2km from the centroid of the Oaza of Chiyoda ward; high-density corresponds to locations with employment densities in the top 10 percent; mid-density corresponds to locations with employment densities from the 50-90th percentiles; low-density corresponds to locations with employment densities from the 0-50th percentiles.

Table H.1.2: Actual and Counterfactual Changes in Employment by Workplace Following the Shift to WFH

		CBD	High	Medium	Low
(1)	Data	-0.15	-0.14	0.04	0.05
(2)	Model: Baseline	-0.12	-0.07	0.03	0.03
(3)	Model: Only direct consumption trips	0.00	0.00	0.00	0.00
(4)	Model: At most two stays	-0.11	-0.07	0.03	0.03
(5)	Model: At most one nonwork stay	-0.11	-0.07	0.03	0.03
(6)	Model: Baseline (Change parameters)	-0.11	-0.07	0.03	0.03
(7)	Model: Baseline (Smoothed commuting)	-0.10	-0.06	0.03	0.03
(8)	Model: Baseline ($\gamma_A = 0.1$)	-0.10	-0.06	0.02	0.03

Note: "(1) Data" corresponds to the observed patterns of users' work locations in our smartphone data; "(2) Model: Baseline" is our estimated model of travel itineraries; "(3) Model: Only direct consumption trips" is the special case of a conventional urban model, in which all consumption trips occur directly from home; "(4) Model: At most two stays" restricts travel itineraries to include at most 2 stays during a workday (including work) and 2 stays during a non-workday. ; "(5) At most one nonwork stay" restricts travel itineraries to include at most 2 stays during a workday (including work) and 1 stay during a non-workday; "(6) Model: Baseline (Change parameters)" uses our estimated baseline model of travel itineraries as in (1), but with the travel itinerary parameters estimated using the post-WFH data (2023) reported in Table H.1.3; "(7) Model: Baseline (Smoothed commuting)" uses smoothed commuting flows from estimating a bilateral commuting equation using the Poisson Pseudo Maximum Likelihood (PPML) estimator to address randomness due to granularity in commuting data following Dingel and Tintelnot (2023); "(8) Model: Baseline ($\gamma_A = 0.1$)" uses an alternative value for the productivity spillover parameter γ_A. Panel (B) reports mean log changes in commercial floor space prices; Panel (C) reports mean log changes in the number of non-traded varieties; in Panels (B) and (C), commercial floor space prices and the number of non-traded varieties normalized such that the mean for low-density locations is zero (as indicated by [0.00]); central business district (CBD) corresponds to locations within 2km from the centroid of the Oaza of Chiyoda ward; high-density corresponds to locations with employment densities in the top 10 percent (excluding the CBD); mid-density corresponds to locations with employment densities from the 50-90th percentiles (excluding the CBD); low-density corresponds to locations with employment densities from the 0-50th percentiles (excluding the CBD).

Table H.1.3: Parameter Estimates: Pre- and Post-WFH

Parameter	(1) Pre-Period	(2) Post-Period
σ	4.60	4.60
θ	3.43 (0.08)	3.43 (0.10)
ρ_W	0.85 (0.02)	0.87 (0.03)
ρ_N	0.58 (0.01)	0.57 (0.02)
η_W	1.56 (0.02)	1.56 (0.01)
η_N	2.58 (0.05)	2.58 (0.03)

Notes: Elasticity of substitution (σ) determined using the ratio of variable profits and revenues in the retail sector; travel parameters estimated using SMM, including workday travel cost (ρ_W); non-workday travel cost (ρ_N); workday intermediate-stay cost (η_W); non-workday intermediate-stay cost (η_N); and dispersion of idiosyncratic preferences (θ); moments weighted using the efficient weighting matrix for the SMM estimates; Pre-period and Post-period indicates the estimates using our data in 2019 and 2023, respectively.

H.2 Travel Itineraries and Agglomeration

In this subsection, we report additional analysis for our counterfactual simulations in which we shut down travel itineraries by assuming that intermediate-stay costs become prohibitive, as discussed in Section 7.2 of the paper.

In Figure H.2.1, we report the results of alternative counterfactual simulations. Starting from the observed equilibrium in the data in 2019, we undertake counterfactuals in which we assume that intermediate-stay costs become prohibitive, such that agents instead make one direct trip to consume non-traded services and another direct trip to commute into work. "Shut Down All Travel Itinerary" shows the fitted values from local polynomial regressions of counterfactual log changes in employment on log travel time to the CBD from our baseline specification in Figure 6 in the paper. As discussed in the paper, we find that travel itineraries play an important role in the concentration of employment in Central Tokyo. In the most central parts of the city (within 20 minutes from the CBD), overall employment by workplace falls by up to around one half, which is compensated by an increase in (nontraded sector) employment in the inner suburbs (within 55 minutes from the CBD), and an increase in (traded sector) employment in the city's outskirts (outside 55 minutes from the CBD).

To further understand the forces underlying these results, we now examine the role of complementarities between work and non-stays within the same travel itinerary versus complementarities between multiple non-work stays. We undertake two additional counterfactual simulations that limit the number of non-work stays relative to our baseline specification (which allows for up to four non-work stays within the same travel itinerary). "Restrict Travel Itinerary to At Most One Non-Work Stay" restricts travel itineraries to include at most 2 stays during a workday (including work) and 1 stay during a non-workday. "Restrict Travel Itinerary to At Most Two Stays" restricts travel itineraries to include at most 2 stays during a workday (including work) and 2 stays during a non-workday.

In each of these counterfactuals, we find modest declines in employment in the CBD, and an increase in employment in the inner and outer suburbs (the two sets of counterfactual predictions are close together such that they are hard to distinguish visually). The magnitudes of these employment changes are substantially smaller than for our baseline counterfactual shutting down all travel itineraries. This pattern of results is consistent with the idea that the complementarity between work and non-work stays within the same travel itinerary is a key agglomeration force. In contrast, the complementarity between multiple non-work stays within the same travel itinerary plays a smaller role, in part because of the substitution of expenditure across non-work stays within the same travel itinerary. As a result, we find smaller reallocations of employment when we limit the number of non-work stays within the same travel itinerary than when we shut down all travel itineraries.

Figure H.2.1: Counterfactual Employment Changes Without Travel Itineraries and Limiting the Number of Non-work Stays Within the Same Travel Itinerary

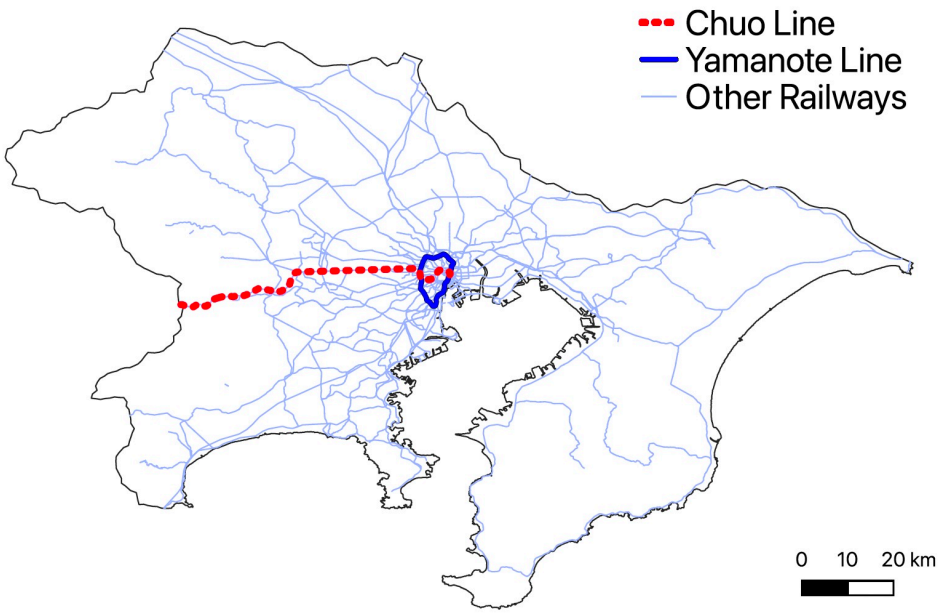
Note: "Shut Down All Travel Itinerary" indicates a counterfactual in which we assume that intermediate-stay costs become prohibitive, such that agents instead make one direct trip to consume non-traded services and another direct trip to commute into work. "Restrict Travel Itinerary to At Most One Non-Work Stay" restricts travel itineraries to include at most 2 stays during a workday (including work) and 1 stay during a non-workday. "Restrict Travel Itinerary to At Most Two Stays" restricts travel itineraries to include at most 2 stays during a workday (including work) and 2 stays during a non-workday. Three lines show fitted values from local polynomial regressions of counterfactual log changes in employment on log travel time to the CBD; solid, dashed and dotted lines represent values for the change in employments under, no travel itinerary, at most one non-work stay, and at most two non-work stays respectively; histograms show the distribution of employment (green) and residents (blue) in the initial equilibrium by distance bin from the CBD.

H.3 Place-Based Infrastructure Policies

In this subsection, we report additional sensitivity analysis for our counterfactual simulations for place-based transportation infrastructure investment in Section 7.3 of our paper.

We undertake counterfactuals for two specific transport improvements: (i) the construction of the Yamanote Railway Line, which follows a circular route connecting subcenters of central Tokyo; (ii) the construction of the Chuo Railway Line, which follows a radial route that connects an outer suburb of Tokyo to the center of the city (as shown in Figure H.3.1 below). First, we recompute bilateral travel times between locations within Tokyo without each of these railway lines. Second, starting from the observed equilibrium in the data in our baseline sample in 2019, we solve for an exact-hat algebra counterfactual equilibrium without each railway line. Third, we compute the percentage welfare gain from constructing each of these railway lines, which equals welfare in the actual equilibrium with these lines divided by welfare in the counterfactual equilibrium without each line. We abstract from construction costs, because they are the same across all of the different model specifications that we consider.

Figure H.3.1: Routes of the Chuo Line and Yamanote Line within the Tokyo Metropolitan Area



Note: Map showing the boundaries of the Tokyo Metropolitan Area and the Routes of the Chuo Line, Yamanote Line, and Other Railway Lines.

Table H.3.1 reports the estimated welfare gains from these counterfactuals. The first three rows reproduce the results reported in Table 6 of the paper. Row (1) corresponds to our baseline estimated travel itinerary model; row (2) corresponds to a conventional urban model, in

which all travel (commuting and consumption) occurs through direct trips from home, using the parameter estimates $\{\theta, \rho_W, \rho_N\}$ under this model specification (Table 2; Column 2); row (3) corresponds to a conventional urban model with only direct trips from home, using the parameter estimates $\{\theta, \rho_W, \rho_N\}$ from our baseline model (Table 2; Column 1).

As discussed in the paper, abstracting from travel itineraries leads to an undervaluation of the welfare gains from these transport improvements for two main reasons. First, it underestimates the travel cost parameters (because of the assumption that all travel occurs from home). Second, it undercounts travel that occurs within multi-stay travel itineraries. Together, these two forces lead to a larger undervaluation of the welfare gains from the urban circular link of the Yamanote Line, because it is disproportionately used for multi-stay travel itineraries (stopping off at intermediate destinations while commuting from home to work). Therefore, taking travel itineraries into account changes the relative welfare gains from alternative transportation infrastructure interventions, in favor of investing in urban commercial centers.

To further understand these results, we again examine the role of complementarities between work and non-stays within the same travel itinerary versus complementarities between multiple non-work stays. We undertake two additional counterfactual simulations that limit the number of non-work stays relative to our baseline specification (which allows for up to four non-work stays within the same travel itinerary). Row (4) "At Most Two Stays" restricts travel itineraries to include at most 2 stays during a workday (including work) and 2 stays during a non-workday. Row (5) "At Most One Non-Work Stay" restricts travel itineraries to include at most 2 stays during a workday (including work) and 1 stay during a non-workday. In both cases, we find that welfare gains are smaller than our baseline specification (row 1), especially for the Yamanote Line. Therefore, abstracting from trips that occur within multi-stay travel itineraries that involve more than two stays is consequential for the underprediction of the welfare gains from urban transportation infrastructure improvements, and particularly so for the Yamanote Line, which is used extensively for multi-stay travel itineraries.

The remaining rows of the table report additional sensitivity analyses. In row (6), instead of calibrating our model using the observed bilateral commuting flows Ω_{hjk} in the initial equilibrium, we use smoothed commuting flows from estimating a bilateral commuting equation using the Poisson Pseudo Maximum Likelihood (PPML) estimator. We study this sensitivity analysis to address the potential randomness from granularity in commuting data, as suggested by Dingel and Tintelnot (2023). Our results are broadly similar to our baseline specification (row 1), with differences of up to 15 percent.

In row (7), we use an alternative value for the productivity spillover parameter of $\gamma_A = 0.1$, instead of our baseline value of $\gamma_A = 0.26$. In row (8), we use an alternative value for the amenity spillover parameter of $\gamma_B = 0.1$, instead of our baseline value of $\gamma_B = 0$. Across the

board, our results are broadly similar to our baseline specification (row 1). Therefore, compared to the omission of travel itineraries, we find that technological productivity or amenity spillovers have a limited influence on the evaluation of these transportation infrastructure improvements in our empirical setting.

Overall, regardless of the alternative model specifications considered here, we find that the consumption externalities implied by travel itineraries are quantitatively relevant for evaluating alternative place-based infrastructure projects, and favor investments in central areas where travel itineraries disproportionately occur.

Table H.3.1: Welfare Gains from Alternative Transport Improvements
(Additional Specifications)

(A) Yamanote Line (Circular Lines within the Central City)

Specification	Welfare Gains (%)	Relative to Baseline (%)
(1) Baseline	0.44	100
(2) Only Direct Consumption Trips	0.27	61
(3) Only Direct Consumption Trips (Baseline θ, ρ_W, ρ_S)	0.36	80
(4) At Most Two Stays	0.39	87
(5) At Most One Nonwork Stay	0.37	83
(6) Baseline, Smoothed Commuting	0.38	85
(7) Baseline, $\gamma_A = 0.1$	0.45	100
(8) Baseline, $\gamma_B = 0.1$	0.45	100

(B) Chuo Line (Radial Lines between the Central City and Suburbs)

Specification	Welfare Gains (%)	Relative to Baseline (%)
(1) Baseline	0.71	100
(2) Only Direct Consumption Trips	0.49	68
(3) Only Direct Consumption Trips (Baseline θ, ρ_W, ρ_S)	0.66	93
(4) At Most Two Stays	0.65	91
(5) At Most One Nonwork Stay	0.63	88
(6) Baseline, Smoothed Commuting	0.77	108
(7) Baseline, $\gamma_A = 0.1$	0.71	100
(8) Baseline, $\gamma_B = 0.1$	0.70	98

Note: Welfare gains from the construction of the Chuo Railway Line (radial route between a suburb and the center) and Yamanote Railway Line (circular route connecting subcenters of central Tokyo); table reports welfare in the actual equilibrium (with these railway lines) relative to welfare in the counterfactual equilibrium (without each railway line); row (1) corresponds to our baseline estimated travel itinerary model; row (2) corresponds to a conventional urban model, in which all travel (commuting and consumption) occurs through direct trips from home, using the parameter estimates $\{\theta, \rho_W, \rho_N\}$ under this model specification (Table 2; Column 2); row (3) corresponds to a conventional urban model with only direct trips from home, using the parameter estimates $\{\theta, \rho_W, \rho_N\}$ from our baseline model (Table 2; Column 1); row (4) restricts travel itineraries to include at most 2 stays during a workday (including work) and 2 stays during a non-workday; row (5) restricts travel itineraries to include at most 2 stays during a workday (including work) and 1 stay during a non-workday; row (6) uses smoothed commuting flows from estimating a bilateral commuting equation using the Poisson Pseudo Maximum Likelihood (PPML) estimator to address randomness due to granularity in commuting data following Dingel and Tintelnot (2023); row (7) uses an alternative value for the productivity spillover parameter γ_A; and row (8) uses an alternative value for the amenity spillover parameter γ_B. Second column reports percentage welfare gains from each transport improvement; third column reports the welfare gain from each transport improvement relative to that our baseline estimated travel itinerary model (as a percentage).

References

Ackerberg, Daniel A. (2009) "A New Use of Importance Sampling to Reduce Computational Burden in Simulation Estimation," *Quantitative Marketing and Economics*, 7(4), 343–376.

Agarwal, Sumit, Ferdinando Monte and J. Bradford Jensen (2020) "Consumer Mobility and the Local Structure of Consumption Industries," Georgetown University, mimeograph.

Aksoy, Cevat Giray, Jose Maria Barrero, Nicholas Bloom, Steven J. Davis, Mathias Dolls and Pablo Zarate (2022) "Working from Home Around the World," *NBER Working Paper*, 30446.

Althoff, Lukas, Fabian Eckert, Sharat Ganapati and Conor Walsh (2022) "The Geography of Remote Work," *Regional Science and Urban Economics*, 93, 103770.

Anas, Alex (2007) "A Unified Theory of Consumption, Travel and Trip Chaining," *Journal of Urban Economics*, 62(2), 162–186.

Barrero, José María, Nicholas Bloom and Steven J. Davis (2023) "The Evolution of Work from Home," *Journal of Economic Perspectives*, 37(4), 23-50.

Benmelech, Efraim, Nittai Bergman, Anna Milanez and Vladimir Mukharlyamov (2019) "The Agglomeration of Bankruptcy," *Review of Financial Studies*, 32(7), 2541–2586.

Bernstein, Shai, Emanuele Colonnelli, Xavier Giroud, and Benjamin Iverson (2019) "Bankruptcy Spillovers," *Journal of Financial Economics*, 133(3), 608–633.

Borusyak, K., X. Jaravel, and J. Spiess (2024) "Revisiting Event Study Designs: Robust and Efficient Estimation," *Review of Economic Studies*, 91(6), 3253–3285.

CATS (1956) Chicago Area Transportation Study, 1956 Base Year Statistics, Vol. 1.

Callaway, Brantly and Pedro H.C. Sant'Anna (2021) "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 225(2), 200-230.

Couture, Victor (2016) "Valuing the Consumption Benefits of Urban Density," Vancouver School of Economics, University of British Columbia, mimeograph.

Couture, Victor, Jonathan I. Dingel, Allison Green, Jessie Handbury and Kevin R. Williams (2022) "Measuring Movement and Social Contact with Smartphone Data: a Real-time Application to COVID-19," *Journal of Urban Economics*, JUE Insight, 127, 103328.

Dekle, Robert, Jonathan Eaton and Samuel Kortum (2007) "Unbalanced Trade," *American Economic Review*, 97(2), 351–355.

Dingel, Jonathan and Brent Neiman (2020) "How Many Jobs Can be Done at Home?" *Journal of Public Economics*, 189, 104235.

Einav, Liran, Peter J. Klenow, Jonathan D. Levin, and Raviv Murciano-Goroff (2021) "Customers and Retail Growth," *NBER Working Paper*, 29561.

Gardner, John (2021) "Two-stage Difference-in-Differences," Working Paper.

Gupta, Arpit, Vrinda Mittal and Stijn Van Nieuwerburgh (2022) "Work from Home and the Office Real Estate Apocalypse," *NBER Working Paper*, 30526.

Kloek, T. and H. K. van Dijk (1978) "Bayesian Estimates of Equation System Parameters: An Application of Integration by Monte Carlo," *Econometrica*, 46(1), 1-19.

Knight, Samsun (2022) "Retail Demand Interdependence and Chain Store Closures," University of Toronto, mimeo.

Koster, H. R. A., I. Pasidis and J. van Ommeren (2019) "Shopping Externalities and Retail Concentration: Evidence from Dutch Shopping Streets," *Journal of Urban Economics*, 114 (Au- gust), 103194.

Kreindler, Gabriel and Yuhei Miyauchi (2023) "Measuring Commuting and Economic Activity inside Cities with Cell Phone Records," *Review of Economics and Statistics*, 105(4), 899-909.

Monte, Ferdinando, Charly Porcher and Esteban Rossi-Hansberg (2023) "Remote Work and City Structure," *NBER Working Paper*, 31494.

Oh, Ryungha and Jaeeun Seo (2023) "What Causes Agglomeration of Services? Theory and Evidence from Seoul," Yale University, mimeograph.

O'Kelly, M and E. Miller (1984) "Characteristics of Multistop, Multipurpose Travel: An Empirical Study of Trip Length," *Transportation Research Record*, 976, 33-9.

Okubo, Toshihiro (2022) "Telework in the Spread of Covid-19," *Information Economics and Policy*, 60, 100987.

Okubo and NIRA (2022) "The 7th Survey on Telework," Technical Report, September 2022.

Oster, C. (1973) "Household Tripmaking to Multiple Destinations: The Overlooked Urban Travel Pattern," *Traffic Quarterly*, 32, 511-29.

Pollmann, Michael (2020) "Causal Inference for Spatial Treatments," *arXiv:2011.00373 [econ, stat]*.

Primerano, Frank, Michael AP Taylor, Ladda Pitaksringkarn and Peter Tisato (2008) "Defining and Understanding Trip Chaining Behaviour," *Transportation*, 35 (1), 55–72.

Qian, Franklin, Qianyang Zhang and Xiang Zhang (2024) "Identifying Agglomeration Spillovers: Evidence from Grocery Store Openings," Princeton University, mimeo.

Ramani, Arjun and Nicholas Bloom (2021) "The Donut Effect of Covid-19 on Cities," *NBER Working Paper*, 28876.

Relihan, L. E. (2022) "Is Online Retail Killing Coffee Shops? Estimating the Winners and Losers of Online Retail using Customer Transaction Microdata," *Working Paper*.

Roth, Jonathan and Pedro H. C. Sant'Anna (2021) "Efficient Estimation for Staggered Rollout Designs," *arXiv*, arXiv:2102.01291.

Roth, Jonathan, Pedro H.C. Sant'Anna, Alyssa Bilinski and John Poe (2023) "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature," *Journal of Econometrics*, 235(2), 2218-2244.

Saiz, Albert (2010) "The Geographic Determinants of Housing Supply," *Quarterly Journal of Economics*, 125(3), 1253-1296.

Shoag, Daniel and Stan Veuger (2010) "Shops and the City: Evidence on Local Externalities and Local Government Policy from Big-Box Bankruptcies," *Review of Economics and Statistics*, 100(3), 440-453.

Sun, Liyang and Sarah Abraham (2021) "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects," *Journal of Econometrics*, 225(2), 175-199.

Thill, Jean-Claude and Isabelle Thomas (1987) "Toward Conceptualizing Trip Chaining: A Review," *Geographical Analysis*, 19(1), 1-17.

Transportation Research Board (2006) *Commuting in America*, Washington DC: Department of Transportation.