# Consumption Access and Agglomeration: Evidence from Smartphone Data*

Yuhei Miyauchi†
*Boston University*

Kentaro Nakajima‡
*Hitotsubashi University*

Stephen J. Redding§
*Princeton University, NBER and CEPR*

October 2, 2020

**Abstract**

We provide new theory and evidence on the role of consumption access in understanding the agglomeration of economic activity. We combine smartphone data that records user location every 5 minutes of the day with economic census data on the location of service-sector establishments to measure commuting and non-commuting trips within the Greater Tokyo metropolitan area. We show that non-commuting trips are frequent, more localized than commuting trips, and strongly related to the availability of nontraded services. Guided by these empirical findings, we develop a quantitative urban model that incorporates travel to work and to consume non-traded services. Using the structure of the model, we estimate theoretically-consistent measures of consumption and workplace access, and show that consumption access is almost as important as workplace access in explaining the observed variation in residents and land prices across locations. Undertaking counterfactuals for changes in travel costs, we show that abstracting from consumption trips leads to a substantial underestimate of the welfare gains from a transport improvement (because of the undercounting of trips) and leads to a distorted picture of changes in travel patterns within the city (because of the different geography of comuting and non-commuting trips).

Keywords: Agglomeration, Urbanization, Transportation

JEL Classification: O18, R12, R40

*We are grateful to Boston University, Hitotsubashi University, and Princeton University for research support. Thanks to Don Davis, Joan Monras, Daniel Sturm, and seminar participants at Boston University, Nihon University, UC Berkeley and OSUS for helpful comments. We are grateful to Takeshi Fukasawa and Peter Deffebach for excellent research assistance. The usual disclaimer applies. "Konzatsu-Tokei (R)" Data refers to people flow data constructed from individual location information sent from mobile phones under users' consent, through applications provided by NTT DOCOMO, INC (including mapping application *Docomo Chizu NAVI*). Those data is processed collectively and statistically in order to conceal private information. Original location data is GPS data (latitude, longitude) sent every five minutes, and it and does not include information to specify individual. The copyrights of all tables and figures presented in this document belong to ZENRIN DataCom CO., LTD. This research was the result of the joint research with Center for Spatial Information Science, the University of Tokyo (No. 954) and used Real Estate Database provided by At Home Co., Ltd and Zmap-TOWN II Digital Building Map Data provided by Zenrin Co. We thank Research Center for Property Assessment System for providing land price data.

†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 WWS, Julis Romo Rabinowitz Building, Princeton, NJ 08544. Tel: 1-609-258-4016. Email: reddings@princeton.edu.

# 1 Introduction

Understanding the agglomeration of economic activity is one of the most central challenges in economics. Traditional theories of agglomeration emphasize increasing returns in production and the costs of workers commuting between their workplace and residence. However, much of the travel that occurs within urban areas is related not to commuting but rather to the consumption of nontraded services, such as trips to restaurants, coffee shops and bars, shopping expeditions, excursions to cinemas, theaters, music venues and museums, and visits to professional service providers. Although a growing number of writers have emphasized the idea of the "consumer city" and the role of consumption in agglomeration, two major challenges faced by research in this area are a limited ability to measure consumption trips within cities and the absence of a widely-accepted theoretical model of agglomeration in consumption.[1] In this paper, we provide new theory and evidence on the role of consumption and workplace access in understanding agglomeration. We combine smartphone data including high-frequency location information with spatially-disaggregated economic census data to measure commuting and non-commuting trips within the Greater Tokyo metropolitan area. Guided by our empirical findings, we develop a quantitative urban model that incorporates both workplace and consumption access. We use the model to evaluate the role of consumption access as a source of agglomeration and for explaining the observed variation in land prices. We show that incorporating consumption access is quantitatively relevant for evaluating transport infrastructure improvements.

We first use our smartphone data to provide fine resolution evidence on travel within the Greater Tokyo metropolitan area. Our data come from a major smartphone mapping application in Japan (*Docomo Chizu NAVI*), which records the Geographical Positioning System (GPS) location of each device every 5 minutes. In July of 2019, the data covers about 545,000 users, with 1.4 billion data points. We measure each location visited by a user using a "stay," which corresponds to no movement within 100 meters for 15 minutes. Based on this definition, we measure each anonymized user's home location as her most frequent location (defined by groups of geographically contiguous stays) and her work location as her second most frequent location. We allocate non-commuting trips to other locations into different types using spatially-disaggregated census data on employment by sector. We validate our smartphone commuting measures by comparing them with official census data. We show that our measures of the shares of residents and workers in each municipality are strongly correlated with those from census data. We also demonstrate similar bilateral commuting patterns between municipalities as in census data.

Having validated our smartphone data using census commuting data, we show that focusing solely on these commuting trips provides a misleading picture of travel within the Tokyo metropolitan area. First, we show that non-commuting trips are more frequent than commuting trips, so that concentrating solely on commuting trips substantially underestimates the amount of travel within urban areas. This finding using our smartphone data is consistent with other evidence from travel surveys. A key advantage of our smartphone data is that they reveal bilateral patterns of travel within the city at a fine level of spatial disaggregation, which is central to our quantitative analysis of the model below. Second, we find that non-commuting trips have destination closer to home than commuting trips, with elasticities of travel flows to travel times that are larger in absolute view than those for commuting trips. Therefore, focusing solely on commuting trips also yields a misleading picture of bilateral patterns of travel within cities.

[1] Early research on the "consumer city" is Glaeser, Kolko, and Saiz (2001) and an influential popular discussion is Florida (2009). A growing number of empirical studies provide empirical evidence of endogenous amenities, including Ahlfeldt, Redding, Sturm, and Wolf (2015) and Diamond (2016), for which we show consumption access provides microfoundations.

Furthermore, as workers choose their residence taking into account access to surrounding locations, the fact that non-commuting trips are concentrated closer to home makes them particularly relevant for these residential location decisions. Third, using our spatially-disaggregated data on employment by sector, we show that these non-commuting trips are closely related to the availability of nontraded services, which is consistent with our modelling of them as travel to consume non-traded services.

We next develop quantitative theoretical of internal city structure that incorporates both consumption and workplace access. We consider a city that consists of a discrete set of blocks that differ in productivity, amenities, supply of floor space and transport connections. Consumer preferences are defined over consumption of a traded good, a number of different types of nontraded services, and residential floor space. The traded good and nontraded services are produced using labor and commercial floor space. We assume that workers choose their residence before observing match-specific productivities for each workplace and idiosyncratic preferences for the nontraded services provided by each location. After choosing their residence, workers select their preferred workplace and locations for consuming each nontraded service. Population mobility implies that workers must obtain the same expected utility from all locations with positive population.

We show that the model implies gravity equations for commuting and non-commuting trips, which provide good approximations to the observed data, and can be used to estimate theoretically-consistent measures of workplace access and consumption access. We use the model's population mobility condition to derive a sufficient statistic for the relative attractiveness of locations, which incorporates both the residential population share and the price of floor space. We show that this sufficient statistic for the relative attractiveness of locations can be decomposed into three components: (i) workplace access, (ii) consumption access, and (iii) residential amenities. Workplace access captures proximity to surrounding employment opportunities and depends on wages and travel times. Consumption access captures proximity to surrounding nontraded services and depends on the prices of these non-traded services and travel times. Residential amenities are a residual and capture unobserved characteristics that make a location more attractive as a residence, such as leafy streets and scenic views. We find that the contribution from consumption access is comparable to the contribution from workplace access (25 percent compared to 45 percent) and that controlling for consumption access substantially reduces the contribution from unobserved amenities (from 50 percent to 30 percent). Taken together, this pattern of results is consistent with the idea that much economic activity in urban areas is concentrated in the service sector, and that access to these services varies within urban areas.

We show how the model can be used to undertake a counterfactual for a transport infrastructure improvement, such as the construction of a new subway line, using the initial values of the endogenous variables and the exact-hat algebra approach of Dekle, Eaton, and Kortum (2007). In addition to initial shares of commuting trips, the predictions of these counterfactuals now also depend on initial shares of non-commuting trips. As a result, frameworks that focus solely on commuting trips generally underestimate the welfare gains from transport infrastructure improvements, because they undercount the number of passenger journeys that benefit from the reduction in travel costs. Furthermore, these frameworks generate different predictions for the impact of the new transport infrastructure on the spatial distribution of economic activity, because of the different bilateral patterns of commuting and non-commuting trips. Undertaking counterfactuals for the construction of a new subway line, we show that taking consumption access into account is quantitatively relevant for the impact of this new transport infrastructure on travel patterns, the spatial organization of economic activity within the city, and welfare.

Our paper is related to a number of different strands of research. First, our paper connects with the broad theoretical and empirical literature on agglomeration, including Henderson (1974), Fujita, Krugman, and Venables (1999), Fujita and Thisse (2002), Davis and Weinstein (2002) and Kline and Moretti (2014), as reviewed in Rosenthal and Strange (2004), Duranton and Puga (2004), Moretti (2011) and Combes and Gobillon (2015). Although the vast majority of this existing research emphasizes agglomeration economies of production, a key focus of our analysis is the way in which access to nontraded services provides an agglomeration force in consumption.

Second, a key part of this agglomeration literature is concerned with the internal structure of cities. Early contributions assumed monocentric organizations of economic activity, including Alonso (1964), Mills (1967) and Muth (1969). Subsequent research explored the conditions under which non-monocentric organizations of economic activity emerge in stylized settings such as a linear city or a symmetric circular city, including Fujita and Ogawa (1982), Fujita and Krugman (1995) and Lucas and Rossi-Hansberg (2002). More recent research has developed quantitative models of internal city structure that allow for asymmetries across locations and yet remain amenable to quantitative analysis, including Ahlfeldt, Redding, Sturm, and Wolf (2015), Allen, Arkolakis, and Li (2017), Monte, Redding, and Rossi-Hansberg (2018), Tsivanidis (2018), Dingel and Tintelnot (2020), and Owens, Rossi-Hansberg, and Sarte (2020), as reviewed in Redding and Rossi-Hansberg (2017).[2] All of these studies emphasize commuting and the separation of workplace and residence. In contrast, one of our main contributions is to highlight the importance of travel to consume nontraded services in shaping agents' location decisions.

Third, our findings relate to recent research within the agglomeration literature on endogenous amenities and social and spatial frictions. Evidence of endogenous amenities has been provided in the context of spatial sorting (Diamond 2016 and Almagro and Domínguez-Iino 2019), gentrification and neighborhood change within cities (Couture and Handbury 2019, Couture, Dingel, Green, and Handbury 2019, Hoelzlein 2020 and Allen, Fuchs, Ganapati, Graziano, Madera, and Montoriol-Garriga 2020), and industry clustering (Leonardi and Moretti 2019). Evidence that both spatial and social frictions are consequential for shaping consumption location decisions has been provided using restaurant choice data (Davis, Dingel, Monras, and Morales 2019), credit card data (Agarwal, Jensen, and Monte 2020 and Dolfen, Einav, Klenow, Klopack, Levin, Levin, and Best 2019), travel surveys and ride sharing data (Gorback 2020 and Zárate 2020) and cellphone data (Couture, Dingel, Green, and Handbury 2019, Athey, Ferguson, Gentzkow, and Schmidt 2018, Kreindler and Miyauchi 2019, Gupta, Kontokosta, and Van Nieuwerburgh 2020 and Büchel, Ehrlich, Puga, and Viladecans 2020). Relative to these existing studies, we incorporate consumption trips into a quantitative urban model of internal city structure, and use high-frequency and spatially-disaggregated data on these consumption trips to evaluate their implications for the strength of agglomeration forces and the impact of transport infrastructure improvements.

Fourth, our work contributes to research on transport infrastructure and the spatial distribution of economic activity. One strand of empirical research has used quasi-experimental variation to provide evidence on the causal impact of transport infrastructure improvements, including Chandra and Thompson (2000), Baum-Snow (2007), Michaels (2008), Duranton and Turner (2011, 2012), Faber (2014), Storeygard (2016), Baum-Snow, Brandt, Henderson, Turner, and Zhang (2017), and Couture, Duranton, and Turner (2018). A second line of work has used quantitative spatial models to evaluate general equilibrium impacts of transport infrastructure investments, including Anas and Liu (2007),

[2]The broader literature on quantitative spatial models across cities or regions includes Allen and Arkolakis (2014), Caliendo, Parro, Rossi-Hansberg, and Sarte (2018), Fajgelbaum and Gaubert (2020), Ramondo, Rodríguez-Clare, and Saborío-Rodríguez (2016), and Redding (2016)

Donaldson (2018), Donaldson and Hornbeck (2016), Heblich, Redding, and Sturm (2020), Tsivanidis (2018), Severen (2019) and Balboni (2019). A third group of papers has compared actual and optimal transport networks, including Allen and Arkolakis (2017) and Fajgelbaum and Schaal (2020). While all of this existing research focuses on the costs of transporting goods and commuting costs, a key feature of our work is to highlight the role of the transport network in providing access to nontraded services.

The remainder of the paper is structured as follows. Section 2 introduces our data. Section 3 presents reduced-form evidence on travel patterns that motivates the theoretical model that we develop below. Section 4 introduces our quantitative urban model that incorporates a role for both consumption access and workplace access in influencing location choices. Section 5 uses the model to the quantify the relative importance of consumption and workplace access for explaining the spatial concentration of economic activity. Section 6 shows that incorporating consumption access is quantitatively relevant for evaluating the counterfactual impact of transport infrastructure improvements, such as the construction of a new subway line. Section 7 concludes.

# 2 Data Description

In this section, we introduce our main smartphone data and the other data used in the quantitative analysis of the model. In Subsection 2.1, we discuss our smartphone data and explain how we use it to identify home location, work location, commuting trips and non-commuting trips. In Subsection 2.2, we discuss the spatially-disaggregated economic census data by sector and location that we use to distinguish between different types of non-commuting trips, as well data on land values and other location characteristics. In Subsection 2.3, we report validation checks of the commuting measures from our smartphone data using official census data on employment by residence, employment by workplace and bilateral commuting flows.

## 2.1 Smartphone GPS Data

Our main data source is one of the leading smartphone mapping applications in Japan: *Docomo Chizu NAVI*. Upon installing this application, individuals are asked to give permission to share location information in an anonymized form. Conditional on this permission being given, the application collects the Geographical Positioning System (GPS) coordinates of each smartphone device every 5 minutes whenever the device is turned on. An important advantage of these data over other sources of smartphone data is that location information is collected regardless of what application the user has open, as long as the device is turned on. These "big data" provide an immense volume of high-frequency and spatially-disaggregated information on the geographical movements of users throughout each day. For example for the month of July 2019 alone, the data include 1.4 billion data points on 545,000 users (about 0.5 percent of the Japanese population).[3]

The raw unstructured geo-coordinates are pre-processed by the cell phone operator: NTT Docomo Inc. to construct measures of "stays," which correspond to distinct geographical locations visited by a user during a day. In particular, a stay corresponds to the set of geo-coordinates of a given user that are contiguous in time, whose first and last data points are more than 15 minutes apart, and whose geo-coordinates are all within 100 meters from the

[3]The mapping application does not send location data points if the smartphone does not sense movement, in which case it is likely that the user has not moved from the last reported location. For this reason, the data points are less frequent than 5 minutes intervals in practice.

centroid of these points.[4] We have access to the data on the sequence of stays of anonymized users with the necessary level of spatial aggregation to deidentify individuals. Our data comprise a randomly selected sample of 80 percent of users in Japan, where the randomization is again to deidentify individuals.

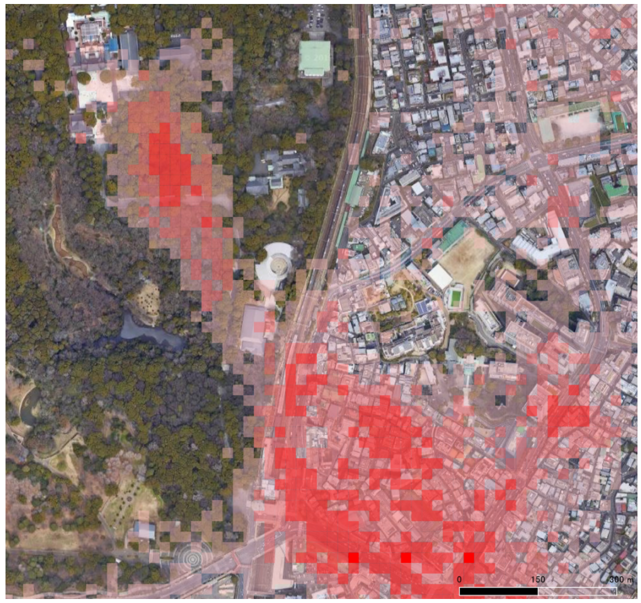
This pre-processing also categorizes all stays in each month into three categories of home, work and other locations for each anonymized user. “Home" location and “work" locations are defined as the centroid of the first and second most frequent locations of geographically contiguous stays, respectively. To ensure that these two locations do not correspond to different parts of a single property, we also require that the “work" location is more than 600 meters away from “home" location. In particular, if the second most frequent location is within 600 meters of the “home" locations, we define the “work” location as the third most frequent location. To abstract from noise in geo-coordinate assignment, all stays within 500 meters of the home location are aggregated with the home location. Similarly, all stays within 500 meters of the work location are aggregated with the work location. We assign “Work” location as missing if the user appears in that location less less than 5 days per month, which applies for about 30 percent of users. These users likely include irregular workers with unstable job locations, those who work at home, and those with limited number of data observations due to infrequent smartphone use.[5] Stays which are neither assigned as home or work are classified as “other.” We distinguish between different types of these “other” stays, such as visits to restaurants and stores, using spatially-disaggregated data on economic activity by sector and location from the economic census, as discussed further in Section 2.2 below.

As an illustration of our data, Figure 1 displays the “stays” recorded in our data for the example of a *Meiji Shrine* in the Shibuya municipality of Tokyo over the period from December 2017 to February 2018. Each red-shaded rectangle corresponds a 20 meter by 20 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 of the neighborhood. 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 an extremely fine level of spatial resolution. Second, we find a sharp discountinuity 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.

[4]See Patent Number “JP 2013-89173 A" and “JP 2013-210969 A 2013.10.10" for the detailed proprietary algorithm. This algorithm involves processes to offset the potential noise in measuring GPS coordinates.

[5]In Subsection 2.3 below, we report validation checks on our classification of home and work locations using commuting data from the population census. As an additional check on our classification of home and work locations, we also show in Figure 6 that users are typically at their home location during the evening, night and early morning of weekdays, and are typically at their work location during the daytime of weekdays.

Figure 1: Example of Stays Around a Meiji Shrine in the Shibuya Municipality of Tokyo



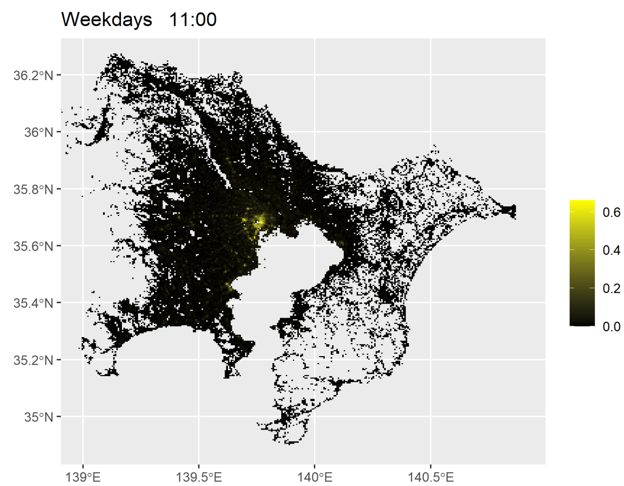
Note: The map shows the geographic location of “stays" around a Meiji Shrine. Each red-shaded rectangle corresponds a $20 \times 20$ meter grid cell. The darkness of the color represents the number of stays in each grid cell between December 2017 and February 2018. The building towards the top-left surrounded 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.

For most of our subsequent analysis, we focus on the sample of users in the month of April 2019 who have home and work locations in the Tokyo Metropolitan Area (which includes the four prefectures of Tokyo, Chiba, Kanagawa, and Saitama). To abstract from overnight trips, we also focus on the sample of user-day observations for which the first and last stay of the day is the user’s home location. In Figure 2, we show the spatial pattern of “stays" in the entire Tokyo Metropolitan Area for this sample. For each user and on each hour of the clock for each day (e.g. at 11am on Monday), we first assign a user’s location based on their most recent stay. Using this assignment, we next compute the density function of users, as the share of users in each location. Finally, we take averages by hour across days, separately for weekdays and weekends.

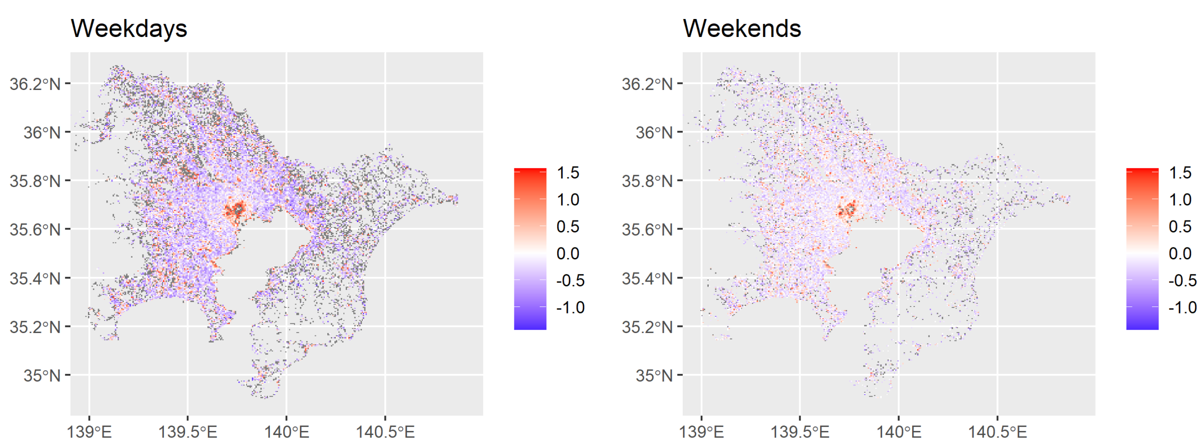
Panel (A) plots the geographic density of smartphone users at 11am on weekdays, where we use brighter yellow colors to indicate higher densities of users. Consistent with an approximately monocentric structure of economic activity, we find that the density has a clear peak at the city center of the Tokyo metropolitan area. Panel (B) plots the log difference of the density of users at 11am and 11pm. On both weekdays and weekends, the center of the city gains population and the suburbs lose population during the day-time relative to the night-time. But these differences between day- and night-time populations are larger on weekdays than weekends, consistent with people staying closer to their residential locations during the weeekend. Both panels confirm that our assignment of home and work locations captures an intuitive spatial distribution of users within the metropolitan area.

Figure 2: Stays in the Tokyo Metropolitan Area

(A) Day-time Population




(B) Log Difference of Day- and Night-time Population




Notes: Panel (A): Density of smartphone users at 11am on weekdays by $500 \times 500$ meter grid cell for our baseline sample for the Tokyo metropolitan area in April 2019. Panel (B): Difference in the density of smartphone users between 11am and 11pm by $500 \times 500$ meter grid cell. Density equals number of users in a $500 \times 500$ meter grid cell divided by the total number of users in our baseline sample.

## 2.2 Other Data Sources

We combine our smartphone data with a number of complementary spatially-disaggregated data sources. We use these additional data sources for the validation of our smartphone data in the next subsection as well as for the calibration and estimation of the model in later sections of the paper.

**Spatial units.** Data are available for the Tokyo metropolitan area at three main levels of spatial aggregation. At the highest level, the metropolitan area includes the four prefectures of Tokyo, Chiba, Kanagawa and Saitama. Each prefecture can be disaggregated into municipalities, of which there are 242 in the Tokyo metropolitan area as a whole (excluding island municipalities). Each municipality can be further disaggregated into *Oaza*, of which there are 9,956 in the Tokyo metropolitan area. Each Oaza has an area of around 1.30 squared kilometers and an average 2011 population of around 3,600.

**Population Census.** We measure residential population, employment by workplace and bilateral commuting flows using the population census, which is conducted by the Statistics Bureau, Ministry of Internal Affairs and Communications every five years. We use the publicly-available data from the 2015 population census that are reported on their website. We use these population census data in validation checks on our smartphone data and in our quantitative analysis of the model. Residential population and total employment are available at the finest level of spatial disaggregation of 250-meter grid cells. Demographic information used in some of our specification is available at a slightly more aggregated level. Bilateral commuting flows are reported between pairs of municipalities.

**Economic Census.** We use the Economic Census to distinguish between different types of non-commuting trips, as well as to capture industry-specific employment for the quantification of the model. Economic Censuses are conducted every 2 to 3 years by the Statistics Bureau, Ministry of Internal Affairs and Communications, and the Ministry of Economy, Trade and Industry. We use the publicly-available data from the 2016 Economic Census 2016 on total employment and the number of establishments by one-digit industry for each 500-meter grid cell in the Tokyo metropolitan area. We use also data on total revenue and factor inputs that are available at the municipality level.

**Building Data.** We measure floor space in each city block using the Zmap-TOWN II Digital Building Map Data for 2008, which is accessible through the Center for Spatial Information Science at the University of Tokyo. This data set contains polygons for all buildings in Japan, with their precise geo-coordinates and information on building use and characteristics. We measure floor space using the number of stories and land area for each building.

**Land Price Data.** We measure the residential land price for each city block using the evaluated land price that is used for the calculation of property tax. Local governments, typically at the municipality level, calculate these land prices for each road segment throughout the city. We take a simple average of these values to construct the average land prices per unit of land at the Oaza level whenever there is an observation. We obtain the consolidated data on these land values from the Research Center for Property Assessment System.

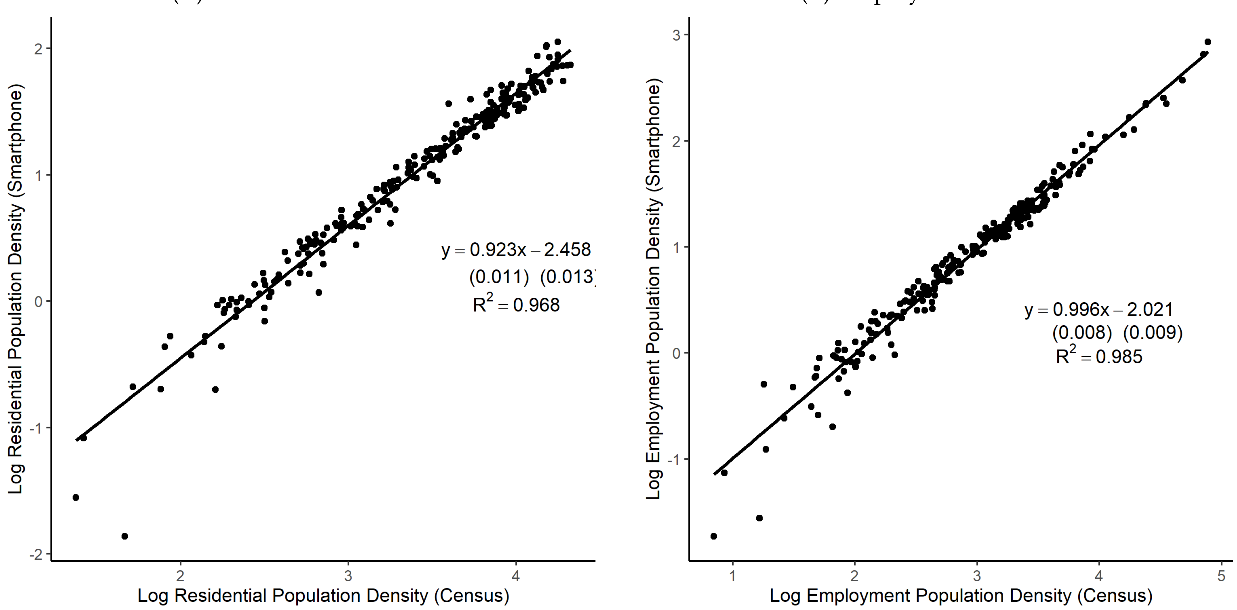
**Municipality Income Tax Base Data.** We measure the average income of the residents in each municipality using official data on the tax base for that municipality.

## 2.3 Validation of Smartphone Commuting Data Using Census Commuting Data

We now report an external validation exercise, in which we compare our measures of "home" location, "work location" and "commuting trips" from the smartphone data to the corresponding measures available from official census data. As the most disaggregated spatial units for which these official census data are available are municipalities, we aggregate our smartphone data to the municipality level in order to undertake this comparison. In the left panel of Figure 3, we display the log number of residents in each municipality in our smartphone data against log population in the census data. As our smartphone data cover only a fraction of the total population, the levels of the two variables necessarily differ from one another. Nevertheless, we find a tight and approximately log linear relationship between them, with a slope coefficient of 0.923 (standard error 0.011) and a R-squared of 0.968. The coefficient is slightly less than one, indicating that the smartphone data has higher coverage in less dense areas. In the right panel of Figure 3, we show the log number of workers in each Tokyo municpality in our smartphone data against log employment by workplace in the census data. Again, the levels of the two variables necessarily differ from one another, but we find a close and approximately log linear relationship between them, with a slope coefficient of 0.996 (standard error 0.008) and a R-squared of 0.985. Taken together, these findings provide strong evidence in support of our measures of home and work location from the smartphone data. Furthermore, the fact that these relationships are so tight across municipalities with very different levels of economic activity suggests that the probability of inclusion in our sample is not strongly correlated with the population or employment of the municipality.[6]

Figure 3: Representativeness of Smartphone Users

(A) Residential Location (B) Employment Location



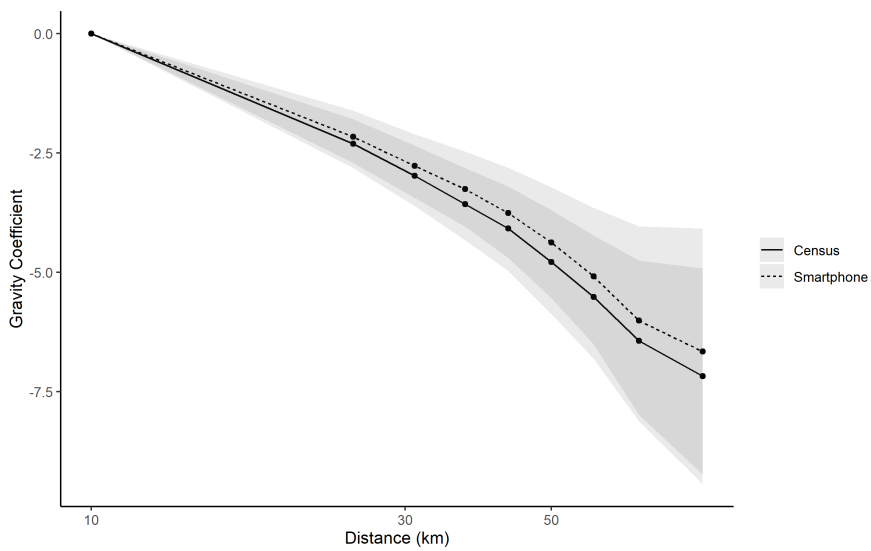

Note: Each dot represents a municipality in the Tokyo metropolitan area. In the left panel, the vertical axis is the log of the number of smartphone users with a home location in the municipality, and the horizontal axis is the log of the number of residents in that municipality from the Population Census in 2011. In the right panel, the vertical axis is the log of the number of smartphone users with a work location in the municipality, and the horizontal axis is the log of employment by workplace in that municipality from the Population Census in 2011. The definitions of home and work location in the smartphone data are discussed in the text of Subsection 2.1 above.

As a further check on the ability of our smartphone data to successfully capture commuting patterns, we now show that the same pattern of spatial decay of bilateral commuting flows with geographical distance is observed in the smartphone data as in the official census data. In each case, we regress the log of bilateral commuting flows between Tokyo municipalities on residence fixed effects, workplace fixed effects and a set of indicator variables for deciles of log

[6]In Appendix Figure A.1 and A.2, we provide further evidence on he representativeness of our smartphone data by comparing the coverage by the characteristics of residential locations (income, age, distance to city center) and employment density (industry characteristics and distance to city center).

bilateral distance. In Figure 4, we display the estimated coefficients on the indicator variables for both the smartphone and official census data and the 95 percent confidence intervals. As sample size is smaller in our smartphone data than in the official census data, we find marginally largely confidence intervals using the smartphone data, particularly for bilateral distances of more than 50 kilometers for which there are relatively few commuters. Nonetheless, for distances of less than 50 kilometers, which account for the vast majority of all commuters in both datasets, we find that the estimates in the smartphone and census data are lie extremely close to one another.

Figure 4: Gravity Equation Estimates for Bilateral Commuting Flows Using Smartphone GPS and Official Census Data




Note: Gravity equation estimation including workplace fixed effects, residence fixed effects and indicator variables for deciles of bilateral distance between workplace and residence; solid line and dark gray shading shows point estimates and 95 percent confidence intervals respectively for the distance decile indicators using the official census data; dashed line and light gray shading shows point estimates and 95 percent confidence intervals respectively for the distance decile indicators using our smartphone GPS data.

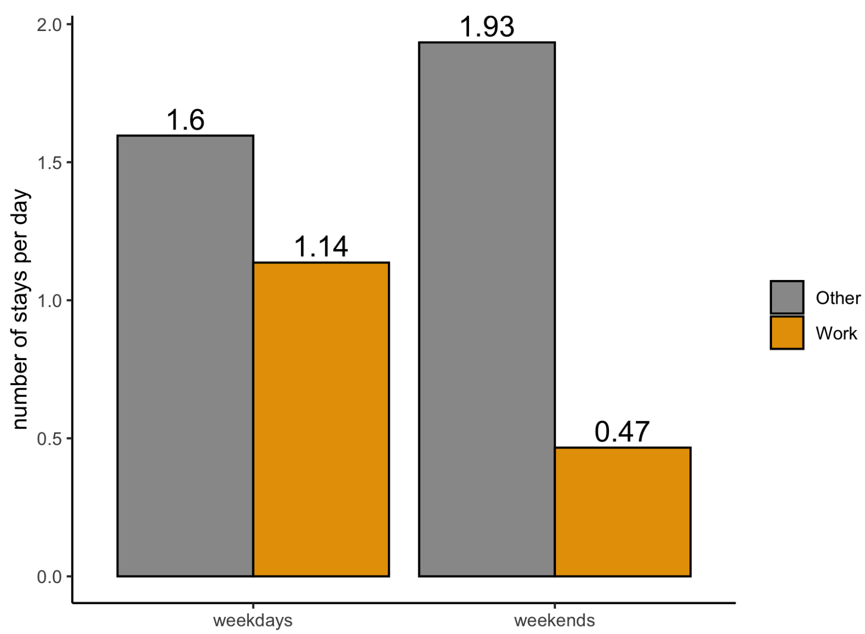
Taken together, the results of this section suggest that our smartphone data is relatively successful in identifying home locations, workplace locations and bilateral commuting patterns compared to official census data. However, a key advantage of our smartphone data relative to the official census data is that we can measure not only commuting trips but also the many non-commuting trips that individuals undertake and that are potentially consequential for our understanding of the spatial distribution of economic activity within cities.

# 3 Reduced-Form Evidence

In this section, we provide reduced-form evidence on patterns of commuting and non-commuting trips trips that guides the theoretical model that we develop below. First, we show that non-commuting trips are more frequent than commuting trips, so that concentrating only on commuting trips underestimates the amount of travel within urban areas. Second, we demonstrate that non-commuting trips are closely-related to the availability of non-tradeable services, which is consistent with them playing an important role in determining consumption access. Third, we show that non-commuting trips exhibit different spatial patterns from commuting trips, so that abstracting from non-commuting trips yields a misleading picture of bilateral patterns of travel within urban areas.

**Fact 1. Non-commuting trips are pervasive.** In Figure 5, we display the average number of stays per day for work and non-work locations (excluding home locations) for our baseline sample of users with home and work locations in the Tokyo Metropolitan Area during April 2019. Note that 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 attending a lunch meeting outside their workplace). Similarly, the average number of work stays can be greater than zero at the weekend, because some workers may be employed during the weekend (e.g. in restaurants and stores). As apparent from the figure, even during weekdays, we find that non-commuting trips are more frequent than commuting trips, with an average of 1.6 non-work stays per day compared to 1.14 work stays per day. This pattern is magnified at weekends, with an average of 1.93 non-work stays per day compared to 0.47 work stays per day. These results are consistent with evidence from travel surveys, in which commuting is only one of many reasons for travel, as found for example in Couture, Duranton, and Turner (2018). A key advantage of our smartphone data over travel surveys is that they reveal bilateral patterns of travel at a fine level of spatial disaggregation within the urban area, which plays a central role in the quantitative analysis of our model below.

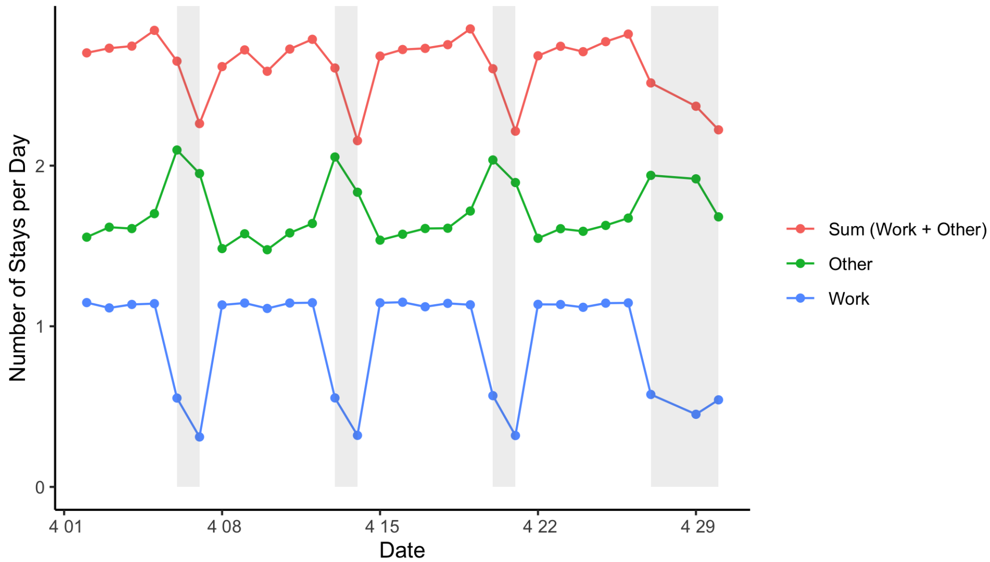
Figure 5: Frequency of Stays at Work and Other Locations (Excluding Home Locations)



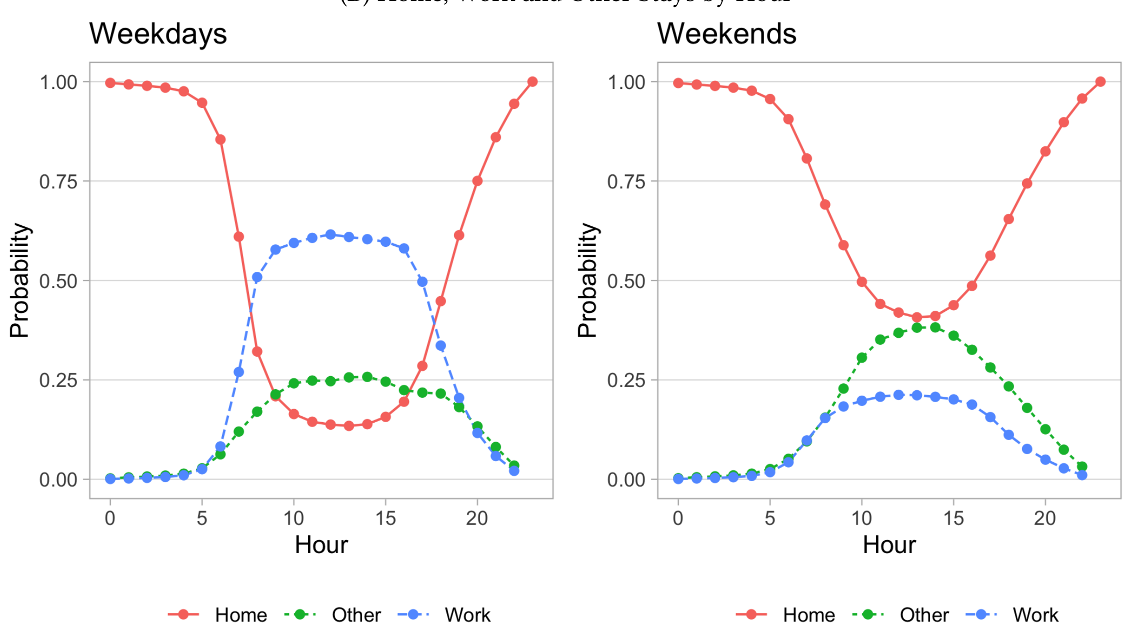

Note: Average number of work and other stays per day for weekdays and weekends (excluding stays at home locations) for our baseline sample users in the metropolitan area of Tokyo in April 2019. See Section 2 above for the definitions of home, work and other stays.

Figure 6: Work and Other Stays by Day and Hour

(A) Work and Other Stays by Day (Excluding Home Locations)




(B) Home, Work and Other Stays by Hour




Note: Panel (A): Average number of work and other stays per day (excluding stays at home locations) for our baseline sample of users in the Tokyo metropolitan area in April 2019. Gray shared areas indicate weekends and holidays in Japan. Panel (B): shows the probability that each user stays at home, work or other locations by each hour of the day, where these three probabilities sum to one. To construct Panel (B), 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 2 above for the definitions of home, work and other stays.

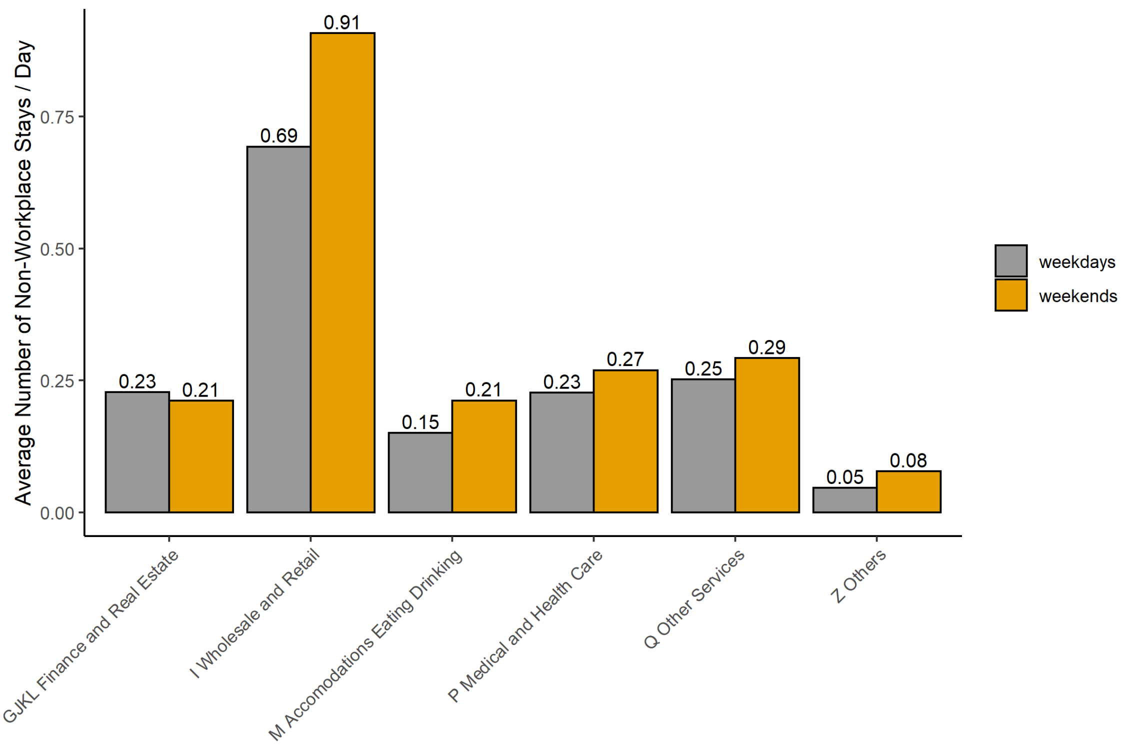
In Panel (A) of Figure 6, we provide further evidence on travel patterns by reporting the average number of work and non-work stays by day of the week from 1-29 April 2019. Consistent with the patterns discussed above, we find that non-commuting trips are more frequent than commuting trips for each day of the week, with non-commuting and commuting trips increasing and decreasing respectively at weekends. In Panel (B) of Figure 6, we show the average probability that a user stays at home, work or other locations by hour. A key difference from Panel (A) is that stays in Panel (B) are implicitly weighted by the length of time that a user spends at each stay. This weighting explains why work stays have a higher probability than other stays in the middle of day during weekdays in Panel (B), even though there is a larger average number of other stays than of work stays in Panel (A). The three probabilities in Panel (B) sum to one, since home, work, and other stays are mutually exclusive and sum to the total number of stays. Even after weighting by time, other stays are quantitatively relevant compared to work stays during both weekdays and weekends. Comparing across hours of the day, we find the expected pattern that home stays fall and both work and other stays rise during the daytime (from around 6am-9pm). During weekdays, the probability of a stay rises more rapidly during the waking hours for work stays than for other stays. During weekends, we find the opposite pattern, with the probability of stay rising more rapidly during the waking hours for other stays than for work stays.

**Fact 2. Non-commuting trips are closely related to consumption.** We now show that non-commuting trips are closely related to consumption by combining our GPS smartphone data with spatially-disaggregated census data on employment by sector. In particular, we stochastically assign other 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 $500 \times 500$ 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 and Real Estate", “Wholesale and Retail", “Accommodations, Eating, Drinking", “Medical and Health Care", and “Other Services". For each other 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 other stay to the category "Z Others."

As a check on this probabilistic assignment of other stays, Figure A.3 in the online appendix, displays the density of each type of other stay by hour and day, as a share of all stays for our baseline sample for the Tokyo metropolitan area in April 2019. We find that our probabilistic assignment captures the expected pattern of these different service-sector activities over the course of the week. First, we typically find a higher density of other stays during the middle of the day at weekends than during weekdays, which is in line with the fact that many of these services are consumed more intensively during leisure time. The one exception is “Finance and Real Estate," which displays the opposite patter, consistent with the fact that establishments providing these services are often closed at the weekends in Japan. Second, we find that the peak densities of stays for “Wholesale and Retail” and “Accommodations, Eating, Drinking” occur at around 6pm on weekdays, corroborating the fact that these activities are typically concentrated after work during the week. For “Accommodations, Eating, Drinking,” we find a smaller peak around noon on weekdays, as expected from the typical timing of lunch in Japan. Third, and finally, both of these activities are more concentrated in the middle of the day on weekends than during the week, which again is in line with workers having greater leisure time in the middle of day at weekends.

Figure 7: Frequency of Non-Commuting Trips and Employment of Consumption Sector

(A) Number of Non-Work Stays in Each Sector




(B) Number of Non-Work Stays and Employment in Each Sector

| **Industry** | **Weekdays** | | **Weekends** | | **Employment Share in Service (%)** | |
|---|---|---|---|---|---|---|
| | Stays / Day | Share (%) | Stays / Day | Share (%) | Total | Average (500m Grids) |
| GJKL Finance and Real Estate | 0.23 | 14.3 | 0.21 | 10.7 | 11.9 | 23.2 |
| I Wholesale and Retail | 0.69 | 43.4 | 0.91 | 46.1 | 32.0 | 28.7 |
| M Accomodations Eating Drinking | 0.15 | 9.4 | 0.21 | 10.8 | 13.2 | 13.2 |
| P Medical and Health Care | 0.23 | 14.2 | 0.27 | 13.7 | 18.7 | 15.2 |
| Q Other Services | 0.25 | 15.8 | 0.29 | 14.8 | 24.3 | 19.8 |
| Z Others | 0.05 | 2.9 | 0.08 | 3.9 | | |

Note: Panel (A): Average number of each type of other stay per day for weekdays and weekends (excluding stays at home locations) for our baseline sample users in the metropolitan area of Tokyo in April 2019. Other stays are allocated probabilistically to each of these five categories using their shares 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 in the total number of other 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 500 $\times$ 500 meter grid cells. See Section 2 above for the definitions of home, work and other stays.

In Panel (A) of Figure 7, we show the average number of these different types of other stays per day during the working week and at weekends. We find that "Wholesale and Retail" stays are by far the most frequent, with an average of 0.69 per day on weekdays and 0.91 per day on weekends. To provide a point of comparison, Panel (B) of Figure 7 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 $\times$ 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 43.4 percent of weekday stays and 46.1 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 other stays are disproportionately targeted towards locations with relatively high shares of the "Wholesale and Retail" sector in employment, which is consistent with these other stays capturing access to consumption opportunities. 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, Eaton and Drinking," "Finance and Real Estate," and "Medical and Health Care" all accounting for around 10 percent of more of the total number of stays.

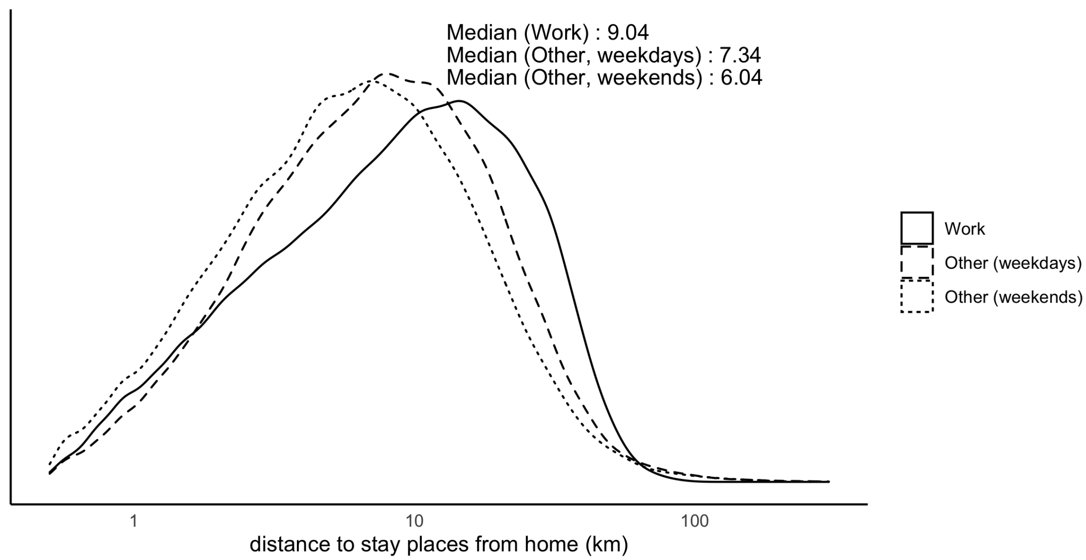
**Fact 3. Non-commuting trips are closer to home.** We now show that non-commuting trips exhibit different spatial patterns from commuting trips, such that observed bilateral commuting flows provide an incomplete picture of patterns of travel within urban areas. In Panel (A) of Figure 8, we display the distribution of distances travelled to work and other stays for our baseline sample of users in the Tokyo metropolitan area in the month of April 2019. We find that other stays are concentrated closer to home than work stays, with average distances travelled of 7.34 and 9.04 kilometers respectively during weekdays. This difference is even greater at the weekend, with an average distance travelled of 6.04 kilometers for other stays, which is consistent with users remaining closer to their residential locations at weekends. Given that users choose where to live taking into account their access to surrounding locations, this clustering of other stays closer to home highlights the relevance of these non-commuting trips for residential location decisions, as explored further in the quantitative analysis of our model below.

In Panel (B) of Figure 8, we display the distribution of distances travelled for each type of other stay separately. Comparing across the different categories, we find that "Wholesale and Retail" and "Accommodations, Eating, Drinking" stays are concentrated closer to home than "Finance and Real Estate" and "Other Services stays." This clustering of these two categories of shopping and visits to bars, restaurants and cafes close to home again highlights the relevance of access to these consumption opportunities for users' residential location decisions. More generally, these differences in bilateral travel patterns for these different types of economic activities suggest that omitting non-consumption trips not only undercounts travel journeys but can also yield misleading inference about the effects of changes in travel costs on bilateral patterns of travel, as explored more formally in later sections.

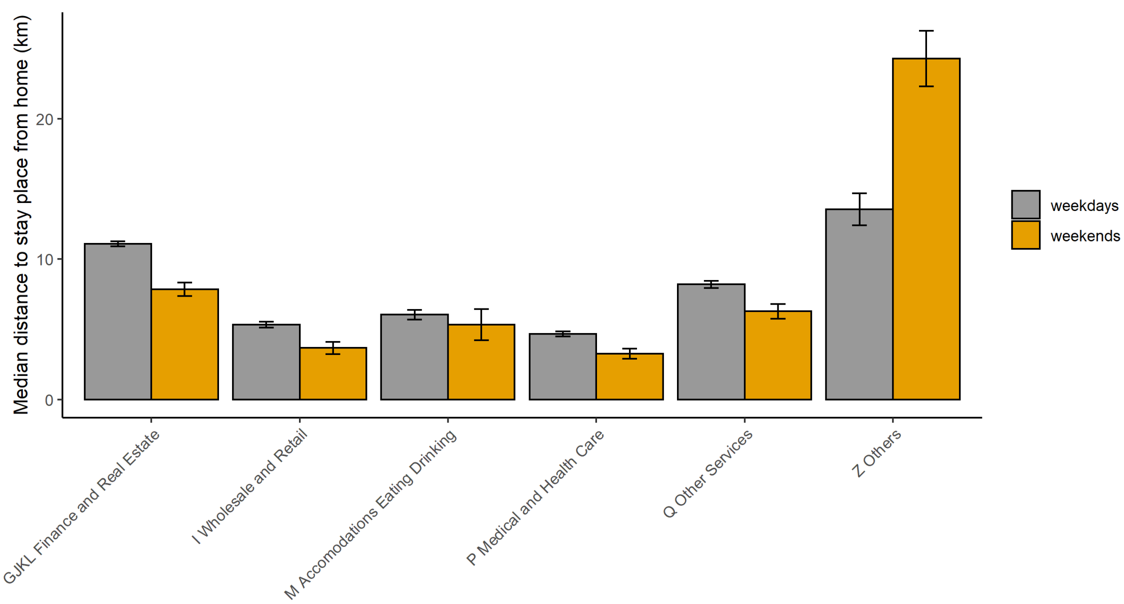
Taking the findings of this section as a whole, we have shown that non-commuting trips are frequent, are closely related to consumption, and exhibit different spatial patterns from commuting trips. Each of these three features of our smartphone data guides our theoretical modelling of commuting and non-commuting trips in the next section.

Figure 8: Distances of Commuting and Non-Commuting Trips

(A) Distribution of Distances of Work and Other Stays from Home Locations




(B) Average Distances of Different Types of Other Stays from Home Locations




Note: Panel (A): Distributions of distance in kilometers of work locations from home location and of other stays from home locations during weekdays and weekends. Panel (B): Distributions of distance in kilometers of each type of other stay from home locations during weekdays and weekends. Distributions computed for our baseline sample of users in the Tokyo metropolitan area in April 2019.

# 4 Theoretical Framework

In this section, we develop our quantitative urban model of internal city structure that incorporates both commuting and non-commuting trips within the city. We show that the model rationalizes the reduced-form features of the smartphone data established above. In Section 5 below, we use the model to quantify the contributions of consumption access, workplace and unobserved amenities in explaining the observed spatial concentration of economic activity. In Section 6, we show how the model can be used to undertake counterfactuals for the impact of transport improvements using only the observed values of the endogenous variables of the model in an initial equilibrium. We demonstrate that the omission of consumption access leads to an underestimation of the welfare gains from these transport improvements and distorted predictions for their impact on bilateral patterns of travel throughout the city. The derivations for all theoretical results in this section are reported in Section B of the online appendix.

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

A continuous measure of workers ($\bar{L}$) choose a residence, a workplace and a set of locations to consume non-traded services in the city.[7] We assume the following timing or nesting structure for workers' location decisions. First, each worker observes her idiosyncratic preferences or amenities ($b$) for each location within the city, and chooses her residence $n$. Second, given a choice of residence, each worker observes her idiosyncratic productivities ($a$) for each workplace $i$ and sector $g$, and idiosyncratic qualities ($d$) for each type of non-traded service $k$ available in each location $j$. After observing these idiosyncratic productivity and quality draws, she chooses her sector and location of employment and her consumption location for each type of non-traded service. We choose this nesting structure because it permits a particularly transparent decomposition of residents and land prices into the contribution of workplace and consumption access.[8] In the open-city specification, population mobility ensures that the expected utility from living in the city equals the reservation utility in the wider economy.

[7] We assume for simplicity a continuous measure of workers in the model, which ensures that the expected values of variables equals their realized values. Nonetheless, it is straightforward to consider the case of a finite number of workers, as in Dingel and Tintelnot (2020), and we allow for this possibility when we estimate the model's commuting and consumption gravity equations.

[8] In the online appendix, we consider alternative potential nesting structures, such as allowing workers to simultaneously choose residence, workplace and consumption locations, and allowing for chains of trips between multiple locations. Although these extensions yield less transparent decompositions, both workplace and consumption access remain important for the spatial concentration of economic activity.

## 4.1 Preferences

The indirect utility for worker $\omega$ who chooses residence $n$, works in location $i$ and sector $g$, and consumes non-traded service $k$ in location $j(k)$ is assumed to take the following Cobb-Douglas form:

$$\begin{aligned} U_{nig\{j(k)\}}(\omega) = & \left\{B_n b_n(\omega)(P_{n,T})^{-\alpha_T} Q_n^{-\alpha_H}\right\}\left\{a_{i,g}(\omega) w_{i,g} e^{-\kappa\tau_{ni}}\right\}, \\ & \times \left\{\prod_{k\in K_S}\left[P_{j(k),k}/\left(d_{j(k),k}(\omega)\exp\left(-\kappa\tau_{nj(k)}\right)\right)\right]^{-\alpha_k}\right\} \\ & 0<\alpha_T,\alpha_H,\alpha_k<1, \qquad \alpha_T+\alpha_H+\sum_{k\in K_S}\alpha_k = 1, \end{aligned} \tag{1}$$

where we use the notation $j(k)$ to indicate that that non-traded service $k$ is consumed in a single location $j$ that is an implicit function of the type of non-traded service $k$; $K_S \subset K$ is the subset of sectors that are non-traded; the first term in brackets captures a residence component of utility; the second term in brackets corresponds to a workplace component of utility; and the third term in brackets reflects a non-traded services component of utility.

The first, residence component includes amenities ($B_n$) that are common for all workers in residence $n$; the idiosyncratic amenity draw for residence $n$ for worker $\omega$ ($b_n(\omega)$); the price of the traded good ($P_{n,T}$); and the price of residential floor space ($Q_n$). We allow the common amenities ($B_n$) to be either exogenous or endogenous to the surrounding concentration of economic activity in the presence of agglomeration forces, as discussed further below. The second, workplace component comprises the wage per efficiency unit in sector $g$ in workplace $i$ ($w_{i,g}$); the idiosyncratic draw for productivity or efficiency units of labor for worker $\omega$ in sector $g$ in workplace $i$ ($a_{i,g}(\omega)$); and travel costs between residence $n$ and workplace $i$ that are an exponential function of bilateral travel time $\tau_{ni}$ with parameter $\kappa$.[9] The third, non-traded services component depends on the price of the non-traded service $k$ in the location $j(k)$ where it is supplied ($P_{j(k),k}$ for $k \in K_S$); the idiosyncratic draw for quality for that service in that location ($d_{j(k),k}(\omega)$ for $k \in K_S$); and travel costs between residence $n$ and consumption location $j(k)$ that are again an exponential function of travel time ($\tau_{nj(k)}$) with parameter $\kappa$. Consistent with travel costs primarily being determined by the transport network rather than the reason for travel, we assume the same elasticity of travel costs with respect to travel time for all economic activities ($\kappa$). Motivated by our empirical finding above that consumption trips are concentrated closer to home that commuting trips, we also assume that workers travel to consume to non-traded service $k$ in location $j(k)$ from their residence $n$, although this assumption can be relaxed.[10]

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

$$\begin{aligned} G_n^B(b) &= \exp\left(-T_n^B b^{-\theta_B}\right), & T_n^B &> 0,\ \theta_B > 1, \\ G_{i,g}^W(a) &= \exp\left(-T_{i,g}^W a^{-\theta_W}\right), & T_{i,g}^W &> 0,\ \theta_W > 1, \\ G_{j(k),k}^S(d) &= \exp\left(-T_{j(k),k}^S d^{-\theta_{S,k}}\right), & T_{j(k)}^S &> 0,\ \theta_{S,k} > 1, k \in K_S \end{aligned} \tag{2}$$

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

[10] In the online appendix, we show that this assumption can be relaxed to allow workers to travel to consume non-traded service $k$ in location $j(k)$ from either their residence or workplace $i$, or to consume them along the route between residence $n$ and workplace $i$.

where the scale parameters $\{T_n^B,\, T_{i,g}^W,\, T_{j(k),k}^S\}$ control the average draws and the shape parameters $\{\theta_B,\, \theta_W,\, \theta_{S,k}\}$ regulate the dispersions of amenities, productivity and quality respectively. The smaller these dispersion parameters, the greater the heterogeneity in idiosyncratic draws, and the less responsive worker decisions to economic variables.[11]

Using our assumption about the timing or nesting structure, the worker location choice problem is recursive and can be solved backwards. First, given a choice of residence, we characterize the probability that a worker chooses each consumption and workplace location, and hence expected consumption and workplace access conditional on choosing that residence. Second, we characterize the probability that a worker chooses each residence, and hence expected utility from living in the city.

## 4.2 Consumption Choices

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

$$\gamma_{nj(k),k}\left(\omega\right) = \left[P_{j(k),k} / \left(d_{j(k),k}\left(\omega\right) \exp\left(-\kappa \tau_{nj(k)}\right)\right)\right]^{-\alpha_k}, \qquad k \in K_S. \tag{3}$$

Using our independent extreme value assumption for idiosyncratic quality, a key prediction of the model is that bilateral travel to consume non-traded services is characterized by a gravity equation. In particular, the probability that a worker in residence $n$ consumes non-traded service $k$ in location $j(k)$ ($\lambda_{nj(k),k|n}^S$) takes the following form:

$$\lambda_{nj(k),k|n}^S = \frac{T_{j(k),k}^S \left(P_{j(k),k}\right)^{-\theta_{S,k}} \exp\left(-\theta_{S,k}\kappa\tau_{nj(k)}\right)}{\sum_{\ell \in N} T_{\ell(k),k}^S \left(P_{\ell(k),k}\right)^{-\theta_{S,k}} \exp\left(-\theta_{S,k}\kappa\tau_{n\ell(k)}\right)}, \qquad k \in K_S, \tag{4}$$

which we term the conditional consumption probability, since it is computed conditional on living in residence $n$. This probability depends on destination characteristics (the price of non-traded services $P_{j(k),k}$ and their average quality $T_{j(k),k}^S$ in the numerator); the bilateral travel costs (as determined by travel times $\tau_{nj(k)}$ in the numerator); and origin characteristics (as captured by the travel-time weighted average of destination characteristics in the denominator). Therefore, as in gravity equation models in international trade, bilateral travel flows to consume non-traded services depend not only on bilateral resistance (travel costs between a pair of locations) but also on multilateral resistance (travel costs between all pairs of locations). We show below that this gravity equation prediction provides a good approximation to the observed data on bilateral travel to consume non-traded services

Using our independent extreme value assumption for idiosyncratic quality, we can also compute the expected contribution to utility from consuming non-traded service $k$ conditional on living in residence $n$. This expectation corresponds to a measure of residence $n$'s consumption access for non-traded service $k$, and depends on the travel-time weighed average of destination characteristics:

$$\mathbb{S}_{nk} \equiv \mathbb{E}_{nk}\left[\gamma_{nj(k),k}\right] = \vartheta_k^S \left[\sum_{\ell \in N} T_{\ell(k),k}^S \left(P_{\ell(k),k}\right)^{-\theta_{S,k}} \exp\left(-\theta_{S,k}\kappa\tau_{n\ell(k)}\right)\right]^{\frac{\alpha_k}{\theta_{S,k}}}, \qquad k \in K_S. \tag{5}$$

where $\vartheta_k^S \equiv \Gamma\left(\frac{(\theta_{S,k}/\alpha_k)-1}{(\theta_{S,k}/\alpha_k)}\right)$ and $\Gamma(\cdot)$ is the Gamma function.

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

Using our assumption that idiosyncratic quality is independently distributed across the non-traded sectors, we can also compute the expected overall contribution to utility from consuming all types of non-traded services. This expectation corresponds to an overall measure of consumption access for residence $n$ and depends on the travel-time weighted average of destination characteristics across all types of non-traded services:

$$\mathbb{S}_n \equiv \prod_{k \in K_S} \mathbb{S}_{nk} = \prod_{k \in K_S} \vartheta_k^S \left[ \sum_{\ell \in N} T_{\ell(k),k}^S \left( P_{\ell(k),k} \right)^{-\theta_{S,k}} \exp\left( -\theta_{S,k} \kappa \tau_{n\ell(k)} \right) \right]^{\frac{\alpha_k}{\theta_{S,k}}}. \tag{6}$$

## 4.3 Workplace Choice

We next turn to the worker's choice of workplace. Conditional on living in residence $n$, each worker chooses the workplace $i$ and sector $g$ that offers the highest income, taking into account the wage per efficiency unit ($w_{i,g}$), the idiosyncratic draw for productivity ($a_{i,g}(\omega)$) and commuting costs (as determined by travel times $\tau_{ni}$):

$$v_{ni,g}(\omega) = w_{i,g} a_{i,g}(\omega) \exp\left( -\kappa \tau_{ni} \right). \tag{7}$$

Using our independent extreme value assumption for idiosyncratic productivity, the model also implies a gravity equation for bilateral commuting, such that the probability that a worker in residence $n$ commutes to workplace $i$ in sector $g$ ($\lambda_{ni,g|n}^W$) is as follows:

$$\lambda_{ni,g|n}^W = \frac{T_{i,g}^W w_{i,g}^{\theta_W} \exp\left( -\theta_W \kappa \tau_{ni} \right)}{\sum_{\ell \in N} \sum_{m \in K} T_{\ell,m}^W w_{\ell,m}^{\theta_W} \exp\left( -\theta_W \kappa \tau_{n\ell} \right)}, \tag{8}$$

which we term the conditional commuting probability, since again it is computed conditional on living in residence $n$. Bilateral commuting flows also depend on destination characteristics (the wage $w_{i,g}$ and average efficiency units $T_{i,g}^W$ in the numerator); bilateral travel costs (as determined by travel times $\tau_{ni}$ in the numerator); and origin characteristics (as captured by the travel-time weighted average of destination characteristics across sectors in the denominator). Consistent with prior research on commuting, we show that the model's gravity equation predictions also provide a good approximation to the commuting patterns observed in our smartphone data.

Taking expectations over idiosyncratic productivity, we measure workplace access using expected income conditional on living in residence $n$, which depends on the travel-time weighted sum of workplace characteristics:

$$\mathbb{W}_n = \mathbb{E}_n \left[ v_{ni,g} \right] = \vartheta^W \left[ \sum_{\ell \in N} \sum_{m \in K} T_{\ell,m}^W w_{\ell,m}^{\theta_W} \exp\left( -\theta_W \kappa \tau_{n\ell} \right) \right]^{\frac{1}{\theta_W}}, \tag{9}$$

where $\vartheta^W \equiv \Gamma\left( \frac{\theta_W - 1}{\theta_W} \right)$ and $\Gamma(\cdot)$ is the Gamma function.

An implication of our extreme value assumption for idiosyncratic productivity is that expected income conditional on living in residence $n$ and being employed in sector $g$ and workplace $i$ in equation (9) is the same across all choices of sector $g$ and workplace $i$ and equal to the overall expected income conditional on living in residence $n$ (as denoted by workplace access $\mathbb{W}_n$). The intuition is as follows. On the one hand, a higher wage per efficiency unit in a sector $g$ and workplace $i$ raises expected income for that sector and workplace. On the other hand, this higher wage per efficiency unit attracts workers with lower realizations for productivity, which reduces expected income. With a Fréchet distribution for idiosyncratic productivity, these two effects exactly offset one another, such that expected income conditional on choosing each sector $g$ and workplace $i$ is the same across all choices of sector and workplace for a worker living in a given residence $n$.

### 4.4 Residence Choice

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

$$U_n(\omega) = B_n b_n(\omega) P_{n,T}^{-\alpha_T} Q_n^{-\alpha_H} \mathbb{W}_n \mathbb{S}_n,$$

Using our independent extreme value assumption for idiosyncratic amenities, the probability that each worker chooses residence $n$ ($\lambda_n^B$) depends on its attractiveness in terms of workplace access ($\mathbb{W}_n$), consumption access ($\mathbb{S}_n$) and residential characteristics ($B_n$, $P_{T,n}$ and $Q_n$) relative to the attractiveness of all other locations within the city:

$$\lambda_n^B = \frac{T_n^B B_n^{\theta_B} \mathbb{W}_n^{\theta_B} \mathbb{S}_n^{\theta_B} P_{T,n}^{-\alpha_T \theta_B} Q_n^{-\alpha_H \theta_B}}{\sum_{\ell \in N} T_\ell^B B_\ell^{\theta_B} \mathbb{W}_\ell^{\theta_B} \mathbb{S}_\ell^{\theta_B} P_{\ell,T}^{-\alpha_T \theta_B} Q_\ell^{-\alpha_H \theta_B}}. \tag{10}$$

Taking expectations over idiosyncratic amenities, expected utility from living in the city depends on the workplace access, consumption access and other residential characteristics of all locations within the city:

$$\mathbb{E}[u] = \vartheta^B \left[ \sum_{\ell \in N} T_\ell^B B_\ell^{\theta_B} \mathbb{W}_\ell^{\theta_B} \mathbb{S}_\ell^{\theta_B} P_{\ell,T}^{-\alpha_T \theta_B} Q_\ell^{-\alpha_H \theta_B} \right]^{\frac{1}{\theta_B}}. \tag{11}$$

where $\vartheta^B \equiv \Gamma\left(\frac{\theta_B - 1}{\theta_B}\right)$ and $\Gamma(\cdot)$ is the Gamma function.

An implication of our extreme value assumption for idiosyncratic amenities is that expected utility conditional on choosing a given residence is the same across all residences in the city and equal to expected utility for the city as a whole. The intuition is similar to that for workplace access above. On the one hand, a lower price of floor space or higher workplace or consumption access in a given residence raises utility for a given realization of idiosyncratic amenities ($b$), which raises the expected utility from living there. On the other hand, a lower price of floor space or higher workplace or consumption access in a given residence attracts workers with lower realizations for idiosyncratic amenities, which reduces the expected utility from living there. With a Fréchet distribution for idiosyncratic amenities, these two effects exactly offset one another, such that expected utility conditional on choosing a given residence is the same across all residences in the city. In the open-city specification, population mobility ensures that this common expected utility is equal to the reservation level of utility in the wider economy.

The residential choice probabilities (10) highlight that the attractiveness of a residential location depends not only on its own characteristics but also on its access to surrounding employment and consumption opportunities. In our quantitative analysis below, we use these residential choice probabilities to decompose the observed spatial variation in economic activity into the contributions of workplace access, consumption access and a composite measure of amenities, without taking a stand on production technology and market structure in the traded and non-traded sectors. As a result, this quantitative analysis holds in an entire class of quantitative urban models with different specifications for production technology and market structure. We show that workers' observed choices of residence, workplace and consumption can be used to reveal their relative valuations of these locations and compute measures of workplace and consumption access that hold in an entire class of quantitative urban models.

Having solved for the residential choice probabilities (10) and expected income conditional on living in each residence (9), we can recover the demand for residential floor space in each location, using the implication of Cobb-Douglas utility that expenditure on residential floor space is a constant share of income:

$$H_{n,U} = \frac{\alpha_H \mathbb{W}_n R_n}{Q_n}, \tag{12}$$

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

### 4.5 Production

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

$$P_{i,k} = \frac{1}{A_{i,k}} w_{i,k}^{\beta_k} Q_i^{1-\beta_k}, \qquad 0 < \beta_k < 1, \qquad k \in K, \tag{13}$$

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

$$H_{i,k} = \frac{1-\beta}{\beta} \left( \frac{P_{i,k} A_{i,k}}{Q_i} \right)^{\frac{1}{\beta}} \tilde{L}_{i,k}, \tag{14}$$

where we use the tilde to denote the adjustment of labor input for effective units of labor.

We allow productivity in equations (13) and (14) to be either exogenous or endogenous to the surrounding concentration of economic activity in the presence of agglomeration forces, as discussed further below. We assume no-arbitrage between residential and commercial floor space, and across the different sectors in which commercial floor space is used, such that there is a single price for floor space within each location ($Q_i$) in equation (13). In general, the wage per efficiency unit ($w_{i,k}$) differs across both sectors and locations in equation (13), because workers draw efficiency units for each combination of sector and location pair, and hence each sector and location pair faces an upward-sloping supply function for effective units of labor. Finally, we assume that the traded good is costlessly traded within the city and wider economy and choose it as our numeraire such that:

$$P_{i,T} = 1 \qquad \forall\, i \in N. \tag{15}$$

### 4.6 Market Clearing

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

$$P_{j,k} A_{j,k} \left( \frac{\tilde{L}_{j,k}}{\beta} \right)^{\beta_k} \left( \frac{H_{j,k}}{1-\beta} \right)^{1-\beta_k} = \alpha_k \sum_{n \in N} \lambda_{nj,k|n}^S R_n \mathbb{W}_n, \qquad j \in K_S, \tag{16}$$

where expenditure on the right-hand side equals the sum across locations of workers travelling to consume non-traded service $k$ in location $j$; $R_n$ is the measure of residents in location $n$; recall that $\lambda^S_{njk|n}$ is the conditional consumption probability and $\mathbb{W}_n$ is expected income in residence $n$; and we have suppressed the implicit dependence of location $j(k)$ on the type of non-traded service $k$ in the conditional consumption probability to streamline notation.

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

$$L_{j,k} = \sum_{n \in N} \lambda^W_{nj,k|n} R_n, \tag{17}$$

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

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

$$H_i = H_{i,U} + \sum_{k \in K} H_{i,k}. \tag{18}$$

### 4.7 General Equilibrium with Exogenous Location Characteristics

We begin by considering the case in which productivity ($A_{i,k}$), amenities ($B_i$) and the supply of floor space ($H_i$) are exogenously determined. The general equilibrium of the model is referenced by the price for floor space in each location ($Q_i$), the wage in each sector and location ($w_{i,k}$), the price of the non-traded good in each service sector and location ($P_{i,k}$) the conditional consumption probabilities ($\lambda^S_{nj,k|n}$), the conditional commuting probabilities ($\lambda^W_{nj,k|n}$), the residence probabilities ($\lambda^B_n$), and the total measure of workers living in the city ($\bar{L}$), where we focus on the open-city specification, in which the total measure of workers is endogenously determined by population mobility with the wider economy. Given these seven equilibrium variables, we can solve for all other endogenous variables of the models. These seven equilibrium variables are determined by the system of seven equations given by the land market clearing condition for each location (18), the labor market clearing condition for each location (17), the non-traded goods market clearing condition for each location and service sector (16), the conditional consumption probabilities (4), the conditional commuting probabilities (8), the residence probabilities (10), and the population mobility condition that equates expected utility in the city (11) to the reservation level of utility in the wider economy ($\bar{U}$).

### 4.8 General Equilibrium with Agglomeration Forces and Endogenous Floor Space

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

**Agglomeration in Production.** In both the traded and non-traded sector, we allow productivity ($A_{i,k}$) to depend on production fundamentals and production externalities. Production fundamentals ($a_{i,k}$) capture features of physical geography that make a location more or less productive independently of neighboring economic activity (e.g. access to natural water). Production externalities capture productivity benefits from the density of employment across all sectors ($L_i/K_i$), where employment density is measured per unit of geographical land area.[12]

$$A_{i,k} = a_{i,k} \left( \frac{L_i}{K_i} \right)^{\eta_W} \tag{19}$$

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

**Agglomeration in Residents.** Similarly, we allow residential amenities ($B_n$) to depend on residential fundamentals and residential externalities. Residential fundamentals ($b_n$) capture features of physical geography that make a location a more or less attractive place to live independently of neighboring economic activity (e.g. green areas). Residential externalities capture the effects of the surrounding density of residents ($L_i/K_i$) and are modeled symmetrically to production externalities:[13]

$$B_n = b_n \left( \frac{R_n}{K_n} \right)^{\eta_B} \tag{20}$$

where $\eta^B$ parameters the strength of residential externalities.

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

$$H_i = M_i^{\mu} K_i^{1-\mu}, \qquad 0 < \mu < 1. \tag{21}$$

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

$$\mu Q_i H_i = \mathbb{P} M_i, \tag{22}$$

where $\mathbb{P}$ is the common user cost of building capital. Using the construction technology (21) to substitute for building captial ($M_i$) in equation (22) linking payments for floor space and building capital, we obtain a constant elasticity supply function for floor space as in Saiz (2010), with the inverse supply function given by:

$$Q_i = \psi_i H_i^{\frac{1-\mu}{\mu}} \tag{23}$$

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

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

$$Q_i = \left( \frac{\mathbb{P}}{\mu} \right)^{\mu} \left( \frac{\tilde{Q}_i}{1-\mu} \right)^{1-\mu} \tag{24}$$

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

Given this specification of agglomeration forces and endogenous floor space, the determination of general equilibrium remains the same as above with exogenous location characteristics above, except that productivity ($A_n$), amenities ($B_n$) and the supply of floor space ($H_n$) are now endogenously determined by equations (19), (20) and (23).

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

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

# 5 Quantitative Analysis

In this section, we use our theoretical model to quantify the contributions of workplace access and consumption access to the observed uneven spatial distribution of economic activity. In particular, given values for the model's parameters, the observed consumption and commuting probabilities in our smartphone data can be used to reveal the relative valuation placed by users on different locations as consumption and workplace locations. In Section 5.1, we show that the model's gravity equation for commuting can be used to estimate a theoretically-consistent measure of workplace access. In Section 5.2, we show that the model's gravity equation for consumption trips can be used to estimated an analogous theoretically-consistent measure of consumption access. In Section 5.3, we discuss the estimation and calibration of the model's parameters. Finally, in Section 5.4, we use the model's residential choice probabilities to quantify the contributions of workplace and consumption access relative to unobserved amenities in explaining the observed spatial concentration of economic activity.

## 5.1 Estimating Workplace Access

We start by using the observed bilateral commuting flows between workplace and residence to estimate workplace access. From the conditional commuting probabilities (8), we obtain the following log linear gravity equation for bilateral commuting between residence $n$ and workplace $i$:

$$\lambda^W_{ni,g|n} = \frac{\xi^W_i \exp\left(-\phi_W \tau_{ni}\right) \exp\left(u^W_{ni}\right)}{\zeta^W_n}, \tag{25}$$

where $u^W_{ni}$ is a stochastic error that captures idiosyncratic determinants of bilateral commuting costs not captured in bilateral travel times ($\tau_{ni}$), such as how pleasant the commute is.

We estimate this commuting gravity equation (25) using the Poisson Pseudo Maximum Likelihood (PPML) estimator of Santos Silva and Tenreyro (2006). This estimator yields theoretically-consistent estimates of the fixed effects (as shown in Thibault 2015) and allows for zeros in bilateral commuting flows (as discussed further in Head and Mayer 2014).[14] The estimated elasticity of commuting flows with respect to travel times ($\phi_W$) in equation (25) is a composite of the elasticity of commuting flows with respect to commuting costs ($\theta_W$) and the elasticity of commuting costs with respect to travel times ($\kappa$) in equation (8), such that $\phi_W = \theta_W \kappa$. Additionally, the estimated workplace fixed effect ($\xi^W_i$) in equation (25) captures the average attractiveness of workplace $i$ across sectors in terms of its wage ($w_{i,g}$) and productivity draws ($T^W_{i,g}$) in equation (8), such that:

$$\xi^W_i = \sum_{m \in K} T^W_{i,m} w^{\theta_W}_{i,m}. \tag{26}$$

Finally, the estimated residence fixed effect ($\zeta^W_n$) in equation (25) corresponds to the denominator in the conditional commuting probability in equation (8) and captures the overall attractiveness of residence $n$ in terms of its travel-time weighted access to all workplaces:

$$\zeta^W_n = \sum_{\ell \in N} \sum_{m \in K} T^W_{\ell,m} w^{\theta_W}_{\ell,m} \exp\left(-\theta_W \kappa \tau_{n\ell}\right). \tag{27}$$

[14] One rationalization for these zeros is that the assumption of a continuous measure of workers in the model is an approximation to the discrete and finite number of workers in the data. Consistent with this rationalization, we follow Dingel and Tintelnot (2020) in estimating the gravity equation using the PPML estimator.

Given an estimated or calibrated dispersion parameter for idiosyncratic productivity draws ($\theta_W$), we can recover a theoretically-consistent estimate of workplace access ($\mathbb{W}_n$) from this estimated residence fixed effects ($\zeta_n^W$):

$$\mathbb{W}_n = \Gamma\left(\frac{\theta_W - 1}{\theta_W}\right)\left[\sum_{\ell \in N}\sum_{m \in K} T_{\ell,m}^W w_{\ell,m}^{\theta_W} \exp\left(-\theta_W \kappa \tau_{n\ell}\right)\right]^{\frac{1}{\theta_W}} = \Gamma\left(\frac{\theta_W - 1}{\theta_W}\right)\left(\zeta_n^W\right)^{\frac{1}{\theta_W}}. \tag{28}$$

Alternatively, we can obtain another estimate of workplace access by summing the estimated workplace fixed effects ($\zeta_i^W$) weighted using the estimated bilateral travel costs ($\exp\left(-\theta_W \kappa \tau_{ni}\right)$). If the model corresponds exactly to the data generating process and bilateral travel times are perfectly symmetric, we should obtain the same measure of workplace access from these two different approaches. Empirically, we find that these two estimates are extremely highly correlated with one another, as discussed further below.[15]

Intuitively, under the assumptions of the model, the observed conditional commuting probabilities in the smartphone data are sufficient to reveal the relative attractiveness of each workplace as a commuting destination, and the overall attractiveness of each residence in terms of its access to these workplaces.

## 5.2 Estimating Consumption Access

We next use the observed bilateral consumption trips in our smartphone data to estimate consumption access. From the conditional consumption probabilities (4), we obtain the following log linear gravity equation for bilateral consumption trips for service $k$ between residence $n$ and consumption destination $j(k)$:

$$\lambda_{nj(k),k|n}^S = \frac{\xi_{j(k),k}^S \exp\left(-\phi_{S,k}\tau_{nj(k)}\right)\exp\left(u_{nj(k),k}^S\right)}{\zeta_{n,k}^S}, \tag{29}$$

where $u_{nj(k),k}^S$ is a stochastic error that captures idiosyncratic determinants of bilateral travel costs not captured in bilateral travel times (e.g. how pleasant the journey is).

Following the same approach as above, we estimate this services gravity equation separately for each type of non-traded service using the Poisson Pseudo Maximum Likelihood (PPML) estimator. The estimated elasticity of consumption trips with respect to travel times ($\phi_{S,k}$) in equation (29) is a composite of the elasticity of consumption trips with respect to travel costs ($\theta_{S,k}$) and the elasticity of travel costs with respect to travel times ($\kappa$) in equation (4), such that $\phi_{S,k} = \theta_{S,k}\kappa$. The estimated consumption destination fixed effect ($\xi_{j(k),k}^S$) in equation (29) captures the average attractiveness of consumption destination $j(k)$ for service $k$ in terms of its price for that non-traded service ($P_{j(k),k}$) and quality draws ($T_{j(k),k}^S$) in equation (4), such that:

$$\xi_{j(k),k}^S = T_{j(k)}^S\left(P_{j(k),k}\right)^{-\theta_{S,k}}. \tag{30}$$

The estimated residence fixed effect in equation (29) corresponds to the denominator in the conditional consumption probability in equation (4) and captures the overall attractiveness of residence $n$ in terms of its travel-time weighted access to all consumption locations $\ell(k)$ for service $k$:

$$\zeta_{n,k}^S = \sum_{\ell \in N} T_{\ell(k),k}^S\left(P_{\ell(k),k}\right)^{-\theta_{S,k}} \exp\left(-\theta_{S,k}\kappa\tau_{n\ell(k)}\right). \tag{31}$$

Given estimated or calibrated values for the dispersion parameter for idiosyncratic quality draws ($\theta_{S,k}$) and the share of consumer expenditure on each type of non-traded service ($\alpha_k$), we can recover a theoretically-consistent estimate

[15]In practice, bilateral travel times need not be perfectly symmetric, because for example of one-way streets or differences in rail timetables in two different directions on a given bilateral route.

of consumption access for each type of non-traded service ($\mathbb{S}_{n,k}$) from the estimated residence fixed effects ($\zeta^S_{n,k}$) for that type of non-traded service:

$$\mathbb{S}_{n,k} = \Gamma\left(\frac{\theta_{S,k}/\alpha_k - 1}{\theta_{S,k}/\alpha_k}\right)\left[\sum_{\ell\in N} T^S_{\ell(k),k}\left(P_{\ell(k),k}\right)^{-\theta_{S,k}}\exp\left(-\theta_{S,k}\kappa\tau_{n\ell(k)}\right)\right]^{\frac{\alpha_k}{\theta_{S,k}}}, \tag{32}$$

$$= \Gamma\left(\frac{\theta_{S,k}/\alpha_k - 1}{\theta_{S,k}/\alpha_k}\right)\left(\zeta^S_{n,k}\right)^{\frac{\alpha_k}{\theta_{S,k}}}. \tag{33}$$

Alternatively, we can again obtain another estimate of consumption access for each type of non-traded service by summing the estimated consumption destination fixed effects ($\zeta^S_{j(k),k}$) weighted using the estimated bilateral travel times ($\exp\left(-\theta_{S,k}\kappa\tau_{ni}\right)$). As discussed above, if the model corresponds exactly to the data generating process and bilateral travel times are perfectly symmetric, we should obtain the same measure of consumption access each type of non-traded service from these two different approaches. As for workplace access, we find that in practice these two estimates are extremely highly correlated with one another, as discussed further below.

Aggregating across the different types of non-traded services $k$ using the Cobb-Douglas functional form of the consumption goods price index, we obtain overall estimated consumption access for each residence ($\mathbb{S}_n$):

$$\mathbb{S}_n = \prod_{k\in K^S}\Gamma\left(\frac{\theta_{S,k}/\alpha_k - 1}{\theta_{S,k}/\alpha_k}\right)\exp\left(\zeta^S_{n,k}\right)^{\frac{\alpha_k}{\theta_{S,k}}}. \tag{34}$$

Intuitively, under the assumptions of the model, the observed conditional commuting probabilities in the smartphone data are sufficient to reveal the relative attractiveness of each consumption destination for each type of non-traded service, and hence the overall attractiveness of each residence in terms of its access to these consumption destinations for all types of non-traded services.

## 5.3 Parameter Estimation and Calibration

We now discuss our estimation and calibration of the model parameters. We proceed in a number of steps. First, we estimate the commuting and consumption gravity equations ((26) and (31) respectively), which yields estimates of the commuting and consumption fixed effects ($\xi^W_i$, $\zeta^W_n$, $\xi^S_{j(k),k}$, $\zeta^S_{n,k}$) and the composite elasticities of commuting and consumption trips with respect to travel times ($\theta_W\kappa$ and $\theta_{S,k}\kappa$). Second, we use the structure of the model and the estimated commuting workplace fixed effects ($\xi^W_i$) to recover the dispersion parameter for workplace idiosyncratic draws ($\theta_W$). Third, we use this estimated dispersion parameter for workplace idiosyncratic draws ($\theta_W$) and the composite elasticities for commuting and consumption trips ($\theta_W\kappa$ and $\theta_{S,k}\kappa$ respectively) to recover the dispersion parameter for services idiosyncratic draws ($\theta_{S,k}$). Fourth, and finally, we calibrate the remaining parameters of the model using the observed data and central values from the existing empirical literature.

### 5.3.1 Gravity Equation Estimation (Step 1)

In our first step, we estimate our gravity equations for consumption and commuting for our baseline sample of users in the Tokyo metropolitan area in April 2019. Column (1) of Table 1 reports our estimates of the commuting gravity equation (25). Observations correspond to bilateral pairs of municipalities within the Tokyo metropolitan area. Consistent with the reduced-form evidence discussed above, we find that the model's gravity equation predictions provide a good approximation to the observed bilateral commuting flows. We estimate a negative and statistically significant semi-elasticity of commuting flows with respect to travel time of $-\phi_W = -\theta_W\kappa = -0.054$.

Columns (2)-(7) report our estimates of the consumption gravity equation (29) for each of our different types of non-traded services: "Finance and Real Estate"; "Wholesale and Retail"; "Accommodation, Eating and Drinking"; "Medical, Welfare and Health Care"; "Other Services"; and "Open space". We find that the model's gravity equation predictions also provide a good approximation to the observed bilateral patterns of consumption trips in our smartphone data. We estimate a negative and statistically significant semi-elasticity for consumption trips ($-\phi_{S,k} = -\theta_{S,k}\kappa$) that is typically somewhat larger than that for commuting trips. This estimated semi-elasticity is relatively constant across the different types of consumption trips, ranging from -0.039 to -0.069, with the most localized consumption trips observed for "Wholesale and Retail" and "Medical, Welfare and Health Care".

Our findings that both commuting and consumption trips decline sharply with bilateral travel time suggest that access to surrounding employment and consumption are both salient for workers' residential location choices. In Subsection 5.4 below, we use our estimates of the composite elasticities of commuting and consumption trips with respect to travel time ($-\phi_W = -\theta_W\kappa$ and $-\phi_S = -\theta_{S,k}\kappa$), the commuting residence and destination fixed effects ($\zeta_n^W$ and $\xi_i^W$), and the consumption residence and destination fixed effects ($\xi_{j(k),k}^S$ and $\zeta_{n,k}^S$) to quantify the contributions of workplace and consumption access to residential location decisions.

Table 1: Gravity Equation Estimation Results (Commuting and Consumption Trips)

| | Commuting | Finance real estate communication professional | Wholesale retail | Accomodations eating drinking | Medical welfare healthcare | Other services | Open space |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Travel Time (Minutes) | -0.054 | -0.065 | -0.064 | -0.063 | -0.069 | -0.059 | -0.039 |
| | (0) | (0.002) | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) |

Note: Observations are bilateral pairs of oazas in the Tokyo metropolitan area. All specifications use the Poisson Pseudo Maximum Likelihood (PPML) estimator. Column (1) regresses the conditional commuting probabilities (8) on bilateral travel time and residence and destination fixed effects. Columns (2)-(7) regress the conditional consumption probabilities (4) for each type of non-traded service on bilateral travel time and residence and destination fixed effects. Standard errors in parentheses are clustered two-way on residence and destination.

### 5.3.2 Dispersion of Workplace Idiosyncratic Draws (Step 2)

In our second step, we use the gravity equation estimates from the first step and the structure of the model to recover the implied dispersion of workplace idiosyncratic draws ($\theta_W$). From equations (9) and (28), the estimated fixed effect for residence $n$ ($\zeta_n^W$) in the commuting gravity equation is equal to $\theta_W$ times the log of expected income conditional on living in residence $n$. Therefore, we estimate the workplace dispersion parameter ($\theta_W$) by regressing these estimated residence fixed effects ($\zeta_n^W$) on the log of average residential income in the census data:

$$\zeta_n^W = \theta_W \log\left(\text{Residential Income}_n\right) + \varepsilon_n, \tag{35}$$

where the regression error ($\varepsilon_n$) captures estimation error in the residence fixed effects, and any departures from model assumptions. Note that since average residential income data is available in municipality level, we estimate the equation by aggregating oaza-level data to municipality level weighting by population density.

In Panel (A) of Table 2, we report the results of estimating this relationship (35) across the municipalities in the Tokyo metropolitan region. In Column (1) including all municipalities, we find a positive and statistically significant relationship, with an estimated dispersion parameter of $\theta_W = 5.00$. In Panel (B), we scatter the estimated residence fixed effect (vertical axis) against average residential income in the data (horizontal axis). Consistent with the predictions of the model, we find an approximately log linear relationship between these variables. At high income levels, we observe three outliers (Minato, Chiyoda, Shibuya Municipalities), where average income in the data is greater than predicted by the linear regression, which could reflect a role for non-labor income that is not captured by the model at high income levels. In our preferred specification in Column (2), we re-estimate equation (35) excluding these three outliers. We find a similar pattern of results, with a marginally larger estimated dispersion parameter of $\theta_W = 6.36$. These estimates are consistent with the range of values for this parameter reported in the existing empirical literature, and compare for example with the estimates of 6.83 in Ahlfeldt, Redding, Sturm, and Wolf (2015), 5.25 in Heblich, Redding, and Sturm (2020), and 7.1 in Kreindler and Miyauchi (2019).

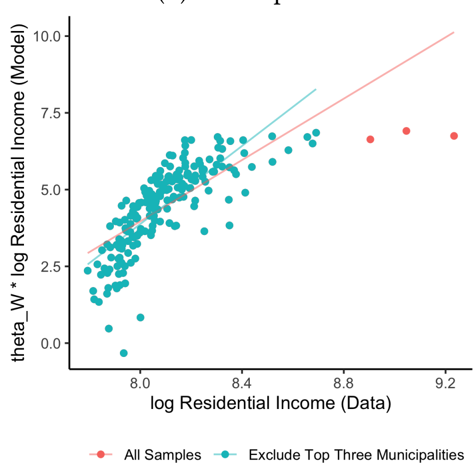
Table 2: Estimation Results of $\theta_W$

(A) regression results

| | (theta_W * log Residential Income (Model)) | |
|---|---|---|
| | (1) | (2) |
| log Residential Income (Data) | 5.00*** | 6.36*** |
| | (0.32) | (0.36) |
| Samples | All | Exclude Top 3 Municipalities |
| Observations | 203 | 200 |
| $R^2$ | 0.55 | 0.61 |

Note: Regression of log estimated residence fixed effect ($\zeta_n^W$) from the commuting gravity gravity equation (26) on log average residential income in the census data; observations are municipalities in the Tokyo metropolitan area; heteroskedasticity standard errors in parentheses.

(B) scatter plots




Note: Scatterplot of log estimated residence fixed effect ($\zeta_n^W$) from the commuting gravity gravity equation (26) on the vertical axis against log average residential income in the income tax base data on the horizontal axis; observations are municipalities in the Tokyo metropolitan area (some municipalities are aggregated due to data availability); straight lines show the fitted values from the regression specifications in Columns (1) and (2) of the table in Panel (A) above.

In both specifications, we find that the estimated residence fixed effects have substantial explanatory power, with a regression R-squared of around 0.60. Although there are several reasons why the estimated residence fixed effects need not coincide exactly with average residential income (including for example non-labor income), this strong relationship between estimated expected worker income in the model and measured average worker income in the data provides an important validation check.

#### 5.3.3 Dispersion of Service Idiosyncratic Draws (Step 3)

In our third step, we use our estimates from the first and second steps, together with the structure of the model, to recover the services dispersion parameter ($\theta_{S,k}$). We recover services dispersion parameter ($\theta_{S,k}$) using the estimated semi-elasticity for consumption trips for non-traded services for each sector $k$ ($\overline{\phi}_{S,k}$) and our estimated workplace dispersion parameter ($\theta_W$) as follows:

$$\theta_{S,k} = \left(\frac{\phi_{S,k}}{\phi_W}\right)\theta_W = \left(\frac{\theta_{S,k}\kappa}{\theta_W\kappa}\right)\theta_W. \tag{36}$$

Table 3 reports our estimates of $\theta_{S,k}$ for each of our different types of non-traded services: "Finance and Real Estate"; "Wholesale and Retail"; "Accommodation, Eating and Drinking"; "Medical, Welfare and Health Care"; "Other Services"; and "Open space". Again, these estimates are obtained by combining the estimated semi-elasticities for commuting trips and consumption trips for each non-tradable sector (Table 1) and our estimated workplace dispersion parameter of $\theta_W = 6.36$ in equation (36). Finally, we can recover the implied travel cost parameter ($\kappa$) from either the consumption trips or commuting trips semi-elasticities, using $\kappa = \phi_W/\theta_W = \phi_{S,k}/\theta_S = 0.0084$.

Table 3: Estimation Results of $\theta_{S,k}$

| Sector | Value |
|---|---|
| GJKL finance realestate communication professional | 7.87 |
| I wholesale retail | 7.57 |
| M accomodations eating drinking | 7.46 |
| P medical welfare healthcare | 8.22 |
| Q other services | 7.03 |
| Z. Open Space | 4.80 |

Note: Estimation results of $\theta_{S,k}$ from equation (36).

#### 5.3.4 Other Model Parameters (Step 4)

We calibrate the other model parameters directly from the data or using central values from the existing empirical literature, as summarized in Table 4 below. First, we calibrate the Cobb-Douglas expenditure shares for residential floor space ($\alpha_H$), the non-traded service sectors ($\alpha_k$), and the traded good ($\alpha_T$) using economy-wide data on expenditure for Tokyo metropolitan area. Second, we calibrate the Cobb-Douglas cost shares for labor in each sector ($\beta_k$) using economy-wide data on production costs by sector for Tokyo metropolitan area. Third, we assume a value for the parameter determining the dispersion of residence idiosyncratic draws of $\theta_B = 5$, which is which is in line with the range of values in the existing empirical literature in Ahlfeldt, Redding, Sturm, and Wolf (2015) and Heblich, Redding, and Sturm (2020). Fourth, we assume a share of land in construction costs of $\mu = 0.25$ following Epple, Gordon, and Sieg (2010) and Combes, Duranton, and Gobillon (2019). Finally, we explore a range of values for the production and residential agglomeration parameters ranging from zero to the values estimated in Ahlfeldt, Redding, Sturm, and Wolf (2015): $\eta_W \in [0, 0.08]$ and $\eta_B \in [0, 0.15]$.

Table 4: Calibrated Parameters

| Parameters | Description | Value |
|---|---|---|
| $\theta_W$ | dispersion of Fréchet shocks for workplace | 6.36 |
| $\theta_{S,k}$ | dispersion of Fréchet shocks for nontradable consumption location | Table 3 |
| $\kappa$ | travel cost per minute | 0.0084 |
| $\alpha_H$ | consumption share of housing | 0.4 |
| $\alpha_k$ | consumption share of nontradeables | 0.48 (total); sales share |
| $\alpha_T$ | consumption share of tradeables | 0.12 |
| $\theta_B$ | dispersion of Fréchet shocks for residence | 5 |
| $\beta_k$ | labor share in production | 0.8 |
| $\eta_W$ | production spillovers | 0.08 |
| $\eta_B$ | residential amenity spillovers | 0.15 |
| $\mu$ | share of capital for floor space production | 0.75 |

Note: Top panel reports estimated parameters; Bottom panel reports calibrated parameters using data for the Japanese economy and central values from the existing empirical literature.

## 5.4 Quantifying the Contributions of Workplace and Consumption Access

Using these parameter values, we now quantify the contributions of workplace access, consumption access, and unobserved amenities in explaining the observed spatial concentration of economic activity. We start by re-writing the residential choice probabilities (10) in the following form:

$$\left(\lambda_n^B\right)^{1/\theta_B} Q_n^{\alpha_H} = \mathbb{B}_n \mathbb{W}_n \mathbb{S}_n, \tag{37}$$

where $\mathbb{B}_n$ is a composite amenities parameter for residence $n$ that includes common amenities ($B_n$), the parameter determining average idiosyncratic amenities ($T_n^B$), the common price of the traded good ($P_n^T = P^T = 1$), and the common reservation level of utility ($\bar{U}$):

$$\mathbb{B}_n \equiv B_n \left(T_n^B\right)^{1/\theta_B} \left(P_n^T\right)^{-\alpha_T} \left(\bar{U}/\vartheta^B\right)^{-1} \tag{38}$$

In these residential choice probabilities (37), we observe the share of residents ($\lambda_n^B$) and the price of floor space ($Q_n$), and we can compute workplace access ($\mathbb{W}_n$) and consumption access ($\mathbb{S}_n$) using our gravity equation estimates in equations (28) and (34). Therefore, we can use these residential choice probabilities (37) to recover the unobserved composite amenities ($\mathbb{B}_n$) that exactly rationalize the observed data as an equilibrium of the model.

Intuitively, the left-hand side of these residential choice probabilities (37) corresponds to a summary measure of the relative attractiveness of locations. If we observe a location that has relatively many residents ($\lambda_n^B$) and high price of floor space ($Q_n$) on the left-hand side, despite having relatively low values of workplace access ($\mathbb{W}_n$) and consumption access ($\mathbb{S}_n$) on the right-hand side, this is rationalized in the model by that location having relatively high unobserved amenities ($\mathbb{B}_n$). These unobserved amenities correspond to a structural residual that ensures that the model is consistent with the observed data. In the special case of our quantitative urban model without consumption trips ($\alpha_k = 0$ for all $k \in K^S$, $\alpha_T = 1 - \alpha_H$, $\lambda^S_{nj(k),k|n} = 0$ and $\mathbb{S}_{nt} = 1$), this structural residual would include the omitted consumption access. In contrast, we use our smartphone GPS to directly estimate consumption access, thereby explicitly modelling this component of residential decisions, and taking it out of the structural residual.

We now decompose the variance of this summary measure of the relative attractiveness of locations into the contributions of workplace access ($\mathbb{W}_n$), consumption access ($\mathbb{S}_n$) and the structural residual of composite amenities ($\mathbb{B}_n$). In particular, we use a regression-based variance decomposition, as implemented in the international trade literature in Eaton, Kortum, and Kramarz (2004). We estimate an ordinary least squares (OLS) regression of each of the components on the right-hand side of the residential choice probabilities (37) on the summary measure of the relative attractiveness of locations from the left-hand side:

$$\begin{aligned}
\ln \mathbb{W}_n &= c_0^W + c_1^W \ln \left( \left(\lambda_n^B\right)^{1/\theta_B} Q_n^{\alpha_H} \right) + u_{nt}^W, \\
\ln \mathbb{S}_n &= c_0^S + c_1^S \ln \left( \left(\lambda_n^B\right)^{1/\theta_B} Q_n^{\alpha_H} \right) + u_{nt}^S, \\
\ln \mathbb{B}_n &= c_0^B + c_1^B \ln \left( \left(\lambda_n^B\right)^{1/\theta_B} Q_n^{\alpha_H} \right) + u_{nt}^B,
\end{aligned} \tag{39}$$

Noting that OLS is a linear estimator with mean zero residuals, and using the residential choice probabilities (37), it follows that $c_0^B + c_0^W + c_0^S = 0$ and $c_1^B + c_1^W + c_1^S = 1$. Implicitly, this variance decomposition allocates the covariance terms equally across each of the three components of the residential choice probailities. The relative values of the slope coefficients $\{c_1^B, c_1^W, c_1^S\}$ provide measures of the relative importance of workplace access ($\mathbb{W}_n$), consumption access ($\mathbb{S}_n$) and composite amenities ($\mathbb{B}_n$) in explaining the observed variation in our summary measure of the relative attractiveness of locations.

In Panel (A) of Table 5, we report the results of estimating equation (39) for oazas in the Tokyo metropolitan area. The price of floor space in each location is constructed from land price data using the assumption of perfectly competitive floor space developpers (24). We estimate positive and statistically significant coefficients for both commuting access and workplace access as well as the structural residual of amenities. Comparing Columns (1) and (2), we find a contribution from consumption access that is around two thirds as large as the contribution from workplace access.[16] Therefore, access to consumption opportunities is almost as important as access to employment opportunities is explaining why some locations with the Tokyo metropolitan area are able to attract more residents and command higher prices for floor space than other locations. If we were to abstract from consumption access and absorb it into the structural residual of amenities, the contribution of this residual would be larger than workplace access ($0.265 + 0.280 = 0.545$). In contrast, once we explicitly model consumption access and separate it from the structural residual of amenities, we find that this structural residual makes a relatively modest contribution (0.280) to the observed variation in the relative attractiveness of locations.

[16] We find a similar pattern of results across a wide range of different specifications, including for example whether we use residence or destination fixed effects to compute workplace and consumption access, as shown in Table A.1 in the online appendix.

Table 5: Variance Decomposition of the Relative Attractiveness of Locations (log $\left[\left(\lambda_n^B\right)^{1/\theta_B} Q_n^{\alpha_H}\right]$)

(A) Commuting Access, Consumption Access and Residual of Composite Amenities

| | log Commuting Access | log Consumption Access | log Amenity |
|---|---|---|---|
| | (1) | (2) | (3) |
| $\log Q_n^{\alpha^H} \left(\lambda_n^B\right)^{1/\theta^B}$ | 0.454*** | 0.265*** | 0.280*** |
| | (0.008) | (0.005) | (0.012) |
| Observations | 4,449 | 4,449 | 4,449 |
| $R^2$ | 0.449 | 0.410 | 0.107 |

(B) Consumption Access for Each Sector

| | Finance real estate communication professional | Wholesale retail | Accommodations eating drinking | Medical welfare healthcare | Other services |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| $\log Q_n^{\alpha^H} \left(\lambda_n^B\right)^{1/\theta^B}$ | 0.111*** | 0.120*** | 0.003*** | 0.012*** | 0.019*** |
| | (0.002) | (0.002) | (0.0001) | (0.0002) | (0.0003) |
| Observations | 4,449 | 4,449 | 4,449 | 4,449 | 4,449 |
| $R^2$ | 0.382 | 0.405 | 0.384 | 0.368 | 0.410 |

Note: Panel (A): Ordinary least squares (OLS) estimates of the regression-based variance decomposition in equation (39); Panel (B) separates consumption access into its different components for each type of non-traded service. The sum of the coefficients across these components in Panel (B) equals the overall coefficient for consumption access in Panel (A). Observations are oazas in the Tokyo metropolitan area. Heteroskedasticity robust standard errors in parentheses.

In Panel (B) of Table 5, we use the property that overall consumption access is a Cobb-Douglas aggregate of the consumption access for each non-traded sector to separate out the overall contribution of consumption access into the contribution from each non-traded sector. Across all sectors, we estimate positive and statistically significant coefficients for consumption access, with "Finance and Real Estate" and "Wholesale and Retail" making the largest contributions, which are roughly equal in magnitude to one another.

# 6 Counterfactuals

In this section, we use our theoretical framework to undertake counterfactuals for changes in travel costs to provide further evidence on the role of consumption trips in shaping the spatial distribution of economic activity. In Section 6.1, we introduce the system of equations for undertaking these counterfactuals. In Section 6.2, we examine the contribution of consumption trips towards agglomeration by undertaking a counterfactual in which we assume no travel costs for commuting and consumption trips, thus eliminating the spatial frictions in commuting or consumption. In Section 6.3, we examine the role of consumption trips in shaping the welfare effects of transport infrastructure improvements by undertaking a counterfactual for the construction of a new subway line.

## 6.1 Exact-hat Algebra Counterfactuals

In this section, we show that we can undertake counterfactuals for the impact of a change in travel costs using only the observed values of the endogenous variables in an initial equilibrium, following the exact-hat algebra approach of Dekle, Eaton, and Kortum (2007). In particular, we rewrite the counterfactual equilibrium conditions in terms of the observed endogenous variables in the initial equilibrium and the unobserved changes in these endogenous variables between the counterfactual and initial equilibria. An important advantage of this approach is that we can undertake these counterfactuals without having to estimate the high-dimensional unobserved location characteristics, such as the Fréchet shape parameters ($T_n^B$, $T_{i,g}^W$, $T_{j(k)}^S$). Instead, we use the observed values of the endogenous variables in the initial equilibrium to capture these unobserved location characteristics.

In our baseline specification, we consider the closed-city specification of the model, in which total population for the city as a whole ($\bar{L}$) is exogenous, and hence the change in travel costs affects worker welfare.[17] We denote the value of a variable in the initial equilibrium by $x$, the value of this variable in the counterfactual equilibrium by $x'$ (with a prime), the relative change in this variable by $\hat{x} = x'/x$ (with a hat), and the absolute change in this variable using the difference operator $\Delta x = x' - x$. Given values for the model parameters ($\alpha_H, \{\alpha_k\}, \theta_S, \theta_W, \theta_B, \kappa_W, \kappa_S, \eta_B, \eta_W, \{\beta_k\}, \mu$), assumed bilateral changes in travel times $\{\Delta\tau_{in}\}$, and observed values for the endogenous variables in the initial equilibrium ($\{\lambda_{ni,g|n}^W, \lambda_{nj,k|n}^S, \lambda_n^B\}$, $\{\tilde{L}_{i,k}\}$, $\{H_{j,k}, H_{n,U}\}$, $\{\mathbb{W}_n\}$), we solve for the counterfactual equilibrium by solving the following system of equations for the general equilibrium of the model.[18]

**(i) Changes in Commuting and consumption probabilities** From equations (4) and (8), the counterfactual changes in conditional commuting probabilities ($\hat{\lambda}_{nig|n}^W$) and conditional consumption probabilities ($\hat{\lambda}_{njk|n}^S$) satisfy:

$$\hat{\lambda}_{ni,g|n}^W = \frac{\hat{w}_{i,g}^{\theta_W} e^{-\theta_W \kappa_W \Delta\tau_{ni}}}{\sum_{\ell \in N} \sum_{m \in K} \hat{w}_{i,m}^{\theta_W} e^{-\theta_W \kappa_W \Delta\tau_{n\ell}} \lambda_{n\ell,m|n}^W}, \tag{40}$$

$$\hat{\lambda}_{nj,k|n}^S = \frac{\hat{P}_{j,k}^{-\theta_{S,k}} e^{-\theta_{S,k}\kappa_S \Delta\tau_{nj}}}{\sum_{\ell \in N} \hat{P}_{\ell,k}^{-\theta_{S,k}} e^{-\theta_{S,k}\kappa_S \Delta\tau_{n\ell}} \lambda_{n\ell,k|n}^S}. \tag{41}$$

Using equations (4), (6), (8) and (9), the corresponding changes in workplace access ($\hat{\mathbb{W}}_n$) and consumption access ($\hat{\mathbb{S}}_n$) can be written in terms of own commuting shares ($\hat{\lambda}_{nn,T|n}^W$) and own consumption shares ($\hat{\lambda}_{nn,k|n}^S$):

$$\hat{\mathbb{W}}_n = \left[\hat{w}_{n,T}^{\theta_W} \exp\left(-\theta_W \kappa_W \Delta\tau_{nn}\right) / \hat{\lambda}_{nn,T|n}^W\right]^{\frac{1}{\theta_W}}, \tag{42}$$

$$\hat{\mathbb{S}}_n = \prod_k \left[\hat{P}_{n,k}^{-\theta_{S,k}} \exp\left(-\theta_{S,k}\kappa_S \Delta\tau_{nn}\right) / \hat{\lambda}_{nn,k|n}^S\right]^{\frac{\alpha_k}{\theta_{S,k}}}. \tag{43}$$

**(ii) Changes in residential location decision** From equation (10), the counterfactual changes in residential probabilities ($\hat{\lambda}_n^B$) satisfy:

$$\hat{\lambda}_n^B = \frac{\hat{\mathbb{W}}_n^{\theta_B} \hat{\mathbb{S}}_n^{\theta_B} \hat{Q}_n^{-\alpha_H \theta_B}}{\sum_{\ell \in N} \hat{\mathbb{W}}_\ell^{\theta_B} \hat{\mathbb{S}}_\ell^{\theta_B} \hat{Q}_\ell^{-\alpha_H \theta_B} \lambda_\ell^B}. \tag{44}$$

[17] It is straightforward to instead consider the open-city specification, in which case total population is endogenous, and the welfare effects of the change in travel costs accrue only to landlords, as in the public finance literature following George (1879).

[18] In the counterfactual simulation, we allow the travel cost parameter $\kappa$ can be different for commuting ($\kappa_W$) and for consumption ($\kappa_S$) to allow for the simulation of shutting down the travel cost for each type of trips.

**(iii) Changes in commercial and residential floor space demand** From equations (12) and (14), the changes in commercial floor space in each sector ($\hat{H}_{i,g}$) and residential floor space ($\hat{H}_{i,U}$) are given by:

$$\hat{H}_{i,g} = \frac{\hat{\tilde{L}}_{i,g}}{\hat{Q}_i}, \tag{45}$$

$$\hat{H}_{i,U} = \frac{\hat{\mathbb{W}}_i \hat{\lambda}_i^B}{\hat{Q}_i}, \tag{46}$$

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

$$\hat{\tilde{L}}_{i,g} = \frac{1}{\tilde{L}_{i,g}} \sum_{n \in N} \mathbb{W}_n^{'} \lambda_{ni,g|n}^{'W} R_n^{'}. \tag{47}$$

**(iv) Changes in the price of floor space** From equation (23), the change in the price of floor space ($\hat{Q}_i$) and the overall quantity of floor space ($\hat{H}_i$) are related as follows:

$$\hat{Q}_i = \hat{H}_i^{\frac{1-\mu}{\mu}}, \tag{48}$$

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

$$\hat{H}_i = \frac{H_{i,U} \hat{H}_{i,U} + \sum_{k \in K} H_{i,k} \hat{H}_{i,k}}{H_{i,U} + \sum_{k \in K} H_{i,k}}. \tag{49}$$

**(v) Changes in endogenous productivities and amenities** From equations (19) and (20), the changes in endogenous productivities ($\hat{A}_{i,k}$) and amenities ($\hat{B}_n$) as a result of agglomeration forces satisfy:

$$\hat{A}_{i,k} = \hat{L}_i^{\eta_W}, \tag{50}$$

$$\hat{B}_n = \hat{R}_n^{\eta_B}. \tag{51}$$

**(vi) Changes in nontraded goods prices** From equation (16), the changes in non-traded goods prices ($\hat{P}_{i,k}$) satisfy:

$$\hat{P}_{i,k} = \frac{1}{\hat{A}_{i,k} \hat{\tilde{L}}_{i,k}^{\beta_k} \hat{H}_{i,k}^{1-\beta_k}} \frac{\sum_{n \in N} \mathbb{W}_n^{'} \lambda_{nj,k|n}^{'S} \lambda_n^{'B}}{\sum_{n \in N} \mathbb{W}_n \lambda_{nj,k|n}^{S} \lambda_n^{B}}. \tag{52}$$

**(viii) Changes in Wages** From the zero-profit condition (13), the changes in wages in each sector and location with positive production ($\hat{w}_{i,k}$) are given by:

$$\hat{w}_{i,k} = \left( \frac{\hat{A}_{i,k} \hat{P}_{i,k}}{\hat{Q}_i^{1-\beta_k}} \right)^{1/\beta_k}. \tag{53}$$

We solve this system of equations (40)-(53), starting with an initial guess of the relative change in each endogenous variable ($\hat{x} = 1$), and updating this initial guess until the solution to this system converges to equilibrium. Using the resulting counterfactual changes in the endogenous variables of the model ($\hat{\lambda}_{ni,g|n}^W$, $\hat{\lambda}_{nj,k|n}^S$, $\hat{\mathbb{W}}_n$, $\hat{\mathbb{S}}_n$, $\hat{\lambda}_n^B$, $\hat{H}_{i,g}$, $\hat{H}_{i,U}$, $\hat{\tilde{L}}_{i,g}$, $\hat{Q}_i$, $\hat{A}_{i,k}$, $\hat{B}_n$, $\hat{P}_{i,k}$, $\hat{w}_{i,k}$), together with equation (11), we can compute the implied change in expected utility ($\widehat{\mathbb{E}[u]}$) induced by the change in travel costs as follows:

$$\widehat{\mathbb{E}[u]} = \left[ \sum_{\ell \in N} \lambda_\ell^B \hat{B}_\ell^{\theta_B} \hat{\mathbb{W}}_\ell^{\theta_B} \hat{\mathbb{S}}_\ell^{\theta_B} \hat{Q}_\ell^{-\alpha_H \theta_B} \right]^{\frac{1}{\theta_B}}, \tag{54}$$

where we have used our choice of numeraire ($P_{\ell,T} = 1$ for all $\ell \in N$).

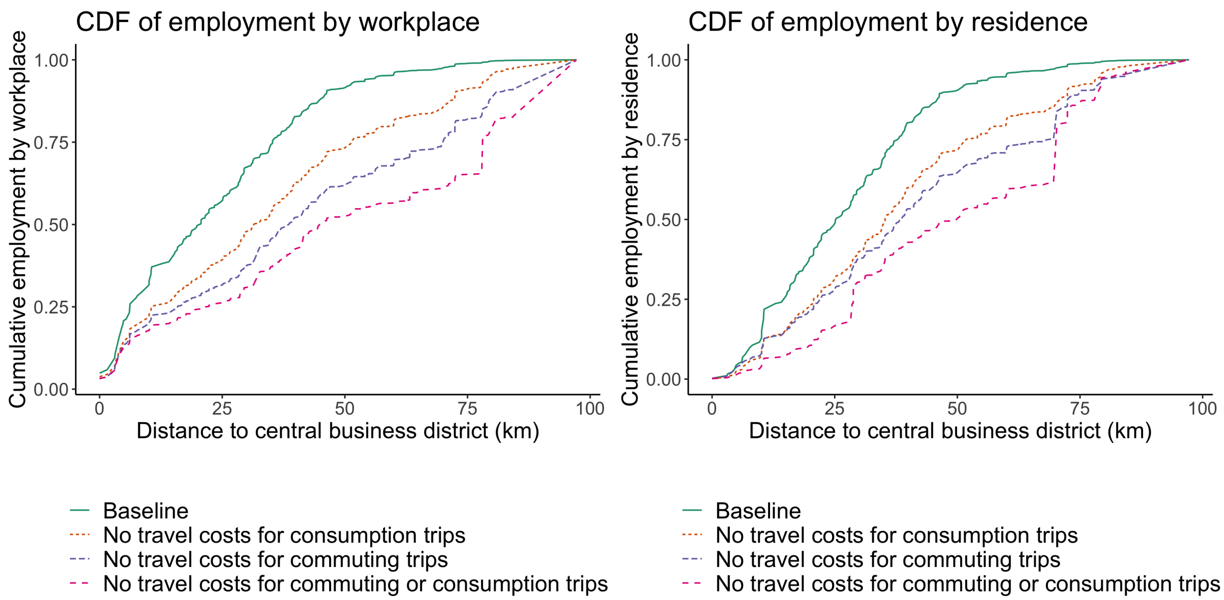
As discussed above, conventional quantitative urban models without consumption trips correspond to a special case of our model, in which $\alpha_k = 0$ for all $k \in K_S$, $\alpha_T = 1 - \alpha_H$, $\lambda^S_{nj(k),k|n} = 0$ and $\mathbb{S}_{nt} = 1$. In this special case, the change in expected utility ($\widehat{\mathbb{E}[u]}$) as a result of the change in travel costs continues to be given by equation (11), but with $\hat{S}_{nt} = 1$ for all $n \in N$. Therefore, if one assumes no consumption trips, whereas in reality they occur, there are two sources of bias in these counterfactual predictions for the impact of changes in travel costs. First, the general equilibrium predictions of the model for the changes in all endogenous variables (such as the price of floor space, the probability of working in each location and the probability of living in each location) are in general different with and without consumption trips. The reason is that these consumption trips change the relative attractiveness of locations as residences (because of their differential access to other locations for consumption) and change relative factor demand across locations (through travel from each location to consume non-traded goods in other locations). Second, even if the general equilibrium predictions for all other endogenous variables were the same, abstracting from consumption trips would lead to a systematic understatement of the welfare gains from reductions in travel costs. The reason is that these reductions in travel costs typically raise consumption access ($\hat{\mathbb{S}}_n > 1$), which acts to increase expected utility in equation (54), whereas the special case of the model without consumption trips involves setting $\hat{\mathbb{S}}_n = 1$ in this expression for the change in expected utility. Intuitively, the special case of the model without consumption trips undercounts the travel journeys that benefit from the reduction in travel costs, which leads to an understatement of the welfare gains from this reduction in travel costs.

## 6.2 Travel Costs and the Spatial Concentration of Economic Activity

In our first set of counterfactuals, we provide further evidence on the role of travel costs for commuting and consumption in shaping the spatial concentration of economic activity, by shutting down each of these sources of spatial frictions. In a first exercise, we set travel costs for commuting trips equal to zero ($\kappa_W = 0$), maintain travel costs for consumption trips equal to their estimated value in the data ($\kappa_S > 0$), and solve for the counterfactual equilibrium distribution of economic activity. In a second exercise, we set travel costs for consumption trips equal to zero ($\kappa_S = 0$), maintain travel costs for commuting trips equal to their estimated value in the data ($\kappa_W > 0$), and solve for the counterfactual equilibrium. Finally, in a third exercise, we set travel costs for both commuting and consumption trips equal to zero ($\kappa_S = \kappa_W = 0$), and solve for the counterfactual equilibrium.

In Figure 9, we display the results of these three counterfactuals. In the left panel, we show the cumulative distribution function (CDF) of employment by workplace against distance in kilometers from the center of Tokyo's central business district (CBD), measured as the centroid of Chiyoda Ward. In the right panel, we present the corresponding cumulative distribution of employment by residence against distance from the CBD. To provide a point of comparison, we begin by displaying these cumulative distributions in the data using our baseline sample (labelled "Baseline"). We next display the counterfactual cumulative distributions for employment by workplace and residence in our first exercise, in which we set travel costs for commuting trips to zero (labelled "no travel costs for commuting trips"). In the initial equilibrium in the data with positive travel costs for both types of trips, workers trade off the lower prices of floor space in outlying locations against the higher commuting costs in those outlying locations. When we set travel costs for commuting trips equal to zero in the counterfactual, we eliminate this difference in commuting costs between central and outlying locations, which increases the relative attractiveness of these outlying locations. As a result, we find a decentralization of employment by residence, as reflected in a rightward shift of the cumulative distribution function in the right panel. Since we maintain positive travel costs for consumption trips, this decentralization of employment by residence increases the demand for non-traded services in outlying locations, which in turn leads to a decentralization of employment by workplace, as reflected in a rightward shift in the cumulative distribution function in the left panel.

Figure 9: Counterfactuals for Eliminating Travel Costs for Commuting and Consumption Trips




Note: Left panel shows cumulative distribution function of employment by workplace against distance to the centroid of Chiyoda Ward, as the central point of Tokyo's central business district (CBD); right panel shows the analogous cumulative distribution function of employment by residence against distance from the CBD; baseline corresponds to the observed distributions for our baseline sample for April 2019; the three counterfactuals assume zero travel costs for consumption trips, for commuting trips, and for both consumption and commuting trips.

We next display these counterfactual cumulative distributions for employment by workplace and residence for our second exercise, in which we set travel costs for consumption trips to zero (labelled "no travel costs for consumption trips"). In the initial equilibrium with positive travel costs for both types of trips in the data, locations close to the central business district have higher employment in non-traded services than outlying locations, as firms trade-off the greater accessibility of central locations against the higher land values in those central locations. When we set travel costs for consumption trips equal to zero in the counterfactual, we eliminate the difference in accessibility to customers between central and outlying locations, which increases the relative attractiveness of outlying locations. As a result, we again find a decentralization of employment by workplace, as reflected in a rightward shift in the cumulative distribution function in the left panel. Since we maintain positive travel costs for commuting trips, this decentralization of employment by workplace increases the demand for labor in outlying locations relative to central locations, which in turn leads to a decentralization of employment by residence, as reflected in a rightward shift in the cumulative distribution function in the right panel. Consistent with the findings of our earlier decomposition in Section 5.4 above, we find that the impact of changes in travel costs for consumption trips on these cumulative distributions of employment is substantial relative to that from changes in travel costs for commuting trips.

Finally, we display these counterfactual cumulative distributions for employment by workplace and residence for our third exercise, in which we set travel costs for both commuting and consumption trips to zero (labelled "no travel costs for commuting or consumption trips"). In the initial equilibrium with positive travel costs for both types of trips in the data, central locations have greater accessibility to both workers and to customers for non-traded services, because the most central locations in the Tokyo metropolitan area have the lowest average travel costs to other locations within the metropolitan area, especially in the presence of a transport network with hub-and-spoke characteristics. In this initial equilibrium, workers and firms trade off the greater accessibility of central locations against their higher land prices. When we set travel costs for both commuting and consumption trips equal to zero in the counterfactual, we eliminate this difference in accessibility between central and outlying locations, which increases the relative attractiveness of outlying locations. As a result, we again find a decentralization of both employment by workplace and employment by residence, as reflected in a rightward shift of the cumulative distribution functions in both panels. Since reductions in each type of travel costs alone lead to a decentralization of economic activity, when we reduce travel costs for both types of trips, we find that this decentralization of economic activity is substantially greater than when we reduce each type of travel cost alone. In our baseline equilibrium, about half of all employment by workplace is located within 20 kilometers of the CBD. When we eliminate travel costs for consumption and commuting trips separately, this distance threshold for half of all employment by workplace rises to 30 and 35 kilometers respectively. When we eliminate travel costs for both types of trips together, this distance threshold rises further to 45 kilometers.

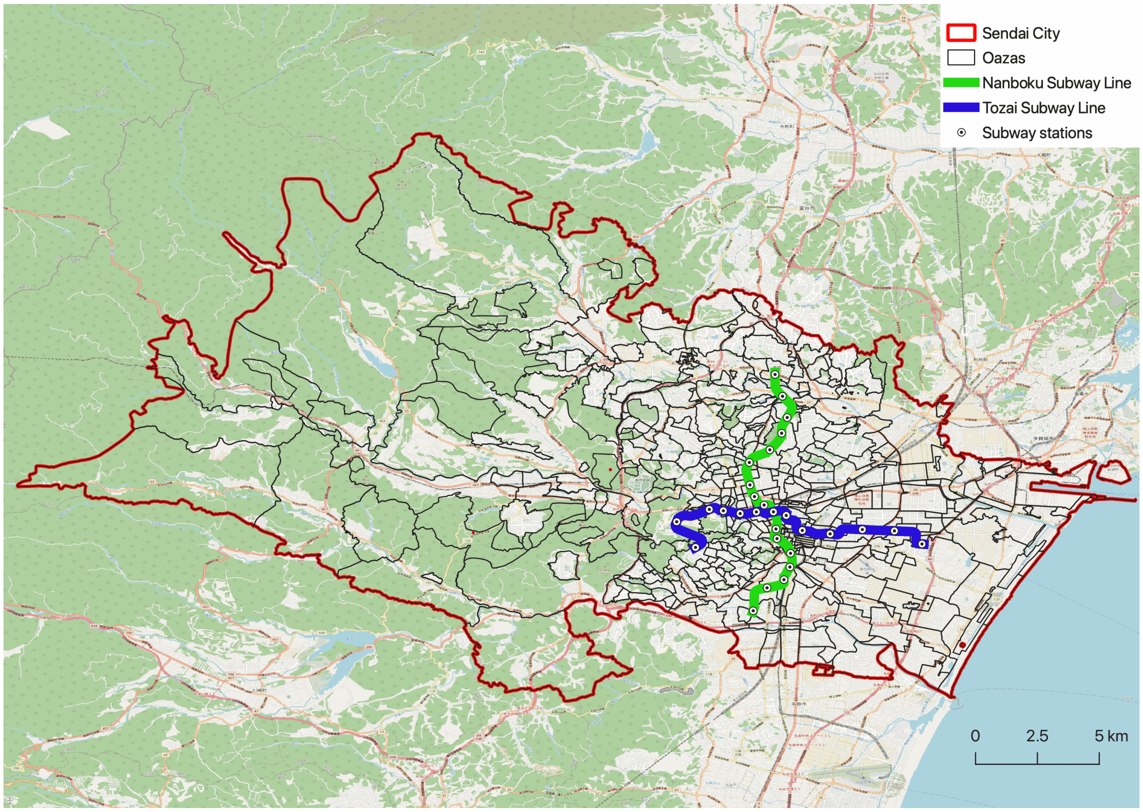
Overall, the results of these counterfactuals for changes in travel costs provide further evidence that consumption access is quantitatively important for the spatial concentration of economic activity in urban areas relative to the workplace access that has received much greater emphasis in previous research.

## 6.3 Transportation Infrastructure

In our second set of counterfactuals, we examine the role of consumption access in shaping the impact of transport infrastructure improvements. In particular, we use our smartphone data and quantitative model to evaluate the impact of the construction of a new subway (underground) line in the city of *Sendai* on welfare. Before the construction of this new line, there had been only one *Nanboku* (North-South) subway line, which had been in operation since 1987. In December 2015, the new *Tozai* (East-West) subway line opened, thereby providing a substantial expansion in the overall subway network, as shown in Figure 10.

Following the opening of this new subway line, we find an increase in stays in our smartphone data around the new subway stations, as shown in Section A.2 of the online appendix. Motivated by these reduced-form findings, we use our model to undertake counterfactuals for the impact of this new subway line. In particular, we first calculate bilateral travel times between locations with and without the new subway line, and hence the resulting bilateral reduction in travel times from the new line. Using these reductions in bilateral travel times, we next solve the system of general equilibrium equations (40)-(53) for the closed-city specification of the model for the counterfactual impact of the new subway line. We compare these counterfactual predictions of our model incorporating consumption trips with those of the special case that abstracts from consumption trips, in which $\alpha_k = 0$ for all $k \in K_S$, $\alpha_T = 1 - \alpha_H$, $\lambda^S_{nj(k),k|n} = 0$ and $\mathbb{S}_{nt} = 1$.

Figure 10: Map of *Sendai* City




In Table 6, we report the counterfactual predictions of the model for the welfare impact of the opening of the *Tozai* (East-West) subway line. Column (1) reports the results for the model with consumption trips and Column (2) reports the results for the special case without consumption trips. As shown in Column (1), we nevertheless find an increase in flow utility from the new subway line of 3.28 percent. Therefore, although there were other modes of transport prior to the opening of the new line (such as buses), we find substantial welfare gains from the reduction in bilateral travel times achieved by this new subway line. As shown in Column (2), we find that these welfare gains are around 35 percent smaller in the special case of the model without consumption trips. As discussed above, these two specifications generate different counterfactual predictions for the spatial distribution of economic activity. Nevertheless, most of this difference in welfare gains is explained by the fact the special case of the model without consumption trips abstracts from the improvement in consumption access (by setting $\hat{\mathbb{S}}_{nt} = 1$ in the change in expected utility in equation (54)). This pattern of results implies that the undercounting of travel journeys by focusing solely on consumption trips is quantitatively important for the evaluation of the welfare effects of actual transport infrastructure improvements. A full cost-benefit analysis of this transport infrastructure improvement would involve a comparison of the net present value of these welfare gains to the construction costs and the net present value of operating profits or losses and maintenance costs. For a marginal project for which the net present values of benefits and costs lie close to one another, an underestimate of the increase in flow utility from a transport improvement of 35 percent could well be consequential in making the case for proceeding with that transport improvement.

Table 6: Increase of Residential Utility by Tozai Subway Line in Sendai City

| (a) With Consumption Trips | (b) Without Consumption Trips |
|---|---|
| 3.28 | 2.40 |
| (1.00) | (0.73) |

Note: The numbers indicate the percentage increase of flow expected utility predicted by the travel time change from the opening of the Tozai (East-West) subway line in the city of Sendai.

# 7 Conclusions

We provide new theory and evidence on the role of consumption access in understanding the spatial concentration of economic activity. We use smartphone data that records the global positioning system (GPS) location of users every 5 minutes to provide an unprecedented level of detail on patterns of travel by hour and day within the Tokyo metropolitan area. Guided by our empirical findings, we develop a quantitative model of internal city structure that captures the fact that much of the travel that occurs within urban areas is related not to commuting but rather to the consumption of non-traded services, such as trips to restaurants, coffee shops and bars, shopping expeditions, excursions to cinemas, theaters, music venues and museums, and visits to professional service providers.

We begin by establishing three key empirical properties of these non-commuting trips. We show that they are more frequent than commuting trips, so that concentrating solely on commuting substantially underestimates travel within urban areas. We find that they are concentrated closer to home and have higher elasticities with respect to travel time than commuting trips, which implies that focusing solely on commuting yields a misleading picture of bilateral patterns of travel within cities. Finally, combining our smartphone data with highly spatially-disaggregated data on employment by sector, we show that these non-commuting trips are closely related to the availability of nontraded sectors, consistent with our modelling of them as travel to consume non-traded services.

We next develop our quantitative theoretical model of internal city structure that incorporates these consumption trips. Workers choose their preferred residence, workplace and consumption locations, taking into account the bilateral costs of travel and idiosyncratic draws for amenities for each residence, productivity for each workplace and service quality for each consumption location. We first use the model's gravity equations for commuting and consumption trips to estimate the relative valuation that users place on different locations and construct theoretically-consistent measures of workplace and consumption access. We next use the model's population mobility condition to decompose the observed variation in residents and the price of floor space into the contributions of workplace access, consumption access and unobserved amenities. We find that the contribution from consumption access is comparable to that from workplace access (25 percent compared to 45 percent) and that controlling for consumption access substantially reduces the contribution from unobserved amenities (from 50 percent to 30 percent).

Finally, we show how the model can be used to undertake counterfactuals for the impact of changes in travel costs on the spatial distribution of economic activity. In a first set of counterfactuals, we eliminate spatial frictions for commuting and consumption trips, and show that both sets of spatial frictions make substantial contributions to the concentration of economic activity. In a second set of counterfactuals, we evaluate the impact of the construction of new transport infrastructure on the spatial distribution of economic activity. We show that abstracting from consumption trips leads to a substantial underestimate of the welfare gains from a transport infrastructure improvement (because of the undercounting of trips) and leads to a distorted picture of changes in travel patterns within the city (because of the different geography of commuting and non-commuting trips).

Taken together, our findings suggest that access to consumption opportunities as well as access to employment opportunities plays a central role in understanding the concentration of economic activity in urban areas.

## References


Agarwal, S., J. B. Jensen, and F. Monte (2020): "Consumer Mobility and the Local Structure of Consumption Industries," *NBER Working Paper*, 23616.

Ahlfeldt, G., S. Redding, D. Sturm, and N. Wolf (2015): "The Economics of Density: Evidence from the Berlin Wall," *Econometrica*, 83(6), 2127–2189.

Allen, T., and C. Arkolakis (2014): "Trade and the Topography of the Spatial Economy," *Quarterly Journal of Economics*, 129(3), 1085–1140.

——— (2017): "The Welfare Effects of Transportation Infrastructure Improvements," Dartmouth College.

Allen, T., C. Arkolakis, and X. Li (2017): "Optimal City Structure," Yale University, mimeograph.

Allen, T., S. Fuchs, S. Ganapati, A. Graziano, R. Madera, and J. Montoriol-Garriga (2020): "Is Tourism Good for Locals? Evidence from Barcelona," Dartmouth College, mimeograph.

Almagro, M., and T. Domínguez-Iino (2019): "Location Sorting and Endogenous Amenities: Evidence from Amsterdam," NYU, mimeograph.

Alonso, W. (1964): *Location and Land Use*. Harvard, Cambridge MA.

Anas, A., and Y. Liu (2007): "A Regional Economy, Land Use, And Transportation Model (RELU-TRAN): Formulation, Algorithm Design, and Testing," *Journal of Regional Science*, 47(3), 415–455.

Athey, S., B. Ferguson, M. Gentzkow, and T. Schmidt (2018): "Experienced Segregation," Stanford University, mimeograph.

Balboni, C. (2019): "In Harm's Way: Infrastructure Investments and the Persistence of Coastial Cities," MIT, mimeograph.

Baum-Snow, N. (2007): "Did Highways Cause Suburbanization?," *Quarterly Journal of Economics*, 122(2), 775–780.

Baum-Snow, N., L. Brandt, J. Henderson, M. A. Turner, and Q. Zhang (2017): "Roads, Railroads and Decentralization of Chinese Cities," *Review of Economics and Statistics*, 99(2), 435–448.

Büchel, K., M. V. Ehrlich, D. Puga, and E. Viladecans (2020): "Calling from the Outside: The Role of Networks in Residential Mobility," *Journal of Urban Economics*, 119(103277).

Caliendo, L., F. Parro, E. Rossi-Hansberg, and P.-D. Sarte (2018): "The Impact of Regional and Sectoral Productivity Changes on the U.S. Economy," *Review of Economic Studies*, 85(4), 2042–2096.

Chandra, A., and E. Thompson (2000): "Does Public Infrastructure affect Economic Activity? Evidence from the Rural Interstate Highway System," *Regional Science and Urban Economics*, 30(4), 457–490.

Combes, P.-P., G. Duranton, and L. Gobillon (2019): "The Costs of Agglomeration: Land Prices in French Cities," *Review of Economic Studies*, 86(4), 1556–1589.

Combes, P.-P., and L. Gobillon (2015): "The Empirics of Agglomeration Economies," in *Handbook of Regional and Urban Economics*, ed. by G. Duranton, J. V. Henderson, and W. C. Strange, vol. 5A, chap. 5, pp. 247–348. Elsevier North Holland, Amsterdam.

Couture, V., J. Dingel, A. Green, and J. Handbury (2019): "Quantifying Social Interactions Using Smartphone Data," Chicago Booth School of Business, mimeograph.

Couture, V., G. Duranton, and M. A. Turner (2018): "Speed," *Review of Economics and Statistics*, 100(4), 725–739.

Couture, V., and J. Handbury (2019): "Urban Revival in America," Wharton, mimeograph.

Davis, D. R., J. I. Dingel, J. Monras, and E. Morales (2019): "How Segregated Is Urban Consumption?," *Journal of Political Economy*, 127(4), 1684–1738.

Davis, D. R., and D. E. Weinstein (2002): "Bones, Bombs, and Break Points: The Geography of Economic Activity," *American Economic Review*, 92(5), 1269–1289.

Dekle, R., J. Eaton, and S. Kortum (2007): "Unbalanced Trade," *American Economic Review*, 97(2), 351–355.

Diamond, R. (2016): "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," 106(3), 479–524.

Dingel, J., and F. Tintelnot (2020): "Spatial Economics for Granular Settings," *NBER working Paper*, 27287.

Dolfen, P., L. Einav, P. J. Klenow, B. Klopack, J. D. Levin, L. Levin, and W. Best (2019): "Assessing the Gains from E-Commerce," *NBER Working Paper*, 25610.

Donaldson, D. (2018): "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*, 108(4-5), 899–934.

Donaldson, D., and R. Hornbeck (2016): "Railroads and American Economic Growth: A Market Access Approach," *Quarterly Journal of Economics*, 131(2), 799–858.

Duranton, G., and D. Puga (2004): "Micro-Foundations of Urban Agglomeration Economies," in *Handbook of Regional and Urban Economics*, ed. by J. V. Henderson, and J.-F. Thisse, vol. 4, pp. 2063–2117. Elsevier, Amsterdam.

Duranton, G., and M. A. Turner (2011): "The Fundamental Law of Road Congestion: Evidence from US Cities," *American Economic Review*, 101(6), 2616–52.

——— (2012): "Urban Growth and Transportation," *Review of Economic Studies*, 79(4), 1407–1440.

Eaton, J., S. Kortum, and F. Kramarz (2004): "Dissecting Trade: Firms, Industries, and Export Destinations," *American Economic Review*, 94(2), 150–154.

Epple, D., B. Gordon, and H. Sieg (2010): "A New Approach to Estimating the Production Function for Housing," *American Economic Review*, 100(3), 905–924.

Faber, B. (2014): "Integration and the Periphery: The Unintended Effects of New Highways in a Developing Country," *Review of Economic Studies*, 81(3), 1046–1070.

Fajgelbaum, P., and C. Gaubert (2020): "Optimal Spatial Policies, Geography and Sorting," *Quarterly Journal of Economics*, 135(2), 959–1036.

Fajgelbaum, P. D., and E. Schaal (2020): "Optimal Transport Networks in Spatial Equilibrium," *Econometrica*, 88(4), 1411–1452.

Florida, R. (2009): *Who's Your City?: How the Creative Economy Is Making Where to Live the Most Important Decision of Your Life*. Basic Books, New York.

Fujita, M., and P. Krugman (1995): "When is the Economy Monocentric? Von Thünen and Chamberlain Unified," *Regional Science and Urban Economics*, 25(4), 505–528.

Fujita, M., P. Krugman, and A. J. Venables (1999): *The Spatial Economy: Cities, Regions, and International Trade*. MIT Press, Cambridge MA.

Fujita, M., and H. Ogawa (1982): "Multiple Equilibria and Structural Transformation of Non-monocentric Urban Configurations," *Regional Science and Urban Economics*, 12(2), 161–196.

Fujita, M., and J.-F. Thisse (2002): *Economics of Agglomeration: Cities, Industrial Location and Regional Growth*. Cambridge University Press, Cambridge.

George, H. (1879): *Progress and Poverty*. Doubleday, Page & Company, New York.

Glaeser, E. L., J. Kolko, and A. Saiz (2001): "Consumer City," *Journal of Economic Geography*, 1(1), 27–50.

Gorback, C. (2020): "Ridesharing and the Redistribution of Economic Activity," National Bureau of Economic Research, mimeograph.

Gupta, A., C. Kontokosta, and S. Van Nieuwerburgh (2020): "Take the Q Train: Value Capture of Public Infrastructure Projects," *NBER Working Paper*.

Head, K., and T. Mayer (2014): "Gravity Equations: Workhorse, Toolkit, and Cookbook," in *Handbook of International Economics*, ed. by E. Helpman, G. Gopinath, and K. Rogoff, vol. 4, pp. 131–196. Elsevier, North Holland, Amsterdam.

Heblich, S., S. Redding, and D. Sturm (2020): "The Making of the Modern Metropolis: Evidence from London," *Quarterly Journal of Economics*, 135(4), 2059–2133.

Henderson, J. V. (1974): "The Sizes and Types of Cities," *American Economic Review*, 64(4), 640–56.

Hoelzlein, M. (2020): "Two-sided Sorting and Spatial Inequality in Cities," University of California, Berkeley.

Hsieh, C., E. Hurst, C. I. Jones, and P. J. Klenow (2019): "The Allocation of Talent and U.S. Economic Growth," *Econometrica*, 87(5), 1439–1474.

Kline, P., and E. Moretti (2014): "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority," *Quarterly Journal of Economics*, 129(1), 275–331.

Kreindler, G., and Y. Miyauchi (2019): "Measuring Commuting and Economic Activity inside Cities with Cell Phone Records," Boston University, mimeograph.

Leonardi, M., and E. Moretti (2019): "Agglomeration of Urban Amenities: Evidence from Milano Restaurants," .

Lucas, R. E., and E. Rossi-Hansberg (2002): "On the Internal Structure of Cities," *Econometrica*, 70(4), 1445–1476.

McFadden, D. (1974): "The Measurement of Urban Travel Demand," *Journal of Public Economics*, 3(4), 303–328.

Michaels, G. (2008): "The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System," *Review of Economics and Statistics*, 90(4), 683–701.

Mills, E. S. (1967): "An Aggregative Model of Resource Allocation in a Metropolitan Centre," *American Economic Review*, 57(2), 197–210.

Monte, F., S. Redding, and E. Rossi-Hansberg (2018): "Commuting, Migration and Local Employment Elasticities," *American Economic Review*, 108(12), 3855–3890.

Moretti, E. (2011): "Local Labor Markets," in *Handbook of Labor Economics*, ed. by D. Card, and O. Ashenfelter, vol. 4b, pp. 1238–1303. Elsevier North Holland, Amsterdam.

Muth, R. (1969): *Cities and Housing*. University of Chicago Press, Chicago.

Owens, R. I., E. Rossi-Hansberg, and P.-D. Sarte (2020): "Rethinking Detroit," *American Economic Journal: Economic Policy*, 12(2), 258–305.

Ramondo, N., A. Rodríguez-Clare, and M. Saborío-Rodríguez (2016): "Trade, Domestic Frictions, and Scale Effects," *American Economic Review*, 106(10), 3159–84.

Redding, S. J. (2016): "Goods Trade, Factor Mobility and Welfare," *Journal of International Economics*, 101, 148–167.

Redding, S. J., and E. Rossi-Hansberg (2017): "Quantitative Spatial Models," *Annual Review of Economics*, 9, 21–58.

Rosenthal, S. S., and W. C. Strange (2004): "Evidence on the Nature and Sources of Agglomeration Economics," in *Handbook of Regional and Urban Economics*, ed. by J. V. Henderson, and J. Thisse, vol. 4. Elsevier, Amsterdam.

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

Santos Silva, J., and S. Tenreyro (2006): "The Log of Gravity," *Review of Economics and Statistics*, 88(4), 641–658.

Severen, C. (2019): "Commuting, Labor, and Housing Market Effects of Mass Transportation: Welfare and Identification," Federal Reserve Bank of Philadelphia.

Storeygard, A. (2016): "Farther on Down the Road: Transport costs, Trade and Urban Growth," *Review of Economic Studies*, 83(3), 1263–1295.

Thibault, F. (2015): "Structural Gravity and Fixed Effects," *Journal of International Economics*, 97(1), 76–85.

Tsivanidis, N. (2018): "The Aggregate And Distributional Effects Of Urban Transit Infrastructure: Evidence From Bogotá's TransMilenio," University of California, Berkeley, mimeograph.

Zárate, R. D. (2020): "Spatial Misallocation, Informality, and Transit Improvements: Evidence from Mexico City," University of California, Berkeley.