

The Economics of Spatial Mobility: Theory and Evidence Using Smartphone Data*

Yuhei Miyauchi[†]
Boston University

Kentaro Nakajima[‡]
Hitotsubashi University

Stephen J. Redding[§]
Princeton University, NBER and CEPR

April 20, 2022

Abstract

Using smartphone geographical positioning systems (GPS) data for Japan, we show that travel within urban areas frequently occurs along trip chains, involving multiple stops as part of a single journey. Motivated by these empirical findings, we develop a tractable theoretical model of travel itineraries, in which agents choose a set and sequence of locations to visit each day. To overcome the resulting high-dimensionality of the choice set, we develop an approach based on importance sampling. We show that trip chains introduce consumption externalities across locations. We show that these consumption externalities are central to explaining the collapse in foot traffic in downtown areas following the shift to remote working during the Covid-19 pandemic.

Keywords: Spatial Mobility, Cities, Economic Geography
JEL Classification: O18, R12, R40

*This paper absorbs and replaces material that previously circulated under the title “Consumption Access and the Spatial Concentration of Economic Activity: Evidence from Smartphone Data.” Thanks to Gabriel Ahlfeldt, Milena Almagro, Daniel Sturm, Gabriel Kreindler, Tobias Salz and conference and seminar participants for helpful comments. We are grateful to Takeshi Fukasawa, Peter Deffebach, and Yun-Ting Yeh 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 are processed collectively and statistically in order to conceal private information. Original location data is GPS data (latitude, longitude) sent every five minutes (minimum), and it does not include information to specify individual. The copyrights of all tables and figures presented in this document belong to ZENRIN DataCom CO., LTD. We also acknowledge Yaichi Aoshima at Hitotsubashi University for coordinating the project with ZENRIN DataCom Co., LTD.; Murata Foundation, Heiwa Nakajima Foundation, The Kajima Foundation, Obayashi Foundation, JSPS KAKENHI (Grant No. 21H00703), and the Hitotsubashi University for their financial support; CSIS at the University of Tokyo for the joint research support (Project No. 954).

[†]Dept. Economics, 270 Bay State Road, Boston, MA 02215. Tel: 1-617-353-5682. Email: miyauchi@bu.edu.

[‡]Institute of Innovation Research, 2-1 Naka, Kunitachi, Tokyo 186-8603, Japan. Tel: 81-42-580-8417. E-mail: nakajima.kentaro@gmail.com

[§]Dept. Economics and SPIA, JRR Building, Princeton, NJ 08544. Tel: 1-609-258-4016. Email: red-dings@princeton.edu.

1 Introduction

Each day, people make millions of travel journeys throughout urban areas, with an estimated 40 million daily trips using the rail system in Greater Tokyo alone.¹ Frequently this travel involves trip chains, in which people make multiple stops as part of a single journey. We show that a key implication of these trip chains is consumption externalities across locations, where having a good reason to visit one location makes it more attractive to visit other locations that are nearby or along the way. We demonstrate that these consumption externalities shape both firm and worker location decisions, and hence influence the internal structure of cities. We show that these consumption externalities are central to understanding the collapse in foot traffic, and hence the demand for non-traded services, in downtown areas following the shift to remote working during the Covid-19 pandemic.

Analyzing these trip chains and consumption externalities raises both theoretical and empirical challenges. From a theoretical perspective, modelling trip chains is difficult, because of the high-dimensionality of the choice set. If a city consists of many locations, and if agents can choose to visit any subset of these locations in any sequence, it soon becomes computationally challenging to evaluate payoffs for every combination of choices. From an empirical perspective, analyzing trip chains is demanding, because the locations visited are often nearby, the time spent at each location can be brief, and the sequence in which these locations are visited matters for travel costs. Therefore, one requires data on travel within urban areas that has a high level of spatial and temporal resolution, and in which one can track the path followed by agents over time. This high temporal resolution becomes all the more important for acute events such as the Covid-19 pandemic, for which administrative data on travel behavior can appear with lags of months or years.

In this paper, we provide new theory and evidence on spatial mobility within urban areas to address these challenges. We develop a tractable theoretical model of travel itineraries, in which agents choose a set of locations and the sequence in which to visit these locations each day. To overcome the resulting high-dimensionality of the choice set, we develop an approach based on importance sampling. We allow for rich patterns of spatial mobility, with agents visiting an endogenous subset of locations in an endogenous sequence. Our framework is thus well suited to analyzing the host of new data sources on spatial mobility that are emerging with the increasing availability of high-resolution Geographical Positioning System (GPS) information (e.g. smartphones, GPS transponders).

We combine our theoretical model with novel data from a major smartphone mapping

¹See for example https://www.mlit.go.jp/kisha/kisha07/01/010330_3/01.pdf. Each segment of a trip that involves transfers between different operators is counted separately.

application in Japan (*Docomo Chizu NAVI*), which records the GPS location of each device up to every 5 minutes each day. As of July, 2019, the raw data cover about 545,000 users, with 1.4 billion raw data points. Although these smartphone data are detailed, in the sense that they contain high-resolution information about spatial mobility, this mobility information alone does not directly tell us about the relative attractiveness of destinations. We overcome this limitation by combining our smartphone data with our theoretical model of travel itinerary decision to reveal the relative attractiveness of home, work and other destinations. We also supplement our smartphone data with administrative data on other economic characteristics, such as sectoral employment and floor space.

Using smartphone data, we structurally estimate the parameters of agents' travel itinerary decisions. We further embed this specification of travel itinerary decisions in a quantitative urban model that determines the remainder of the general equilibrium. We use the natural experiment of the Covid-19 pandemic as a specification check on our quantitative model's predictions. We find that our model incorporating trip chains is quantitatively successful in explaining the decline in foot traffic in downtown areas following the shift to remote working. In contrast, the explanatory power of the model deteriorates substantially if we abstract from trip chains, and the resulting consumption externalities, by assuming that all travel begins and ends from home. We use our estimated model to show that incorporating trip chains is also consequential for the evaluation of public policies in urban areas, such as the construction of new transport infrastructure.

This paper proceeds in the following order. We begin by using our smartphone data to establish a number of stylized facts about travel in urban areas. We measure the locations visited by users using "stays," which correspond to occurrences of no movement within 100 meters for 15 minutes. Based on this definition, we measure a 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 refer to other locations as "other stays" and trips to these other locations as "non-commuting trips." 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 the official census data. We find similar bilateral commuting patterns between municipalities in our smartphone data as in the official census data.

Having validated our smartphone data using the census commuting data, we exploit the high spatial and temporal frequency of our data to establish a number of stylized facts about spatial mobility within urban areas. First, non-commuting trips are more frequent than commuting trips, so that focusing solely on commuting trips understates travel within urban areas. Second, non-commuting trips are closely related to the availability of non-traded services.

Third, non-commuting trips exhibit different spatial patterns from commuting trips, such that they are not well approximated by commuting trips. Fourth, trips chains are a pervasive feature of the data, where individuals stop at multiple destinations along a single journey starting from and ending at home. Finally, we show that there is a decline in the frequency and travel length of trips during the Covid-19 pandemic, which results in a disproportionate reduction in foot traffic in downtown Tokyo.

To rationalize these observed features of the data, we develop a tractable theoretical model of travel itineraries. We consider a nested decision problem, in which agents in a given residential location first choose a workplace, taking into account commuting costs and access to consumption opportunities. Having chosen their workplace, they next choose a travel itinerary each day, which corresponds to the set of locations to visit and the sequence in which to visit these locations. Agents make this decision based on the net utility from the bundle of non-traded services at each destination, the travel costs incurred by the travel itinerary, and an idiosyncratic preference draw for each itinerary. This preference draw captures the many idiosyncratic reasons that individuals can choose a particular route (e.g., the scenery or the preferred time of arrival at each stop). We characterize the probability that each agent chooses a particular itinerary and the expected utility derived from that itinerary.

Our travel-itinerary specification considerably generalizes conventional models of consumption within urban areas. In our specification, the market access of both workers and firms depends on the frequency with which travel routes are chosen. This property introduces a consumption externality across locations: if one location becomes a more attractive destination, this makes other locations that are nearby or along the way more attractive. This consumption externality crucially affects location decisions of firms and workers, and hence the internal structure of the city. First, employment concentrations in traded sectors (e.g., manufacturing) attract employment concentrations in non-traded services (e.g., restaurants serving manufacturing-sector workers). Second, agents' choice of locations in which to consume these non-traded services depends on their residence, workplace and the route followed from home to work. Third, non-traded service firms (e.g., retail stores) cluster together to attract the customers of their competitors, as they pass by along the way to visit those competitors (as in the classic example of clusters of shoe shops).

To assess the quantitative importance of these consumption externalities for the location decisions of firms and workers, and hence the internal structure of cities, we develop a framework to estimate and simulate our model. One important challenge for the estimation and simulation of agents' travel itinerary decisions is the high-dimensionality of the choice set, which involves many different combinations of the locations to visit and the order in which to visit them (e.g., [Bowman and Ben-Akiva 2001](#); [Anas 2007](#)). To overcome this high-dimensionality

of the choice set, we develop a new method based on importance sampling (e.g., [Kloek and van Dijk 1978](#)). This method uses a Monte Carlo simulation from an auxiliary distribution, and adjusts the sampling rate based on the likelihood ratio between the true distribution and the auxiliary distribution. We are thus able to simulate itinerary choice probabilities using this Monte-Carlo simulation, while avoiding the curse of dimensionality. We estimate the model parameters that govern agents' travel itinerary choices using a method-of-moments estimation procedure, which minimizes the distance between moments in the simulation and the observed data. We show that our estimated model provides a good fit to observed patterns of spatial mobility for both targeted and untargeted moments. We also embed our specification of travel itineraries in a conventional quantitative urban model of residence choice following [Ahlfeldt, Redding, Sturm, and Wolf \(2015\)](#), where wages, firm entry, prices, and rents are endogenously determined in the general equilibrium.

We use the natural experiment of the Covid-19 pandemic as a further specification check on our model's predictions. First, we calibrate our model using data from before the Covid-19 pandemic (April 2019). Second, we use data in the early stage of the Covid-19 pandemic (April 2020) to estimate two structural parameters (the change in the frequency of traveling to work and the change in the cost of travelling per minute). Third, we use the estimated model to predict changes in travel patterns within the Tokyo Metropolitan Area. Despite the tight parameterization of our model, we show that it is quantitatively successful in capturing the decline in foot traffic, and hence the demand for non-traded services, in downtown areas following the shift to remote working. We compare our model's predictions to the special case in which we abstract from trip chains (and the resulting consumption externalities) by assuming that all travel begins and ends from home. We show that this special case substantially underestimates the decline in foot traffic in downtown areas following the shift to remote working, highlighting the relevance of consumption externalities in understanding the internal structure of cities.

Having validated our model using the Covid-19 pandemic, we show that these consumption externalities are also quantitatively relevant for understanding the impact of public policies, such as transport infrastructure improvements. Starting at the observed equilibrium in the data in April 2019, we undertake a counterfactual for the removal of all overground and underground railways constructed in Tokyo from 1960 to 2019. We compare the predictions of our model incorporating travel itineraries with two special cases: (i) only commuting trips from home to work; (ii) both commuting and non-commuting trips, but no trip chains, such that all travel is assumed to begin and end at home.

We find that frameworks that focus solely on commuting trips generally underestimate the welfare gains from transport improvements. This is in part because models that abstract from

these non-commuting trips undercount the amount of travel that benefits from reduced travel costs. Similarly, models that abstract from trip chains also typically underestimate the welfare gains from these transport improvements. By assuming that all travel starts and ends at home, these models do not fully capture all the ways in which agents can adjust their travel patterns in response to reductions in travel costs. Therefore, whether we use within-sample variation from a natural experiment or undertake counterfactuals for out-of-sample shocks, we find that accurately measuring trip chains and resulting consumption externalities is quantitatively relevant for understanding the spatial distribution of economic activity within cities.

The remainder of the paper is structured as follows. Section 2 discusses the relationship between our research and the existing literature. Section 3 introduces our data. Section 4 presents the stylized facts about spatial mobility that motivate our theoretical model of travel itineraries. Section 5 introduces and estimates this travel itinerary model. Section 6 embeds this travel itinerary decision in a general equilibrium quantitative urban model. Section 7 uses quasi-experimental variation from the shift to remote working during the Covid-19 pandemic as a specification check on our model’s predictions of consumption externalities across locations. Section 8 undertakes counterfactuals to examine the implications of these consumption externalities for the impact of transport improvements. Section 9 concludes.

2 Related Literature

Our research is related to a number of different strands of existing work. First, our paper contributes to the theoretical and empirical literature on the internal structure of cities. Traditional theories of urban economics emphasize commuting between workplace and residence, including monocentric city models (Alonso 1964; Mills 1967; Muth 1969), polycentric city models (Fujita and Ogawa 1982; Lucas and Rossi-Hansberg 2002), and more recent quantitative urban models (Ahlfeldt, Redding, Sturm, and Wolf 2015; Allen, Arkolakis, and Li 2017; Monte, Redding, and Rossi-Hansberg 2018; Tsivanidis 2019; Dingel and Tintelnot 2020). While this research focuses on travel from home to work, we provide evidence of richer patterns of spatial mobility, including non-commuting trips, and develop a tractable theoretical framework to rationalize these observed patterns of spatial mobility.

Second, we contribute to recent research on endogenous amenities within cities (Glaeser, Kolko, and Saiz 2001; Florida 2009; Diamond 2016; Gechter and Tsivanidis 2020). One strand of this literature studies how endogenous amenities affect residential income segregation, gentrification and demographic sorting (Couture, Gaubert, Handbury, and Hurst 2022; Hoelzlein 2020; Balboni, Bryan, Morten, and Siddiqi 2021; Almagro and Domínguez-Iino 2021). Another strand of this literature analyses the determinants of consumption location choices and the

role played by spatial and social frictions (Couture 2016; Davis, Dingel, Monras, and Morales 2019; Su 2022; Hausman, Samuels, Cohen, and Sasson 2021; Tan and Lee 2021). Relative to this research, we develop a quantitative theoretical model in which consumption location decisions are part of a wider travel itinerary choice, which involves choosing the optimal set of locations to visit and the sequence in which to visit these locations.

Third, our work contributes to research on the role of consumer mobility in influencing firm location choices through shopping externalities. Theoretical studies have considered models with stylized geographies, in which these shopping externalities provide a mechanism for agglomeration, including Eaton and Lipsey (1982); Claycombe (1991); and Ushchev, Sloev, and Thisse (2015). Empirical research has provided evidence of shopping externalities using quasi-experimental variation from the bankruptcies of large retail chains and online shopping, including Shoag and Veuger (2018); Benmelech, Bergman, Milanez, and Mukharlyamov (2019); Koster, Pasidis, and van Ommeren (2019); and Relihan (2022).² Our theoretical model of travel itineraries incorporates these shopping externalities, while remaining sufficiently tractable as to be taken directly to observed data on cities. We provide evidence that these shopping externalities are quantitatively relevant for understanding the impact of the Covid-19 pandemic and for the evaluation of transport infrastructure investments.

Fourth, our use of smartphone data is related to the growing literature that use spatially-granular, high-frequency, large-scale data to measure spatial mobility patterns. Researchers have used credit card data (Agarwal, Jensen, and Monte 2020; Dolfen, Einav, Klenow, Klopach, Levin, Levin, and Best 2022; Allen, Fuchs, Ganapati, Graziano, Madera, and Montoriol-Garriga 2021); ride-sharing data (Gorback 2021; Bucholz, Doval, Kastl, Matejka, and Salz 2021); car navigation data (Hausman, Samuels, Cohen, and Sasson 2021) and cellphone data (Couture, Dingel, Green, and Handbury 2019; Athey, Ferguson, Gentzkow, and Schmidt 2021; Kreindler and Miyauchi 2022; Gupta, Kontokosta, and Van Nieuwerburgh 2022; Büchel, Ehrlich, Puga, and Viladecans 2020; Atkin, Chen, and Popov 2022). Our work also connects with related research that has used travel diary surveys to measure commuting and non-commuting trips within urban areas, including Couture, Duranton, and Turner (2018) and Zárte (2021). An advantage of our smartphone data is that the data is collected at a high level of spatial and temporal resolution of minutes, hours and days, which allows us to examine the impact in real time of events such as the Covid-19 pandemic.³

Fifth, to overcome the high-dimensionality of the choice set of travel itineraries, we make use of importance sampling methods following Kloek and van Dijk (1978) and Ackerberg

²Brancaccio, Kalouptsi, and Papageorgiou (2020) develop a model of the transportation sector, in which there are externalities across locations because of search frictions for ships: as one destination becomes more attractive for ships, this increases the supply of ships for locations that are nearby or along the way.

³For example, in the United States, the National Household Travel Survey is conducted every 6 years.

(2009). While we apply our method to travel itinerary choice, it is broadly applicable to discrete choice models with high-dimensional state spaces, such as bundled products and store choices (Thomassen, Smith, Seiler, and Schiraldi 2017), establishment location choices (Jia 2008; Oberfield, Rossi-Hansberg, Sarte, and Trachter 2020), import choice problems (Antràs, Fort, and Tintelnot 2014), or route-choice problems for trucking and shipping (Allen and Arkolakis 2021; Allen, Atkin, Cantillo, and Hernandez 2021; Braccaccio, Kalouptsi, and Papageorgiou 2020). In each of these settings, existing research either makes additional assumptions, such as global supermodularity or submodularity (Jia 2008; Antràs, Fort, and Tintelnot 2014; Arkolakis, Eckert, and Shi 2022); relies on analytically-tractable functional form assumptions or limiting parameter values (Allen and Arkolakis 2021 and Oberfield, Rossi-Hansberg, Sarte, and Trachter 2020); or imposes restrictions on the dimensionality of the choice set (Thomassen, Smith, Seiler, and Schiraldi 2017). In comparison, our method is typically more computationally intensive, because it involves Monte Carlo simulation.⁴ But an advantage is that we are not required to make these additional assumptions, which is relevant for our application, where for example global supermodularity or submodularity need not be satisfied.

Finally, our analysis of spatial mobility during the Covid-19 pandemic relates to the literature that has examined the impact of this pandemic on spatial interactions, including Antràs, Redding, and Rossi-Hansberg (2020); Glaeser, Gorbach, and Redding (2020); Alvarez, Argente, and Lippi (2021); Argente, Hsieh, and Lee (2022); Fajgelbaum, Khandelwal, Kim, Mantovani, and Schaal (2021); Delventhal and Parkhomenko (2022); Giannone, Paixao, and Pang (2022); and Couture, Dingel, Green, Handbury, and Williams (2022). In particular, Althoff, Eckert, Ganapati, and Walsh (2022) provides evidence that the shift to remote work reduced the local demand for non-traded services in central cities. We show that our theoretical model of travel itineraries, and the consumption externalities to which it gives rise, are quantitatively successful in rationalizing these observed declines in foot traffic in downtown areas.

3 Data Description

In this section, we introduce our smartphone data and other data sources.⁵ In Subsection 3.1, we explain how we use our smartphone data to identify home location, work location, commuting trips and non-commuting trips. In Subsection 3.2, we discuss the other administrative data that we combine with our smartphone data. In Subsection 3.3, we report validation checks

⁴While we use Monte Carlo simulation based on importance sampling to overcome the high-dimensionality of travel itinerary decisions, Dingel and Tintelnot (2020) uses Monte Carlo simulation to address small sample variation (granularity) with spatially-disaggregated data. As part of our importance sampling approach, we allow in our Monte Carlos for this small sample variation (granularity).

⁵See Online Appendix A for further details on our smartphone data.

on our smartphone commuting data using official census data on employment by residence, employment by workplace and bilateral commuting flows.

3.1 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 up to every 5 minutes whenever the device is turned on (regardless of whether the application is being used).⁶ 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).

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 centroid of these points.⁷ We have data on the sequence of stays of anonymized users with the necessary level of spatial aggregation 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 the “home” location. In particular, if the second most frequent location is within 600 meters of the “home” locations, we define the “work” location as the third most frequent location. To abstract from noise in geo-coordinate assignment, all stays within 500 meters of the home location are aggregated with the home location. Similarly, all stays within 500 meters of the work location are aggregated with the work location. We assign “Work” location as missing if the user appears in that location for less than 5 days per month, which applies for about 30 percent of users in our baseline sample during April 2019. These users primarily

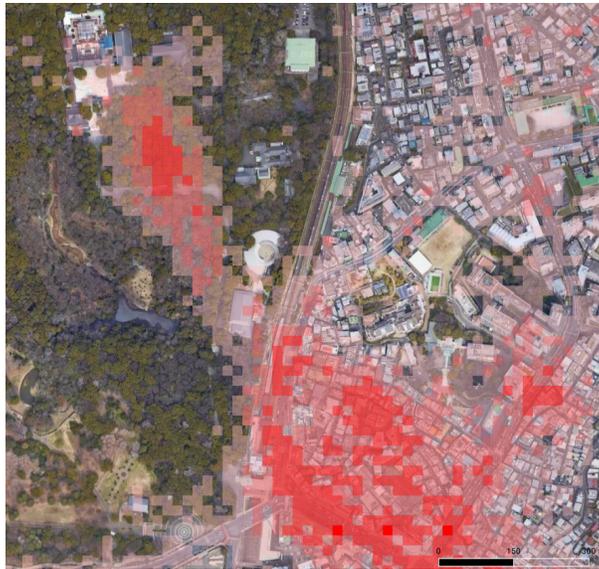
⁶The mapping application does not send location data points if the smartphone does not sense movement, in which case it is likely that the user has not moved from the last reported location. For this reason, the data points are less frequent than 5 minutes intervals in practice.

⁷See 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.

include those with limited number of data observations due to infrequent smartphone use, and also include irregular workers with unstable job locations and those who work at home.⁸ In Subsection 3.3 below, we report validation checks on our classification of home and work locations using commuting data from the population census. Stays which are neither assigned as home or work are classified as “other.”

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). When we analyze the change of spatial mobility during the Covid-19 pandemic, we use data from February to May 2020 (at the beginning of the Covid-19 pandemic) along with the same months in 2019 (before the Covid-19 pandemic). To abstract from overnight trips, we focus on the sample of user-day observations for which the first and last stay of the day is the user’s home location.

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 grid cell of the size of approximately 25×25 meters. The darkness of the color represents the number of stays in each grid cell between December 2017 and February 2018. The building towards the top-left surrounded by trees is the main building of the shrine. The stays are concentrated tightly along the path that runs from the road to the main building of the shrine, consistent with them accurately capturing patterns of movement within the city.

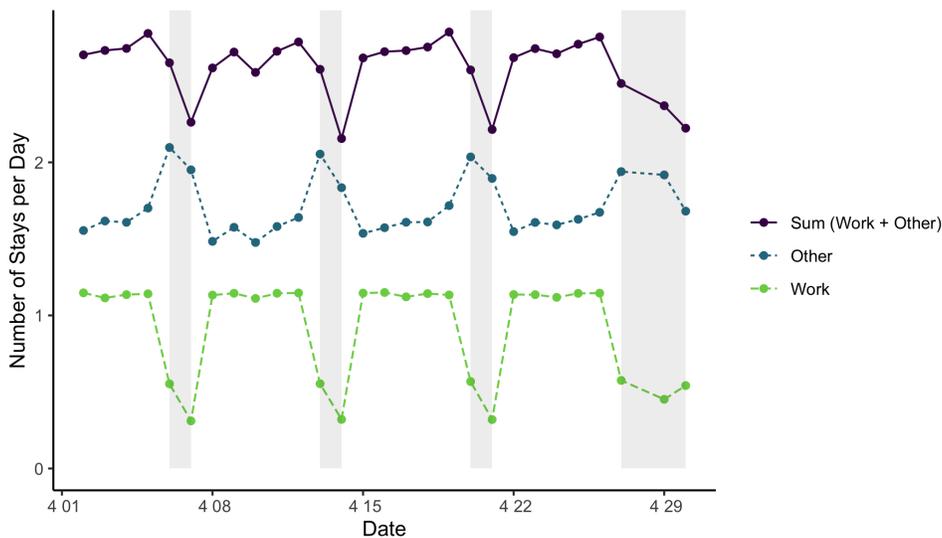
As an illustration of our data, Figure 1 displays the “stays” recorded in our data for a *Meiji*

⁸In Online Appendix A.4, we show that the devices with missing “work” locations have significantly fewer number of active days (even at home locations), indicating that these devices are less likely to be actively used. We also show that the probability of assigning missing “work” locations is uncorrelated with the observable characteristics of the municipality of residence.

Shrine in the Shibuya municipality of Tokyo over the period from December 2017 to February 2018. Each red-shaded rectangle corresponds to a 25 meter by 25 meter grid cell. The darker the red shading of the grid cell, the larger the number of stays in that grid cell. We have overlaid these grid cells on a satellite photograph. In this photograph, the building towards the top-left of the image surrounded by trees corresponds to the main building of the Meiji shrine. Several features of our data are apparent from this image. First, we observe movement within the city at a high level of spatial resolution. Second, we find a sharp discontinuity in the density of stays at the road that separates the wooded area surrounding the shrine to the left from the developed area to the right, suggesting that the stays accurately capture the density of movement. Third, in the middle of this wooded area, the stays are concentrated tightly along the path that runs from the road to the main building of the shrine, again confirming the ability of our data to capture the main pathways of movement through the city.

As an illustration of the temporal dimension of our data, Figure 2 displays the average number of work and non-work stays by day of the week from 1-30 April 2019. For each day of the week, non-commuting trips are more frequent than commuting trips. At weekends, we find the expected pattern that non-commuting trips increase and commuting trips decrease, confirming the ability of our smartphone data to capture this temporal variation.

Figure 2: Work and Other Stays by Day and Hour



Note: 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 shaded areas indicate weekends and holidays in Japan. In Online Appendix A.3, we provide a further validation check by showing that our data captures the expected temporal pattern of home, work and other stays across the hours of the day for both weekdays and weekends. See the main text above for the definitions of home, work and other stays.

3.2 Other Data Sources

We combine our smartphone data with a number of other complementary data sources. These complementary data sources are used for the purpose of validating our smartphone data or for calibrating our model.

Spatial units: We focus on the Tokyo Metropolitan Area, which comprises the four prefectures of Tokyo, Chiba, Kanagawa and Saitama. Together these prefectures cover an area of about 13,500 square kilometers and include around 36 million residents. These prefectures are in turn further disaggregated into 242 municipalities (excluding islands).

Population Census: We measure residential population, employment by workplace and bilateral commuting flows using the 2015 population census, which is conducted by the Statistics Bureau, Ministry of Internal Affairs and Communications every five years. Residential population and total employment are available at the finest level of spatial disaggregation of 250-meter grid cells. Bilateral commuting flows are reported between pairs of municipalities.

Economic Census: We use data from the 2016 Economic Census on total employment and the number of establishments by one-digit industry for each 500-meter grid cell in the Tokyo Metropolitan Area, the finest level of disaggregation from publicly available data. We also use 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. 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. We take a simple average of these values to construct the average land prices per unit of land at the Municipality level.

Travel Time Data: We measure travel time by public transportation using the web-based route choice service, *Eki-spert* API.⁹ Eki-spert API provides the minimum travel times between any pairs of coordinates using public transport, including suburban rail, subway, and bus, and walking. We use the extracted travel time data from October 2, 2020 (weekday timetable).

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.

3.3 Validation of Smartphone 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 official census

⁹See <https://roote.ekispert.net/en> for details.

data that are available at the municipality level. In the left panel of Figure 3, we display the log density of residents in each municipality in our smartphone data against log population density 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 density of workers in each Tokyo municipality in our smartphone data against log employment density by workplace in the census data. Again, 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.

Figure 3: Representativeness of Smartphone Users



Note: Each dot is 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 divided by its geographic area, and the horizontal axis is the log of the number of residents in that municipality from the Population Census in 2011 divided by its geographic area. In the right panel, the vertical axis is the log of the number of smartphone users with a work location in the municipality divided by its geographic area, and the horizontal axis is the log of employment by workplace in that municipality from the Population Census in 2011 divided by its geographic area. The definitions of home and work in the smartphone data are discussed in the text of Subsection 3.1 above.

In Online Appendix A.1, we provide further evidence on the representativeness of our smartphone data by comparing the coverage by residence characteristics (income, age and distance to city center) and workplace characteristics (employment by industry and distance to city center). In Online Appendix A.2, we show that the commuting flows from our smartphone data have a similar rate of spatial decay with geographic distance as in the official census data.

In Online Appendix [A.3](#), we show that home stays tend to occur during nighttime (outside 6am-9pm) and both work and other stays occur during the daytime (from 6am-9pm), providing further validation of our home and work classification using our smartphone data.

4 Patterns of Spatial Mobility

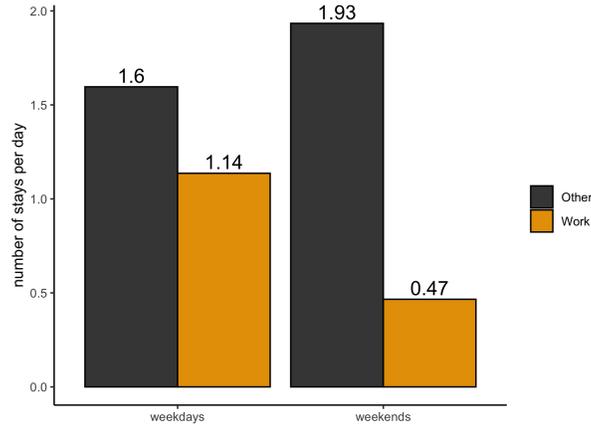
In this section, we establish a number of stylized facts about the patterns of spatial mobility that guide our theoretical model below. First, non-commuting trips are more frequent than commuting trips, so that focusing solely on commuting trips understates travel within urban areas. Second, non-commuting trips are closely-related to the availability of non-traded services. Third, non-commuting trips exhibit different spatial patterns from commuting trips, such that they are not well approximated by commuting trips. Fourth, trip chains are a pervasive feature of the data, where individuals stop at multiple destinations along a single journey starting from and ending at home. Finally, we show that there is a decline in the frequency and the travel length of commuting and non-commuting trips during the Covid-19 pandemic in 2020, which results in a disproportionate reduction in foot traffic in the downtown area.

Fact 1. Non-commuting trips are pervasive. In Figure 4, 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 lunch elsewhere). Similarly, the average number of work stays can be greater than zero on the weekend, because some workers can be employed during the weekend (e.g., in restaurants and stores). As 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.¹⁰

Fact 2. Non-commuting trips are closely related to nontradable service availability. Figure 5 shows that there is a tight relationship between consumer foot traffic and the density of nontradable service-sector establishments. The figure plots the logarithm of the number of non-commuting stays for each 1 kilometer by 1 kilometer grid cell (on the y-axis) against the number of establishments in nontradable service sectors in the cell (on the x-axis). The

¹⁰In Online Appendix [A.5](#), we show that this pattern of more frequent non-commuting stays than commuting stays holds in separate travel survey data in Tokyo Metropolitan Area, which are available for weekdays only for more aggregated spatial units after a substantial time lag.

Figure 4: 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 home stays) for our baseline sample users in the metropolitan area of Tokyo in April 2019. See Section 3 above for the definitions of home, work and other stays.

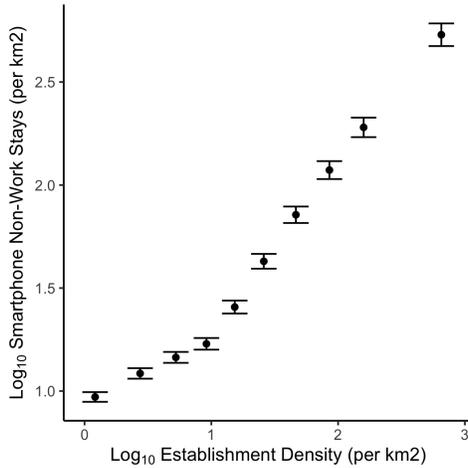
nontradable service sector is defined to include “Finance, Real Estate, Communication, and Professional”, “Wholesale and Retail” (category G, J, K, L of the Japanese Standard Industrial Classification (JSIC)), “Accommodations, Eating, Drinking” (category I), “Medical and Health Care” (category P), and “Other Services” (category Q). We find a strong, positive and increasing relationship that holds throughout the distribution of the two variables.¹¹

Fact 3. Non-commuting trips are closer to home. We now show that non-commuting trips exhibit different spatial patterns from commuting trips. In Figure 6, we display the distribution of distances from home locations to work locations and from home locations to other stays for our baseline sample of users in the Tokyo Metropolitan Area in 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, with an even larger difference in distances travelled on the weekend. 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 on weekends.

Fact 4. Non-commuting trips frequently occur as a part of trip chains. Figure 7 shows that non-commuting trips frequently occur along trip chains, in which people stop at multiple destinations as part of a single journey starting from and ending at home. We display both the number of stays outside home and the number of trip chains, where a trip chain involves more than one stay outside the home location. On weekdays (weekends), we find that users

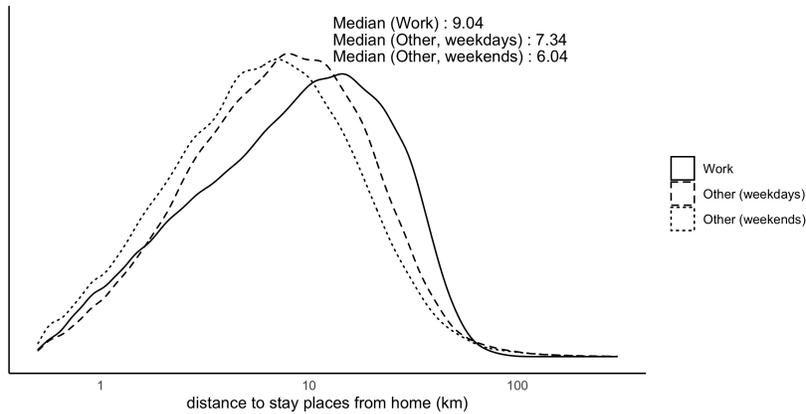
¹¹In Online Appendix A.6, we use spatially-disaggregated economic census data on employment by sector to distinguish between different types of non-commuting stays.

Figure 5: Non-Commuting Stays and Retail Establishment Density



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

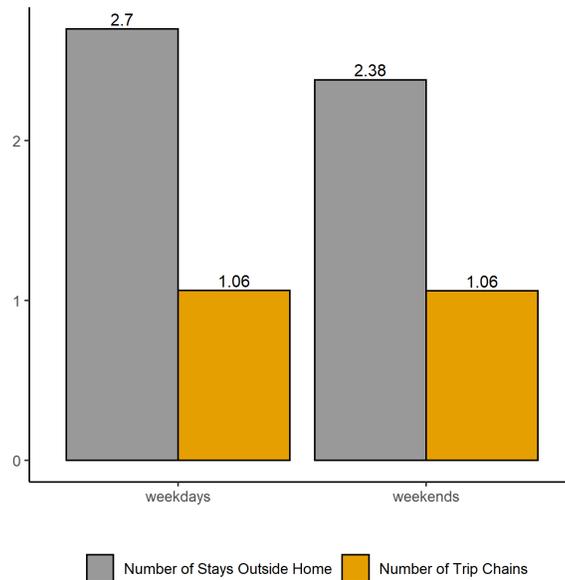
Figure 6: Distances of Commuting and Non-Commuting Trips



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

make 2.7 (2.38) stays outside the home per day, and undertake 1.06 (1.06) trip chains per day. This pattern of results suggests that the typical user’s travel behavior involves one trip chain per day, stopping at intermediate destinations along the way. In Appendix A.7, we further validate these findings by showing that trip chains typically include a user’s workplace during weekdays, but typically exclude the user’s workplace on weekends.

Figure 7: Frequency of Trip Chains



Note: Number of stays per day and number of trip chains per day per user, disaggregated by weekdays and weekends. Trip chains are defined as a journey starting from and ending at home that includes more than one stay outside of the home location. See Section 3 above for the definition of home, work and other stays. Results for our baseline sample of users in the Tokyo Metropolitan Area in April 2019.

Fact 5. Both non-commuting and commuting trips became less frequent and shorter during the Covid-19 pandemic. The outbreak of the Covid-19 pandemic starting in early 2020 had a dramatic impact on spatial mobility in cities around the world. In an effort to prevent a surge of cases, national and city governments in many places imposed lockdowns and mobility restrictions. Even in places where there were no explicit mobility restrictions, people chose on their own initiative to make fewer trips to reduce the risk of infection. In the remainder of this subsection, we provide evidence on how patterns of commuting and non-commuting trips in Tokyo changed during the Covid-19 pandemic.

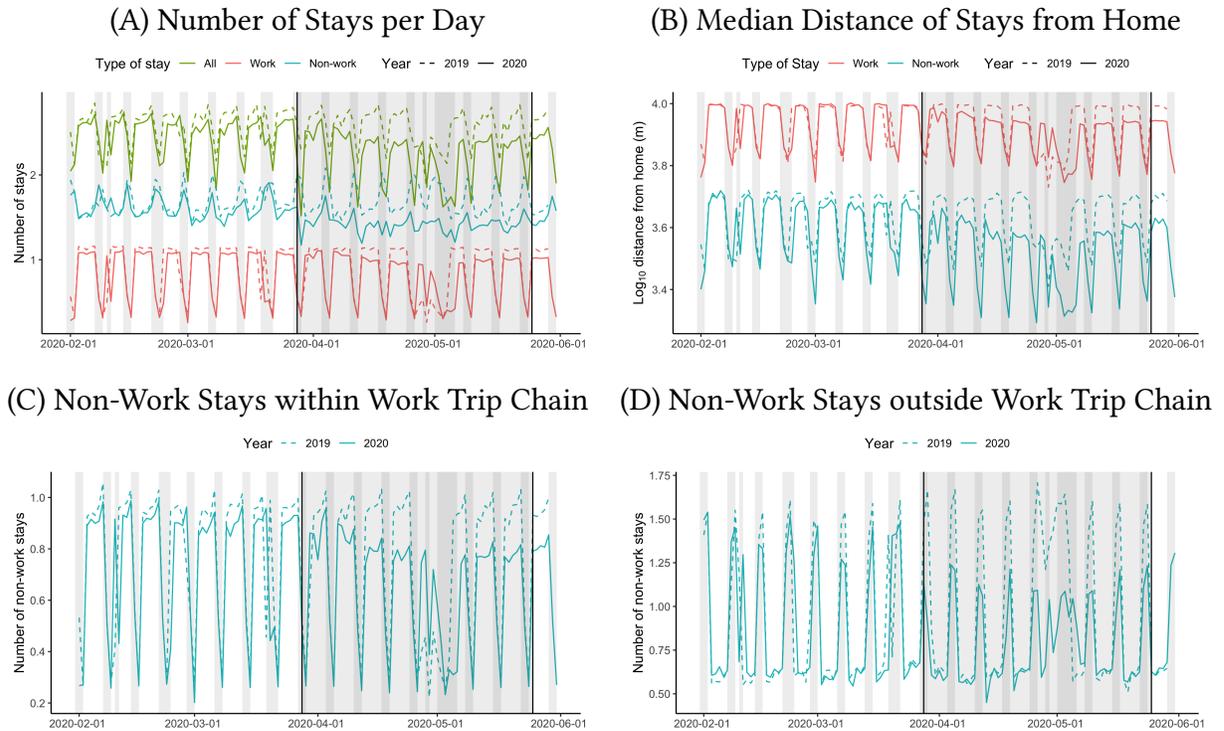
Compared to other developed countries, Japan experienced relatively low infection rates. Partly because of this, Japanese cities did not impose a strict form of lockdown. However, governments issued multiple waves of “emergency orders” that asked residents to stay home unless travel was “absolutely necessary.”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

Figure 8 shows the frequency of trips and length of travel from February-May 2020 (including the period of the emergency order) compared to the same months in 2019 (before the Covid-19 pandemic). Panel A shows the number of stays outside the home location per day;

Panel B gives the median distance of these stays from the home location; Panel C reports the number of non-work stays away from home that occur within trip chains including workplace; Panel D displays the number of non-work stays away from home that take place outside trip chains including workplace. In all figures, the dashed lines correspond to 2019, and the solid lines represent 2020. In Panels A and B, the red lines denote work stays, whereas the blue lines indicate non-work stays. In Panel A, the green lines represent all stays away from home (work plus non-work).

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

Figure 8: Commuting and Non-Commuting Trips from February-May in 2020 (including the period of the Emergency Order) and 2019 (before the Covid-19 Pandemic)

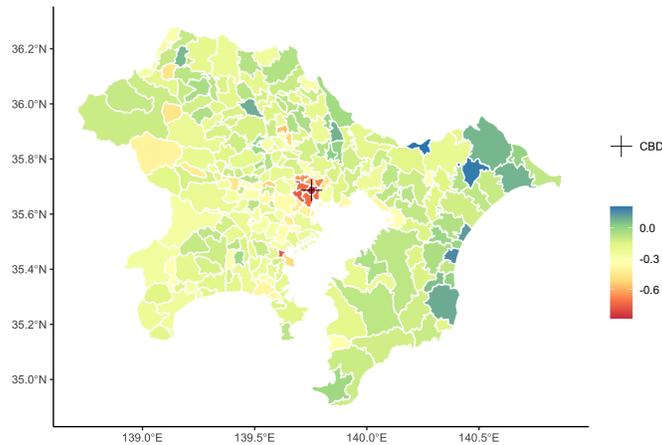


Note: Each figure shows the average number of stays away from home for different types of stays (Panels A, C and D) and the median distance from home (Panel B) for each day from February-May in 2020 (including the period of the emergency order) and in 2019 (before the Covid-19 pandemic). For the 2019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

In Figure 9, we show the change in the geography of spatial mobility in Tokyo between

April 2019 (before the Covid-19 pandemic) and April 2020 (during the period of the emergency order). We display a map of the log changes in the total number of other stays (i.e., non-commuting trips to locations that are neither the home nor the workplace) in each municipality in the Tokyo Metropolitan Area over this twelve-month period. We find the largest declines in non-commuting trips to downtown areas, consistent with the decline in commuting during the period of the emergency order reducing the demand for non-traded services in these downtown areas. In contrast, we find a much smaller decline in non-commuting trips in outlying suburbs, with some areas even experiencing an increase in non-commuting trips. This pattern of results is consistent with people travelling to consume non-traded services closer to home during the period of the emergency order.

Figure 9: Log Changes in Other Stays in each Tokyo Municipality Between April 2019 (before the Covid-19 Pandemic) and April 2020 (during the Emergency Order)



Note: This figure plots log changes in the number of other stays (i.e., non-commuting trips to a location that is neither the residence nor the workplace of a user) in each municipality in the Tokyo Metropolitan Area from April 2019 (before the Covid-19 pandemic) to April 2020 (during the period of the emergency order). CBD (Central Business District) corresponds to the centroid of Chiyoda Ward in Tokyo Prefecture. Darker red colors indicate greater reductions in other stays; darker blue colors indicate greater increases in other stays. See Section 3 above for the definition of home, work and other stays.

Taking the results of this section as a whole, we find that non-commuting trips are frequent, are closely related to consumption, exhibit different spatial patterns from commuting trips, often occur as part of trip chains, and are affected by the Covid-19 pandemic. Each of these five features of our smartphone data guides our theoretical model of travel itineraries that we develop in the next section.

5 Travel Itinerary Decisions

In this section, we introduce and estimate our model of agents' travel itinerary decisions. In Subsection 5.1, we develop our theoretical specification of travel itineraries given an

agent’s choice of residence. In Subsection 5.2, we develop a method to simulate these travel itinerary decisions based on importance sampling, which allows us to overcome the high-dimensionality of the choice set. In Subsection 5.3, we estimate the parameters that govern agents’ travel itinerary decisions. In Subsection 5.4, we show that our estimated model provides a good fit to observed patterns of spatial mobility and the distribution of sales of non-traded services across locations. In Section 6, we embed this specification of travel itinerary decisions in a quantitative urban model of residence decisions. In Sections 7 and 8, we provide evidence on the impact of the consumption externalities introduced by travel itineraries on the spatial distribution of economic activity.

5.1 Theoretical Framework

We consider a city that consists of a set of locations: $N \equiv \{1, \dots, n\}$. We characterize the travel decision of an agent residing in location $h \in N$. In this section, we assume that agents’ residential location is given, while we endogenize this choice of residential location in Section 6 below. We assume that the agent’s choices are nested in two steps. First, given her residential location, she decides her workplace location $j \in N$. Second, she decides the sequence of destinations to visit to consume nontradable services each day, separately for workdays (when she has to go to j) and non-workdays (when she does not have to go to j). We begin by characterizing the choice of non-commuting trips for a given workplace, before turning to the choice of workplace.

Choice of non-commuting trips for a given workplace. Each day, the agent decides the sequence of trip destinations (i.e., the number of locations to visit and in what order) for nontradable services (e.g., bars, restaurants, grocery shopping). More specifically, she chooses the subset of locations to visit for nontradable service consumption, which we denote by $C \subset N$. She also chooses the itinerary I as the ordered set of locations to visit, where I is the ordered set of $C \cup \{j\}$ for workdays and C for non-workdays. For expositional convenience, we sometimes use $j = \emptyset$ to refer to non-workdays, and we denote $C(I)$ as the set of locations to consume nontradable services included in itinerary I . To streamline notation, we do not index C and I by workday and non-workday, but we take it as understood that we allow a different set of locations to consume nontradable services (C), and a different sequence in which to visit them (I), for workdays and non-workdays.

Nontradable services are assumed to be horizontally differentiated, with a constant elasticity of substitution of $\sigma > 1$. Nontradable service producers compete under conditions of monopolistic competition. In each location $n \in N$, there is an endogenous measure of producers (M_n). For simplicity, we assume that all producers in location n have the same production

technology, with a constant marginal cost. Given these assumptions on preferences, market structure and production technology, all producers in location n charge an identical price p_n . We embed agents' travel itinerary decisions in general equilibrium in Section 6 below, where we solve explicitly for M_n and p_n using profit maximization and free entry. We assume that agents can only consume the nontradable services in location n if they physically visit that location such that $n \in C$. Under these assumptions, agent ω chooses the itinerary I_ω that maximizes the following nontradable services component of her indirect utility:

$$I_\omega = \max_{I \in \mathcal{I}_{hj}} \left(\sum_{n \in C(I)} P_n^{1-\sigma} \right)^{-\frac{1}{1-\sigma}} \tau_{I|hj}^{-1} \epsilon_{\omega I},$$

where \mathcal{I}_{hj} is the set of possible itineraries given residence h and workplace j (recall that j has to be included in I on workdays but not on non-workdays). We use $P_n = M_n^{1/(1-\sigma)} p_n$ to denote the price index of the bundle of varieties of nontradable services provided in location n , such that $\left(\sum_{n \in C(I)} P_n^{1-\sigma} \right)^{-\frac{1}{1-\sigma}}$ is the price index for the nontradable services provided in all locations included in itinerary I . We use $\tau_{I|hj} \geq 1$ to denote the travel cost of following itinerary I given residence h and workplace j . Importantly, we accommodate the possibility of travel cost savings from trip chains, such that visiting multiple locations $1, 2 \in N$ as part of a single journey can be less costly than visiting each location through two separate journeys: i.e., $\tau_{\{1\}|hj} \tau_{\{2\}|hj} > \tau_{\{1,2\}|hj}$. Finally, $\epsilon_{\omega I}$ is an idiosyncratic preference draw for individual itineraries, which is specific to individual ω , and captures all the idiosyncratic reasons why individuals may choose a particular subset and sequence of destinations.

We assume that this idiosyncratic preference draw for each individual ω and itinerary I ($\epsilon_{\omega I}$) is drawn from an independent Fréchet distribution, with a dispersion parameter θ that governs the heterogeneity of idiosyncratic preferences: $\epsilon_{\omega I} \sim \exp(-\varsigma \epsilon_{\omega I}^{-\theta})$, with $\varsigma > 0$.¹² Under this assumption, the probability that an agent with residence h and workplace j chooses itinerary I takes the following form:

$$\Lambda_{I|hj} = \frac{\left[\left(\sum_{n \in C(I)} P_n^{1-\sigma} \right)^{-\frac{1}{1-\sigma}} \tau_{I|hj}^{-1} \right]^\theta}{\sum_{\ell \in \mathcal{I}_{hj}} \left[\left(\sum_{n \in C(\ell)} P_n^{1-\sigma} \right)^{-\frac{1}{1-\sigma}} \tau_{\ell|hj}^{-1} \right]^\theta}. \quad (1)$$

We can also derive a closed-form solution for a measure of “consumption access,” which corresponds to the inverse of the expected price index net of the idiosyncratic shocks for a

¹²We introduce the shock to the itinerary I , instead of an individual location $n \in C(I)$, to obtain an analytical characterization of the itinerary choice probability $\Lambda_{I|hj}$ and consumption access \mathbb{A}_{hj} . This specification is similar in spirit to the shipment route choice problem in [Allen and Arkolakis \(2021\)](#), but distinct in that agents choose both an endogenous number of locations to visit and an endogenous route between them each day.

given residence h and workplace j :

$$\mathbb{A}_{hj} = \varrho \left[\sum_{\ell \in \mathcal{I}_{hj}} \left(\sum_{n \in C(\ell)} P_n^{1-\sigma} \right)^{-\frac{\theta}{1-\sigma}} (\tau_{\ell|hj})^{-\theta} \right]^{-\frac{1}{\theta}}, \quad (2)$$

where $\varrho = \Gamma\left(\frac{\theta-1}{\theta}\right)$ and $\Gamma(\cdot)$ is the Gamma function.

This travel-itinerary specification for consuming nontradable services considerably generalizes conventional models of consumption within cities. In [Ahlfeldt, Redding, Sturm, and Wolf \(2015\)](#), agents consume a single traded good that is costlessly traded across locations. In [Allen, Arkolakis, and Li \(2017\)](#), [Monte, Redding, and Rossi-Hansberg \(2018\)](#), [Couture, Gaubert, Handbury, and Hurst \(2022\)](#) and [Hoelzlein \(2020\)](#), agents consume varieties from all locations at their residence, subject to bilateral travel costs. In [Couture \(2016\)](#), [Davis, Dingel, Monras, and Morales \(2019\)](#), [Hausman, Samuels, Cohen, and Sasson \(2021\)](#) and [Su \(2022\)](#), agents can visit one location at a time from either home or work. In [Almagro and Domínguez-Iino \(2021\)](#), agents consume a measure of nontradable services in a single residential location.

In contrast, our specification allows agents to visit an endogenous subset of locations along an endogenous route to consume nontradable services in those locations. As a result, we are able to capture the geography of spatial mobility observed in our smartphone data, including trip chains. A key implication of these trip chains is the existence of consumption externalities across locations: as one location becomes more attractive destination for consuming nontradable services, this increases the attractiveness of other locations that are nearby or along the way. We show below that these consumption externalities are quantitatively important for the counterfactual effects of changes in mobility costs, whether through natural experiments (e.g., the Covid-19 pandemic) and/or policy interventions (e.g., transport improvements).

Workplace decision. Anticipating these travel itinerary decisions for nontradable services consumption, each agent chooses her workplace and sector of employment. We denote sectors by $k \in \{T, S\}$, where T denotes the tradable sector, and S denotes the nontradable service sector. The workplace component of indirect utility for an agent ω residing in location h and working in location j and sector k depends on her income, expected consumption access, commuting costs, and an idiosyncratic preference draw:

$$\begin{aligned} U_{jk\omega|h} &= w_{jk} \tilde{\mathbb{A}}_{hj}^{\alpha^S} (\tau_{hj}^W)^{-1} \epsilon_{jk\omega|h}^W, \\ \tilde{\mathbb{A}}_{hj}^{\alpha^S} &= \xi \mathbb{A}_{hj}^{\alpha^S} + (1 - \xi) \mathbb{A}_{h\emptyset}^{\alpha^S}, \end{aligned} \quad (3)$$

where w_{jk} is the wage in location j and sector k ; $\tau_{hj}^W \geq 1$ is the iceberg cost of commuting from residence h to workplace j ; $\epsilon_{jk\omega|h}^W$ is an idiosyncratic preference shock draw for workplace j and sector k ; α^S is the expenditure share for nontradable services; $\tilde{\mathbb{A}}_{hj}$ is expected

consumption access, which is a weighted average of consumption access on workdays ($\mathbb{A}_{h,j}$) and non-workdays ($\mathbb{A}_{h,\emptyset}$), as characterized above; ξ is the fraction of workdays during the week;¹³ and we embed this workplace component of utility ($U_{jk\omega|h}$) in overall indirect utility below, when we characterize an agent's choice of residence below.

We assume that the idiosyncratic preference draw for each individual ω , workplace j and sector k is drawn from an independent Fréchet distribution with dispersion parameter ϕ : $\epsilon_{jk\omega|h}^W \sim \exp\left(-\varsigma^W \epsilon_{jk\omega|h}^W - \phi\right)$, with $\varsigma^W > 0$. Under this assumption, the probability that an agent with residential location h chooses workplace j and employment sector k is:

$$\Omega_{jk|h} = \frac{\left(w_{jk} \tilde{\mathbb{A}}_{h,j}^{\alpha^S} (\tau_{h,j}^W)^{-1}\right)^\phi}{\sum_{j'} \sum_{k' \in \{T,S\}} \left(w_{j'k'} \tilde{\mathbb{A}}_{h,j'}^{\alpha^S} (\tau_{h,j'}^W)^{-1}\right)^\phi}. \quad (4)$$

This probability shares some similar features with a conventional commuting gravity equation: bilateral commuting flows depend on residence fixed effects, workplace fixed effects, and bilateral travel times between an agent's residence and workplace ($\tau_{h,j}^W$). A key difference is that bilateral commuting flows also depend on bilateral consumption access ($\tilde{\mathbb{A}}_{h,j}$), which introduces a source of mis-specification into conventional commuting gravity equations. Consumption access varies bilaterally, because agents can choose to travel to consume non-traded services from either their workplace or their residence, or along the route from home to work. Our specification therefore captures the intuitive idea that the desirability of a workplace depends not only on the wage that it offers, but also on the access to local nontradable services that it provides (e.g. theaters and restaurants), both in the area surrounding the workplace itself, and along the route from home to work.

5.2 Importance Sampling Method

The main challenge for the estimation and simulation of our model is the dimensionality of the travel itinerary decision ($\mathcal{I}_{h,j}$). To see this, note that the denominator of the itinerary choice probability ($\Lambda_{I|h,j}$) in equation (1) involves the summation of all elements in $\mathcal{I}_{h,j}$. Even with a moderate number of locations N , the dimension of $\Lambda_{I|h,j}$ is extremely large, since it involves all combinations of the sequence of locations that the agent can potentially choose from.¹⁴ Therefore, computing the exact probability of $\Lambda_{I|h,j}$ is impractical.

To overcome this high-dimensionality of the choice set, we develop a method to simulate $\Lambda_{I|h,j}$ based on the importance sampling method (Kloek and van Dijk 1978). The basic idea

¹³We assume that agents allocate the same share of expenditure on non-traded services on workdays and weekends, such that $\tilde{\mathbb{A}}_{h,j}^{\alpha^S}$ is a weighted average of $\mathbb{A}_{h,j}^{\alpha^S}$ (workdays) and $\mathbb{A}_{h,\emptyset}^{\alpha^S}$ (non-workdays).

¹⁴If we allow agents to visit up to K locations, the set of possible itineraries that agents can choose is of dimension $\prod_{i=0, \dots, K-1} (|N| - i)$.

is to obtain a Monte-Carlo sample from an auxiliary distribution, and adjust the sampling rate based on the likelihood ratio between the true distribution and the auxiliary distribution. Formally, the procedure is described in Algorithm 1.

Algorithm 1 (Importance Sampling) Denote the auxiliary probability distribution of itineraries by agents with residential location h and workplace j by $F_{hj}(I)$, defined over \mathcal{I}_{hj} . The simulated probability $\Lambda_{I|hj}^*$ is defined as follows:

1. Draw R itineraries $\{I_r\}$ from auxiliary distribution $F_{hj}(\cdot)$. Denote the empirical distribution of the simulated draws by $\mathcal{E}_{I|hj} = \frac{1}{R} \sum_{1, \dots, R} 1[I = I_r]$.
2. Weight each draw by the likelihood ratio between $F_{hj}(I)$ and $\Lambda_{I|hj}$ to obtain the simulated probability distribution $\Lambda_{I|hj}^*$:

$$\Lambda_{I|hj}^* = \frac{\mathcal{E}_{I|hj} \Lambda_{I|hj} / F_{hj}(I)}{\sum_{\ell \in \mathcal{I}_{hj}^R} \mathcal{E}_{\ell|hj} \Lambda_{\ell|hj} / F_{hj}(\ell)} = \frac{\mathcal{E}_{I|hj} \left[\left(\sum_{n \in C(I)} P_n^{1-\sigma} \right)^{-\frac{1}{1-\sigma}} \tau_{I|hj}^{-1} \right]^\theta / F_{hj}(I)}{\sum_{\ell \in \mathcal{I}_{hj}^R} \mathcal{E}_{\ell|hj} \left[\left(\sum_{n \in C(\ell)} P_n^{1-\sigma} \right)^{-\frac{1}{1-\sigma}} \tau_{\ell|hj}^{-1} \right]^\theta / F_{hj}(\ell)}, \quad (5)$$

where \mathcal{I}_{hj}^R is the subset of \mathcal{I}_{hj} that are sampled in Step 1, i.e., $\mathcal{E}_{I|hj} > 0$.

The simulated probability distribution in equation (5) has an intuitive interpretation. Compared to the actual itinerary choice probability $\Lambda_{I|hj}$, the Monte-Carlo draws from $F_{hj}(\cdot)$ under-sample itineraries with higher likelihood ratios $\Lambda_{I|hj} / F_{hj}(I)$. Therefore, re-weighting each draw by this likelihood ratio yields a consistent estimator of $\Lambda_{I|hj}$. This likelihood ratio $\Lambda_{I|hj} / F_{hj}(I)$ is proportional to $\left[\left(\sum_{n \in C(I)} P_n^{1-\sigma} \right)^{-\frac{1}{1-\sigma}} \tau_{I|hj}^{-1} \right]^\theta / F_{hj}(I)$, and omits the normalizing constant in the denominator of $\Lambda_{I|hj}$ in equation (1), which involves the summation over all possible itineraries. By abstracting from this normalizing constant, we avoid having to directly compute the denominator of $\Lambda_{I|hj}$ that is subject to the curse of dimensionality.

From equation (5), we can also see that $\Lambda_{I|hj}^* \rightarrow \Lambda_{I|hj}$ as $R \rightarrow \infty$ as long as the support of $F_{hj}(\cdot)$ has common support with $\Lambda_{I|hj}$.¹⁵ Therefore, an advantage of this algorithm is that the choice of the auxiliary distribution $F_{hj}(\cdot)$ does not affect the results asymptotically as $R \rightarrow \infty$. For finite R , the precision of the approximation depends on how close $F_{hj}(\cdot)$ is to the original distribution $\Lambda_{I|hj}$, as discussed by Kloek and van Dijk (1978) and Akerberg (2009). In our application, the following choice of $F_h(\cdot)$ performs well in practice: We assume that each individual chooses a location to visit in sequence myopically. Appendix B.1 formally describes this choice of auxiliary distribution in further detail. Appendix B.2 shows that in practice the

¹⁵To see this, note that $\mathcal{E}_{I|hj} \rightarrow F_{hj}(I)$ as $R \rightarrow \infty$. Therefore, as long as $F_{hj}(I) > 0$ whenever $\Lambda_{I|hj} > 0$, $\Lambda_{I|hj}^* \rightarrow \Lambda_{I|hj}$ in equation (5).

sampling errors are negligible, given our choice of the number of importance samples R for estimation and simulation.¹⁶

We also use importance sampling to simulate the value of consumption access (\mathbb{A}_{hj}). For the same reason that computing the travel itinerary probabilities ($\Lambda_{I|hj}$) is difficult, computing consumption access (\mathbb{A}_{hj}) is also problematic, because it involves a summation over all possible itineraries ($\Lambda_{I|hj}^*$). Instead of directly computing this summation, we use equations (1) and (2) to re-write consumption access (\mathbb{A}_{hj}) as follows:

$$\mathbb{A}_{hj} = \varrho \left[\left(\sum_{n \in C(I)} P_n^{1-\sigma} \right)^{-\frac{1}{1-\sigma}} \tau_{I|hj}^{-1} \right]^\theta / \Lambda_{I|hj}, \quad (6)$$

for any itinerary I , and for workdays ($j \notin \emptyset$) and non-workdays ($j \in \emptyset$) separately. We replace the actual travel itinerary probabilities ($\Lambda_{I|hj}$) in this expression with the simulated probabilities from our importance sampling ($\Lambda_{I|hj}^*$). In particular, we construct consumption access (\mathbb{A}_{hj}) using the itinerary I that is most frequently drawn from our auxiliary distribution.

5.3 Estimation Procedure

In this subsection, we estimate the parameters of the travel itinerary choice problem. We begin by parameterizing the travel cost for non-commuting trips as follows:

$$\tau_{I|hj} = \eta^{|I|} \exp(\rho D_{I|hj}), \quad (7)$$

where $D_{I|hj}$ is the total travel time to follow itinerary I starting from and ending at the home location h ; to avoid double counting, we subtract the direct round trip to workplaces on workdays ($j \notin \emptyset$).¹⁷ The parameter η captures the iceberg cost of visiting a location. The parameter ρ governs the semi-elasticity of travel cost with respect to travel time. Similarly, we also parameterize travel costs for commuting trips as an exponential function of the travel time for the round trip from home to work:

$$\tau_{hj}^W = \exp(\rho (D_{hj} + D_{jh})), \quad (8)$$

¹⁶Another advantage of importance sampling is that researchers do not have to redraw samples from auxiliary distribution for each parameter value during estimation, as emphasized by [Ackerberg \(2009\)](#). Instead, one can obtain the simulated values of $\Lambda_{I|hj}^*$ simply by recalculating the likelihood ratio for each parameter value. This property substantially reduces the computational burden for our estimation and counterfactual simulations.

¹⁷Formally, $D_{I|hj} = (D_{hI_1} + \sum_{i=1, \dots, |I|-1} D_{I_i I_{i+1}} + D_{I_{|I|} h}) - \mathbf{1}[j \notin \emptyset] \times (D_{hj} + D_{jh})$, where I_i is the i -th location of the itinerary I , $D_{n\ell}$ is the travel time distance from location n to location ℓ , $\mathbf{1}[j \notin \emptyset]$ is the dummy variable that takes one if $j \notin \emptyset$, i.e., when the agent does not travel to work during the day. We subtract the travel time for the direct round trip to the agent's workplace to avoid the double counting of the travel cost for commuting in equation (8). We abstract from cases in which agents make more than one trip chain per day, since this is only infrequently observed in our data (see Figure 7).

where $D_{hj} + D_{jh}$ indicates the travel time required for the round trip from home to work.

We calibrate two of the model’s parameters. We set the fraction of days for which agents travel to work as equal to $\xi = 5/7$. We set the expenditure share for nontradable services as $\alpha^S = 0.6$, based on the revenue share in those sectors in the economic census.¹⁸

We next estimate the remaining model parameters: the iceberg cost of visiting a location (η); the semi-elasticity of travel costs to travel times (ρ); the heterogeneity of preferences for nontradable services across locations (θ); the elasticity of substitution between varieties of nontradable services (σ); and the heterogeneity of preferences across workplaces (ϕ).

As part of this estimation, we also recover estimates of the nontradable price indexes $\{P_n\}$ and wages $\{w_{nk}\}$ in each location and sector. In Section 6 below, we show how to endogenously determine these variables in general equilibrium. In this section, we use agents’ travel itinerary choices from the observed equilibrium in the data to reveal the implied equilibrium values of these endogenous variables.

We estimate the parameters $\{\eta, \rho, \theta, \sigma, \phi\}$ and the endogenous variables $\{P_n, w_{nk}\}$ using a step-wise generalized method-of-moments (GMM) procedure. In the first step, we estimate the parameters $\{\theta, \sigma, \phi\}$ for an assumed value of the parameters $\{\eta, \rho\}$. In the second step, we estimate the parameters $\{\eta, \rho\}$. Given these estimates for $\{\eta, \rho\}$ from the second step, we can then immediately recover $\{\theta, \sigma, \phi\}$ from the first step. We now discuss each step in turn.

Step 1a. Estimate $\{P_n\}$ and θ . We estimate the nontradable price indexes $\{P_n\}$ using the subsample of agent-days for which agents do not travel to work and only visit a single consumption location. More specifically, from the itinerary choice probability (1), the probability that an agent chooses a single consumption location n if she has a home location h and does not travel to work that day ($j = \emptyset$) is given by:

$$\Lambda_{\{n\}|h\emptyset}^{\text{single}} = \frac{P_n^{-\theta} \exp(-\rho\theta D_{\{n\}|h\emptyset})}{\sum_{\ell \in N} P_\ell^{-\theta} \exp(-\rho\theta D_{\{\ell\}|h\emptyset})}. \quad (9)$$

This choice probability for a single consumption location follows a conventional gravity equation. We estimate this gravity equation using the Poisson-pseudo maximum likelihood (PPML) estimator following Santos Silva and Tenreyro (2006), which allows for zero bilateral travel flows between locations. In this specification, the destination fixed effect $\{P_n^{-\theta}\}$ is a power function of the nontradable price index $\{P_n\}$, and the semi-elasticity of travel costs with respect to travel time is equal to $(-\rho\theta)$. Given our assumed semi-elasticity of travel costs to

¹⁸The nontradable service sector is defined to include “Finance, Real Estate, Communication, and Professional”, “Wholesale and Retail” (category G, J, K, L of the Japanese Standard Industrial Classification (JSIC)), “Accommodations, Eating, Drinking” (category I), “Medical and Health Care” (category P), and “Other Services” (category Q), as used in Section 4.

travel times (ρ), we can recover the heterogeneity in preferences for nontradable services (θ) from this semi-elasticity. Using this estimate for θ , we can in turn solve for the nontradable price indexes $\{P_n\}$ from the destination fixed effects.

Step 1b. Estimate σ . We estimate the elasticity of substitution across nontradable varieties (σ) using these solutions for nontradable price indexes $\{P_n\}$ and data on the number of varieties of nontradable services M_n . Recall that the CES utility function implies $P_n = p_n M_n^{\frac{1}{1-\sigma}}$. Rewriting this expression, we obtain the following estimating equation:

$$\log P_n - \log p_n = \beta_0 + \beta_1 \log M_n + \epsilon_n, \quad (10)$$

where p_n is the price of each nontradable variety; $\beta_1 = \frac{1}{1-\sigma}$ is the coefficient on the number of nontradable varieties; and ϵ_n captures estimation error in P_n .

Consistent with the model, we assume that each establishment produces a distinct variety of nontradable services, and proxy M_n by the number of establishments in the nontradable service sector. We proxy p_n by the official price index for food at the prefecture level from the Retail Price Survey conducted by the Ministry of Internal Affairs and Communications in 2019. We also control for land prices and the logarithm of travel time from the central business district (CBD) to control for unobserved differences in marginal costs. One potential concern about estimating this specification using OLS is classical measurement error in M_n as a proxy for the number of varieties, which leads to a downward bias in β_1 . To address this concern, we instrument M_n by the number of establishments in the nontradable sector in 1980. Estimating equation (10) using two-stage least squares, we recover $\sigma = 1 - \frac{1}{\beta_1}$.

Step 1c. Estimate $\{w_{jk}\}$ and ϕ . We estimate wages $\{w_{jk}\}$ and the heterogeneity in preferences across workplaces (ϕ) using the bilateral commuting probabilities ($\Omega_{jk|h}$) from residence h to workplace j and employment sector k . These bilateral commuting probabilities ($\Omega_{jk|h}$) share features with the conventional gravity equation: they depend on residence fixed effects ($\zeta_h = \sum_{j'} \sum_{k' \in \{T, S\}} (w_{j'k'} \tilde{\Delta}_{hj'}^{\alpha^S} (\tau_{hj'}^W)^{-1})^\phi$), workplace fixed effects ($\mu_{jk} = w_{jk}^\phi$) and bilateral commuting costs ($(\tau_{hj}^W)^{-\phi} = \exp(-\rho\phi(D_{hj} + D_{jh}))$). However, a key difference from a conventional gravity equation is that these bilateral commuting probabilities now depend on bilateral consumption access ($(\tilde{\Delta}_{hj}^{\alpha^S})^\phi$).

We use an iterative procedure to estimate $\{w_{jk}\}$ and ϕ . We start with an initial guess for the parameter ϕ . Given this initial guess, we construct bilateral consumption access ($\tilde{\Delta}_{hj}$) using the parameters estimated in Steps 1a and 1b, and our importance sampling (equation (6)). We next estimate the commuting gravity equation (3) using the Poisson Pseudo Maximum Likelihood (PPML) estimator following Santos Silva and Tenreyro (2006), which again allows for zero

bilateral travel flows between locations. We control for bilateral consumption access ($\tilde{\Delta}_{hj}^{\alpha^S \phi}$), and include bilateral commuting travel time ($\exp(-\rho\phi(D_{hj} + D_{jh}))$), residence fixed effects (ζ_h) and workplace fixed effects (μ_{jk}). From the estimated coefficient on bilateral commuting travel time $-\rho\phi$, we recover an estimate of ϕ given our assumed value for ρ . If our estimate of ϕ is not equal to our initial guess for this parameter, we update our guess, and repeat this procedure, until our estimate of ϕ and our guess converge to one another. Finally, we recover wages $\{w_{jk}\}$ from the workplace fixed effect of the gravity equation ($w_{jk} = \mu_{jk}^{1/\phi}$) using our estimate for ϕ .

Step 2. Construct moments to identify η and ρ . In our second step, we construct our moments to estimate iceberg cost of visiting a location (η) and the elasticity of travel costs with respect to travel times (ρ). Our first moment is the average number of stays per day for residents in each location. This moment is informative about η , because a lower value of η implies that agents visit more locations during the day. Our second moment is the squared value of residential income, as determined by wages and commuting probabilities. This moment is informative about ρ , because a higher value of ρ implies a lower value of ϕ , which in turn implies a higher variance of wages (w_{jT} and w_{jS}), and hence induces a higher variance of residential income across locations.¹⁹ Using these estimates for $\{\eta, \rho\}$, we recover the other parameters $\{\theta, \sigma, \phi\}$, wages $\{w_{nk}\}$ and price indexes $\{P_n\}$ from Step 1 above.

Estimation Results. Table 1 reports our estimates of the travel itinerary parameters, using the 242 municipalities in Tokyo Metropolitan Area as our spatial units. We use our smartphone data for April 2019 to construct the itinerary choice probabilities ($\Lambda_{I|hj}$). If an agent makes multiple stops within a municipality, we count these stops as a single stay. To reduce the computation burden, we impose a maximum number of stays per day of 5.²⁰ In implementing our importance sampling (Algorithm 1), we take 200 draws for each combination of home and work locations. Online Appendix B.2 shows that this choice of importance sampling draws gives negligible sampling errors.

We estimate an elasticity of substitution across varieties of nontradable services of $\sigma = 5.3$. Although there are relatively few estimates of substitution elasticities for nontradable services, because of a scarcity of available data, this estimate is in line with several other findings in the existing empirical literature. Using consumer panel data on retail store market shares and prices for Mexico, [Atkin, Faber, and Gonzalez-Navarro \(2018\)](#) estimates elasticities of substitution ranging from 2.28-4.36. Using data on Core-Based Statistical Areas (CBSA)

¹⁹In our model, the residential income of location h is given by $E_h = \sum_j \sum_{k \in \{T, S\}} \Omega_{jk|h} w_{jk}$. In the data, we construct residential income from the municipality income tax base data described in Section 3.

²⁰Less than one percent of user-days makes more than 5 stays per day in our sample.

in the United States, [Couture, Gaubert, Handbury, and Hurst \(2022\)](#) assumes an elasticity of substitution across non-traded varieties within neighborhoods of 6.8.

We find a preference heterogeneity parameter across travel itineraries of $\theta = 4.5$, implying substantial substitutability across alternative travel itineraries, but less than across varieties of non-traded services. We obtain a preference heterogeneity parameter across workplaces of $\phi = 3.04$, which is in line with the range of findings in the existing empirical literature. Using a similar moment condition for the variance of wages across locations, [Ahlfeldt, Redding, Sturm, and Wolf \(2015\)](#) estimates a value of 6.83 for this parameter. Using the construction of London’s 19th-century railway network as a source of quasi-experimental variation, [Heblich, Redding, and Sturm \(2020\)](#) estimates a value of 5.25 for this parameter. Using the natural experiment of Bogota’s Bus Rapid Transit (BRT) system, [Tsivanidis \(2019\)](#) obtains values for this parameter ranging from 2 to 2.8. Finally, in an analysis of Los Angeles’s metro system, [Severen \(2022\)](#) find values for this parameter ranging from 2.18 to 2.90.

We estimate the semi-elasticity of travel costs with respect to travel time of $\rho = 0.69$, which implies that one hour of travel is equivalent to a 69 percent increase in the nontradables price index in a given location, and highlights the substantial opportunity cost of travel time. Finally, we find an iceberg cost of visiting an additional location of $\eta = 2.1$, which implies that visiting an additional location is equivalent to $\log(\eta)/\rho \approx 0.5$ hours of additional travel. This pattern of results is consistent with the idea that there are substantial fixed costs of visiting additional destinations, including time spent entering and exiting transit stations and waiting for transit connections.

Table 1: Estimated and Calibrated Travel Itinerary Parameters

Parameters	Values	Estimated Parameters
σ	5.3	Elasticity of substitution
θ	4.5	Dispersion of Fréchet shocks for travel itinerary choice
ϕ	3.04	Dispersion of Fréchet shocks for workplace choice
ρ	0.69	Travel cost per hour
η	2.1	Travel cost of stopping at a location
Parameters	Values	Calibrated Parameters
ξ	0.71	Fraction of workdays
α^S	0.6	Expenditure share of nontradable sector

Note: Top panel reports the estimated travel itinerary parameters using our smartphone data for April 2019 and the stepwise procedure discussed in the main text; bottom panel reports the calibrated parameters.

5.4 Model Fit

We now provide evidence on the fit of our estimated travel itinerary model for targeted and untargeted moments.

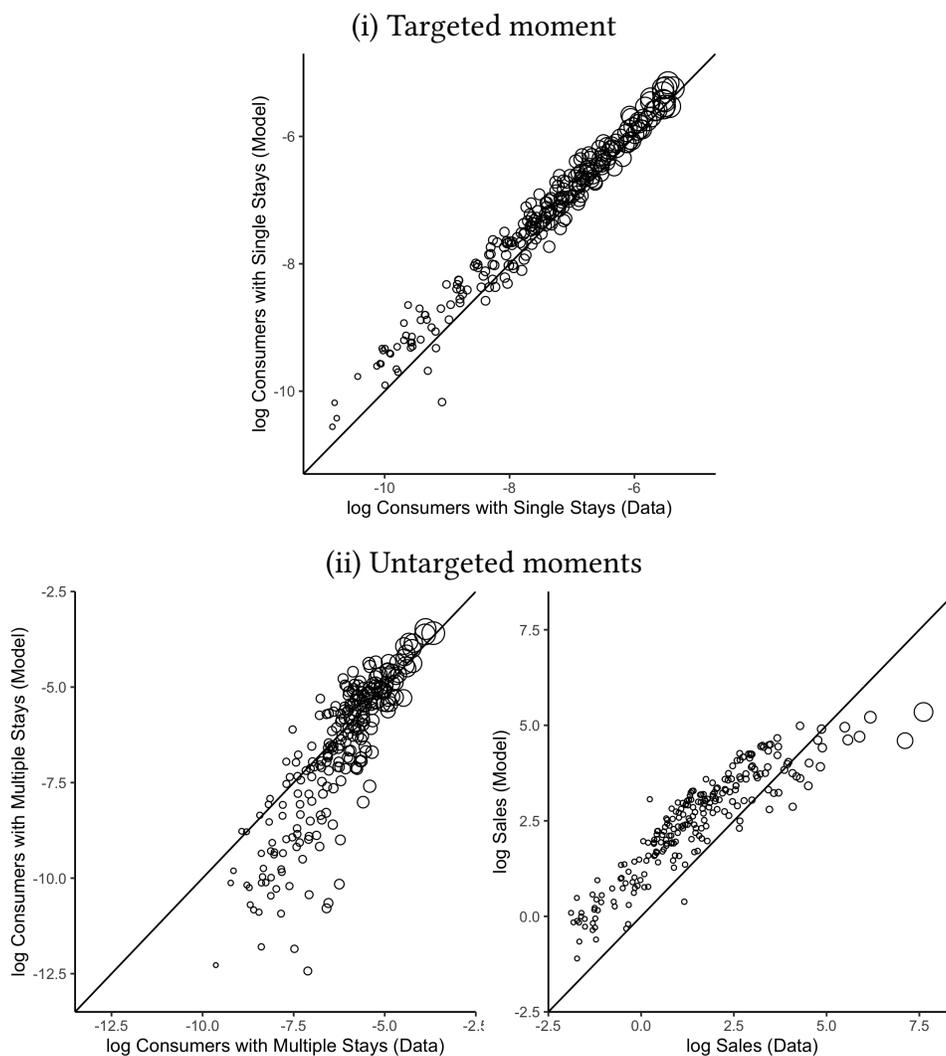
In Panel (i) of Figure 10, we display the log total number of stays in each municipality among agents who visit only one location and do not go to their workplace. We show the model's predictions on the vertical axis and the observed values in the data on the horizontal axis. This statistic is targeted in our estimation procedure through the estimation of the gravity equation (9) in our Step 1a and the average number of stays during each day in our Step 2. We find that our model provides a good fit to this targeted statistic, with the model predictions lying close to the 45-degree line.²¹

In Panel (ii) of Figure 10, we assess our model's fit in terms of untargeted moments. In the left figure, we plot the log total number of stays in each municipality among agents who visit multiple locations during the day. Again we show the model's predictions on the vertical axis and the observed values in the data on the horizontal axis. The model's predictions are clustered around the 45-degree line for municipalities that attract a larger number of stays (a higher value on the x-axis). For smaller municipalities (a lower value on the x-axis), the model's predictions tend to lie below the 45-degree line, indicating that our model underpredicts the number of stays in these municipalities. One possible reason for this departure is that our assumption of constant elasticity of substitution preferences for nontradable services may be an approximation to the data. For example, locations with small numbers of stays could offer less-substitutable nontradable services, which could lead agents to visit these locations more often than our model predicts. Nevertheless, we still find a strong, positive relationship between the model's predictions and the observed data.

In the right figure of Panel (ii), we plot log sales in the nontradable service sector in each municipality in the model (vertical axis) and data (horizontal axis). Our model prediction of the sales of the nontradable service sector is derived from embedding our specification of travel itineraries in a quantitative urban model, as discussed in the next section (see equation (24)). We compare this model prediction against the sales of the nontradable service sector reported in a separate economic census. We again find that the model's predictions are concentrated around the 45 degree line. The model tends to under predict nontradable sales in larger municipalities and over predict nontradable sales in smaller municipalities. One potential explanation could be that in reality some amenities are not priced and hence do not raise

²¹The departures from the 45-degree line arise from the residuals in estimating the gravity equation (9) and the fact that the fraction of consumers visiting only one municipality can be different between the model and data for an individual municipality. Our moment condition for η ensures that they are on average the same, but they can differ for individual municipalities.

Figure 10: Model Fit for our Travel Itinerary Model



Note: Model fit for our travel itinerary model using the estimated and calibrated parameters from Table 1. In each figure, we plot the statistic described on the axes for each municipality. We show our model prediction on the vertical axis and the observed values in the data on the horizontal axis. The size of the dot corresponds to the residential population size of the municipality. Panel (i) plots the log total number of stays in each municipality among agents who visit only one location and do not visit their workplace. The left figure of Panel (ii) plots the log total number of stays in each municipality among agents who visit multiple locations during the day. The right figure of Panel (ii) plots sales in the nontradable service sector. Our model prediction for sales in the nontradable service sector is given by equation (24), from embedding our travel itinerary specification in general equilibrium in Section 6 below. The data on sales in the nontradable service sector are constructed from the economic census.

revenue (e.g., parks). These unpriced amenities could be relatively more important in municipalities with lower levels of economic activity. Nevertheless, we again find a strong positive relationship between our model’s predictions and the observed data.

6 Quantitative Urban Model

We now assess the quantitative implications of trip chains and consumption externalities for the spatial distribution of economic activity, by embedding our specification of travel itineraries in a conventional quantitative urban model following [Ahlfeldt, Redding, Sturm, and Wolf \(2015\)](#), as reviewed in [Redding and Rossi-Hansberg \(2017\)](#).

In our baseline specification, we consider a closed city with an exogenous measure of agents.²² We normalize this measure of agents to one, such that we solve for the probabilities with which agents make different choices. In addition to nontradable services, there is a tradable good that can be costlessly traded across locations. We denote sectors by $k \in \{T, S\}$, where T denotes the tradable sector and S denotes the nontradable service sector. We choose the tradable good as the numeraire, such that its common price across all locations is equal to one ($p^T = 1$). Both the tradable good and nontradable services are produced using labor and commercial floor space according to a constant returns to scale technology. Floor space is supplied by a competitive housing sector.

Residence choice. Agents decide their residential location by considering residential amenities, the price of residential floor space, and expected workplace and consumption access, as determined by the travel itinerary decision in the previous section. The indirect utility of agent ω residing in location h is given by:

$$U_{h\omega} = B_h Q_h^{-\alpha^H} \left(\sum_{j'} \sum_{k' \in \{T, S\}} \left(w_{j'k'} \tilde{\Delta}_{hj'}^{\alpha^S} (\tau_{hj'}^W)^{-1} \right)^\phi \right)^{1/\phi} \epsilon_{h\omega}^R, \quad (11)$$

where B_h denotes residential amenities (e.g., leafy parks and scenic views); Q_h denotes the price of floor space; α^H is the expenditure share for residential floor space; $1 - \alpha^H - \alpha^S$ is the expenditure share for tradable goods; the term in parentheses is the expected utility from both workplace and travel itinerary decisions (from equation (3) above); we have used our choice of numeraire ($p^T = 1$); and $\epsilon_{h\omega}^R$ is an idiosyncratic preference shock that captures all the idiosyncratic reasons why workers choose particular residential locations. We assume that

²²It is straightforward to consider an open-city specification, in which total city population is endogenously determined by population mobility with a wider economy that offers a reservation level of utility \bar{U} .

this idiosyncratic preference shock is drawn from an independent Fréchet distribution with the same dispersion parameter ϕ as for workplace decisions.²³

Under this assumption of a Fréchet distribution for idiosyncratic preferences $\epsilon_{h\omega}^R$, the probability that agents choose residential location h is given by:

$$\Omega_h^R = \frac{\left(B_h Q_h^{-\alpha^H}\right)^\phi \sum_{j'} \sum_{k' \in \{T, S\}} \left(w_{j'k'} \tilde{\mathbb{A}}_{hj'}^{\alpha^S} (\tau_{hj'}^W)^{-1}\right)^\phi}{\sum_{h'} \left(B_{h'} Q_{h'}^{-\alpha^H}\right)^\phi \sum_{j'} \sum_{k' \in \{T, S\}} \left(w_{j'k'} \tilde{\mathbb{A}}_{h'j'}^{\alpha^S} (\tau_{h'j'}^W)^{-1}\right)^\phi}, \quad (12)$$

where Ω_h^R equals with the residential population R_h under our normalization of a unit measure of agents.

Combining this residential choice probability (Ω_h^R) with the workplace choice probability conditional on residence ($\Omega_{jk|h}$) in equation (3), the unconditional probability that an agent chooses residence h , workplace j , and sector k is given by:

$$\Omega_{hjk} = \Omega_{jk|h} \Omega_h^R = \frac{\left(w_{jk} B_h Q_h^{-\alpha^H} \tilde{\mathbb{A}}_{hj}^{\alpha^S} (\tau_{hj}^W)^{-1}\right)^\phi}{\sum_{h', j'} \sum_{k' \in \{T, S\}} \left(w_{j'k'} B_{h'} Q_{h'}^{-\alpha^H} \tilde{\mathbb{A}}_{h'j'}^{\alpha^S} (\tau_{h'j'}^W)^{-1}\right)^\phi}. \quad (13)$$

Summing across residences h , and again using our normalization of a unit measure of agents, the measure of workers employed in workplace j in sector k (L_{jk}) is given by:

$$L_{jk} = \sum_h \Omega_{hjk}. \quad (14)$$

Nontradable services production ($k = S$). We now specify the production technology for nontradable services and solve for the endogenous measure of nontradable varieties in each location (M_n). Nontradable services are produced under conditions of monopolistic competition using labor and floor space. We assume a homothetic production technology, such that fixed and variable costs use the two factors of production with the same intensity. The total costs of producing $x_i(\nu)$ units of output of variety ν in location i are given by:

$$c_i(\nu) = \left(f_{iS} + \frac{x_i(\nu)}{a_{iS}}\right) w_{iS}^{\beta^S} Q_i^{1-\beta^S}, \quad 0 < \beta^S < 1,$$

where a_{iS} is nontradables productivity in location i .

²³This assumption of a common shape parameter for workplace and residence decisions is isomorphic to an alternative specification, in which agents make simultaneous decisions over residence and workplace pairs with an idiosyncratic Fréchet preference shock for each workplace-residence pair and the same shape parameter ϕ , as considered in [Ahlfeldt, Redding, Sturm, and Wolf \(2015\)](#). To see this, if agents make a simultaneous decision of residence h , workplace j , and sector k based on an idiosyncratic Fréchet preference shock $\epsilon_{jkh\omega} \sim \exp(-\varsigma \epsilon_{jkh\omega}^{-\phi})$, we obtain the same unconditional probability Ω_{hjk} as equation (13). Aggregating Ω_{hjk} by residence and workplace leads to isomorphic residence and employment distributions.

Recall that consumers have constant elasticity of substitution (CES) preferences across nontradable services, with elasticity of substitution $\sigma > 1$. Therefore, from the first-order conditions for profit maximization, each nontradable services firm charges a price that is a constant markup over marginal cost: $p_i = \frac{\sigma}{\sigma-1} w_{iS}^{\beta^S} Q_i^{1-\beta^S} / a_{iS}$.

From profit maximization and zero profits, the equilibrium measure of nontradable varieties (M_{iS}) produced in location i is therefore:

$$M_{iS} = \frac{1}{\sigma-1} \frac{1}{f_{iS}} \left(\frac{L_{iS}}{\beta^S} \right)^{\beta^S} \left(\frac{H_{iS}}{1-\beta^S} \right)^{1-\beta^S}, \quad (15)$$

where L_{iS} and H_{iS} are the aggregate inputs of labor and commercial floor space used in nontradable services in location i .²⁴

Using the properties of the CES utility function, the corresponding price index for the nontradable services offered in location in i is given by:

$$P_i = p_i (M_{iS})^{\frac{1}{1-\sigma}} = \frac{1}{A_{iS}} w_{iS}^{\beta^S} Q_i^{1-\beta^S}, \quad (16)$$

where A_{iS} is defined by:

$$A_{iS} = \tilde{a}_{iS} (L_{iS})^{\frac{\beta^S}{\sigma-1}} (H_{iS})^{\frac{1-\beta^S}{\sigma-1}}, \quad (17)$$

and $\tilde{a}_{iS} \equiv \frac{\sigma-1}{\sigma} \left[\frac{1}{\sigma-1} \left(\frac{1}{\beta^S} \right)^{\beta^S} \left(\frac{1}{1-\beta^S} \right)^{1-\beta^S} \right]^{\frac{1}{\sigma-1}} \left(\frac{1}{f_{iS}} \right)^{\frac{1}{\sigma-1}} a_{iS}$.²⁵

Tradables production ($k = T$). Tradable goods are produced under conditions of perfect competition using labor and floor space. We assume a constant returns to scale Cobb-Douglas production technology. Profit maximization and zero profits imply that the price of the tradable good in location i must equal its unit cost if the tradable good is produced:

$$P_i^T = \frac{1}{A_{iT}} w_{iT}^{\beta^T} Q_i^{1-\beta^T}, \quad 0 < \beta^T < 1. \quad (18)$$

Recall that the tradable good is costlessly traded and we choose it as our numeraire such that:

$$P_i^T = 1 \quad \forall i \in N. \quad (19)$$

We incorporate agglomeration forces in tradables by allowing productivity (A_{iT}) to depend on production fundamentals and production externalities. Production fundamentals (a_{iT})

²⁴To derive this result, note that aggregate profits from nontradable services production are given by $\frac{p_i}{\sigma} \left(\frac{L_{iS}}{\beta^S} \right)^{\beta^S} \left(\frac{H_{iS}}{1-\beta^S} \right)^{1-\beta^S}$ and the aggregate fixed cost payment is $((\sigma-1)/\sigma) p_i M_{iS} f_{iS}$. Equating these two objects in the free-entry condition yields the above expression for the measure of nontradable varieties (M_{iS}).

²⁵Although we assume monopolistically-competitive nontradable service firms, this formulation is isomorphic to an alternative specification of perfectly competitive nontradable service firms with agglomeration forces from external economies of scale according to the functional form given by equation (17).

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. Formally,

$$A_{iT} = a_{iT} \left(\frac{L_{iT}}{K_i} \right)^{\eta^W}, \quad (20)$$

where K_i is the geographic area of location i , and η^W is the parameter that governs the strength of production externalities.²⁶

Floor space demand and supply. The price of floor space in location i is determined by the floor space market clearing condition that equates the demand and supply for floor space. The total demand for floor space equals residential demand plus commercial demand in the tradable and nontradable sectors. Under our assumptions on utility and production, we can write this floor space market clearing as:

$$H_i = H_i^U + \sum_{k \in \{T, S\}} H_{ik}, \quad (21)$$

where H_i is the supply of floor space; $H_{ik} = \frac{1-\beta^k}{\beta^k} w_{ik} L_{ik} / Q_i$ is the commercial demand for floor space in each sector k ; $H_i^U = \alpha^H E_i R_i / Q_i$ corresponds to the residential demand for floor space, where $E_i = \sum_{j,k} \Omega_{jk|h} w_{jk}$ is the average income of the residents of location i .

Following the conventional approach in the urban economics literature, as in [Epple, Gordon, and Sieg \(2010\)](#), we assume that floor space is supplied by a competitive construction sector using a Cobb-Douglas technology with land K and capital M as inputs:

$$H_i = M_i^\mu K_i^{1-\mu}, \quad 0 < \mu < 1. \quad (22)$$

Using cost minimization and zero profits, and assuming a common price of capital ($\mathbb{P}_i = \mathbb{P}$), we obtain a constant elasticity inverse supply function for floor space, as in [Saiz \(2010\)](#):

$$Q_i = \psi_i H_i^{\frac{1-\mu}{\mu}}, \quad (23)$$

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

Nontradable services market clearing. In each location, nontradable service supply and demand are equalized. More specifically, revenue equals expenditure on nontradable services in each location n , such that:

$$P_n A_{nS} \left(\frac{L_{nS}}{\beta^S} \right)^{\beta^S} \left(\frac{H_{nS}}{1-\beta^S} \right)^{1-\beta^S} = \alpha^S \sum_{h,j,k} w_{jk} \Omega_{hjk} \tilde{\Lambda}_{I|jh} \Psi_{n|I}, \quad (24)$$

²⁶We 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.

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

Agglomeration spillovers for residential amenities. We allow for agglomeration forces in residential choices by allowing 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 in independent of neighboring economic activity (e.g., green areas). Residential externalities capture the effects of the surrounding density of residents (R_n/K_n) and are modeled symmetrically to production externalities:²⁷

$$B_n = b_n \left(\frac{R_n}{K_n} \right)^{\eta^B}, \quad (25)$$

where η^B is the parameter for the strength of residential externalities.

General equilibrium. The general equilibrium of the model is referenced by consumption access (\mathbb{A}_{h_j}); the travel itinerary choice probabilities ($\Lambda_{I|h_j}$); the residence and workplace choice probabilities (Ω_{h_jk}); the price index for nontradable services in each location (P_i); the price for floor space in each location (Q_i); and the wage in each sector and location (w_{ik}). The equilibrium values of these variables are determined by consumers' optimal itinerary decisions in equations (1) and (2); workers' residential and workplace choice probabilities from equations (4) and (14); firms' optimal decisions in each sector from equations (16), (17), (18) and (19); the floor space market clearing condition (21); the non-traded goods market clearing condition (24); and the productivity and amenity spillovers (20) and (25).

Counterfactual equilibrium. To undertake counterfactuals, we follow an exact-hat algebra approach (Dekle, Eaton, and Kortum 2007), which characterizes the counterfactual equilibrium using a limited set of parameters and baseline equilibrium variables. In particular, given an assumed change in travel costs $\{\tau_{I|h_j}, \tau_{h_j}^W\}$, the counterfactual equilibrium can be computed using the values of the structural parameters $\{\eta, \rho, \theta, \sigma, \phi, \xi, \alpha^S, \alpha^H, \beta^S, \beta^T, \mu, \eta^W, \eta^B\}$, baseline nontradable price indexes and wages $\{P_n, w_{nk}\}$, baseline choice probabilities of residence, workplace, and employment sector $\{\Omega_{h_jk}\}$, and baseline floor space $\{H_{ik}, H_i^U\}$.

²⁷As for production externalities, 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.

Conditional on the values of these variables, the knowledge of unobserved location characteristics, such as productivity and amenities, is not required. We exploit these properties when we undertake counterfactuals for the Covid-19 pandemic and the construction of new transport infrastructure in the next two sections. In Online Appendix C.1, we report the complete system of equations for the counterfactual equilibrium.

Calibration. We now discuss the calibration of the remaining general equilibrium parameters of the model. Recall that we calibrate two of the travel itinerary parameters $\{\xi, \alpha^S\}$, and estimate the other four travel itinerary parameters $\{\eta, \rho, \theta, \sigma, \phi\}$ and the nontradable price indexes and wages $\{P_n, w_{nk}\}$, as discussed in Section 5.3 above.

For the remaining general equilibrium parameters of the model $\{\alpha^H, \alpha^T, \beta^S, \beta^T, \mu, \eta^W, \eta^B\}$, we take central values from the existing empirical literature, as summarized in Table 2 below. We calibrate the Cobb-Douglas expenditure share parameters $\{\alpha^H, \alpha^T\}$ using aggregate data on expenditure shares in Japan. In particular, we set the share of expenditure on residential floor space equal to $\alpha^H = 0.25$, which is consistent with Davis and Ortalo-Magné (2011), and recover the implied share of expenditure on traded goods as $\alpha^T = 1 - \alpha^H - \alpha^S$. We calibrate the share of labor in costs in each sector (β^S, β^T) as 0.8, which is in line with the share of labor in production costs for the Tokyo Metropolitan Area. We assume a standard share of land in construction costs of $\mu = 0.75$ as in Ahlfeldt, Redding, Sturm, and Wolf (2015).

We set the production spillover elasticity for the tradable sector as $\eta^W = \beta^S / (\sigma - 1) = 0.19$, which equals the implied increasing returns to scale in the nontradable sector in equation (17), given our estimated elasticity of substitution across nontradable varieties (σ). We set the amenity spillover elasticity as $\eta^B = 0$, which is conservative among the values reported in the meta-analyses of Melo, Graham, and Noland (2009) and Ahlfeldt and Pietrostefani (2019), because we explicitly model consumption externalities through travel itineraries.

To undertake counterfactual simulations, we additionally need baseline values of choice probabilities of residence, workplace, and employment sector $\{\Omega_{hjk}\}$, and baseline floor space $\{H_{ik}, H_i^U\}$. We proxy $\{\Omega_{hjk}\}$ by combining commuting flows from smartphone data and the sectoral employment share from the economic census, as already used in Section 5.3. We construct floor space $\{H_{ik}, H_i^U\}$ using the floor space for residence and commercial purposes from Building Data described in Section 3.2.²⁸

²⁸We allocate commercial floor space into tradable and nontradable sector proportionally to sector employment, since we do not observe these classifications in our Buildings Data.

Table 2: Calibration of General Equilibrium Parameters

Parameter	Description	Value	Source
α^H	expenditure share for residential floor space	0.25	Data
β^S	labor share in production for nontradable sector	0.8	Data
β^T	labor share in production for tradable sector	0.8	Data
η^W	elasticity of production spillover in tradable sector	0.19	$\beta^S/(\sigma - 1)$
η^B	elasticity of residential amenity spillover	0	Ahlfeldt-Pietrostefani (2019) (Conservative)
μ	share of capital for floor space production	0.75	Ahlfeldt et al (2015)

Note: The list of parameters and their calibrated values for general equilibrium counterfactuals. See Table 1 for our calibrated and estimated travel itinerary parameters, including the elasticity of substitution across nontradable service varieties (σ).

7 Covid-19 Pandemic

In this section, we use the natural experiment of the Covid-19 pandemic as a specification check on our model’s prediction of consumption externalities across locations. In our model of trip chains, the decline in commuting into downtown Tokyo during the period of the emergency order (April 2020) reduces the demand for nontradable services close to or along the route to downtown workplaces. In contrast, in a conventional quantitative urban model without trip chains, in which agents make direct trips to consume goods or services from their residence, there are limited reasons why the shift to remote working would lead to such a reduction in demand for nontradable services downtown.

We begin by estimating the change in two structural parameters that are plausibly affected by the Covid-19 pandemic, using our smartphone data for the month of April 2020 during the period of the emergency order. The first parameter is the value of travel time (ρ). We estimate the change in this parameter from the gravity equation (9). In particular, using the subsample of agent-days for which agents do not travel to work and only visit a single consumption location in April 2020, we estimate the semi-elasticity of non-commuting travel with respect to travel time ($\rho \times \theta$). We then estimate the new value of ρ , holding constant the parameter θ at the value estimated in Section 5 above for April 2019. The second parameter is the probability of travelling to work (ξ). We calibrate the change in this parameter using the observed changes in the probability of travelling to work from April 2019 to April 2020.²⁹

We next examine our model’s quantitative predictions for the impact of the Covid-19 pandemic. Starting from the observed equilibrium in the data in April 2019, we undertake a

²⁹Note that the change of ξ does not affect the number of efficiency units of labor per person in our general equilibrium model. It simply changes the probability that travel itineraries include workplaces. Therefore this parameter has a direct interpretation as a shift to remote working. Existing research finds mixed results of telecommuting on worker productivity during the Covid-19 pandemic depending on sectors and occupations (e.g., Barrero, Bloom, and Davis 2021 and Delventhal and Parkhomenko 2022).

counterfactual using our estimated change in these two structural parameters $\{\rho, \xi\}$. In our baseline specification, we hold employment by residence and workplace $\{R_n, L_n\}$ and prices $\{P_n, w_{nk}, Q_n\}$ constant at their values in April 2019, reflecting the fact that April 2020 is still in the early stages of the pandemic, before agents could react by changing residence or workplace, and before prices started to respond. In Online Appendix Figure C.2.1, we report a robustness test, in which we allow for general equilibrium changes in employment by residence and workplace $\{R_n, L_n\}$ and prices $\{P_n, w_{nk}, Q_n\}$.³⁰

In Table 3, we compare the actual changes in travel patterns from April 2019 to April 2020 to those predicted by our model. In Panel (A), we present aggregate statistics on the changes of the number of stays and distance traveled. In Panel (B), we show changes in consumer foot traffic at destinations (the number of non-work stays outside the home location) by terciles of distance to the Central Business District (CBD). We measure the CBD as the Chiyoda municipality (ward) of the Tokyo Metropolitan Area. As a point of comparison, we also report the predictions of a special case of our model without trip chains, in which agents choose a single location for nontradable services consumption and make a roundtrip from their residence, as in conventional quantitative urban models (labeled as “Model (No Trip Chain)”).³¹

In Rows (a-1) to (a-2) of Panel (A), we report two statistics that we directly target to estimate the changes in the model’s structural parameters in response to the Covid-19 pandemic: the log change in an agent’s probability of travelling to her workplace during the day (for ξ); and the change in the estimated gravity coefficient for non-commuting trips that involve an agent visiting one consumption location that is not their workplace during the day (for ρ). We find that the probability of travelling to work falls and the gravity coefficient becomes more negative. These patterns are consistent with the reduced-form findings in Section 4 above, where we found that both the frequency of travelling to workplaces and the distance traveled for non-work stays declined. By construction, the model’s predictions coincide with the actual data for these two statistics, because they are targeted in our estimation.

In the remaining rows of Panel (A), we present other statistics on changes in travel patterns that are not directly targeted in our estimation. We show that our model also provides a close approximation to these non-targeted moments. In Row (b-1) of Column (1), we find a 6 percent

³⁰During the emergency order, some nontradable sectors such as grocery stores and restaurants were encouraged to close early in a day, which may have affected $\{P_n\}$ even in the short run. Lacking the detailed data on how stores responded during the emergency order, we simply fix $\{P_n\}$ at the value of 2019. Nonetheless, our simulation predicts a collapse of foot traffic in downtown areas of a similar magnitude to that observed in the data, which suggests that the extent of early store closing did not differ systematically across space.

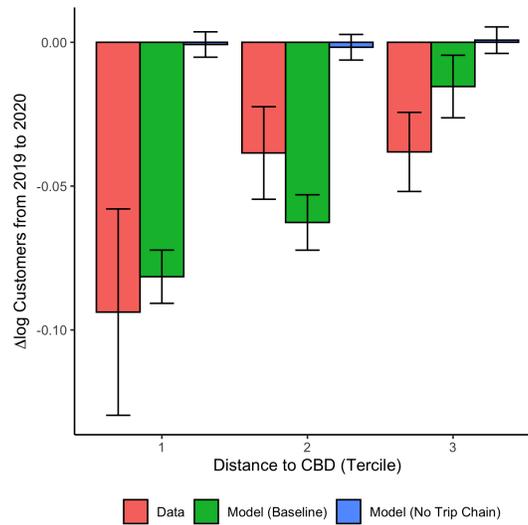
³¹We fit non-commuting travel in this special case of our model using the sub-sample of stays in which agents do not go to their workplace and only visit a single consumption location, as in our estimation of our travel itinerary model in Section 5.3 above. In Appendix Figure C.2.2, we find the same qualitative and quantitative pattern of results if we instead use all non-work stays outside the home location to fit this special case.

Table 3: Predicted and Actual Changes in Travel Patterns from April 2019 to April 2020

(A) Changes in the Number of Stays and Distance Traveled

	(1) Data	(2) Model (Baseline)	(3) Model (No Trip Chain)
(a) Targeted Moments			
(a-1) Δ log probability of work stay per day	-0.04	-0.04	-0.04
(a-2) Δ gravity coefficient conditional on visiting only one location	-0.27	-0.27	-0.27
(b) Untargeted Moments			
(b-1) Δ log number of nonwork stay per user-day	-0.06	-0.08	0.00
(b-2) Δ log number of nonwork stay per user-day given workdays	-0.04	-0.07	0.00
(b-3) Δ log number of nonwork stay per user-day given non-workdays	-0.10	-0.09	0.00
(b-4) Δ log median distance to nonwork stays from home	-0.19	-0.19	-0.06

(B) Changes in Consumer Foot Traffic by Distance to CBD



Note: Actual changes in the data are from April 2019 to April 2020 (during the period of the emergency order), where we adjust the residential population in two periods to control for the change in residential population preceding the Covid-19 pandemic; Predicted changes in the model are based on starting at the observed equilibrium in the data in April 2019 and undertaking a counterfactual for our estimated change in the value of travel time (ρ) and the probability of travelling to work (ξ); In Panel (B), we take Chiyoda municipality (ward) as the Central Business District (CBD) of the Tokyo Metropolitan Area.

reduction in non-work stays in the data, which is somewhat larger than the reduction in work stays in Row (a-1).³² Our travel itinerary model predicts an 8 percent reduction in these non-work stays (Row (b-1) of Column (2)), close to the observed value in the data. In contrast, the special case of our model without trip chains predicts no change in the number of non-work stays (Row (b-1) of Column (3)), because agents are assumed to make a single consumption trip from their residence each day.

We next decompose changes in non-work stays into those that occur as part of trip chains

³²This number is somewhat smaller than the reductions shown in Figure 8. This is mainly because here we aggregate stays to the municipality level and hence avoid double-counting stays at workplaces when workers leave their workplace during the day and return there later the same day (e.g., after lunch elsewhere).

including workplace (Row (b-2)), versus those that occur as part of trip chains excluding workplace (Row (b-3)). In the data, as reported in Column (1), we find a larger reduction in non-work stays for trip chains excluding workplace than for those including workplace. This pattern has an intuitive explanation. Conditional on visiting a workplace, the additional time required for a detour to consume non-traded services is short, in particular when there are attractive consumption locations around the workplace or along the commuting route. In contrast, agents spend a longer time travelling to consume non-traded services if they do not visit a workplace. Therefore, the increase of travel cost during the Covid-19 pandemic has a greater impact on non-work stays for trip chains excluding workplace. Our baseline model successfully captures this pattern, as shown in Column (2).

In Row (b-4) of Panel (A), we report changes in distances travelled to non-work destinations. In the data, we find a 19 percent reduction in the median distance from home to non-work destinations (Column (1)). Our travel itinerary model successfully replicates this 19 percent reduction in distance to non-work destinations (Column (2)). In contrast, in the special case of our model with no trip chains, we predict a 6 percent decline in distance to non-work destinations (Column (3)), which is significantly smaller than the reductions in both the data and our baseline model. The reason is because this special case does not capture the loss of access to consumption opportunities from not travelling to work, since it assumes that all consumption travel occurs from home.

In Panel (B), we examine how these changes in travel patterns affect the spatial distribution of foot traffic within the Tokyo Metropolitan Area. We display average changes in the number of non-work stays for destinations in different terciles of distance from the Central Business District (Chiyoda Ward). In the data, we observe a larger decline in consumer foot traffic in municipalities closer to the CBD, as indicated by the red bars. Our baseline model successfully captures this pattern, as the reduction in commuting trips from the shift to remote working leads to a collapse in local demand for non-traded services in downtown areas, as shown by the green bars. In contrast, the special case of our model with no trip chains fails to capture this pattern in the data, because it assumes that all travel to consume nontradable services occurs from home, and hence does not capture the negative pecuniary externalities for nontradable consumption travel from the reduction in commuting travel into downtown areas.

Taken together, we find strong evidence from the natural experiment of the Covid-19 pandemic for the consumption externalities across locations implied by trip chains. Our estimated model is not qualitatively but also quantitatively successful in explaining the collapse in demand for local non-traded services in downtown areas from the shift to remote work.

8 Transportation Infrastructure

We now show that trip chains and consumption externalities are not only important for understanding the impact of the Covid-19 pandemic on the spatial distribution of economic activity, but are also relevant for evaluating the impact of public policies, such as the construction of new transport infrastructure.

To assess the relevance of trip chains and consumption externalities for the evaluation of public policies, we undertake counterfactuals for the construction of subway and railway lines in the Tokyo Metropolitan area in the period since 1960. In Figure 11, we display the subway and railway network in 1960 and the expansions in this network in subsequent years. The additional subway and railway lines that were constructed during this time period were intended to enhance the public transport access of suburban areas to the city center and to improve the connections between different parts of the city center. To assess the welfare implications of these network expansions, we begin by calibrating our model using our smartphone data in April 2019. We next undertake a counterfactual for the removal of all subway and railway lines that were constructed in the period since 1960 by feeding in the implied increase of travel time throughout the city.³³

Table 4: Counterfactuals for the Removal of all Subway and Railway Lines Built in the Tokyo Metropolitan Area Since 1960

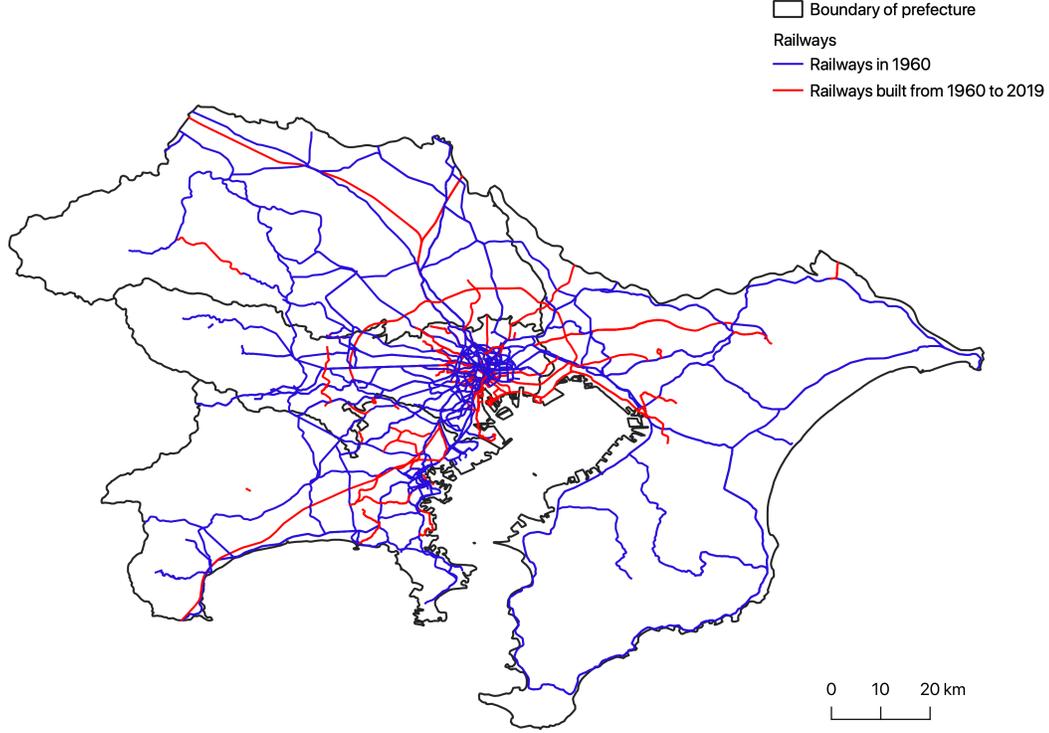
	Δ Welfare (%)	Relative to Baseline (%)
(1) Baseline	-7.4	100
(2) No Consumption Trips	-5.8	78
(3) No Trip Chains (Single Consumption Location from Home)	-6.0	81
(4) No Trip Chains (All Consumption Locations from Home)	-6.5	88

Note: This table presents the welfare losses by removing all subway and railway lines built in the Tokyo Metropolitan Area from 1960-2019. For each model specification stated in the first column, the second column lists the changes in welfare in percentage points from the removal of these subway and railway lines. The third column presents the predicted changes in welfare for each model specification relative to the prediction of our baseline specification (Row (1)). Online Appendix Figure C.3.1 provides further evidence on the spatial pattern of counterfactual changes in rents, wages, price indices, employment by residence, and employment by workplace.

Table 4 reports the counterfactual changes in residential welfare implied by our model. We measure the change of residential welfare using the expected utility of residents in equation

³³We compute the implied change in travel times from the removal of subway and railway lines constructed since 1960 using the following procedure. We first obtain the travel time for all pairs of municipalities in 2019 using Eki-spert API described in Section 3.2. We next compute the ratio of the required travel time using the subway and railway network in 2019 and 1960 using ArcGIS Network Analyst extension, assuming a travel speed of 80 meters per minute for walk, 600 meters per minute for subways and railways, and 150 meters per minute for buses. (We assume that bus networks are unchanged in the counterfactual simulation.) Finally, we multiply this ratio with the travel time in 2019 to obtain the implied travel time in 1960.

Figure 11: Map of Subway and Railway System in Tokyo Metropolitan Area



(11), given by:

$$\mathbb{E}[\max_h U_{h\omega}] = \varrho^R \left[\sum_h B_h^\theta Q_h^{-\theta\alpha^H} \left(\sum_{j'} \sum_{k' \in \{T,S\}} \left(w_{j'k'} \tilde{\mathbb{A}}_{hj'}^{\alpha^S} (\tau_{hj'}^W)^{-1} \right)^\phi \right)^{\theta/\phi} \right]^{1/\theta}, \quad (26)$$

where $\varrho^R = \Gamma(\frac{\theta-1}{\theta})$ and $\Gamma(\cdot)$ is the Gamma function. For each model specification stated in the first column, the second column lists the changes in welfare in percentage points from the removal of the subway and railway lines constructed since 1960. The third column presents the predicted changes in welfare for each model specification relative to the predictions of our baseline specification (Row 1).

As reported in Row (1), our baseline model predicts that these expansions in the subway and railway network increased welfare by around 7.4 percentage points. These findings of substantial welfare gains from transport infrastructure investments are consistent with existing empirical evidence from other settings including the construction of London’s 19th-century railway network (Heblich, Redding, and Sturm 2020), the interstate highway network and road network in the United States (Allen and Arkolakis 2021), and Bogota’s Bus Rapid Transit Sys-

tem (Tsivanidis 2019). For our case of the Tokyo Metropolitan Area, these welfare gains reflect the incremental improvements in transport connectivity relative to what was already a quite extensive subway and railway network in 1960, and they must be offset against the substantial construction costs of these network expansions.

In Row (2), we undertake a counterfactual for the special case of our model in which we shut down non-commuting trips by assuming that $\alpha^S = 0$, i.e., there is no nontradable service consumption and all consumption goods are freely traded within the city, as in the canonical urban model. We find substantially smaller welfare gains of 5.8 percentage points, around 78 percent of the welfare gains in our baseline model. The difference arises in part because abstracting from non-commuting trips mechanically undercounts the amount of travel that benefits from the reduction in travel costs from the transport improvement.

In Row (3), we undertake a counterfactual simulation for the special case of our model in which there is nontradable service consumption but no trip chains, such that agents make a single consumption trip from their residence each day. As in Section 7, we estimate this special case using the subsample of observations for which agents do not travel to work and only visit a single consumption location, consistent with our estimation in Section 5.3. We find that this special case predicts welfare gains of 6.0 percentage points. This number is only slightly larger than the version of the model that completely abstracts from nontradable services consumption in Row (2) and only 81 percent of the welfare gains in our baseline model in Row (1). This pattern of results is consistent with the idea that this special case of the model with consumption of non-traded services from home fails to capture all of the ways in which agents adjust their travel patterns in response to the reduction in travel costs.

As a robustness check, Row (4) shows the welfare gains under the same special case of our model abstracting from trip chains as in Row (3), except that we calibrate non-commuting travel using the full sample of observations regardless of whether agents visit a single location or multiple locations (instead of using the subsample of users visiting only one location during the day in Row (3)). This calibration implies that people travel more in the baseline equilibrium than in Row (3), because people who visit multiple locations tend to travel farther from home. In this robustness check, we find a welfare gain of 6.5 percentage points, which is larger than in Row (3), but again remains smaller than in our baseline specification in Row (1).

Overall, the results from these counterfactuals corroborate our findings from the natural experiment of the Covid-19 pandemic. We find that taking into account trip chains, and the resulting consumption externalities, is quantitatively important for the evaluation of public policy interventions, such as transport infrastructure improvements.

9 Conclusions

We develop a theoretical and empirical framework for analyzing the rich patterns of spatial mobility observed in recent sources of Geographical Positioning System (GPS) data. We allow agents to visit an endogenous subset of locations in an endogenous sequence. We show that these rich patterns of spatial mobility give rise to consumption externalities across locations, where having a good reason to visit one location makes it more attractive to visit other locations that are nearby or along the way.

We use smartphone GPS data for Tokyo that records location at high temporal and spatial resolution each day to establish a number of stylized facts about patterns of spatial mobility in urban areas. First, non-commuting trips are more frequent than commuting trips, such that abstracting from them substantially underestimates travel within urban areas. Second, non-commuting trips are closely related to the availability of non-traded services. Third, non-commuting trips exhibit different spatial patterns from commuting trips, such that they are not well approximated by commuting trips. Fourth, trips chains are a pervasive feature of the data, in which agents make multiple stops as part of a single journey. Lastly, we find a sharp decline in the frequency and travel length of trips during the Covid-19 pandemic, which results in a disproportionate reduction in foot traffic in downtown Tokyo.

To rationalize these observed features of the data, we develop a tractable theoretical model of travel itineraries, in which agents choose a set of locations and a sequence in which to visit them each day. Our travel itinerary specification considerably generalizes conventional models of consumption in urban areas, because the market access of both workers and firms depends on the frequency with which travel routes are chosen. Our framework thus rationalizes the idea that retail stores cluster together to attract the customers of their competitors, as they pass by to visit those competitors (as in the classic example of clusters of shoe shops). To assess the quantitative importance of these consumption externalities, we develop a framework to estimate and simulate this model, which uses importance sampling methods to overcome the high-dimensionality of the choice set implied by travel itineraries.

We show that our model of travel itineraries is quantitatively successful in explaining the observed decline in foot traffic in downtown areas during the Covid-19 pandemic. As workers in traded sectors (e.g., manufacturing) shifted to remote work, this led to a collapse in local demand for non-traded services (e.g., coffee shops and restaurants) in downtown areas. In contrast, models that abstract from trip chains and consumption externalities tend to underestimate the decline in economic activity in downtown areas.

More broadly, our research highlights the relevance of accurately modelling spatial mobility for assessing the impact of public policies, such as transport improvements. Abstracting

from non-commuting trips mechanically undercounts the amount of travel benefiting from reductions in travel costs. Additionally, omitting trip chains fails to fully capture the consumption externalities that occur between locations, when agents endogenously adjust the set and sequence of locations they visit in response to changes in travel costs.

References

- ACKERBERG, D. A. (2009): “A New Use of Importance Sampling to Reduce Computational Burden in Simulation Estimation,” *Quantitative Marketing and Economics*, 7(4), 343–376.
- AGARWAL, S., J. B. JENSEN, AND F. MONTE (2020): “Consumer Mobility and the Local Structure of Consumption Industries,” *NBER Working Paper*, 23616.
- AHLFELDT, G., AND E. PIETROSTEFANI (2019): “The Economic Effects of Density: A Synthesis,” *Journal of Urban Economics*, 111, 97–107.
- 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 (2021): “The Welfare Effects of Transportation Infrastructure Improvements,” *Review of Economic Studies*, forthcoming.
- ALLEN, T., C. ARKOLAKIS, AND X. LI (2017): “Optimal City Structure,” Yale University, mimeograph.
- ALLEN, T., D. ATKIN, S. CANTILLO, AND C. HERNANDEZ (2021): “Trucks,” MIT, mimeograph.
- ALLEN, T., S. FUCHS, S. GANAPATI, A. GRAZIANO, R. MADERA, AND J. MONTORIOL-GARRIGA (2021): “Urban Welfare: Tourism in Barcelona,” Dartmouth College, mimeograph.
- ALMAGRO, M., AND T. DOMÍNGUEZ-IINO (2021): “Location Sorting and Endogenous Amenities: Evidence from Amsterdam,” NYU, mimeograph.
- ALONSO, W. (1964): *Location and Land Use*. Harvard, Cambridge MA.
- ALTHOFF, L., F. ECKERT, S. GANAPATI, AND C. WALSH (2022): “The Geography of Remote Work,” *Regional Science and Urban Economics*, 93, 103770.
- ALVAREZ, F., D. ARGENTE, AND F. LIPPI (2021): “A Simple Planning Problem for Covid-19 Lockdown, Testing and Tracing,” *American Economic Review: Insights*, 3(3), 367–82.
- ANAS, A. (2007): “A unified theory of consumption, travel and trip chaining,” *Journal of Urban Economics*, 62(2), 162–186.
- ANTRÀS, P., T. C. FORT, AND F. TINTELNOT (2014): “The Margins of Global Sourcing: Theory and Evidence from U.S. Firms,” *American Economic Review*, 107(9), 2514–64.
- ANTRÀS, P., S. J. REDDING, AND E. ROSSI-HANSBERG (2020): “Globalization and Pandemics,” *NBER Working Paper*, 27840.
- ARGENTE, D., C.-T. HSIEH, AND M. LEE (2022): “The Cost of Privacy: Welfare Effects of the Disclosure of Covid-19 Cases,” *Review of Economics and Statistics*, 104(1), 176–186.
- ARKOLAKIS, C., F. ECKERT, AND R. SHI (2022): “Combinatorial Discrete Choice,” *Working Paper*.
- ATHEY, S., B. FERGUSON, M. GENTZKOW, AND T. SCHMIDT (2021): “Estimating experienced racial segregation in US cities using large-scale GPS data,” *Proceedings of the National Academy of Sciences*, 118(46), e2026160118.
- ATKIN, D., K. CHEN, AND A. POPOV (2022): “The Returns to Face-to-Face Interactions: Knowledge Spillovers in Silicon Valley,” MIT, mimeograph.
- ATKIN, D., B. FABER, AND M. GONZALEZ-NAVARRO (2018): “Retail globalization and household welfare: Evidence from Mexico,” *Journal of Political Economy*, 126(1), 1–73.
- BALBONI, C., G. BRYAN, M. MORTEN, AND B. SIDDIQI (2021): “Transportation, Gentrification, and Urban Mobility: The Inequality Effects of Place-Based Policies,” MIT, mimeograph.
- BARRERO, J. M., N. BLOOM, AND S. J. DAVIS (2021): “Why Working from Home Will Stick,” Discussion paper, National Bureau of Economic Research.
- BENMELECH, E., N. BERGMAN, A. MILANEZ, AND V. MUKHARLYAMOV (2019): “The Agglomeration of Bankruptcy,” *Review of Financial Studies*, 32(7), 2541–2586.
- BOWMAN, J. L., AND M. E. BEN-AKIVA (2001): “Activity-Based Disaggregate Travel Demand Model System with Activity Schedules,” *Transportation Research Part A: Policy and Practice*, 35(1), 1–28.

- BRANCACCIO, G., M. KALOUPTSIDI, AND T. PAPAGEORGIOU (2020): “Geography, Transportation and Endogenous Trade Costs,” *Econometrica*, 88(2), 657–691.
- 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.
- BUCHOLZ, N., L. DOVAL, J. KASTL, F. MATEJKA, AND T. SALZ (2021): “The Value of Time: Evidence from Auctioned Cab Rides,” Princeton University, mimeograph.
- CLAYCOMBE, R. J. (1991): “Spatial retail markets,” *International Journal of Industrial Organization*, 9(2), 303–313.
- COUTURE, V. (2016): “Valuing the Consumption Benefits of Urban Density,” Vancouver School of Economics, mimeograph.
- COUTURE, V., J. DINGEL, A. GREEN, AND J. HANDBURY (2019): “Quantifying Social Interactions Using Smartphone Data,” Chicago Booth School of Business, mimeograph.
- COUTURE, V., J. I. DINGEL, A. GREEN, J. HANDBURY, AND K. R. WILLIAMS (2022): “JUE Insight: Measuring Movement and Social Contact with Smartphone Data: A Real-Time Application to COVID-19,” *Journal of Urban Economics*, 127, 103328.
- COUTURE, V., G. DURANTON, AND M. A. TURNER (2018): “Speed,” *Review of Economics and Statistics*, 100(4), 725–739.
- COUTURE, V., C. GAUBERT, J. HANDBURY, AND E. HURST (2022): “Income Growth and the Distributional Effects of Urban Spatial Sorting,” *Review of Economic Studies*, forthcoming.
- 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, M. A., AND F. ORTALO-MAGNÉ (2011): “Housing Expenditures, Wages, Rents,” *Review of Economic Dynamics*, 14(2), 248–261.
- DEKLE, R., J. EATON, AND S. KORTUM (2007): “Unbalanced Trade,” *American Economic Review*, 97(2), 351–355.
- DELVENTHAL, M., AND A. PARKHOMENKO (2022): “Spatial Implications of Telecommuting,” *Available at SSRN* 3746555.
- DIAMOND, R. (2016): “The Determinants and Welfare Implications of US Workers’ Diverging Location Choices by Skill: 1980-2000,” *American Economic Review*, 106(3), 479–524.
- DINGEL, J., AND F. TINTELOT (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 (2022): “Assessing the Gains from E-Commerce,” *American Economic Journal: Macroeconomics*, forthcoming.
- EATON, B. C., AND R. G. LIPSEY (1982): “An Economic Theory of Central Places,” *The Economic Journal*, 92(365), 56–72.
- 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.
- FAJGELBAUM, P., A. KHANDELWAL, W. KIM, C. MANTOVANI, AND E. SCHAAL (2021): “Optimal Lockdown in a Commuting Network,” *American Economic Review: Insights*, 3(4), 503–22.
- 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 H. OGAWA (1982): “Multiple Equilibria and Structural Transformation of Non-monocentric Urban Configurations,” *Regional Science and Urban Economics*, 12(2), 161–196.
- GECHTER, M., AND N. TSIVANIDIS (2020): “Spatial Spillovers from Urban Redevelopments: Evidence from Mumbai’s Textile Mills,” University of California, Berkeley.
- GIANNONE, E., N. PAIXAO, AND X. PANG (2022): “JUE Insight: The Geography of Pandemic Containment,” *Journal of Urban Economics*, 127, 103373.
- GLAESER, E. L., C. GORBACK, AND S. J. REDDING (2020): “How Much Does COVID-19 Increase with Mobility? Evidence from New York and Four Other U.S. Cities,” *Journal of Urban Economics*, 1(1), 103292.
- GLAESER, E. L., J. KOLKO, AND A. SAIZ (2001): “Consumer City,” *Journal of Economic Geography*, 1(1), 27–50.

- GORBACK, C. (2021): “Ridesharing and the Redistribution of Economic Activity,” National Bureau of Economic Research, mimeograph.
- GUPTA, A., C. KONTOKOSTA, AND S. VAN NIEUWERBURGH (2022): “Take the Q Train: Value Capture of Public Infrastructure Projects,” *Journal of Urban Economics*, 129, 103422.
- HAUSMAN, N., P. SAMUELS, M. COHEN, AND R. SASSON (2021): “Urban Pull: The Roles of Amenities and Employment,” Harvard University, mimeograph.
- 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.
- HOELZLEIN, M. (2020): “Two-sided Sorting and Spatial Inequality in Cities,” Berkeley, mimeograph.
- JIA, P. (2008): “What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry,” *Econometrica*, 76(6), 1263–1316.
- KLOEK, T., AND H. K. VAN DIJK (1978): “Bayesian Estimates of Equation System Parameters: An Application of Integration by Monte Carlo,” *Econometrica*, 46(1), 1–19.
- KOSTER, H. R. A., I. PASIDIS, AND J. VAN OMMEREN (2019): “Shopping externalities and retail concentration: Evidence from dutch shopping streets,” *Journal of Urban Economics*, 114(August), 103194.
- KREINDLER, G., AND Y. MIYAUCHI (2022): “Measuring Commuting and Economic Activity inside Cities with Cell Phone Records,” *Review of Economics and Statistics*, forthcoming.
- LUCAS, R. E., AND E. ROSSI-HANSBERG (2002): “On the Internal Structure of Cities,” *Econometrica*, 70(4), 1445–1476.
- MELO, P. C., D. J. GRAHAM, AND R. B. NOLAND (2009): “A Meta-analysis of Estimates of Urban Agglomeration Economies,” *Regional Science and Urban Economics*, 39(3), 332–342.
- 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.
- MUTH, R. (1969): *Cities and Housing*. University of Chicago Press, Chicago.
- OBERFIELD, E., E. ROSSI-HANSBERG, P.-D. SARTE, AND N. TRACHTER (2020): “Plants in Space,” Discussion paper, National Bureau of Economic Research.
- REDDING, S. J., AND E. ROSSI-HANSBERG (2017): “Quantitative Spatial Models,” *Annual Review of Economics*, 9, 21–58.
- RELIHAN, L. E. (2022): “Is Online Retail Killing Coffee Shops? Estimating the Winners and Losers of Online Retail using Customer Transaction Microdata,” *Working Paper*.
- 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. (2022): “Commuting, Labor, and Housing Market Effects of Mass Transportation: Welfare and Identification,” *Review of Economics and Statistics*, forthcoming.
- SHOAG, D., AND S. VEUGER (2018): “Shops and the City: Evidence on Local Externalities and Local Government Policy from Big-Box Bankruptcies,” *Review of Economics and Statistics*, 100(3), 440–453.
- SU, Y. (2022): “Measuring the Value of Urban Consumption Amenities: A Time-Use Approach,” *Available at SSRN 3454631*.
- TAN, B., AND K.-H. LEE (2021): “Urban Transit Infrastructure and Inequality: The Role of Access to Non-Tradable Goods and Services,” *Working Paper*.
- THOMASSEN, Ø., H. SMITH, S. SEILER, AND P. SCHIRALDI (2017): “Multi-Category Competition and Market Power: A Model of Supermarket Pricing,” *American Economic Review*, 107(8), 2308–2351.

- TSIVANIDIS, N. (2019): “Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogotá’s Trans-Milenio,” *Working Paper*.
- USHCHEV, P., I. SLOEV, AND J. F. THISSE (2015): “Do We Go Shopping Downtown or in the ‘Burbs?,” *Journal of Urban Economics*, 85, 1–15.
- ZÁRATE, R. D. (2021): “Spatial Misallocation, Informality, and Transit Improvements: Evidence from Mexico City,” University of California, Berkeley.