American Economic Review 2024, 114(10): 3161–3205
https://doi.org/10.1257/aer.20220212

Efficiency and Equity Impacts of Urban Transportation Policies with Equilibrium Sorting†

By Panle Jia Barwick, Shanjun Li, Andrew Waxman, Jing Wu, and Tianli Xia*

We estimate an equilibrium sorting model of housing location and commuting mode choice with endogenous traffic congestion to evaluate urban transportation policies. Leveraging fine-scale data from travel diaries and housing transactions identifying residents' home and work locations, we recover rich preference heterogeneity over both travel mode and residential location decisions. While different policies produce the same congestion reduction, their impacts on social welfare differ drastically. In addition, sorting undermines the congestion reduction under driving restrictions and subway expansion but strengthens it under congestion pricing. The combination of congestion pricing and subway expansion delivers the greatest congestion relief and efficiency gains. (*JEL* H76, O18, P25, R23, R31, R41, R48)

Transportation plays a crucial role in shaping the urban spatial structure and the organization of economic activity. In many developing countries, rapid urbanization and motorization, together with poor infrastructure, have created unprecedented traffic congestion with severe economic consequences (Li 2018a; Akbar et al. 2023).[1] To address these challenges, local governments around the world have implemented a suite of policies, including driving restrictions, public transit

* Barwick: Department of Economics, UW-Madison and NBER (email: pbarwick@wisc.edu). Li: Dyson School of Applied Economics and Management, Cornell University, NBER, and RFF (email: SL2448@cornell.edu). Waxman: LBJ School of Public Affairs, University of Texas-Austin (email: awaxman@utexas.edu). Wu: Hang Lung Center for Real Estate, Tsinghua University (email: ireswujing@tsinghua.edu.cn). Xia: University of Rochester, Simon Business School (email: tianli.xia@rochester.edu). Katja Seim was the coeditor for this article. We thank the editor and four anonymous referees whose comments have greatly improved the paper. We also thank Kelly Bishop, Jan Brueckner, Shengmao Cao, Jack Collison, Gilles Duranton, Fernando Ferreira, Edward Glaeser, Matthew Kahn, Robert Metcalfe, Alvin Murphy, Matías Navarro, Stuart Rosenthal, Kerry Smith, Christopher Timmins, and seminar participants at Boston University, Brazilian School of Economics and Finance, Cornell, Duke, Jinan University, MIT, Peking University, Shanghai University of Finance and Economics, Stanford, Syracuse, University of California Berkeley, University of Illinois, University of Maryland, University of Texas-Austin,UW-Madison, the World Bank, the 2019 Urban Economics Association meeting, and the 2021 NBER Summer Institute for helpful comments. We acknowledge the excellent research assistance from Avralt-Od Purevjav and Ziye Zhang. Barwick and Li acknowledge the financial support from the International Initiative for Impact Evaluation (3ie) under project DPW1.1106 and from the Center for Transportation, Environment, and Community Health at Cornell University. Wu thanks the National Natural Science Foundation of China (72174100 and 71874093) for financial support.

† Go to https://doi.org/10.1257/aer.20220212 to visit the article page for additional materials and author disclosure statement(s).

[1] The TomTom Traffic Index, based on real-time GPS traffic data from 403 cities in 56 countries, shows that the ten most congested cities in 2018 were all from developing and emerging economies. Four cities in China (including Beijing) were among the top 30 list. Beijing's drivers spend nearly 180 extra hours on the road (or 9 percent of

investment, congestion pricing, and gasoline taxes. In the short term, the effectiveness of these policies in alleviating congestion crucially hinges on the substitutability of travel modes and the sensitivity of travel demand to changes in commuting costs. In the medium to long run, these policies are likely to have broader impacts on the urban spatial structure through household adjustment of residential locations. This sorting response, in turn, could mediate the effectiveness of transportation policies on congestion reduction. In addition, many policies that address congestion have distributional consequences. For example, collecting tolls could intensify equity considerations since low-income households spend a larger share of income on transportation. This paper aims to understand the efficiency and equity impacts of urban transportation policies while accounting for sorting responses and endogenous congestion.

Both the theoretical literature (LeRoy and Sonstelie 1983; Brueckner 2007) and empirical analyses (Baum-Snow and Kahn 2000; Jerch et al. 2021) have demonstrated a close connection between transportation policies and the housing market. Motivated by these studies, we develop and estimate an equilibrium model of residential sorting that incorporates preference heterogeneity and allows general equilibrium feedback between housing locations and commuting decisions through endogenous congestion. A key aspect of our model is that congestion, which varies spatially and spills over to other regions, is determined by the commuting modes and residential locations of all households. At the same time, congestion directly affects households' location choices through its impact on the ease of commuting. Once estimated, our model allows us to predict new equilibrium outcomes under different transportation policies in terms of travel mode choices, household locations, congestion level, housing prices, and social welfare.

The empirical context of our study is Beijing, which has a population of 21.5 million and has routinely ranked as one of the most congested and polluted cities in the world. Beijing's municipal government has implemented several policies to aggressively combat traffic congestion and air pollution. It has adopted a driving restriction policy since 2008 that restricts vehicles from driving one weekday per week based on the last digit of the license plate. It also invested a staggering \$100 billion in transportation infrastructure between 2007 and 2018 by adding 16 new subway lines with a total length of 523 km. Beijing's anti-congestion policies (driving restrictions and subway expansion), together with a proposed congestion pricing scheme, represent three broad approaches to regulating the unpriced congestion externality: the first a command-and-control demand-side policy, the second a supply-side policy, and the third a market-based demand-side policy.

Our analysis leverages two unique data sources with fine spatial resolution (street addresses) that allow us to jointly model residential locations and commuting choices. The first dataset is the Beijing Household Travel Survey (BHTS) from 2010 and 2014, a large representative survey that records households' home and work addresses, trips made in a 24-hour window, and other demographic and transportation-related information. We complement this dataset by constructing the historical commuting

working hours) per year relative to the travel time under the free-flow speed. See https://www.tomtom.com/en_gb/traffic-index/ranking. Commuting is one of the most unpleasant uses of time according to subjective well-being assessments (Kahneman et al. 2004).

route, distance, travel time, and pecuniary travel cost for all travel modes (walking, biking, bus, subway, car, or taxi) for each observed trip. The second data source contains housing transactions from a major government-run mortgage program for Beijing residents. Critically for our analysis, the housing data report not only the home addresses but also the work addresses of both the primary and secondary borrowers. Using these home and work addresses, we put together over 13 million hypothetical work-commute and travel-mode combinations for all primary and secondary borrowers. To our knowledge, these datasets constitute the most comprehensive data that link work commutes and housing transactions in the context of equilibrium sorting models.

We estimate the equilibrium sorting model in two steps. First, using household travel surveys, we estimate heterogeneous preferences on travel times and monetary costs, allowing us to recover the distribution of the value of time (VOT). We then utilize the estimated parameters and the work locations of both the primary and secondary borrower to construct an "ease-of-commute" attribute for each commuter in the household and for all properties in the household's choice set. The ease-of-commute attribute reflects the attractiveness of a property's location to its household members' work commutes. These household-property-specific attributes are included in the second step, where we estimate households' preferences for property attributes from observed home purchases. In both steps, we account for a rich set of observed and unobserved household heterogeneity and address potential endogeneity concerns.

The average and median VOT from our preferred specification is 95.6 and 84.6 percent of survey respondents' hourly wage, consistent with estimates in recent literature (Small, Verhoef, and Lindsey 2007; Buchholz et al. 2020; Goldszmidt et al. 2020). Households are willing to pay 18 percent more for an equivalent reduction in the wife's commuting time than the husband's commuting time. We report the income elasticity of housing demand and the income elasticity of the marginal commuting cost, both of which are estimated within a unified framework. They are key determinants of the urban spatial patterns of residential locations.

We next simulate equilibrium residential sorting and transportation outcomes based on three policies of interest: driving restrictions, congestion pricing, and subway expansion, as well as combinations of the three. We decompose the welfare effect along five margins of adjustment. The first margin measures change in welfare when households adjust travel mode in response to increasing commuting costs, holding congestion and residential locations fixed. The second margin considers partial speed adjustments, while the third margin allows the transportation sector to clear; both are commonly studied in the transportation literature. The fourth margin incorporates sorting and simulates general equilibrium outcomes when both the transportation sector and housing market clear. Lastly, we model housing supply adjustments. We are unaware of any prior work in the urban economics literature that combines a structurally estimated model with simulations to decompose all of these margins of adjustment. In addition to decomposing the welfare effect along different margins, we also report welfare benefits from reduced air pollution as a result of less driving.

Our policy simulations yield several key findings. First, while all three policies are designed to reduce congestion, they exhibit different and sometimes opposite impacts on the spatial patterns of residential locations and equilibrium housing prices. The congestion alleviation under the driving restriction disproportionately benefits long commutes and leads to minimal sorting. Conversely, distance-based congestion pricing provides strong incentives for commuters to move closer to their workplaces. In comparison, subway expansion generates the most variable changes in commuting costs across households and triggers the strongest sorting responses, though in the opposite direction of congestion pricing. It disperses households away from the city center and workplaces into locations near new subway stations in the suburbs.

Second, residential sorting can either strengthen or undermine the congestion-reduction potential of transportation policies. Sorting enhances the efficacy of congestion relief for congestion pricing because households, especially those with long commutes, are incentivized to live closer to their work locations and drive less. This magnifies the welfare gain of congestion pricing by as much as 40 percent for high-income households and 16 percent for low-income households. On the other hand, sorting in response to subway expansion leads to further separation between residential and work locations, dampening the congestion reduction and welfare gains from infrastructure investment.

Finally, transportation policies generate different welfare implications in the aggregate and across income groups. Beijing's rapid subway expansion increased consumer surplus and aggregate welfare despite achieving a modest reduction in congestion. In contrast, driving restrictions are welfare-reducing in spite of their larger associated congestion reduction. Congestion pricing is welfare-improving but regressive, creating a significant impediment to its adoption in practice. Nevertheless, such a distributional concern can be addressed via appropriate revenue recycling. Congestion pricing and subway expansion in tandem deliver the largest improvement to traffic speed and net welfare gain—equivalent to 3 percent of the average household's annual income. In addition, the revenue from congestion pricing could fully finance the capital and operating costs of subway expansion, eliminating the need to resort to distortionary taxes. These results showcase the capabilities of our sorting model in capturing various adjustment margins and evaluating different policy scenarios under a unified framework that accounts for general equilibrium effects and preference heterogeneity.

We conduct robustness checks to examine choices of housing demand IVs, measurement errors, and speed endogeneity. In several extensions, we also consider housing supply responses (with a calibrated housing supply elasticity), more granular congestion measures, removing random coefficients, and incorporating migration and consumption access. Incorporating the housing responses allows more people to move to desirable locations and magnifies the role of sorting. Results with more granular congestion measures are similar to the baseline with citywide congestion, as most of Beijing's urban core is severely congested during rush hour. Shutting down random coefficients overestimates the welfare losses from driving restrictions and congestion pricing but underestimates the benefits of subway expansion. Incorporating migration and consumption access does not change the qualitative results of our analysis.

Our study makes three main contributions. First, it is the first study in the empirical sorting literature (Epple and Sieg 1999; Bayer, Ferreira, and McMillan 2007) to jointly model residential locations and travel mode choices and evaluate how these choices simultaneously determine both congestion and distance to work in equilibrium. Most sorting papers (see Kuminoff, Smith, and Timmins 2013 for a review) treat both the distance to work and the level of congestion as exogenous attributes.[2] In our study, the aggregate welfare implications differ qualitatively whether we account for or abstract from endogenous changes in congestion.

Second, our study relates to the recent advances using quantitative spatial equilibrium (QSE) models to explore the role of transportation in urban systems (see Redding and Rossi-Hansberg (2017) for a review). QSE models offer two major improvements over sorting-based approaches such as ours: they can directly incorporate the labor market and model agglomeration forces (Ahlfeldt et al. 2015; Tsivanidis 2023). On the other hand, QSE models have a couple of limitations. Commuting costs are often characterized as iceberg costs and estimated from observed worker flows and wages via gravity equations. In addition, these models rely on tractable distributional assumptions (i.e., Fréchet) with homothetic preferences to ensure analytic tractability. In contrast, we microfound commuting costs by estimating preferences over commuting time (that depends on endogenous congestion) and monetary costs directly from observed mode choices. This allows us to recover the distribution of the value of time (VOT), the single most important object for the welfare effects of transportation policies. In addition, we incorporate nonhomothetic, heterogenous preferences, which matter for equilibrium responses and lead to meaningful changes in welfare evaluations.

A third and final contribution of our paper is that by characterizing the underlying travel and housing choices, our equilibrium sorting framework provides a micro-foundation that bridges the short-run and long-run policy impacts in transportation studies. Studies with a focus on short-run effects often find results that are different or with opposite signs (e.g., Duranton and Turner 2011) from those that examine longer-run implications in terms of travel choices, traffic congestion, and air pollution. As pointed out by Gallego, Montero, and Salas (2013), little has been done to understand the transition from the short- to the medium- or long-run.[3] We illustrate that the equilibrium adjustments in the transportation sector and housing market offset more than one-half of the initial congestion relief of Beijing's subway expansion. Our framework provides a common yardstick to compare actual and counterfactual policies over the medium- to long-run.

The paper proceeds as follows. Section I describes the data and policy background. Section II lays out the equilibrium sorting model and the estimation strategy. The estimation results are presented in Section III. Section IV explains the counterfactual simulation algorithm. Section V examines different transportation policies and compares their welfare consequences. Section VI concludes.

[2] An exception is Kuminoff (2012), which models household decisions in both the work and housing markets and endogenizes the commuting distance but keeps congestion exogenous. In contrast, we treat both congestion and distance to work as endogenous objects that are determined simultaneously in equilibrium.

[3] Our work complements but differs from calibrated computable general equilibrium studies (Yinger 1993; Anas and Kim 1996; Basso and Silva 2014). In our analysis, the estimation of the preference parameters and simulation of policy counterfactuals are internally consistent and based on the same model.

I. Policy Background and Data Description

A. *Policy Background*

The central and municipal governments in China have pursued a series of policies to address growing urban traffic congestion over the past decades. In Beijing, these policies include driving restrictions, vehicle purchase restrictions, and an investment boom in subway and rail transportation infrastructure. The driving restrictions were implemented as part of Beijing's effort to prepare for the 2008 Summer Olympics.[4] Initially, one-half of all vehicles were restricted from driving on a given weekday based on their license plate number. After the Olympics concluded, the restrictions were relaxed to one weekday per week for each vehicle, depending on the last digit of the license plate. Acquiring a second vehicle to avoid the restriction is difficult due to Beijing's quota system that caps the number of new vehicle sales to curb the growth in car ownership. These quotas are allocated via lotteries. The winning odds decreased drastically from 1:10 in early 2012 to nearly 1:2,000 in 2018 as a result of the increasing number of lottery participants and decreasing quota sizes over time (Xiao, Zhou, and Hu 2017; Li, 2018a; Liu, Li, and Shen 2020).[5] These time-series changes in the winning odds provide useful exogenous variation for the housing demand analysis, as we discuss below.

Beyond implementing demand-side policies, the Beijing municipal government also invested heavily in public transportation infrastructure. From 2007 to 2018, 16 new subway lines were built with a combined length of over 500 km (See online Appendix Figure A1 for subway maps over time). By the end of 2019, Beijing had the world's longest and busiest subway system, with a total length of nearly 700 km and daily ridership of over 10 million. This expansion echoed the boom in infrastructure investment across many regions in China. The number of cities with a subway system in mainland China increased from 4 to over 40 between 2000 and 2019, and the total urban rail network reached over 6,700 km by the end of 2019. These expansions were designed, in part, to slow the growth of personal vehicle use by making public transportation more accessible.

Despite these policy efforts, traffic congestion continues to be a pressing issue. The average traffic speed in 2019 was 24.6 km/h during peak hours (7 AM–9 AM and 5 PM–7 PM), according to the 2020 Beijing Transportation Report (Beijing Transport Institute 2020). From a neoclassical perspective, the aforementioned policies fail to address the root cause of traffic congestion: the mispricing of road capacity.[6] The Beijing municipal government recently announced a plan to introduce road pricing

[4] Athens, Greece, implemented the first driving restrictions in 1982. Since then, a dozen other large cities in the world, including Bogotá, Mexico City, and New Delhi, have adopted similar policies. The impacts of these policies on congestion and air pollution have been mixed (Davis 2008; Viard and Fu 2015; Zhang, Lawell, and Umanskaya 2017).

[5] About 20,000 new licenses were distributed each month through nontransferable lotteries from 2011 to 2013. The monthly quota was reduced to 12,000 after 2013. These quotas were considerably lower than the historical vehicle sales in Beijing.

[6] Despite being advocated by economists since Vickrey (1963), congestion pricing is limited in practice due to technical feasibility and especially political acceptability. Several European cities (London, Milan, Stockholm, and Gothenburg) implemented various congestion pricing schemes over the past few decades. The New York state legislature recently approved a congestion pricing plan for New York City, which is set to become the first US city to enact congestion pricing.

in the future while soliciting feedback from experts and the public (Yang, Purevjav, and Li 2020).

B. *Data Description*

We rely on two main datasets for our analysis: (i) the Beijing Household Travel Surveys (BHTS) from 2010 and 2014 and (ii) housing mortgage data from 2006–2014 with detailed information on household demographics and the work addresses of home buyers. The combined commuter- and household-level data provide one of the most comprehensive descriptions of commuting patterns and residential locations in any city to date. Online Appendix Section A provides more details on data construction and how our data compare to other datasets used in recent empirical studies of transportation and urban issues.

Beijing Household Travel Survey.—The two rounds of BHTS were collected by the Beijing Transportation Research Center to inform transportation policies and urban planning. It includes data on individual and household demographics and a travel diary on all trips taken during the preceding 24 hours. The survey reports detailed information for each trip by each commuting member of a household, including the origin and destination, departure and arrival time, trip purpose, and travel mode used.

We focus on six travel modes: *walk*, *bike*, *bus*, *subway*, *car*, and *taxi*, as other modes (e.g., motorcycles, company shuttles, carpooling, multi-modal) collectively account for less than 4 percent of all trips. For each trip, we construct attributes including the travel time, travel distance, and monetary cost for all travel modes in commuters' choice set. The size of the choice set varies across trips depending on vehicle ownership, the driving restriction policy, and bus/subway routes. Online Appendix Section A.3 provides a detailed discussion on the procedures used to calculate the mode-specific travel time and monetary cost for each trip.

Table 1 provides summary statistics of the data. Figure 1 plots for each travel mode the observed share of commuting trips and the constructed travel time, cost, and distance. Panel A presents travel patterns in 2010 and 2014. Walking accounts for a significant share of all commuting trips: 15.0 percent in 2010 and 13.5 percent in 2014, respectively. These trips take 51 and 40 minutes on average, with a distance of 4.9 and 3.7 km. From 2010 to 2014, the shares of walk, bike, and especially bus trips decreased while the shares of car (i.e., driving) and subway trips increased, reflecting rising vehicle ownership and expansion of the subway network. Walking and subway trips are the longest in duration, while subway and car trips are the longest in distance. Car trips have a slightly longer duration and distance than taxi trips but are cheaper. Overall, the trade-off between time and cost is clear: walking trips are the slowest but also the cheapest. Car and taxi trips are faster but more expensive than other trip types. Panel B of Figure 1 contrasts travel patterns between high- and low- (above- and below-median) income households. High-income households are more likely to drive, use the subway, and take taxis and are less likely to use other modes. Car and taxi trips are much more expensive for low-income than for high-income households as a percentage of the hourly wage. There are limited differences in travel distance across the two income groups except for the distance of car trips.

Table 1—Summary Statistics of Household Travel Survey

	2010			2014		
	N	Mean	SD	*N*	Mean	SD
Respondent characteristics						
Income: $<$ ¥ 50k	14,780	0.48	0.50	20,573	0.18	0.38
Income: [¥ 50k, ¥ 100k)	14,780	0.39	0.49	20,573	0.44	0.50
Income: $>=$ ¥ 100k	14,780	0.13	0.34	20,573	0.38	0.49
Having a car (= 1)	14,780	0.44	0.50	20,573	0.62	0.49
Female (= 1)	14,780	0.44	0.50	20,573	0.43	0.50
Age (in years)	14,780	37.59	10.28	20,573	38.47	9.84
College or higher (= 1)	14,780	0.61	0.49	20,573	0.64	0.48
Home within fourth ring (= 1)	14,780	0.51	0.50	20,573	0.41	0.49
Workplace within fourth ring (= 1)	14,780	0.59	0.49	20,573	0.50	0.50
Trip related variables						
Travel time (hour)	30,334	0.87	1.06	42,820	0.74	0.98
Travel cost (¥)	30,334	2.47	5.55	42,820	3.83	6.96
Distance $<$ 2 km	30,334	0.25	0.43	42,820	0.24	0.43
Distance in [2, 5 km)	30,334	0.27	0.45	42,820	0.26	0.44

Notes: The table reports survey respondent demographics and trip attributes of all work commuting trips within the sixth ring road from the 2010 and 2014 Beijing Household Travel Survey. *Travel time* and *travel cost* are constructed as in online Appendix A.3. Trip *distance* is measured by straight-lines. *Distance* $<$ 2 km and *Distance* within 2–5 km flag commuting trips with a short to medium-distance.

Housing Transactions.—The data on housing transactions come from a major government-sponsored mortgage program in Beijing and covers July 2006 to July 2014. Virtually all eligible home buyers apply for mortgages through this program before obtaining commercial loans, as it offers a subsidized interest rate that is more than 30 percent lower than commercial mortgage rates. Each transaction in our data corresponds to a mortgage application, and there are no refinancing loans in the sample.

The final dataset includes 77,696 mortgage transactions, with detailed information on housing attributes such as the property size, age, street address, transaction price, and date when the mortgage was signed.[7] As is reflective of the housing supply in urban China, most housing units are within housing complexes, equivalent to condominiums in the United States. Table 2 provides summary statistics of the data. We observe household demographics including income, age, gender, marital status, residency status (*hukou*), and—critically for our analysis—the work addresses of the primary borrower and the co-borrower if one is present. In our housing demand analysis, we distinguish the borrowers by gender. We geocode the home and work addresses and construct measures of proximate amenities (e.g., schools and parks). The mortgage data represent a subset of housing transactions (not all buyers apply for mortgages) and may be subject to selection issues. To address this concern, we reweight the mortgage data to match the distribution of housing price, size, age, and distance to the city center among more representative housing transactions (obtained from separate datasets) by using entropy balancing (Hainmueller 2012). All of the empirical analyses, as well as the counterfactual analyses, use the weighted sample. Results using the unweighted sample are similar. Online Appendix Section A.4 discusses the reweighting procedure in more detail and describes additional data patterns.

[7] We remove transactions with a missing or zero reported price, a price lower than ¥5,000/m^2 (the average price is ¥19,800/m^2), buyers with no reported income, and addresses outside the sixth ring road.

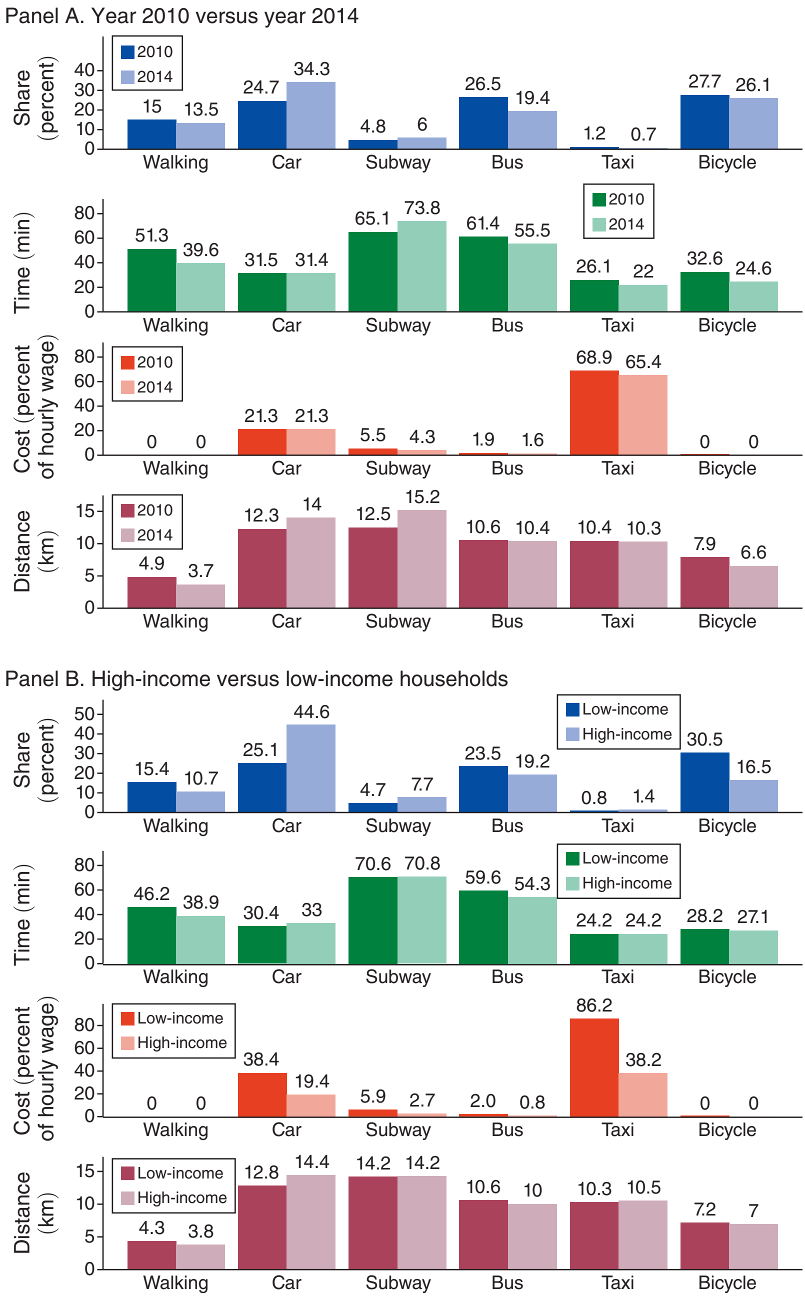

Figure 1—Travel Patterns for Commuting Trips from Beijing Household Travel Survey

Notes: This figure plots the trip share, time, costs, and average distance by travel modes for work commuting trips in the Beijing Household Travel Survey of 2010 and 2014. There are six main trip modes: walk, bike, bus, subway, car, and taxi. Bus and subway trips could include segments with other modes but we characterize them as bus and subway trips. Trips using both bus and subway are rare (less than 3 percent in the data and are dropped in the analysis). Travel time, cost (defined as percent of hourly wage), and distance are constructed as in online Appendix A.3. High-income households are defined as households whose income is greater than the median in the survey year.

Table 2—Summary Statistics of Housing Data

	Mean	SD	Min	Max
Housing attributes				
Transaction year	2011	1.89	2006	2014
Price (¥ 1,000/m^2)	19.83	9.56	5.00	68.18
Unit size (m^2)	92.68	40.13	16.71	400.04
Household annual income (¥1,000)	159.71	103.34	6.24	2,556.90
Primary borrower age	33.99	6.62	20.00	62.00
Housing complex attributes				
Distance to key school (km)	6.05	5.61	0.03	23.59
Complex vintage	2004	8	1952	2017
Green space ratio	0.32	0.06	0.03	0.85
Floor area ratio	2.56	1.12	0.14	16.00
Number of units	1,972	1,521	24	13,031
Home-work travel variables				
Walking distance (km)	14.10	9.51	0.00	62.92
Driving distance (km)	16.13	10.87	0.00	85.22
Home to subway distance (km)	2.13	2.31	0.04	28.37
Subway route distance (km)	15.17	10.70	0.00	68.40

Notes: This table reports statistics from the 2006–2014 mortgage dataset. The number of housing transactions is 79,884, all of which are within the sixth ring road. The dataset is weighted to match the population of all home sales. A housing complex consists of a group of buildings in the same development. Distance to key school is the distance to the nearest signature elementary school. *Home to subway distance* is the distance to the nearest subway station. *Subway route distance* is the distance between the two subway stations that are closest to home and work locations.

Figure 2 shows the spatial pattern of housing and household attributes based on mortgage transactions from 2006 to 2014, with a warmer color representing a higher value. Housing prices tend to be higher, and the distance to work is shorter near the city center. The outskirts of Beijing have larger homes with a lower unit price, reflecting the classic distance-housing size trade-off. There are some exceptions: the high-tech center in northwestern Beijing outside the fifth ring road has high housing prices and short commutes that are comparable to places in the city center. The northern parts of the city have better amenities (schools and parks) and more work opportunities and attract high-income households.

Our equilibrium sorting model requires us to construct home buyers' choice sets. In theory, these choice sets could consist of *all* properties listed on the market. Researchers need to construct hypothetical commuting attributes of different travel modes for all properties in these choice sets. This is technically infeasible because the number of potential home-work-mode combinations exceeds hundreds of billions. Large choice sets are a common empirical challenge in the housing demand literature. To reduce the computational burden, we follow a choice-based sampling strategy as in Bayer, Keohane, and Timmins (2009). The choice set for a household is assumed to include the purchased home and a 1 percent sample of houses randomly chosen from those sold during a two-month window around the purchase date.[8] Section IIIB conducts robustness analyses using different choice

[8]Choice-based sampling has been shown to yield consistent results in Wasi and Keane (2012) and Guevara and Ben-Akiva (2013). We choose a two-month window because Beijing's real estate market was fluid during our sample period. The median (average) number of days on the market was only 8 (22) and 13 (38) in 2013 and 2014, respectively.

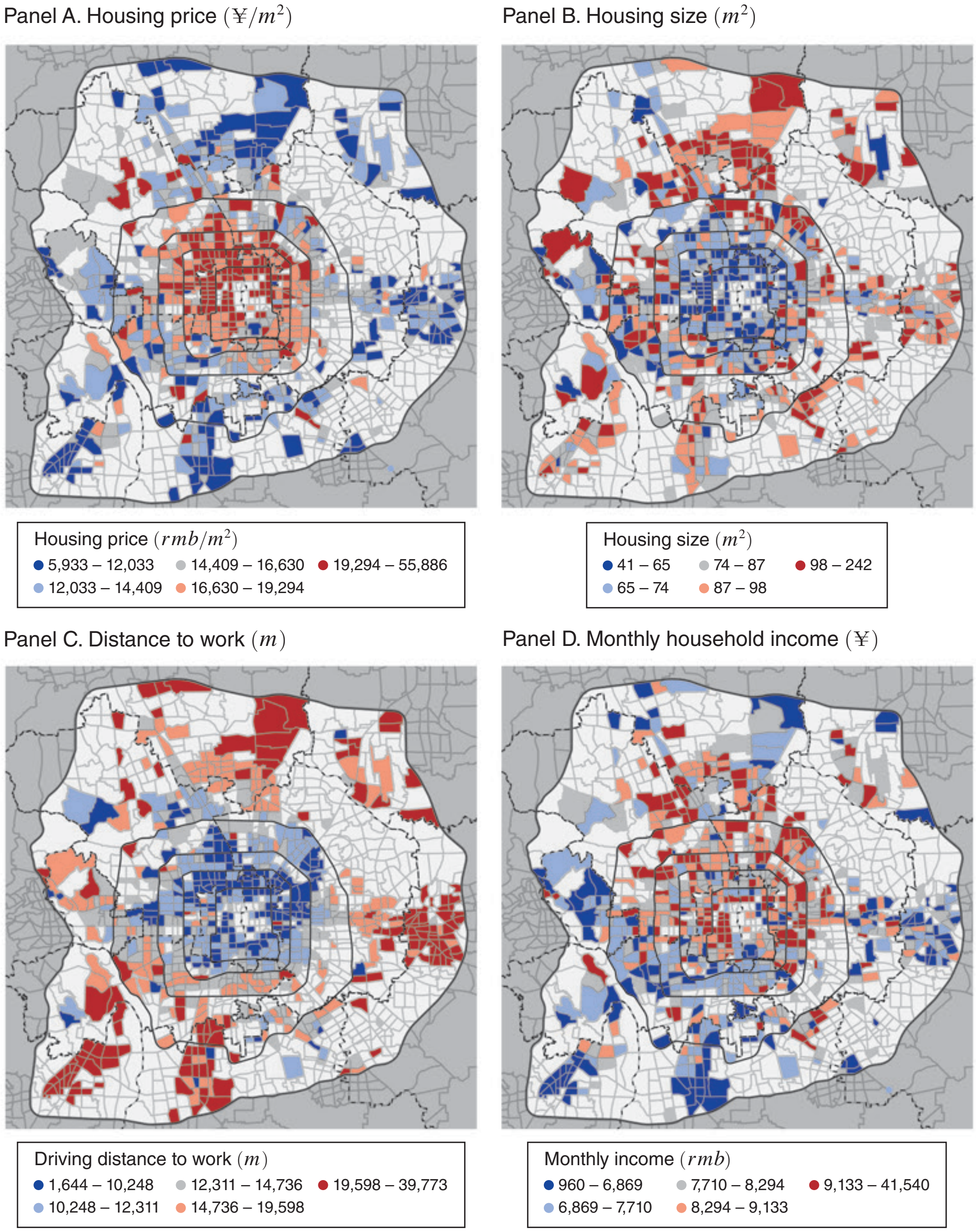

Figure 2. Housing and Household Attributes from Mortgage Data

Notes: This figure plots the average housing price and size and household commuting distance and monthly income by Traffic Analysis Zones (TAZ) based on the 2006–2014 mortgage data. TAZs are standardized spatial units used by transportation planners. There were 2,050 TAZs in Beijing in 2014. Distance to work is the driving distance for all borrowers in the data (including primary and secondary borrowers when both are present). Monthly household income is measured at the time of purchase. Warmer colors (red and orange) correspond to larger values while colder colors (dark and light blue) correspond to lower values. TAZs with no observations are blank.

sets and finds little impact on our estimates. For each property in a household's choice set, we construct the travel mode attributes for both the primary borrower's and co-borrower's work commute. The construction of travel mode attributes involves over 13 million route-mode combinations.

C. *Commuting Route, Speed, and Congestion*

The travel survey does not report the actual commuting routes chosen by individuals. We assume households follow the routes recommended by the Baidu (2019) and Gaode (2019) APIs, which may differ between years (e.g., 2010 and 2014). In counterfactual analyses, commuting routes are held fixed under driving restrictions and congestion pricing and only reoptimized for subway trips with subway expansions. In other words, we do not endogenize commuting routes in the model, which is computationally infeasible with a complex road network.

Driving speeds constructed using Baidu APIs vary by commuting routes and exhibit significant variation among households living in the same neighborhood but traveling to different workplaces, reflecting the varying degrees of congestion in parts of the city.

Congestion is measured by traffic density, constructed as the mileage-weighted number of vehicles on the road. Due to the fluid nature of congestion (which spreads across areas), determining the appropriate level of spatial granularity for measuring congestion is a complex empirical matter. We adopt three measures: (i) citywide congestion, (ii) congestion within each ring-road band, and (iii) congestion within each ring-road quadrant in our policy simulations. Citywide congestion allows for the maximum spillover but may be too coarse. The ring-road-quadrant measure strikes a reasonable balance between the need for localized measures and the complexities associated with the network effects of congestion.[9] To construct region-level congestion, commuting routes are divided into the corresponding segments by region. Congestion within a region equals the sum of mileage-weighted driving demands of individuals from different neighborhoods whose commuting paths cross the focal area. Correspondingly, a driver's commuting time is the sum of driving time for each trip segment within the relevant area. This approach provides a good representation of how congestion affects commuting time and allows congestion in one area to spill over to other areas with a complex road network.

The effect of congestion on speed is governed by the speed-density elasticity, which states the percentage improvement in speed when congestion decreases by one percent. The speed-density relationship embodies Beijing's existing transportation technology. We use the same traffic data as Yang, Purevjav, and Li (2020b) to estimate this relationship, which contain real-time traffic volume and speed at two-minute intervals (aggregated to hourly intervals) from over 1,500 remote traffic sensors covering all major roads throughout Beijing in 2014. Results in Section IIIC indicate limited heterogeneity in speed-density elasticity across regions during rush hour.

[9] There are 15 ring-road-quadrant areas based on the intersection of ring roads and quadrants (NW, SW, NE, SE, see Figure A2).

D. *Reduced-Form Evidence*

Due to space constraints, we refer readers to the event studies in a concurrent study Jerch et al. (2021) that demonstrate the impact of Beijing's driving restriction policy on the housing market. Online Appendix Section B.1 presents additional evidence on sorting where neighborhoods that gain access to new subway lines account for a greater fraction of aggregate property transactions, suggesting that Beijing households actively sort across residential areas in response to transportation policies. We now turn to a structural model integrating the transportation sector with the housing market.

II. Empirical Equilibrium Sorting Model

The sorting model characterizes how household members choose commuting modes and residential locations.[10] It also specifies the joint equilibrium conditions for the transportation sector and the housing market. On the one hand, residential locations determine households' commute distances, which affect driving demand and contribute to traffic congestion. On the other hand, traffic congestion impacts the desirability of different residential locations and directly influences housing demand. Once estimated, the model allows counterfactual simulations and provides a direct comparison of congestion levels, residential locations, housing prices, and social welfare across different policies.

Work locations are fixed ex ante and do not change. This assumption is motivated by several observations. First, event studies in online Appendix Section B.2 suggest that job changes tend to predate home purchases, not the other way around.[11] Second, employment opportunities in the same industry tend to be clustered in Beijing. Hence, switching jobs may not entail meaningful changes in work locations. Third, the mortgage data only report current employment but contain no information on alternative job opportunities available to each household. Finally, adding the labor market component would significantly complicate our empirical analysis with rich individual-level observed and unobserved preference heterogeneity.

A. *Housing Demand*

We specify a characteristic-based housing demand model, where preferences over housing units are parameterized as a function of both observed and unobserved property attributes and household characteristics (Berry, Levinsohn, and Pakes 1995). Households are denoted by i and a housing unit/property is indexed by j. A home or work location refers to a specific address. Our data are longitudinal, but we suppress time t to ease exposition. Online Appendix Table A1 tabulates all mathematical notations used in the order of appearance.

Conditioning on work locations, the utility for household i choosing housing unit j is specified as

$$\max_{\{j \in \mathcal{J}_i\}} U_{ij} = \alpha_i p_j + \mathbf{x}_j' \boldsymbol{\beta}_i + \sum_k \phi_{ik} EV_{ijk}(v_{ijk}) + \xi_j + \varepsilon_{ij}, \tag{1}$$

[10] A more detailed version of the model can be found in Barwick et al. (2021).

[11] About 60 percent of home buyers in our data changed jobs within three years prior to purchasing the home.

where $\mathcal{J}_i$ is the choice set for household i, p_j denotes home price, and $\mathbf{x}_j$ denotes a vector of observed housing attributes such as property size and age.[12] Commuting members within household i are denoted by $k \in \{male\ borrower, female\ borrower\}$.[13] $EV_{ijk}(v_{ijk})$ is the expected commuting utility that member k in household i derives from the optimal commuting mode. It characterizes home j's attractiveness in terms of member k's work commute. Our notation makes it explicit that the commuting utility depends on the driving speed v_{ijk} of member k's work commute (which is affected by congestion) in addition to the travel cost. As shown in Section IIB below, this commuting utility is our key innovation relative to traditional residential sorting models. Transportation policies generate different impacts on the commuting utility. They can affect either commuting time (such as driving restrictions and subway expansion) or costs (such as congestion pricing), which interact with households' heterogeneous commuting preferences. These changes then influence individuals' travel modes, which collectively determine congestion and further affect their commuting utility. The variable ξ_j represents unobserved housing attributes, and ε_{ij} is an i.i.d. error term with the type I extreme value distribution that reflects unobserved preferences over each housing choice.

The household-specific price coefficient α_i is related to the logarithm of household income y_i:

$$\alpha_i = \alpha_1 + \alpha_2 \times \ln(y_i).$$

Household preferences over housing attributes are denoted as $\boldsymbol{\beta}_i$, which consists of a household-specific component and a population average. For each element ℓ in β_i:

$$\beta_{i\ell} = \bar{\beta}_\ell + \mathbf{z}_i' \boldsymbol{\beta}_\ell,$$

where $\mathbf{z}_i$ is household demographics such as age and income. The ease-of-commute preference ϕ_{ik} differs across household members and is characterized by random coefficients:

$$\phi_{ik} = \bar{\phi}_k + \phi_k \zeta_{ik}, \quad k \in \{Male\ borrower, Female\ borrower\},$$

where ζ_{ik} is i.i.d. normal. The probability that household i chooses home j is denoted by

$$P_{ij}(\mathbf{p}, \mathbf{v}) = h(\mathbf{EV}(\mathbf{v}), \mathbf{p}, \mathbf{x}, \boldsymbol{\xi}, \mathbf{z}_i), \tag{2}$$

where $\mathbf{p}$ and $\mathbf{v}$ denote the vector of prices for all properties and driving speeds for all individuals' commuting trips, and $\mathbf{EV}(\mathbf{v})$ is a vector of the ease-of-commute utility for different properties given household i's work location. The triplet $\mathbf{x}$, $\boldsymbol{\xi}$, and $\mathbf{z}_i$

[12] As is standard in the sorting literature, we do not consider an outside option and assume all households live in a property.

[13] The fraction of male and female borrowers in our sample is 60 percent and 40 percent, respectively.

denote observed housing attributes, unobserved housing quality, and household i's demographics, respectively.[14]

B. *Choice of Travel Mode*

Utility-maximizing individuals in a household choose from six commuting modes (walk, bike, bus, subway, car, and taxi) based on the trip time and financial costs. With slight abuse of notation, we use i to denote an individual in a household rather than the whole household in this subsection. Individual i's utility of commuting from home j to work using mode choice m is specified as

$$(3) \qquad \max_{m \in \mathcal{M}_{ij}} u_{ijm} = \theta_{im} + \gamma_{1i} \cdot time_{ijm}(v_{ij}) + \gamma_2 \cdot cost_{ijm}/y_i + \mathbf{w}'_{ijm}\boldsymbol{\eta} + \varepsilon_{ijm},$$

where $\mathcal{M}_{ij}$ is the set of transportation modes available to individual i's work commute.[15]The mode-specific random coefficients, θ_{im}, have a normal distribution with mean μ_m and variance σ_m. Without loss of generality, the random coefficient for walking is normalized to zero. These random coefficients capture unobserved heterogeneous preferences that vary across individuals, such as the enjoyment of driving a car, the perceived environmental friendliness of using public transportation, scheduling or inconvenience costs that vary across individuals but do not scale with the time or distance traveled, and the health benefits of biking and walking.

Variable $time_{ijm}$ denotes the commute duration between i's work location and home j via mode m. The driving time ($time_{ij,car}$ and $time_{ij,taxi}$) depends on the driving speed v_{ij}, which is ultimately determined by the congestion level.[16] The monetary cost of the trip is denoted as $cost_{ijm}$ and household income as y_i. The variable $\mathbf{w}_{ijm}$ includes a rich set of interactions between mode dummies and year-fixed effects, trip attributes, and commuter demographics to control for time-varying and location-specific factors by travel mode (such as changes in public transportation) as well as preference heterogeneity. Finally, ε_{ijm} is an i.i.d. error term with a type I extreme value distribution.

The key parameters in the travel demand analysis are γ_{1i} and γ_2. The time preference γ_{1i} follows a chi-squared distribution with mean μ_γ as in Petrin (2002).[17] An individual's sensitivity to the monetary costs of commuting is assumed to decrease in income: γ_2/y_i. VOT, the most important preference parameter for transportation decisions (Small 2012), is measured by $\frac{\gamma_{1i}}{\gamma_2} \cdot y_i$ and directly linked with the hourly wage. Our specification of VOT follows the transportation literature. Its theoretical foundation is the time allocation models of Becker (1965) and Small (1982), where

[14] Equation (1) does not control for neighborhood composition, such as race and ethnicity (Bayer, Ferreira, and McMillan 2007), as 96 percent of Beijing's population is Han Chinese (2010 census). In addition, as Figure 2 illustrates, Beijing is characterized by co-mingling among households with different socioeconomic statuses and exhibits high income heterogeneity within small neighborhoods.

[15] The size of the choice set varies across commuters and trips. Walk, bike, taxi, and subway are available for all trips. Driving is only available for car owners on nonrestricted days. Bus availability is determined by home and work locations.

[16] Road congestion affects the travel time for bus trips in addition to car and taxi trips. We treat buses as if they run in dedicated lanes unaffected by congestion. We also abstract away from capacity constraints for buses and subways. Hence, we may overpredict bus and subway mode shares in simulations with high congestion levels.

[17] A chi-squared distribution ensures all individuals have a positive value of time and is computationally tractable.

the value of time spent on commuting is a function of the lost wage income and scheduling preferences.

Conditional on home location j, the probability that individual i chooses mode m to commute to work is

$$R_{ijm}(v_{ij}) = r(\mathbf{time}_{ij}(v_{ij}), \mathbf{cost}_{ij}/y_i, \mathbf{w}_{ijm}), \tag{4}$$

where $\mathbf{time}_{ij}(v_{ij})$ and $\mathbf{cost}_{ij}/y_i$ denote the vector of travel time and cost (as a share of individual i's hourly wage) for all travel modes. Vector $\mathbf{w}_{ijm}$ captures all other individual- and trip-mode-specific characteristics.

The ex ante expected commuting utility (before the realization of travel shocks) is defined as

$$\begin{aligned} EV_{ij}(v_{ij}) &= \mathbb{E}_{\varepsilon_{ijm}}\left[\max_{m \in \mathcal{M}_{ij}} u_{ijm}(v_{ij})\right] \\ &= \log\left(\sum_{m \in \mathcal{M}_{ij}} \exp\left\{\theta_{im} + \gamma_{1i}\, time_{ijm}(v_{ij}) + \gamma_2\, cost_{ijm}/y_i + \mathbf{w}'_{ijm}\boldsymbol{\eta}\right\}\right), \end{aligned} \tag{5}$$

which is the ease-of-commute variable discussed in equation (1).

C. *Market-Clearing Conditions and the Sorting Equilibrium*

In the housing market, choices of individual households aggregate to total housing demand, and housing prices adjust to equate demand and supply. In the transportation sector, the equilibrium congestion level and, hence, driving speed is jointly determined by driving demand through all individuals' travel mode choices and road capacity. The housing market and the transportation sector interact in two dimensions: the spatial locations of households affect the distance of work commutes and the choice of travel mode and hence congestion and driving speeds in the transportation sector. At the same time, the level of traffic congestion that is determined in the transportation sector affects the attractiveness of residential locations through the commuting utility as discussed above, which, in turn, determines households' sorting decisions and shapes their spatial distribution.

Housing Market.—Households' choice probabilities P_{ij} give rise to the aggregate housing demand:

$$D_j(\mathbf{p}, \mathbf{v}) = \sum_i P_{ij}(\mathbf{p}, \mathbf{v}), \quad \forall j,$$

which depends on the vector of housing prices $\mathbf{p}$ as well as the vector of the driving speeds $\mathbf{v}$ (through the ease-of-commute utility). We consider two scenarios for housing supply. In the first case, the housing supply is fixed: $S_j(\mathbf{p}) = 1$ (the supply of each property is one). In the second case, housing supply has a constant elasticity and adjusts at the neighborhood level in response to the average price within the neighborhood. This assumption mimics developers' considerations to build more properties in desirable neighborhoods.

Transportation Sector.—Demand for driving is determined by both housing locations and travel mode choices. Intuitively, mode choices determine the extensive margin (whether to drive), while housing locations determine the intensive margin (the commuting distance). Traffic density, or congestion, is the aggregation over all households' driving demand (the notation makes it explicit that traffic density depends on the vector of housing prices $\mathbf{p}$ and driving speeds $\mathbf{v}$):

$$(6)\quad D_{T,r}(\mathbf{p},\mathbf{v}) \equiv \sum_i \sum_j \mathbf{1}\big\{\{i \rightarrow j\} \cap r\big\} \cdot P_{ij}(\mathbf{p},\mathbf{v}) \cdot \Big(\big[R_{ij,car}(\mathbf{v}) \cdot \mathrm{dist}_{ijr,car}\big] + \big[R_{ij,taxi}(\mathbf{v}) \cdot \mathrm{dist}_{ijr,taxi}\big]\Big).$$

The subscript T denotes the transportation sector, and r defines the spatial granularity of the traffic density measure. The indicator function $\mathbf{1}\big\{\{i \rightarrow j\} \cap r\big\}$ takes value 1 when the route from i to j intersects region r. P_{ij} is the probability that household i chooses property j. $R_{ij,car}$ and $R_{ij,taxi}$ are the probabilities that household i living in property j drives and takes taxi, and $\mathrm{dist}_{ijr,car}$ and $\mathrm{dist}_{ijr,taxi}$ are the distance traveled by car and taxi within region r. Given the empirical nature of the appropriate level of spatial granularity for traffic density/congestion r, we consider three levels: citywide congestion, congestion at the ring-road-band level, and congestion at the ring-road-quadrant level as discussed in Section IC.

The supply side of the transportation sector describes the relationship between the traffic density $S_{T,r}$, the number of vehicles on the road in region r, and the travel speed $\mathbf{v}$ that can be sustained given Beijing's transportation technology and road capacity. We assume the speed and density relationship has a constant elasticity that differs across regions. This supply relationship characterizes the nature of congestion externality: drivers on the road reduce other drivers' speed.

Sorting Equilibrium.—A sorting equilibrium is defined as a vector of housing prices, $\mathbf{p}^*$, and a vector of driving speeds, $\mathbf{v}^*$, such that:

(i) The housing market clears for all properties:

$$(7)\qquad D_j = \sum_i P_{ij}(\mathbf{p}^*,\mathbf{v}^*) = S_j(\mathbf{p}^*), \qquad \forall j.$$

When housing supply adjusts at the neighborhood level, demand and supply are equal for each neighborhood n: $\sum_{j\in n} D_j(\mathbf{p}^*,\mathbf{v}^*) = S_n(\mathbf{p}^*), \quad \forall n.$

(ii) The transportation sector clears for every region r, where aggregate driving demand at speeds $\mathbf{v}^*$ is equal to the traffic density that can be sustained under the existing road capacity at speed $\mathbf{v}^*$:

$$(8)\qquad D_{T,r}(\mathbf{p}^*,\mathbf{v}^*) = S_{T,r}(\mathbf{v}^*), \qquad \forall r.$$

Our model follows the class of equilibrium sorting models with local spillovers studied in Bayer and Timmins (2005) and more closely in Bayer, Ferreira, and McMillan (2007), where the local spillover in our context is traffic congestion from personal vehicles. If the error terms in both the housing demand equation (1) and commuting mode choice equation (3) are from continuous distributions, then the system of equations (2), (4), (7), and (8) is continuous. The existence of a sorting equilibrium follows Brouwer's fixed point theorem. Intuitively, a unique vector of housing prices (up to a constant) $\mathbf{p}$ solves the system of equations defined by equations (2) and (7), conditional on traffic speed $\mathbf{v}$. At the same time, equations (4) and (8) define a continuous mapping of traffic speed $\mathbf{v}$ on a compact and convex set given $\mathbf{p}$. The fixed point of the system of equations (2), (4), (7), and (8) defines the equilibrium housing prices and traffic speeds $\{\mathbf{p}^*, \mathbf{v}^*\}$.[18]

III. Estimation Results

This section discusses how we estimate the parameters in the travel mode choice, housing demand, and the speed-traffic density relationship. Online Appendix Section C includes further details.

A. *Commuting Mode Choice*

The parameters of the travel mode choices are estimated via simulated maximum likelihood estimation (MLE) using household travel surveys. The key parameters of interest are time and monetary cost preferences. We assume that the error term ε_{ijm} in equation (3) is uncorrelated with commuting trips' monetary costs and travel time. Monetary costs are likely to be exogenous because Beijing's transportation bureau sets bus and subway fares uniformly across all routes. Gas prices are determined by the National Development and Reform Council and adjusted periodically, and taxi fares are regulated by the Beijing government.[19] Hence, monetary costs do not vary by the congestion level or service quality. Travel time is determined by congestion. We include a rich set of interactions of travel mode with demographics, time, and spatial fixed effects to absorb shocks that are common across households. Given these controls, ε_{ijm} reflects idiosyncratic considerations at the individual home-workplace-mode level that are unlikely to be correlated with travel time.

Table 3 presents parameter estimates for six specifications. The first three specifications control for demographics but do not have random coefficients. The last three specifications include random coefficients on travel time and travel mode dummies. Column 1 controls for interactions between the year dummies (2010 or 2014) and mode fixed effects to capture changes in public transportation service over time. The implied VOT is 75.7 percent of the hourly wage. Column 2 adds the interactions between mode fixed effects and trip characteristics, which include trip distance bins ($\leq$ 2 km, 2–5 km, $\geq$ 5 km), origin ring road fixed effects (e.g., whether the origin is within the fourth ring road), and destination ring road fixed effects. These controls

[18] The proof of equilibrium existence closely follows Bayer and Timmins (2005) and is available upon request. In our model with spatially varying congestion responses and preference heterogeneity for endogenous attributes, uniqueness is not guaranteed. One sufficient condition for a unique equilibrium requires exogenous attributes of housing and commuting to be "sufficiently explanatory" of demand relative to endogenous ones, as pointed out by Bayer, Ferreira, and McMillan (2007). To address the possibility of multiple equilibria, we simulate our model with 100 different starting values. The simulation analyses always converge to the same equilibrium outcomes, providing empirical evidence for uniqueness in our setting.

[19] For gas price regulations, see page 8 of https://www.globalpetrolprices.com/articles/43/. For taxi fare regulations, see http://fgw.beijing.gov.cn/bmcx/djcx/cxldj/202003/t20200331_1752789.htm.

TABLE 3—ESTIMATION RESULTS FOR TRAVEL MODE CHOICES

	logit			Random coefficient		
	(1)	(2)	(3)	(4)	(5)	(6)
Travel time (γ_1)	−1.194	−0.270	−0.191			
	(0.082)	(0.006)	(0.006)			
Travel cost/hourly wage (γ_2)	−1.578	−0.788	−0.565	−1.411	−1.424	−2.531
	(0.324)	(0.028)	(0.034)	(0.041)	(0.052)	(0.065)
Random coefficients on travel time (μ_γ)						
Travel Time				−0.955	−0.885	−0.931
				(0.008)	(0.008)	(0.012)
Random coefficients on mode dummies (σ_m)						
Driving					3.394	3.391
					(0.049)	(0.054)
Subway						4.470
						(0.142)
Bus						3.851
						(0.056)
Bike						3.887
						(0.054)
Taxi						4.203
						(0.353)
Mode × year FE	Yes	Yes	Yes	Yes	Yes	Yes
Mode × trip related FE		Yes	Yes	Yes	Yes	Yes
Mode × demographic FE			Yes	Yes	Yes	Yes
log-likelihood	−116,287	−109,929	−91,119	−87,353	−85,099	−77,706
Implied mean VOT	0.757	0.342	0.339	1.760	1.615	0.956
Implied median VOT	0.757	0.342	0.339	1.557	1.429	0.846

Notes: The number of observations is 73,154. The specifications include an increasingly rich set of fixed effects interacted with travel model dummies. Trip related FEs include trip distance bins (e.g., whether the distance is shorter than 2 km or between 2–5 km), origin ring road dummies (e.g., whether the trip origin is within the fourth ring road), and destination ring road dummies. Demographics FEs include a respondent's age, age squared, gender, education, car ownership, and whether the household has more than one commuter. The first three specifications are multinomial logit while the last three add random coefficients. The distribution of preference on travel time is specified as a chi-square distribution (winsorized at the fifth and ninety-fifth percentiles) with three degrees of freedom to allow for long tails. The random coefficients on travel mode dummies (driving, subway, bus, bike, and taxi) are assumed to have a normal distribution with a standard deviation of σ_m. The last two rows report the implied mean and median value of time (VOT). Robust standard errors based on the sandwich formula are reported in parentheses.

account for local shocks to travel demand and significantly improve the model fit. For example, the literature has documented that drivers value the reliability of travel time (Small, Winston, and Yan 2005). Uncertainty in travel time likely scales with the trip distance and is partially absorbed by the mode and trip-distance bin fixed effects. Ring road dummies for trip origins and destinations capture differences in the frequency and quality of public transit services as well as congestion. Column 3 further includes interactions of mode fixed effects with household demographics (age, age squared, gender, education, vehicle ownership, and number of workers). These variables help explain different mode choices across demographic groups (e.g., wealthier households' greater propensity of driving and using taxis) and improve the model fit.

Columns 4 to 6 use a chi-squared distribution with three degrees of freedom to approximate heterogeneous travel time preferences following Petrin (2002).[20] In

[20] As in Petrin (2002), we winsorize the top and bottom 5 percent of the draws to minimize the impact of extreme draws, as VOT should be finite. The distribution with three degrees of freedom (DF) provides the best fit, though results are similar with two or four DF.

addition to the random coefficient on travel time, column 5 incorporates a random coefficient on the mode of driving. Column 6 further includes random coefficients for all travel modes (with walking as the reference group), capturing the impact of unobserved preferences on mode choices. For example, some commuters choose driving or taxi not because of a high VOT but because of scheduling constraints. The dispersion of these preference parameters is economically large and statistically significant, suggesting significant preference heterogeneity.

Column 6 is our preferred specification. Adding travel time and mode-specific random coefficients leads to a stronger sensitivity to travel costs and delivers a much more reasonable estimate of VOT. Online Appendix Figure A3 depicts the VOT estimate histogram. The average and median VOT is 95.6 and 84.6 percent of the hourly wage, respectively, which is within the range typically found in the recent literature.[21]

Once we have estimated parameters from travel mode choices, we construct the commuting utility EV_{ij} as defined in equation (5) for both the male borrowers and female borrowers based on their work locations.[22] These variables are included as part of the (buyer-specific) housing attributes.

Identification and Robustness.—The identification of the preference parameters follows the standard identification arguments of random coefficient models. The parameters are identified by the variation in commuters' characteristics and route attributes, as well as the correlation between these attributes and the chosen travel mode. Mode-specific random coefficients are identified from differences in choice sets across individuals (e.g., some do not have easy access to public transportation). Additionally, the parametric assumptions on the functional form and distributions also contribute to the identification.

One common issue encountered in the estimation of travel mode choices pertains to the simultaneous relationship between equilibrium mode choices and travel times. If households in a neighborhood share similar preferences for driving, it would result in a high driving share and at the same time low driving speed due to congestion. Table 3 incorporates a rich set of interaction terms between modes, trip attributes, and household characteristics to address this. To further investigate the issue of simultaneity, we include the interaction of driving with fine spatial controls in online Appendix Table A2. These regressors absorb correlated preferences in local areas and alleviate endogeneity concerns. Reassuringly, the resulting estimates closely resemble those of the baseline. The parameters for travel time, costs, and VOT remain unchanged even with these fine spatial controls. The model's overall fit only improves marginally. These results suggest that the extensive set of controls included in our baseline model adequately addresses potential endogeneity concerns.

[21] The VOT estimates typically range between 30 and 100 percent of hourly income (Small, Verhoef, and Lindsey 2007). Using a discrete choice framework similar to ours, Small, Winston, and Yan (2005) estimate the median VOT at 93 percent of the hourly wage for commuters in Los Angeles. The US Department of Transportation recommends using 50 percent of the hourly income as the VOT for local *personal* trips (e.g., work commute and leisure but not business trips) to estimate the value of time savings for transportation projects (USDOT 2015).

[22] We set $EV_{ijk} = 0$ for unemployed family members. Note that the calculation of EV_{ij} requires us to construct the travel time and cost for all available travel modes for every property in households' choice sets, as described in Section IB.

Apart from the conventional concern of endogeneity, an additional, more nuanced issue arises regarding measurement. During the survey years (2010 and 2014), real-time GPS applications were not widely accessible and individuals were generally unaware of idiosyncratic factors that impacted travel time when selecting commuting mode. This mitigates the simultaneity concern, consistent with the findings in online Appendix Table A2. On the other hand, households were likely making decisions based on *anticipated* travel times rather than the actual travel times or the travel times we constructed using Baidu/Gaode.

To address measurement errors in travel times, we conducted several robustness checks beyond controlling for a rich set of trip-related fixed effects. First, we construct alternative travel time variables based on the average speed either at the ring-road-band level or at the ring-road-quadrant level. The average local speed might better reflect households' expectations. Both parameter estimates and VOT are comparable to the baseline (online Appendix Table A3). Second, we repeat the exercise using self-reported travel times for the chosen mode in online Appendix Table A4. The VOT estimate (at 124 percent of the hourly wage) is higher than the baseline (at 96 percent). This is driven by the fact that self-reported values tend to underestimate the actual travel times, the so-called recall biases as shown in online Appendix Section A.3, and therefore inflate the implied VOT. Overall, our findings remain robust to measurement errors.

B. *Housing Location Choice*

We now turn to the estimation of housing demand using the mortgage data. The ease-of-commute variable EV_{ij} that is derived from travel mode choices enters the housing demand equation (1) as an observed housing attribute. Similar approaches that nest the expected utility as a choice attribute have been used by Capps, Dranove, and Satterthwaite (2003) and Phaneuf et al. (2008) to estimate healthcare and recreational demand, respectively, though the application to residential sorting is new to the best of our knowledge. Because $\widehat{EV}_{ij}$ is estimated separately, we bootstrap the standard errors for housing demand parameters.

Nonlinear Parameters.—Housing demand is estimated using a two-step procedure: the first step uses simulated MLE with a nested contraction mapping to estimate household-specific preference parameters (nonlinear parameters), and the second step uses linear IV for coefficients in the mean utility (linear parameters). Specifically, we reorganize household i's utility of purchasing property j into a sum of household-specific utility μ_{ij} and population-average utility δ_j, which absorbs the unobserved housing attribute ξ_j (we suppress time subscript t to ease exposition):

$$U_{ij} = \mu_{ij} + \delta_j + \varepsilon_{ij} \tag{9}$$

$$\mu_{ij} = \alpha_2 \ln(y_i) p_j + \sum_l x_{jl} \cdot \mathbf{z}_i' \boldsymbol{\beta}_l + \sum_k \phi_{ik} EV_{ijk} \tag{10}$$

$$\delta_j = \alpha_1 p_j + \mathbf{x}_j' \bar{\boldsymbol{\beta}} + \xi_j. \tag{11}$$

Table 4—Housing Demand-Nonlinear Parameters from Simulated MLE

	No EV (1)		With EV (2)		EV and random coef. (3)	
	Para	SE	Para	SE	Para	SE
Demographic interactions						
Price (¥mill.) × ln(*income*)	0.965	0.007	1.005	0.014	1.030	0.016
Age in 30–45 × ln(*distance to key school*)	−0.329	0.004	−0.391	0.011	−0.420	0.013
Age > 45 × ln(*distance to key school*)	−0.074	0.009	−0.111	0.025	−0.123	0.026
Age in 30–45 × ln(*home size*)	1.343	0.014	1.443	0.026	1.486	0.030
Age > 45 × ln(*home size*)	2.394	0.028	2.665	0.070	2.746	0.060
EV_{Male}			0.708	0.015	0.755	0.016
EV_{Female}			0.833	0.017	0.893	0.019
Random coefficients						
$\sigma(EV_{Male})$					0.379	0.019
$\sigma(EV_{Female})$					0.482	0.018
log-likelihood	−206,829		−170,057		−168,808	

Notes: This table reports MLE estimates of the nonlinear parameters in housing demand using mortgage data from 2006–2014 with 77,696 observations. The ease-of-commuting utility (EV) is constructed using column 6 of Table 3 via equation (5). The first specification does not include EV, the second specification does, and the third specification further incorporates random coefficients on EV terms. Standard errors are bootstrapped to account for the estimation errors in EV terms.

In the first step, we search for nonlinear parameters in equation (10) to maximize simulated MLE while inverting the population-average utilities δ_j. In the second step, we regress the population-average utilities δ_j on prices instrumented by IVs to recover linear parameters.

Table 4 reports three specifications: without the EV terms (the ease-of-commute utility), with the EV terms, and with random coefficients on the EV terms. The coefficient estimates are similar across specifications. As expected, high-income households tend to be less price-sensitive.[23] We interact the age group dummies with the distance to the nearest signature elementary school. Enrollment in these top schools is restricted to residents in the corresponding school district, and houses in these districts command a high premium. The baseline group is primary borrowers younger than age 30. The interaction coefficients in all specifications are negative and highly significant, though borrowers between ages 30 and 45 exhibit the strongest preference for proximity to key schools, as they are most likely to have school-aged children. Household size is not reported in the mortgage data. Instead, we use the age of the primary borrower as a proxy for household size and interact age group dummies with the property size. Older households have a stronger preference for large houses. The group over 45 has the strongest large-house preference, likely due to the presence of children and grandparents in the same household, a common household structure in China.

The EV terms for both household members have significant explanatory power and are associated with a sizable increase in the log-likelihood. Both working family members prefer homes with easier commutes. To evaluate households' willingness to pay (WTP) for a one minute shorter commute, we search for changes in housing prices that would keep households' utility constant. According to our

[23] The price coefficient is $\alpha_1 + \alpha_2 \times \ln(y_i)$. Since α_1 is negative, a positive α_2 means the absolute level of price sensitivity is lower for higher-income households.

preferred specification in column 3, an average household is willing to pay an additional ¥18,525 for a home that shortens the male member's daily work commute by one minute and ¥21,885 for a similar reduction in the female member's commute time.[24, 25] The gap suggests that households prioritize the convenience of female members' commute in housing choices, consistent with descriptive evidence that women live closer to work locations (online Appendix Figure A4) and the existing literature (Le Barbanchon, Rathelot, and Roulet 2020).

Household preferences for shorter commutes vary significantly. The interquartile range of WTP for a one-minute reduction in the male member's commute is ¥11,000 and ¥24,400, and for the female member, it is ¥15,000 and ¥34,500. Both demographic factors and random coefficients reflecting unobserved preferences contribute to this preference heterogeneity. Utilizing estimates from column 3 of Table 4 but fixing the random coefficients for *EV* terms at the mean, the interquartile range narrows to [¥14,800, ¥21,000] for males and [¥17,400, ¥24,700] for females, about a 53 percent reduction for males and 63 percent for females.

Linear Parameters.—Table 5 reports the coefficient estimates on the population-average utility in equation (11) for the specification in column 3 of Table 4. Columns 1 and 2 use OLS while columns 3–6 use IVs. All regressions include the month-of-sample by district fixed effects to capture time-varying changes in local market conditions and amenities that could vary across districts in Beijing. Columns 2–6 also include neighborhood fixed effects to capture unobserved time-invariant neighborhood amenities.

We use three sets of IVs for housing prices: (i) the number of properties that are located in a different complex and within 3 km of unit j and sold within a two-month window around property j's sale; (ii) the average attributes of these properties; (iii) and the interaction between the average attributes and the odds of winning the license plate lottery (see Section IA for its policy background). The first two sets of IVs are sometimes called "donut instruments" in the housing literature (Bayer, Ferreira, and McMillan 2007), because the instruments are constructed from properties that are located between concentric circles around a given property. Our preferred specification is column 6, with a first-stage F-statistic of 14.2.

The price coefficient estimate is negative and statistically significant across all columns. The IV estimates are larger in magnitude than the OLS estimates, consistent with the finding in the demand literature that unobserved product attributes bias OLS estimates toward zero. The average price elasticities derived from the OLS estimates are of the wrong sign, as the population average price coefficient is not negative enough to offset the positive coefficient of the income-price interaction. The signs on the other coefficient estimates from the IV regressions in columns 3–6

[24] This is derived by $\frac{\phi_{ik}\partial EV_{ijk}}{\partial\, travel\ time_{ijm}} \times \frac{1}{\alpha_i} \times 10^6$, as housing price is measured in millions of yuan.

[25] The value of one-minute reduction in commute time is approximately ¥1.1, given that VOT is estimated to be 96 percent of the hourly wage (at ¥67.6 on average). In order for the WTP estimates based on housing demand to align with the VOT estimate, households would need to expect an average of 500 commuting trips per year for a duration of thirty to forty years.

Table 5—Housing Demand-Linear Parameters

	OLS (1)	OLS (2)	IV1 (3)	IV2 (4)	IV2 + IV3 (5)	ALL (6)
Price (¥mill.)	−2.24	−2.191	−7.091	−6.283	−6.454	−6.596
	(0.186)	(0.184)	(1.640)	(0.867)	(0.583)	(0.534)
ln(*home size*)	−3.648	−3.797	4.721	3.331	3.631	3.879
	(0.257)	(0.261)	(2.927)	(1.505)	(1.022)	(0.969)
Building age	−0.043	−0.029	−0.144	−0.125	−0.129	−0.132
	(0.007)	(0.006)	(0.040)	(0.020)	(0.014)	(0.013)
Floor area ratio	−0.006	−0.009	−0.019	−0.023	−0.023	−0.023
	(0.034)	(0.025)	(0.036)	(0.032)	(0.033)	(0.034)
ln(*dist. to park*)	0.21	0.074	−0.475	−0.389	−0.408	−0.424
	(0.069)	(0.057)	(0.222)	(0.117)	(0.101)	(0.103)
ln(*dist. to key school*)	0.95	0.782	0.210	0.323	0.304	0.288
	(0.080)	(0.137)	(0.213)	(0.139)	(0.121)	(0.118)
District-month-of-sample FE	Y	Y	Y	Y	Y	Y
Neighborhood FE		Y	Y	Y	Y	Y
First-stage Kleinberg-Paap *F*			9.88	10.48	14.22	14.22
p-value: overidentification test				0.03	0.10	0.19
Avg. housing demand price elasticity			−2.42	−1.40	−1.61	−1.79

Notes: The number of observations is 77,696. The dependent variable is the population-average utilities recovered using parameter estimates in column 3 of Table 4. The first two columns are OLS estimates and the last four are IV estimates. The *floor area ratio* of a residential complex is total floor area over the complex's parcel size and measures complex density. *Distance to key school* is the distance to the nearest key elementary school. IV1 is the number of homes that are within 3 km from a given home, outside the same complex, and sold in a two-month window. IV2 is the average attributes of these homes (building size, age, log *distance to park*, and log *distance to key school*). IV3 is the interaction between IV2 and the winning odds of the license lottery. The winning odds decreased from 9.4 percent in January 2011 to 0.7 percent by the end of 2014. Column 6 uses all IVs. Standard errors are bootstrapped to account for the estimation errors in *EV* terms and clustered at the neighborhood level.

are as expected: households prefer larger properties and those closer to key schools but dislike older buildings and housing units far from parks.[26]

Incorporating the commuting utility *EV* not only improves the model fit but also has implications for other parameter estimates, especially the price coefficient and price elasticity. Online Appendix Table A5 reports the linear parameter estimates without the *EV* terms. Both the price coefficient and price elasticities are smaller in magnitude, consistent with the downward bias arising from the omission of important attributes (*EV* terms), as also shown in Timmins and Murdock (2007).

Based on the parameter estimates from our preferred specification (the last set of results in Tables 4 and 5), the income elasticity of marginal driving costs is 0.78 and income elasticity of housing size is 0.10, respectively.[27] To our knowledge, these are the first such estimates for Chinese households. The elasticity of marginal driving costs is largely consistent with other estimates in the literature (LeRoy and Sonstelie 1983; Glaeser, Kahn, and Rappaport 2008), while the elasticity for housing size is somewhat smaller than estimates based on US data. Using the 2003 American Housing Survey, Glaeser, Kahn, and Rappaport (2008) find the elasticity of lot size

[26] While the coefficient of distance to key schools is positive for the base group (borrowers under 30), the coefficients for borrowers in other age groups are negative and significant, as they are more likely to have school-aged children.

[27] To calculate these elasticities, we increase household income, re-solve the equilibrium for both the transportation sector and the housing market (holding housing supply fixed), and calculate the changes in driving costs and housing size.

to be from 0.25 to 0.5. They argue that these estimates provide an upper bound on the income elasticity of land demand. In comparison, our elasticity of housing demand is in terms of the condo interior size rather than the lot size, which might explain the lower values of our estimates.

Identification and Robustness.— The identification of linear and nonlinear housing demand parameters closely follows the demand literature. The choice of nonlinear parameters in Table 4 is largely driven by the model's goodness of fit and our preference for a parsimonious model for computational reasons. Online Appendix Table A6 examines increasingly saturated models, including full interactions between all housing attributes (including ease-of-commute) and all observed demographics. The key parameters such as price elasticities and random coefficients for ease-of-commute and model fitness are very similar to our preferred specification, indicating limited explanatory power for these additional controls.

Turning to the estimation of linear parameters in Table 5, the first IV on the number of nearby properties and the second set of IVs on these properties' attributes are the "BLP instrument" (Berry, Levinsohn, and Pakes 1995). The third set of IVs exploits citywide shocks induced by exogenous policy changes (time-varying winning odds of license lotteries) that make areas close to city centers and job clusters more attractive.[28] The last column that uses all three sets of IVs is our preferred specification.

To examine the robustness to the choice of IVs, online Appendix Table A7 presents parameter estimates for all combinations of IVs, the test statistics for weak-IV and over-identification tests, as well as the average price elasticities. The parameter estimates are robust across all columns, with the same sign and similar magnitudes. Results from the weak-IV and over-identification tests support our IV strategy.

We do not model policy-induced changes in amenities as we lack appropriate measures on retail shops, restaurants, and entertainment facilities. Any time-varying amenities are absorbed by δ_{jt}, the population-average utility for property j, and do not affect nonlinear parameters that govern household preference heterogeneity, such as the random coefficient for ease-of-commute. We use neighborhood fixed effects and district-month-of-sample fixed effects to control for policy-induced amenities in estimating linear parameters. Online Appendix Section C.4 provides suggestive evidence that amenities might have improved after the subway expansion, and Section VD discusses the implications of improved amenities.

Lastly, to examine the robustness of our results to the choice sampling method (as described in Section IB), we repeat the housing demand estimation with a 0.5 percent instead of 1 percent random sample to construct households' choice sets. The parameter estimates implied that WTP for housing attributes and housing demand elasticity are similar across these two samples. See online Appendix C.5 for more discussions on the sampling method and additional estimates.

[28]The effect of Beijing's license lottery on the housing market has been documented in Lyu (2022), which validates the lottery winning odds as an IV. Similar findings have also been reported for Singapore (Huang, Li, and Ross 2018).

C. *Speed-Density Elasticity*

To recover the speed-density elasticity (the supply side of the transportation sector), we use hourly data from remote traffic microwave sensors covering all major roads throughout Beijing in 2014. We focus on observations during peak hours with traffic density higher than 35 cars per lane-km, which are more relevant since we focus on commuting trips. We assume that the relationship between speed and density (the right-hand side of equation (8)) has a constant elasticity and estimate the following:

$$\ln(v_{st}) = e_{T,r} \times \ln(\textit{Traffic Density}_{st}) + \mathbf{x}_{st}^{'}\boldsymbol{\beta}_T + \varepsilon_{st}, \tag{12}$$

where the unit of observation is a road segment s by hour t, v_{st} is segment s's speed in km/h, and $\textit{Traffic Density}_{st}$ is measured by the number of vehicles per lane-km. The key parameter is $e_{T,r}$, the speed-density elasticity that differs across region r, the ring-road-bands. Vector $\mathbf{x}_{st}$ includes weather-related variables (e.g., temperature, wind speed) and time and spatial fixed effects (e.g., hour-of-day, day-of-week, road segment).

As the regressor $\ln(\textit{Traffic Density}_{st})$ could be correlated with the residual due to accidents, road construction, and major events, we construct IVs based on Beijing's driving restriction policy following Yang, Purevjav, and Li (2020b). We construct a dummy for days when vehicles with a license number ending in 4 or 9 are restricted from driving. The policy generates exogenous variation in traffic density, as far fewer vehicles have license numbers ending in the digit 4 due to a commonly held superstition.

To examine potential differences in the speed-density elasticity across regions, we split our sample into four groups: between the second and third ring roads, the third and fourth ring road band, the fourth and fifth ring road band, and the fifth and sixth ring road band. Online Appendix Table A8 shows that the IV estimates are twice as large in magnitude as the OLS estimates, but Heterogeneity across regions is limited.[29] In the counterfactual analysis below, we use the citywide elasticity estimate of -1.1 in counterfactual analyses with citywide congestion, and ring-road level elasticity for ring-road congestion and ring-road-quadrant congestion, though results do not change much regardless of which elasticity is used.

IV. Counterfactual Simulation Algorithm

Here we briefly explain the algorithm for the counterfactual analyses. Online Appendix D provides more details.

[29] We do not report IV results for column 4, the fifth to sixth ring road group, since driving restrictions are only implemented for roads within the fifth ring road. We also estimate this relationship in semi-elasticity form (speed level against log traffic density) following Vickrey(1965) and report them in panel B, but do not find qualitatively meaningful differences.

A. *Simulating the Counterfactual Equilibrium*

Algorithm.—The counterfactual equilibrium is defined as new vectors of housing prices and travel speeds $\{\mathbf{p}^*, \mathbf{v}^*\}$ that satisfy the market-clearing conditions (equations (7) and (8)). For each counterfactual analysis, we iterate equations (7) and (8) sequentially to find the fixed point $\{\mathbf{p}^*, \mathbf{v}^*\}$.

The iteration process requires us to update the driving speed vector that can be sustained given the existing road capacity at new traffic density levels. To do so, we use the following formula:

$$\frac{\tilde{v}_{r,ij} - v^o_{r,ij}}{v^o_{r,ij}} = e_{T,r} \times \frac{\tilde{D}_{T,r} - D^o_{T,r}}{D^o_{T,r}}, \tag{13}$$

where r denotes the spatial scope of congestion, either citywide (in baseline simulations), ring-road specific, or at the ring-road-quadrant level. $\tilde{v}_{r,ij}$ is the counterfactual driving speed for the segment of household i's work commute in region r, $v^o_{r,ij}$ is the observed driving speed (see Section IB for its construction), $e_{T,r}$ is the speed-density elasticity estimate for region r from equation (12), and $\tilde{D}_{T,r}$ and $D^o_{T,r}$ are the counterfactual and observed traffic density for region r. We update speed, travel mode choices, ease-of-commute, housing choices, and traffic density via equations (13), (4), (5), (2), and (6) until we find the fixed point $\{\mathbf{p}^*, \mathbf{v}^*\}$.[30]

Environmental Considerations.—The housing demand model does not include local pollution as a neighborhood attribute. Air pollution in Beijing varies less across locations within the urban core than it does from day to day (Chen, Tang, and Zhao 2015). In addition, past work has shown that public awareness of air pollution was limited before the installation of air quality monitors in 2013 (Barwick et al. 2024). We control for pollution indirectly using district-and-month-of-sample fixed effects as well as neighborhood fixed effects.

Given the importance of air pollution as a motivating factor in anti-congestion policies, we report welfare benefits from reduced air pollution as a result of lower congestion. Specifically, the expected air pollution damage caused by household i in counterfactual simulations can be measured by

$$B_i = \sum_j \Pr(\textit{Household i buys property j}) \times B_{ij},$$

with

$$B_{ij} = \sum_{k=1}^{K} VKT_{ij} \times EF_{ijk} \times MD_k,$$

where B_{ij} is the pollution damage if household i resides in property j. It consists of three terms: VKT_{ij} denotes the commuting distance, EF_{ijk} is the emissions factor

[30] While we cannot rule out multiple equilibria in $\{\mathbf{p}^*, \mathbf{v}^*\}$ as a theoretical possibility, our simulation analyses always converge to the same equilibrium outcomes for a given policy, providing empirical evidence for uniqueness in our setting.

that converts the kilometers driven into grams of pollutant k (such as CO_2, NOx, and $PM_{2.5}$), and MD_k indicates the marginal damage per gram of pollutant k, which are derived using an intake fraction approach following the air pollution literature (Apte et al. 2012). We calculate the decrease in environmental damages resulting from reduced household driving under various policies, aggregate the impacts across households, and divide them by the number of Beijing households to obtain the benefit per household.[31] See online Appendix D.4 for more details.

Fiscal Balance.—We account for capital and operating costs for subway construction and congestion pricing. Online Appendix D.3 presents the source of these numbers. We assume that funds to cover subway costs are raised via a uniform head tax, and that when toll revenues are redistributed, they are redistributed by a uniform lump sum, following the standard practice of nondistortionary distributions in the literature.

B. *Welfare Decomposition*

We now consider the underlying channels that govern welfare changes. Households' ex ante welfare is

$$W_i = \mathbb{E}_{\varepsilon_{ij}}\left[\max_{j \in \mathcal{J}_i} U_{ij}(\mathbf{p}, \mathbf{v}, \mathbf{cost})\right],$$

where $\mathbf{p}, \mathbf{v}, \mathbf{cost}$ are vectors of housing prices, travel speeds, and commuting costs, respectively. Transportation policies directly affect commuting costs. The total derivative of household welfare with respect to commuting costs consists of five elements, corresponding to different margins of adjustment:

$$\frac{dW}{dcost} = \underbrace{\left.\frac{\partial W}{\partial cost}\right|_{\mathbf{p}=\mathbf{p}_0, \mathbf{v}=\mathbf{v}_0}}_{\text{(1) direct policy effect}} \tag{14}$$

$$+ \underbrace{\underbrace{\left.\frac{\partial W}{\partial \mathbf{v}'}\frac{\partial \mathbf{v}}{\partial cost}\right|_{\tilde{\mathbf{v}}}}_{\text{(2) partial speed effect}} + \underbrace{\left.\frac{\partial W}{\partial \mathbf{v}'}\frac{\partial \mathbf{v}}{\partial cost}\right|_{D_T(\mathbf{v}^*)=S_T(\mathbf{v}^*)} - \left.\frac{\partial W}{\partial \mathbf{v}'}\frac{\partial \mathbf{v}}{\partial cost}\right|_{\tilde{\mathbf{v}}}}_{\text{(3) rebound effect}}}_{\text{(2) + (3) equil. speed effect}}$$

$$+ \underbrace{\left.\frac{\partial W}{\partial \mathbf{p}'}\frac{\partial \mathbf{p}}{\partial cost}\right|_{D(\mathbf{p}^*,\mathbf{v}^*)=1}}_{\text{(4) equil. sorting effect}} + \underbrace{\left.\frac{\partial W}{\partial \mathbf{p}'}\frac{\partial \mathbf{p}}{\partial cost}\right|_{D(\mathbf{p}^*,\mathbf{v}^*)=S} - \left.\frac{\partial W}{\partial \mathbf{p}'}\frac{\partial \mathbf{p}}{\partial cost}\right|_{D(\mathbf{p}^*,\mathbf{v}^*)=1}}_{\text{(5) housing supply effect}}.$$

The first channel, the direct policy effect, measures changes in household welfare when commuters change their travel mode in response to increasing commuting costs. The housing price, traffic speed, and household residential locations are fixed

[31] This approach does not account for additional particulate matters produced by road dust.

at their initial values. The second channel captures the partial speed effect, where the traffic speed adjusts one time from $\mathbf{v}^0$ to $\tilde{\mathbf{v}}$ via equation (13) as households reoptimize their travel mode choices, without imposition of the transportation sector's clearing condition. For example, the driving restriction moves 20 percent of drivers off the road, which leads to an initial 22 percent improvement in traffic speed. These first two channels correspond to short-run effects in some empirical studies that measure the effectiveness of transportation policies for congestion reduction via the product of implied driving time savings and VOT (Anderson 2014a; Hanna, Kreindler, and Olken 2017; Adler and van Ommeren 2016; Bauernschuster, Hener, and Rainer 2017).

The third channel quantifies the additional change in welfare when traffic speeds adjust to clear the transportation sector. As travel speed improves with driving restrictions, people are more likely to drive on days when their vehicle usage is not restricted, which partially offsets the initial speed gains. This channel is analogous to the rebound effects found in more recent reduced-form papers that account for equilibrium responses in the transportation sector (Bento, Hall, and Heilmann 2020). In the analysis below, we refer to the first channel as the direct effect, the third one as the rebound effect, and the second and third channels together as the equilibrium speed effect. The fourth channel, the equilibrium sorting effect, incorporates sorting and evaluates changes in welfare when households relocate in response to changes in the commuting utility, with the housing supply held fixed. The last channel allows housing supply to adjust in response to neighborhood housing price changes.

Before we present the simulation results, we validate the structural model by comparing its predictions with results in the literature. We simulate the market equilibrium under the 2008 subway network with and without driving restrictions and examine changes in the model-predicted housing price gradient with respect to subway access. Online Appendix Table A9 shows that the model-predicted change in the price gradient from driving restrictions is -0.034, consistent with the reduced-form evidence in Jerch et al. (2021).[32] This suggests that our analysis replicates well the documented pattern of price changes from driving restrictions.

V. Counterfactual Results

We compare five policy scenarios: driving restrictions, congestion pricing, subway expansion, and combinations of subway expansion with one of the other two policies. The first two scenarios represent two types of demand-side policies: a command-and-control approach and a market-based approach, respectively. The third one is a supply-side policy, while the last two represent combinations of demand- and supply-side policies. The first scenario follows the actual driving restriction policy implemented in Beijing: a vehicle is prohibited from driving one day a week. For congestion pricing, we calibrate a distance-based charge (at ¥1.13/km) to achieve the same level of congestion reduction as that from the driving

[32] The coefficient of −0.034 is somewhat smaller in absolute magnitude than the reduced-form analysis in Jerch et al. (2021). This is partly because the reduced-form result reflects a short-run response while the structural simulation incorporates long-run equilibrium adjustments (especially the rebound effects) and partly because the data sources and periods are different.

restriction to facilitate comparison. The subway expansion simulation compares the subway networks in 2008 and 2014. During this period, the length of the subway network increased from 100 km to 486 km. We conduct the counterfactual analysis using the 2014 cohort to allow for maximum coverage of the subway expansion. We first simulate equilibrium outcomes without any transportation policy and gradually introduce different policies.

Sections VA to VC analyze the congestion reduction, sorting patterns, and social welfare in the baseline scenario when we fix the housing supply and use a citywide congestion measure. These results are summarized in Table 6. Column 1 presents the scenario with no policies. Columns 2–6 describe the differences in outcomes from each of the five policy scenarios relative to column 1. All results are shown separately for households with income above or below the median (high- versus low-income) to reflect distributional considerations. The average across the two groups delivers the welfare effect per household.

Section VD and online Appendix E consider extensions and robustness checks, including variable housing supply, more granular measures of congestion, removing random coefficients (differences between models with and without unobserved heterogeneity), accounting for migration and consumption access, and speed variability across regions.

A. *Mode Choice and Congestion Reduction*

Driving Restriction.— Panel A of Table 6 examines changes in the travel mode choices and congestion.[33] The driving restriction policy entails two countervailing forces. On the one hand, it moves households off the road on 20 percent of workdays when driving with personal vehicles is restricted, forcing them to switch to slower modes. This reduces congestion and increases the driving speed. On the other hand, the improved travel speed from less congestion induces households whose vehicle usage is not restricted to drive more, especially among those with a long commute. This rebound effect dampens the congestion reduction from the direct policy effect. On average, driving restrictions increase traffic speed by 18 percent from 21.5 km to 25.3 km/h.

Congestion Pricing.—Congestion pricing is levied on a per-kilometer basis. There are three key differences between congestion pricing and driving restrictions. First, congestion pricing imposes a higher monetary cost of driving that scales with the distance traveled while driving restrictions lead to a longer travel time. Second, congestion pricing reduces driving much more than driving restrictions among low-income households and much less among high-income households, as low-income households are more sensitive to travel costs. Third, even though both policies lead to the same congestion reduction, a larger share of commuters drive under congestion pricing. This is because congestion pricing induces a stronger sorting response (more on this below), with households from both income

[33] The mode choice shares are slightly different between column 1 of Table 6 and Figure 1, because the former reports mode choices among home buyers in the no-policy scenario while the latter reports mode choices for all residents (including non-homeowners, or 27.8 percent of residents in 2010) under existing policies.

Table 6—Simulation Results with Household Sorting

	2008 Subway Network						2014 Subway Network					
	No Policy (1)		Driving restriction Δs from (1) (2)		Congestion pricing Δs from (1) (3)		Subway expansion Δs from (1) (4)		+Driving restriction Δs from (1) (5)		+Congestion pricing Δs from (1) (6)	
Income relative to the median	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low
Panel A. Travel mode shares in percentage points and average speed												
Drive	41.65	21.44	−7.17	−3.4	−3.48	−5.39	−2.14	−1.66	−8.52	−4.62	−5.2	−6.4
Subway	9.02	10.77	1.29	0.7	0.84	0.96	4.62	6.06	5.79	6.44	5.24	6.83
Bus	22.44	30.47	1.78	0.6	0.57	1.24	−1.54	−2.53	0.31	−1.57	−0.76	−1.03
Bike	15.96	24.01	1.6	0.8	0.77	1.78	−0.8	−1.64	0.52	−0.94	−0.13	−0.13
Taxi	2.2	1.32	1.19	0.55	0.63	0.57	−0.16	−0.11	0.89	0.36	0.39	0.36
Walk	8.74	11.99	1.31	0.74	0.67	0.83	0.02	−0.13	1.01	0.32	0.46	0.37
Avg. Speed (km/h)	21.49		3.83		3.83		1.49		5.08		5.29	
Panel B. Sorting outcomes												
Distance to work (km)	18.56	15.66	0.01	0.01	−0.17	−0.06	0.36	0.18	0.41	0.17	0.15	0.12
Distance to subway (km)	5.33	4.3	−0.03	0.03	−0.03	0.03	−4.14	−3.44	−4.14	−3.44	−4.14	−3.44
Panel C. Welfare changes per household (thousand ¥)												
Consumer surplus (+)			−227.1	−32.7	−98.2	−73.1	220.3	100	−14	64	108.7	28.7
Toll revenue (+)					137.4	137.4					127.7	127.7
Subway costs (−)							103	103	103	103	103	103
Pollution reduction (+)			4.25	4.25	4.25	4.25	1.69	1.69	5.79	5.79	6.03	6.03
Net welfare			−222.8	−28.4	43.5	68.6	119.0	−1.3	−111.2	−33.2	139.4	59.4

Notes: Simulations use the 2014 cohort (households who purchased homes in 2014) and are based on parameters reported in column 6 of Table 3, column 3 of Table 4, and column 6 of Table 5. We incorporate household sorting but fix the housing supply and use a citywide congestion index. Consumer surplus estimates are based on housing demand and travel mode preference estimates and should be interpreted as total consumer surplus over a property's life span. Toll revenue is net of the capital and operating costs of revenue collection. Subway costs include construction and operation costs. Welfare benefits from pollution reduction arise from reduced tailpipe emissions. Toll revenue, subway costs, and environmental benefits are the discounted sum over thirty years (which is approximately the lifespan of a property) and allocated uniformly across households. Net welfare is consumer surplus plus toll revenue and environmental benefits minus subway costs. Column 1 reports results when no policy was in place. Columns 2 to 6 present differences from column 1. Driving restriction prohibits driving in one of five work days. Congestion pricing is set at ¥1.13 per km to generate the same speed improvement as the driving restriction. High-income households are those with income above the median. See online Appendix Section D for more details on the simulation procedure.

groups and especially high-income households moving closer to work. In contrast, the commuting distance under driving restrictions barely changes. Thus, there are fewer long commutes but more people on the road under congestion pricing than under the driving restriction.

Subway Expansion.—Despite the immense scale of Beijing's subway expansion over 2008–2014, it leads to the smallest congestion reduction among the three policies. Column 4 shows that traffic speed increases by 7 percent, only 40 percent of that under the first two policies. The reason for this muted response is twofold. First, the reduction in the driving share is smaller under subway expansion than the first two policies. Second, and more importantly, both high- and low-income households move farther away from work and commute longer distances as the subway network expands. We discuss the spatial sorting responses in Section VB.

These results, and especially the one on sorting, point to important channels beyond what has been examined in empirical studies on the short-run congestion-reduction impact of the subway system (Anderson 2014; Gu et al. 2021). Our findings are consistent with the prior literature that (i) with sufficient time, induced travel demand increases one-for-one with capacity expansion (Downs 1962; Duranton and Turner 2011), and (ii) subway expansion itself lowers the cost associated with the commuting distance and increases urban sprawl (Gonzalez-Navarro and Turner 2018; Heblich, Redding, and Sturm 2020).[34] Nonetheless, the subway expansion dramatically increased subway access: the distance between home and the nearest subway station declined by about 80 percent on average. Subway ridership increased significantly by 51 and 56 percent among high- and low-income groups, respectively.

Policy Combinations.—Column 5 represents Beijing's actual transportation policy which combines subway expansion with driving restrictions. Column 6 offers an alternative policy of subway expansion combined with congestion pricing. We are interested in examining which demand-side policy exhibits stronger complementarity with the supply-side policy. The improvement in driving speed under the policy combinations is close to the sum of the speed improvements under the individual policies. For example, the speed improvement is 3.83 km/h under the driving restriction, 1.49 km/h under the subway expansion, and 5.08 km/h under the combination of both policies. There may be two countervailing forces at play. First, the supply-side policy could complement the demand-side policy in that a larger subway network makes substitution away from driving easier. Indeed, as the subway becomes more attractive, driving restrictions lead to an 8.52 percentage point reduction in high-income households' driving probability under the 2014 subway network in comparison to a 7.17 percentage point reduction under the 2008 network. On the other hand, there could be policy redundancy: some of the driving trips could be reduced under either the supply-side or the demand-side policy, leading to a smaller aggregate impact than the sum of individual policy impacts.

Congestion pricing exhibits stronger complementarity with subway expansion than driving restrictions and is more effective in moving people off the road: the speed improvement is 5.29 km/h in column 6, higher than the 5.08 km/h improvement in column 5. Congestion pricing affects both the extensive margin (whether to drive) and the intensive margin (how far to drive); both effects could be reinforced by subway expansion. In contrast, the driving restriction primarily operates along the extensive margin.

B. *Sorting and the Housing Price*

Sorting and Household Spatial Distribution.— Panel B of Table 6 examines the impacts of transportation policies on households' spatial distribution. Transportation

[34] Anas (2024) provides theoretical evidence that road capacity enhancements may increase road usage (VKT) but still lower congestion even in the long-run.

policies directly affect commuting costs. These direct changes set in motion a series of behavioral responses whereby households substitute across different travel modes and adjust their residential locations. We report the average distance to work and to the subway for both the high- and low-income groups.[35] With a fixed housing supply, the average distance to the subway system across all households remains fixed under both driving restrictions and congestion pricing. If rich households move closer to subway stations, poor households would be displaced and move farther away from the subway by construction. We report the average distance to the subway separately for the two income groups to illustrate this displacement effect.

Different policies lead to different sorting patterns. There are two countervailing forces under driving restrictions. On the one hand, the policy would incentivize households to live closer to their workplaces. On the other hand, driving speed improves but it disproportionately benefits long-distance trips. The correlation between the driving probability and driving distance increases from 0.28 to 0.33 under the driving restriction. That is, the speed effect undermines households' incentive to move closer to their workplaces. On net, the commuting distance remains approximately the same as before, with minimal sorting responses.

In comparison, congestion pricing is distance-based and causes a much higher increase in commuting costs for longer trips. Hence, both income groups move closer to work, as shown in column 3. However, high-income households exhibit a much stronger sorting response for three reasons. First, 41.65 percent of high-income households drive to work, in comparison to 21.44 percent of low-income households. As a result, high-income households are much more affected by congestion pricing. Second, high-income households have a higher WTP for commuting convenience as their VOT is higher. Lastly, properties closer to employment centers command a housing premium. They are more affordable for high-income than for low-income households.

Subway expansion generates the strongest sorting responses among the three policies, and these responses run in the opposite direction to those from congestion pricing. The direct policy effect moves people off the road as they substitute toward subways. The improved driving speed and expanded subway system reduce the cost of long-distance commuting by both driving and subway. As a result, both the direct policy effect and the equilibrium speed effect work in the same direction and disperse households from the city center into the suburbs and locations near the new subway stations.

Figure 3 plots changes in the average commuting distance for residents in each TAZ relative to the distance in the no-policy scenario. Driving restrictions lead to modest commuting distance changes that are often in opposite directions across neighborhoods. The commuting distance is reduced in almost all TAZs under congestion pricing, suggesting better spatial matches between work and housing locations. The reduction in commuting distance is most pronounced for TAZs outside the fourth ring road, where the average commuting distance is 21.9 km, versus 11.9 km for households living inside the fourth ring road. Subway expansion increases the commuting distance in most TAZs, especially along the new subway lines, exacerbating

[35] We report straight-line distances, which are not affected by travel mode changes.

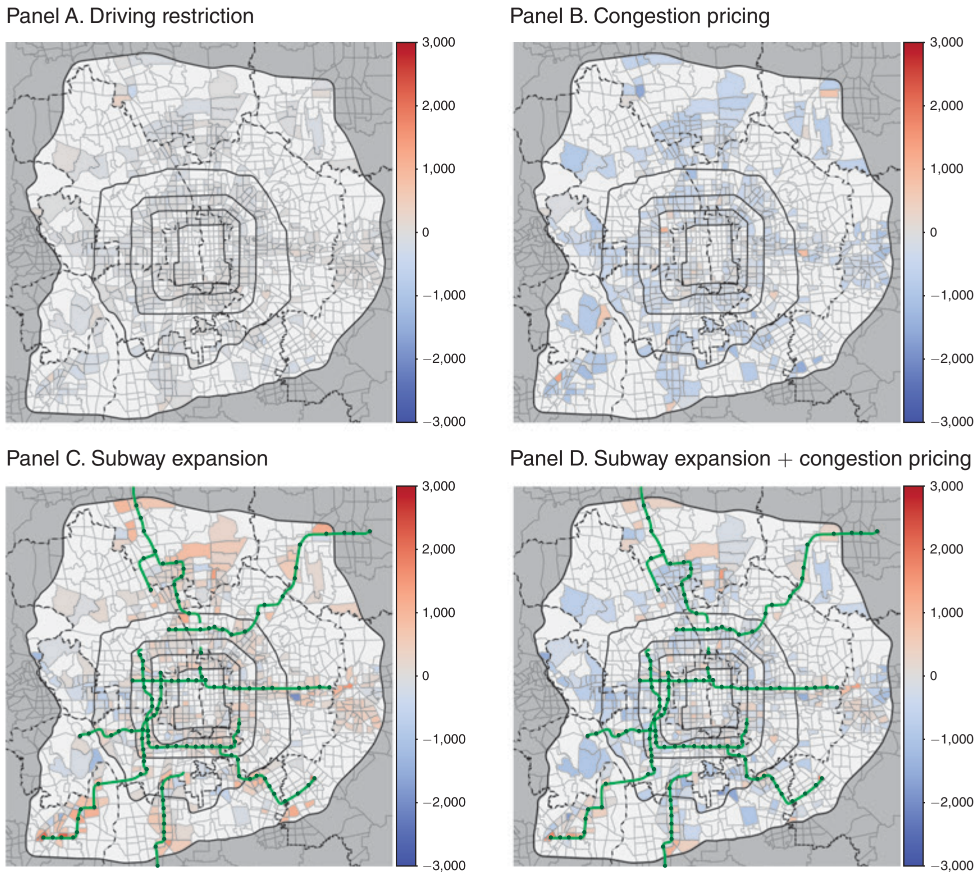

Figure 3. Changes in Commuting Distances from Sorting in Counterfactual Simulations (in Meters)

Notes: This figure illustrates simulated changes in commuting distances in meters across TAZs under different counterfactual policies (relative to the no policy scenario). The results are based on the simulations in Table 6 that allow for household sorting, fix housing supply, use estimates including random coefficients, and use a single citywide congestion index. Warmer colors correspond to increases in commuting distance while colder colors represent decreases. Green lines represent new subway lines built between year 2008 and 2014.

"wasteful" commuting. This further separation of workplace and residence following subway expansion is consistent with the evidence in Gonzalez-Navarro and Turner (2018) and Heblich, Redding, and Sturm (2020).

Both driving restrictions and congestion pricing result in high-income households moving closer to and low-income households moving further away from the subway in comparison to the baseline as depicted in online Appendix Figure A5. This reflects transit-based gentrification, where lower-income households are priced out of locations closer to the subway. Beijing's subway expansion, on the other hand, drastically reduced the distance to subway stations for both groups: the average distance to the nearest subway station dropped from 5.33 km to 1.19 km for high-income households and from 4.3 km to 0.86 km for low-income households.

To further examine the sorting pattern, we regress changes in a property owner's household income due to subway-expansion induced sorting (as depicted in panel C of online Appendix Figure A5) on a property's distance to the nearest subway station built between 2008 and 2014. The regression results from online Appendix Table A10 suggest that properties that gained access to the expanded subway network experienced significant increases in household income, suggesting that new owners are richer. These results indicate that people actively sort, and high-income households outbid low-income households to move closer to newly built subway stations.

Housing Price.—Changes in housing prices closely mirror the sorting patterns. Online Appendix Figure A6 exhibits the housing price responses across neighborhoods. Both driving restrictions and congestion pricing increase the prices of homes closer to job centers (such as locations inside the fourth ring road and close to the tech and financial centers), but the impact is stronger under congestion pricing. Subway expansion generates opposite spatial impacts: housing prices depreciate near the city center and appreciate in city suburbs along the new subway lines where public transportation was poor prior to the expansion.[36] With both subway expansion and congestion pricing, the price impacts of subway expansion dominate.

To further illustrate the differential impact of subway expansion on home prices, online Appendix Figure A7 plots the housing price gradient with respect to the subway distance separately for the 2008 and 2014 subway networks. The bid-rent curve is steeper under the 2014 network ($-¥1900/m^2$ per km) than under the 2008 network ($-¥700/m^2$ per km) because the 2014 network is larger and hence the proximity to this network is more valuable to commuters. The bid-rent curve under the 2014 network shifts down, reflecting the composition change of homes whereby the subway expansion reaches cheaper homes farther away from the city center.

C. *Welfare Analysis*

Panel C of Table 6 presents the welfare results, including changes in consumer surplus, toll revenue, costs of subway expansion, and environmental benefits. Figure 4 decomposes changes in consumer surplus following equation (14) to illustrate different adjustment margins. Note that the consumer surplus estimates are based on preference estimates from the housing demand. As a result, they should be interpreted as total consumer surplus over a property's life span.[37] Housing supply is fixed in this subsection. We consider variable housing supply and other extensions in Section VD.

[36] Under congestion pricing, housing prices in northwestern Beijing would increase by about ¥2,000/m^2, while those in some southeastern areas would decrease by ¥2,000/m^2 from a baseline average price at ¥24,022/m^2. Under subway expansion, housing prices increase by about ¥4,000/m^2 in the southwest, where the subway expansion is greatest and historical prices have been lowest.

[37] We assume that a property lasts for thirty years, which is consistent with the revealed WTP estimates for shorter commutes documented in Section IIIB. The thirty-year assumption is also motivated by the US Internal Revenue Service's rules that residential properties depreciate 3.6 percent per year and last 27.5 years. We use the US reference as we could not find relevant depreciation rates for Beijing. Transportation policies are assumed to last over a property's life span, namely thirty years. The thirty-year assumption does not affect estimates of consumer surplus. It is relevant for the calculation of toll revenue (net of the expenses of toll collection), the costs of operating the subway system, and environmental benefits, all of which are discounted over thirty years at a discount rate of 0.98.

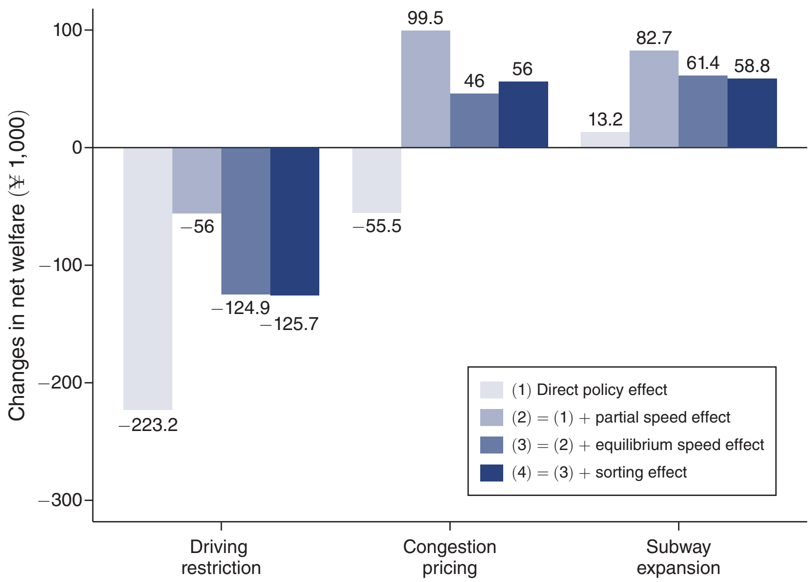

Figure 4. Welfare Decomposition for Different Transportation Policies

Notes: This figure decomposes welfare changes per household along four adjustment margins. All simulations use a citywide congestion index and fix housing supply. For each policy, the bars display the cumulative welfare changes incorporating previous margins. The direct policy effect measures changes in social welfare when commuters change travel mode in response to increasing commuting costs, holding housing prices, traffic speed, and residential locations fixed. The partial speed effect allows the driving speeds to adjust one-time via equation (13), but does not impose the transportation sector's clearing condition. The third bar additionally incorporates the full equilibrium speed effect and additional changes in welfare when traffic speeds adjust further to clear the transportation sector. The last bar includes household sorting in addition to the three channels above. The welfare calculations account for subway costs (both construction and operating costs), toll revenues (net of capital and operating costs for the congesting pricing system), and environmental benefits.

Welfare Loss from Driving Restrictions.—First, despite their effectiveness in congestion reduction, driving restrictions generate a total welfare loss of ¥125,700 per household, though high-income households experience a much steeper reduction than low-income households.[38] Driving restrictions force drivers to switch to slower modes and significantly increase commuting time, especially for households with long commutes. On average, a household spends 16.8 more minutes commuting each day due to driving restrictions.

Figure 4 decomposes the welfare loss along different adjustment margins, and we present the same analysis for speed changes in online Appendix Figure A8. The direct policy effect of driving restrictions is large and negative at ¥223,200 per household since it distorts commuting choices and forces households to substitute toward inferior travel modes. As commuters switch to nondriving travel modes, traffic speeds and commuting time improve, mitigating the welfare loss of

[38] The welfare effect per household is the average between high- and low-income households. According to the Beijing Statistics Bureau, the average household income in Beijing in 2014 was ¥128,000, and ¥172,000/¥84,000 for high-/low-income households, respectively. The annualized loss is roughly 4 percent of the household income, assuming a thirty-year period.

the direct policy effect, as shown by the second and third bars in Figure 4. The second bar highlights the effect of a partial, or short-run, speed adjustment, while the third bar represents welfare changes where the driving speed (hence congestion) and travel mode choices are in equilibrium and clear the transportation sector following equation (8). The disparity between the welfare loss linked to the second channel at ¥56,000 and the third channel at ¥124,900 underscores the importance of incorporating the rebound effect and allowing for the full equilibrium adjustment of the transportation sector. Otherwise, the welfare losses could be underestimated by 55 percent for driving restrictions and welfare gains overstated by 36–116 percent for other policies. The fourth bar further incorporates the sorting effect. With all four channels incorporated, the welfare loss is at ¥125,700 per household.

Welfare Gain from Congestion Pricing.—Second, before revenue recycling (i.e., toll revenues counted as government surplus), low-income households experience a greater loss under congestion pricing than under driving restrictions. This reflects the fact that low-income households are more affected by increases in monetary costs from congestion pricing than they are by longer commuting times under driving restrictions. However, when the toll revenue is uniformly recycled across groups (i.e., toll revenues redistributed uniformly among households), congestion pricing leads to *welfare gains* for both groups: consumer surplus increases by ¥43,500 and ¥68,600 for high- and low-income households, respectively. Low-income households witness a larger consumer surplus increase than the high-income group partly because they pay a smaller amount in congestion charges but receive 50 percent of the revenue. This highlights the role of revenue recycling to abate distributional concerns from congestion pricing.

In terms of the underlying channels, the direct effect of congestion pricing (with revenue recycling) reduces welfare by ¥55,500 per household. The equilibrium speed effect (the partial speed and rebound effect) reverses the welfare loss to yield a welfare gain of ¥46,000 per household. As sorting works in the same direction as the speed effect and moves households closer to their places of work, the welfare gain further increases to ¥56,000 per household. Residential sorting enhances the welfare gain from congestion pricing by 22 percent, consistent with the result based on 98 US cities in Langer and Winston (2008).

Welfare Gain from Subway Expansion.—Third, while the subway expansion from 2008 to 2014 resulted in limited congestion reduction relative to that under the other two policies, it led to a larger increase in consumer surplus. Much of this increase comes from improved subway access: the distance to the nearest subway station declined by 80 percent on average, and subway ridership increased by more than 50 percent. Although switching from nondriving modes to the subway does not alleviate traffic congestion, it improves consumer welfare by offering better commuting choices. After accounting for the fixed and operating costs of subway expansion, net welfare is almost halved for high-income at ¥119,000 and close to zero for low-income households.

Looking at the different margins of adjustments, the direct effect of subway expansion generates a welfare gain of ¥13,200 per household as a result of improved subway accessibility. The improved driving speed in equilibrium increases consumer welfare by ¥47,800 per household, with the overall welfare gain reaching ¥61,000 per household. Sorting induces households to move away from their workplaces and the city center, which increases congestion and dampens the welfare gain to ¥58,800 per household.

Welfare Superiority of Policy Combination.—Fourth, the combination of congestion pricing and subway expansion achieves the largest congestion reduction (with a 25 percent speed improvement) and generates the highest welfare gain at ¥99,400 per household across all policy scenarios. The annualized welfare gains are equivalent to 3 percent of household annual income. The revenue from congestion pricing at ¥127,700 per household could fully cover the costs of subway expansion at ¥103,000 per household. This finding has broader applicability for the design of transportation infrastructure outside the context of Beijing.[39] While it is distinct from results in prior work on the role of self-financing toll roads (Mohring and Harwitz 1962; Winston 1991; Verhoef and Mohring 2009), its policy implications may be equally relevant. Our analyses suggest that welfare gains from infrastructure improvements could be mitigated by induced congestion. Pairing these investments with pricing instruments such as congestion pricing is critical to successfully address preexisting and induced congestion and finance the cost of infrastructure investment to increase social welfare.

The welfare magnitudes discussed above, such as the annualized gains at 3 percent of household annual income for the combined policy of subway expansion and congestion pricing, are within the spectrum of welfare effects for large-scale transportation policies reported in the literature (2.36 percent in Tsivanidis 2023 and 2.3 percent in Kreindler and Miyauchi 2023).

Environmental Benefits.—Lastly, both reduced driving and improved speeds generate environmental benefits. This is because fewer vehicle kilometers traveled directly reduce tailpipe emissions and improve air quality. In addition, improved driving speeds increase fuel efficiency and lead to lower emissions per kilometer traveled. The estimated environmental benefits vary between ¥1,690 to ¥6,030 per household across policy scenarios. While these benefits are nontrivial, they are much smaller than changes in consumer surplus and do not affect the relative comparisons across transportation policies.

Importance of Sorting and Endogenous Congestion.— Panel A of Table 7 reports changes in driving speed and welfare when we shut down sorting and endogenous congestion. To facilitate comparison, the congestion price is kept the same as before at ¥1.13/km. The first row of panel A in Table 7 reproduces the baseline results from Table 6 with sorting. The second row allows the transportation sector to clear but does not allow households to relocate by shutting down sorting (See online Appendix Table A11 for the full results). Sorting amplifies the effectiveness of

[39] As an example, transportation funds allocated through the US American Reinvestment and Recovery Act of 2008 required several pilot pricing projects to reuse toll revenues to enhance affected corridors, including public transit (GAO 2012).

Table 7—Importance of Sorting, Endogenous Congestion, and Extensions: Changes in Speed and Welfare per Household

	Driving restriction			Congestion pricing			Subway expansion		
	ΔSpeed	Δ Welfare (¥1000)		Δ Speed	Δ Welfare (¥1000)		Δ Speed	Δ Welfare (¥1000)	
Income relative to the median	(km/h)	High	Low	(km/h)	High	Low	(km/h)	High	Low
Panel A. Sorting and endogenous congestion									
With sorting (main results)	3.83	−222.8	−28.4	3.83	43.5	68.6	1.49	119.0	−1.3
Without sorting	3.82	−223.1	−26.8	3.61	32.1	59.6	1.76	106.7	16.0
With sorting but without endogenous congestion	5.47	−107.3	−4.4	5.15	118.1	81.0	2.36	158.8	6.4
Panel B. Extensions and robustness checks									
With sorting and housing supply response	3.85	−225.5	−27.9	4.02	56.3	72.8	0.95	64.0	−16.0
With ring-road-quadrant-level traffic density	3.46	−239.4	−31.8	3.38	24.7	67.5	1.25	104.9	−4.8
Without random coefficients	4.59	−1,447.6	−338.0	4.59	−406.9	−3.6	1.61	15.5	−42.5
With migration	4.65	−193.4	−22.6	4.63	77.1	75.7	0.73	82.0	−8.5
With consumption access	3.83	−297.8	−43.5	3.83	56.4	89.8	1.49	191.6	31.6

Note: This table examines speed and welfare implications under various extensions. Each cell reports changes relative to the no-policy scenario. Speed without any policy is 21.49 km/h. The unit of welfare changes per household is ¥1,000. Congestion pricing is fixed at ¥1.13 per km as in Table 6. The first row in panel A summarizes results in Table 6. Panel A examines the importance of sorting and endogenous congestion. "Without sorting" holds residential locations fixed and does not impose the housing market clearing condition. "With sorting but without endogenous congestion" keeps sorting but shuts down endogenous congestion. To do so, we adjust traffic speed once in response to households' travel mode changes via equations (4) and (13), but do not impose the transportation sector's equilibrium condition. Panel B relaxes modeling and simulation assumptions. "With sorting and housing supply response" assumes that the housing supply responds to neighborhood average price changes at a constant price elasticity of 0.53. "With ring-road-quadrant-level traffic density" solves the endogenous traffic density/congestion at the ring-road-quadrant level as described in Section IC. "Without random coefficients" reestimates the model and repeats the counterfactual analyses with no random coefficients. "With migration" assumes 5 percent more vehicles (in-migration) under the subway expansion and 5 percent fewer vehicles (out-migration) under the driving restriction and congestion pricing. "With consumption access" incorporates an additional 33 percent of changes in consumer surplus through consumption access (easy access to amenities) following Miyauchi, Nakajima, and Redding (2021).

congestion pricing but undermines that of subway expansion on congestion reduction. In addition, sorting increases the welfare gain from congestion pricing by as much as 40 percent for high-income households and 16 percent for low-income households, consistent with Figure 4.

Sorting also has important distributional implications. This is most evident under subway expansion, where sorting improves the welfare of high-income households at the cost of low-income households. This reflects transit-based gentrification: both high- and low-income households prefer places near the subway, but higher WTP from high-income households raises prices and displaces low-income households. In contrast, under congestion pricing, both groups are better off with sorting. This is because work locations differ across income groups. Thus, sorting, in the form of moving closer to work, moderates housing market competition across income groups, alleviates congestion, and increases welfare. The opposing effect of sorting under congestion pricing and subway expansion and the unequal distributional impacts highlight the importance of accounting for sorting. Otherwise, we risk not only overestimating or underestimating the welfare gains but also getting the signs wrong and making inappropriate policy recommendations.

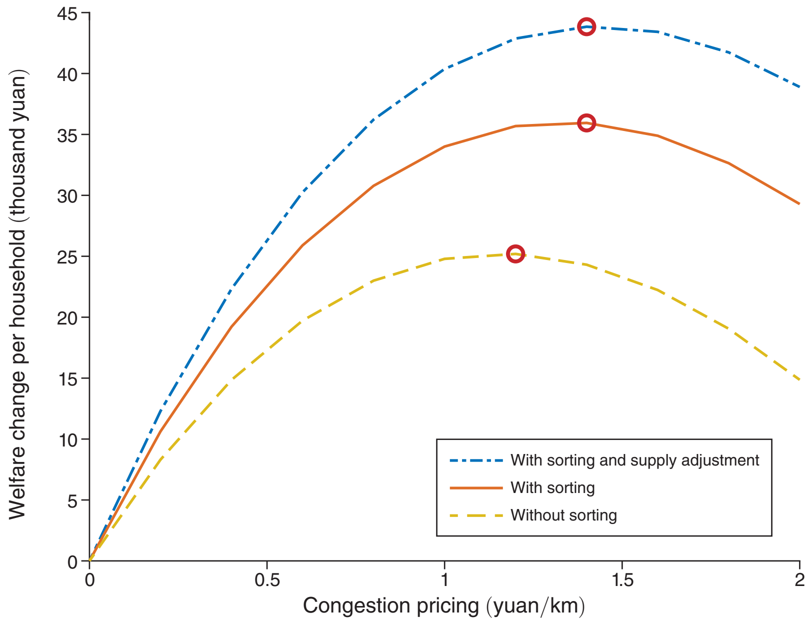

Figure 5. Optimal Congestion Pricing under the 2014 Subway Network

Notes: The figure plots welfare changes over different congestion prices under the 2014 subway network, without household sorting (yellow dotted line), with sorting (orange solid line), and sorting together with housing supply adjustments (blue dashed line). The welfare calculation uses a citywide congestion index, accounts for subway costs (both construction and operating costs) and toll revenues (net of capital and operating costs for the congesting pricing system), but excludes pollution reduction benefits. The optimal congestion pricing is ¥1.2/km without sorting and ¥1.4/km with sorting. Sorting increases consumer welfare by 20 –30 percent, and housing supply adjustment contributes another 10–20 percent for most congestion pricing levels.

The third row of panel A in Table 7 keeps sorting but shuts down endogenous congestion. To do so, we adjust the traffic speed once in response to households' travel mode changes (the second channel in equation (14)) but do not impose the transportation sector's equilibrium condition (8). In other words, we do not incorporate the rebound effect (the third channel in equation (14)) and do not allow the full equilibrium adjustment of the traffic speed. Households sort according to the one-time traffic speed adjustment. The results echo the point that we made above: without incorporating the full equilibrium adjustment, the speed improvement would be overestimated by 43–58 percent, and the welfare benefit would be inflated even more.

Optimal Congestion Pricing.—Figure 5 plots changes in welfare as the congestion charge varies. The optimal congestion charge is ¥1.2 per km without sorting, ¥1.4 per km with sorting, and approximately the same when we incorporate both sorting and housing supply. For most levels of congestion charges shown in the figure, sorting increases consumer welfare by 20–30 percent, and housing supply adjustment contributes to another 10–20 percent of welfare gain. In addition, changes in social welfare are positive for a wide range of congestion charges ($<$ ¥2.5/km). This indicates that congestion pricing is likely to be an effective and welfare-improving policy even when governments cannot gauge the exact optimal pricing level *a priori*.

D. *Robustness Checks*

Panel B of Table 7 presents five robustness checks based on alternative modeling assumptions and online Appendix E provides a detailed discussion. First, allowing neighborhood-level adjustment of housing supply in response to policy-induced price changes mitigates the overall price effect and enables more households to move into desirable neighborhoods, thus magnifying the role of sorting. Second, increasing the granularity of congestion (such as using ring-road-quadrant-level congestion) and traffic speed responses does not meaningfully affect our results. Third, excluding unobserved consumer preference heterogeneity (i.e., random coefficients) from housing demand and travel model choices generates counterintuitive substitution patterns and erroneous welfare estimates. The estimated effect of subway expansion on ridership is substantially muted, and the estimated welfare losses (gains) under driving restrictions (subway expansion) are off by an order of magnitude with a wrong sign for the welfare effect of congestion pricing. Lastly, allowing for urban population growth and incorporating improved consumption access into welfare analysis does not affect our findings qualitatively.

VI. Conclusion

Transportation plays a critical role in determining residential locations. At the same time, household location choices help determine the efficacy and efficiency of urban transportation policies. This study provides a unified equilibrium sorting framework with endogenous congestion to evaluate the efficiency and equity impacts of various urban transportation policies. The framework incorporates rich preference heterogeneity and equilibrium feedback effects between the transportation sector and the housing market.

Our analysis delivers several important takeaways. First, including the utility from the ease-of-commuting in housing demand dramatically improves the model fit. Having flexible preference heterogeneity, incorporating sorting responses, and modeling the joint equilibrium of the transportation sector and housing market all have important implications for the welfare and distributional outcomes. Second, compared to driving restrictions, congestion pricing better incentivizes residents to live closer to their work locations. Subway expansion does the opposite by increasing the separation between residences and workplaces. Third, different policies generate drastically different efficiency and equity impacts. While driving restrictions reduce social welfare due to the large distortion in travel choices, congestion pricing is welfare improving for both the high- and low-income groups with a uniform revenue recycling. The combination of congestion pricing and subway expansion stands out as the best policy: it delivers the largest congestion reduction and the highest welfare gains. In addition, the revenue from congestion pricing can fully cover the cost of subway expansion.

Our analysis does not consider the potential implications for the labor market and firm locations, two additional channels that affect the long-term urban spatial structure. Incorporating these margins would require additional data and computational resources. We leave this task for future research.

REFERENCES

Adler, Martin W., and Jos N. van Ommeren. 2016. "Does Public Transit Reduce Car Travel Externalities? Quasi-Natural Experiments' Evidence from Transit Strikes." *Journal of Urban Economics* 92: 106–119.

Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm, and Nikolaus Wolf. 2015. "The Economics of Density: Evidence from the Berlin Wall." *Econometrica* 83 (6): 2127–89.

Akbar, Prottoy A., Victor Couture, Gilles Duranton, and Adam Storeygard. 2023. "Mobility and Congestion in Urban India." *American Economic Review* 113 (4): 1083–111.

Anas, Alex. 2024. "'Downs's Law' Under the Lens of Theory: Roads Lower Congestion and Increase Distance Traveled." *Journal of Urban Economics* 139: 103607.

Anas, Alex, and Ikki Kim. 1996. "General Equilibrium Models of Polycentric Urban Land Use with Endogenous Congestion and Job Agglomeration." *Journal of Urban Economics* 40 (2): 232–56.

Anderson, Michael L. 2014. "Subways, Strikes, and Slowdowns: The Impacts of Public Transit on Traffic Congestion." *American Economic Review* 104 (9): 2763–96.

Anonymous Agency. "Beijing Mortgage Data 2006-2014." Accessed 2019-03-02.

Apte, Joshua S., Emilie Bombrun, Julian D. Marshall, and William W Nazaroff. 2012. "Global Intraurban Intake Fractions for Primary Air Pollutants from Vehicles and Other Distributed Sources." *Environmental Science & Technology* 46 (6): 3415–3423.

Baidu Maps. 2019. "Commuting Distances and Times." Accessed 2019-01-12.

Barwick, Panle Jia, Shanjun Li, Liguo Lin, and Eric Zou. 2024. "From Fog to Smog: The Value of Pollution Information." *American Economic Review* 114 (5): 1338-1381.

Barwick, Panle Jia, Shanjun Li, Andrew R. Waxman, Jing Wu, and Tianli Xia. 2021. "Efficiency and Equity Impacts of Urban Transportation Policies with Equilibrium Sorting." NBER Working Paper 29012.

Panle Jia Barwick, Shanjun Li, Andrew Waxman, Jing Wu, and Tianli Xi. 2024. *Data and Code for: "Efficiency and Equity Impacts of Urban Transportation Policies with Equilibrium Sorting."* Nashville, TN: American Economic Association; distributed by Inter-university Consortium for Political and Social Research, Ann Arbor, MI. https://doi.org/10.3886/E198573V1.

Basso, Leonardo J., and Hugo E. Silva. 2014. "Efficiency and Substitutability of Transit Subsidies and Other Urban Transport Policies." *American Economic Journal: Economic Policy* 6 (4): 1–33.

Bauernschuster, Stefan, Timo Hener, and Helmut Rainer. 2017. "When Labor Disputes Bring Cities to a Standstill: The Impact of Public Transit Strikes on Traffic, Accidents, Air Pollution, and Health." *American Economic Journal: Economic Policy* 9 (1): 1–37.

Baum-Snow, Nathaniel, and Matthew E Kahn. 2000. "The Effects of New Public Projects to Expand Urban Rail Transit." *Journal of Public Economics* 77 (2): 241–63.

Bayer, Patrick, and Christopher Timmins. 2005. "On the Equilibrium Properties of Locational Sorting Models." *Journal of Urban Economics* 57 (3): 462–77.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan. 2007. "A Unified Framework for Measuring Preferences for Schools and Neighborhoods." *Journal of Political Economy* 115 (4): 588–638.

Bayer, Patrick, Nathaniel Keohane, and Christopher Timmins. 2009. "Migration and Hedonic Valuation: The Case of Air Quality." *Journal of Environmental Economics and Management* 58 (1): 1–14.

Becker, Gary S. 1965. "A Theory of the Allocation of Time." *Economic Journal* 75 (299): 493–517.

Beijing Infrastructure Investment Co., LTD. 2015. "Subway Costs." Accessed 2015-06-01.

Beijing Municipal Commission of Housing and Urban-Rural Development. 2020. "Home Registration Records." Accessed 2020-03-13.

Beijing Transport Institute. 2010. "Beijing Transport Annual Report." http://www.bjtrc.org.cn/JGJS.aspx?id=5.2&Menu=GZCG. Formerly Beijing Transportation Research Center. Accessed 2019-03-2.

Beijing Transport Institute. 2020 "Beijing Transport Annual Report." http://www.bjtrc.org.cn/JGJS.aspx?id=5.2&Menu=GZCG Formerly Beijing Transportation Research Center. Accessed 2020-10-11.

Beijing Transport Institute. 2010, 2014. "Beijing Household Travel Surveys 2010, 2014." Formerly Beijing Transportation Research Center. Accessed 2019-03-02.

Beijing Transport Institute. 2015. "Beijing Shapefiles." Formerly Beijing Transportation Research Center. Accessed 2015-05-21.

Beijing Transport Institute. 2018. "Beijing Transportation Cost Data." Formerly Beijing Transportation Research Center. Accessed 2018-06-01.

Beijing Transport Institute. 2019. "Historical Subway Lines." Formerly Beijing Transportation Research Center. Accessed 2019-03-02.

Beijing Transport Institute. 2020. "Roadway Sensor Data." Formerly Beijing Transportation Research Center. Accessed 2020-02-12.

Beijing Transport Institute. 2018. "Traffic Congestion Index for Beijing." Formerly Beijing Transportation Research Center. Accessed 2018-08-06.

Bento, Antonio, Jonathan Hall, and Kilian Heilmann. 2020. "Estimating Congestion Externalities Using Big Data." Unpublished.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 63 (4): 841–890.

BJSubway.com. 2020. "Beijing Subway Timetables." Https://Www.Bjsubway.Com/. Accessed 2020-06-16.

Brueckner, Jan K. 2007. "Urban Growth Boundaries: An Effective Second-Best Remedy for Unpriced Traffic Congestion?" *Journal of Housing Economics* 16 (3-4): 263–273.

Buchholz, Nicholas, Laura Doval, Jakub Kastl, Filip Matejka, and Tobias Salz. 2020. "The Value of Time: Evidence from Auctioned Cab Rides." Unpublished.

Capps, Cory, David Dranove, and Mark Satterthwaite. 2003. "Competition and Market Power in Option Demand Markets." *RAND Journal of Economics* 34 (4): 737–763.

Chen, Wei, Hongzhao Tang, and Haimeng Zhao. 2015. "Diurnal, Weekly and Monthly Spatial Variations of Air Pollutants and Air Quality of Beijing." *Atmospheric Environment* 119: 21–34.

Davis, Lucas W. 2008. "The Effect of Driving Restrictions on Air Quality in Mexico City." *Journal of Political Economy* 116 (1): 38–81.

Downs, Anthony. 1962. "The Law of Peak-Hour Expressway Congestion." *Traffic Quarterly* 16 (3): 393–409.

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

Epple, Dennis, and Holger Sieg. 1999. "Estimating Equilibrium Models of Local Jurisdictions." *Journal of Political Economy* 107 (4): 645–81.

Gallego, Francisco, Juan-Pablo Montero, and Christian Salas. 2013. "The Effect of Transport Policies on Car Use: Evidence from Latin American Cities." *Journal of Public Economics* 107: 47–62.

GAO. 2012. "Report to the Subcommittee on Transportation, Housing, and Urban Development and Related Agencies, Committee on Appropriations, House of Representatives: TRAFFIC CONGESTION: Road Pricing Can Help Reduce Congestion, But Equity Concerns May Grow." Technical Report 2012.

Gaode. 2019. "Bus Commuting Distances and Times." Accessed 2019-01-01.

Glaeser, Edward L., Matthew E. Kahn, and Jordan Rappaport. 2008. "Why Do the Poor Live in Cities? The Role of Public Transportation." *Journal of Urban Economics* 63 (1): 1–24.

Goldszmidt, Ariel, John A. List, Robert D. Metcalfe, Ian Muir, V. Kerry Smith, and Jenny Wang. 2020. "The Value of Time in the United States: Estimates from a Nationwide Natural Field Experiment." Unpublished.

Gonzalez-Navarro, Marco, and Matthew A. Turner. 2018. "Subways and Urban Growth: Evidence from Earth." *Journal of Urban Economics* 108: 85–106.

Gu, Yizhen, Jiang Chang, Junfu Zhang, and Ben Zou. 2020. "Subways and Road Congestion." *American Economic Journal: Applied Economics* 13 (2): 83-115.

Guevara, C. Angelo, and Moshe E. Ben-Akiva. 2013. "Sampling of alternatives in Logit Mixture models." *Transportation Research Part B: Methodological* 58: 185 – 98.

Hainmueller, Jens. 2012. "Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies." *Political Analysis* 20 (1): 25-46.

Hanna, Rema, Gabriel Kreindler, and Benjamin A. Olken. 2017. "Citywide effects of high-occupancy vehicle restrictions: Evidence from "three-in-one" in Jakarta." *Science* 357 (6346): 89–93.

Heblich, Stephan, Stephen J. Redding, and Daniel M. Sturm. 2020 "The Making of the Modern Metropolis: Evidence from London." *Quarterly Journal of Economics* 135 (4): 2059–133.

Housing Provident Fund (**HPF**)**.** 2023. "Linked Employer-Employee History." Accessed 2023-04-18.

Huang, Naqun, Jing Li, and Amanda Ross. 2018. "The Impact of the Cost of Car Ownership on the House Price Gradient in Singapore." *Regional Science and Urban Economics* 68: 160–71.

Jerch, Rhiannon, Panle Jia Barwick, Shanjun Li, and Jing Wu. 2021. "Road Rationing Policies and Housing Markets." Unpublished.

Kahneman, Daniel, Alan B. Krueger, David A. Schkade, Norbert Schwarz, and Arthur A Stone. 2004. "A Survey Method for Characterizing Daily Life Experience: The Day Reconstruction Method." *Science* 306 (5702): 1776–80.
Kreindler, Gabriel E., and Yuhei Miyauchi. 2023. "Measuring Commuting and Economic Activity Inside Cities with Cell Phone Records." *Review of Economics and Statistics* 105 (4): 899-909.
Kuminoff, Nicolai V. 2012. "Partial Identification of Preferences in A Dual-Market Sorting Equilibrium." Unpublished.
Kuminoff, Nicolai V., V. Kerry Smith, and Christopher Timmins. 2013. "The New Economics of Equilibrium Sorting and Policy Evaluation Using Housing Markets." *Journal of Economic Literature* 51 (4): 1007–62.
Langer, Ashley, and Clifford Winston. 2008. "Toward a Comprehensive Assessment of Road Pricing Accounting for Land Use." In *Brookings-Wharton Papers on Urban Affairs,* edited by gary Burtless and Janet Rothenberg Pack, 127–175. Washington DC, Brookings Institute Press.
Le Barbanchon, Thomas, Roland Rathelot, and Alexandra Roulet. 2020. "Gender Differences in Job Search: Trading off Commute against Wage." *Quarterly Journal of Economics* 136 (1): 381–426.
LeRoy, Stephen F., and Jon Sonstelie. 1983. "Paradise Lost and Regained: Transportation Innovation, Income, and Residential Location." *Journal of Urban Economics* 13 (1): 67–89.
Li, Fangzheng, Fengyi Li, Shuangjin Li, and Ying Long. 2020. "Replication Package For: Deciphering the Recreational Use of Urban Parks: Experiments Using Multi-Source Big Data for All Chinese Cities." *Science of the Total Environment* 701: 134896–134909. Accessed 2024-06-18.
Li, Shanjun. 2018a. "Better Lucky Than Rich? Welfare Analysis of Automobile License Allocations in Beijing and Shanghai." *Review of Economic Studies* 85 (4): 2389–2428.
Li, Shanjun. 2018b. "Replication Package for: Better Lucky Than Rich? Welfare Analysis of Automobile License Allocations in Beijing and Shanghai." https://www.jstor.org/stable/get_asset/10.2307/26543968?supp_index=1.
Liu, Youming, Shanjun Li, and Caixia Shen. 2020. "The Dynamic Efficiency in Resource Allocation: Evidence from Vehicle License Lotteries in Beijing." NBER Working Paper 26904
Lyu, Xueying. 2022. "Car Restriction Policies and Housing Markets." *Journal of Development Economics* 156: 102850.
Miyauchi, Yuhei, Kentaro Nakajima, and Stephen J. Redding. 2021 "Consumption Access and the Spatial Concentration of Economic Activity: Evidence from Smartphone Data." NBER Working Paper 28497.
Mohring, Herbert, and Mitchell Harwitz. 1962. *Highway Benefits: An Analytical Framework.* Evanston, IL: Northwestern University Press.
National Oceanic and Atmospheric Association (**NOAA**)**.** 2020. "Integrated Surface Database-Lite Dataset." Accessed 2020-03-23.
Petrin, Amil. 2002. "Quantifying the Benefit of New Products: The Case of Minivan." *Journal of Political Economy* 110 (4): 705–29.
Phaneuf, Daniel J., V. Kerry Smith, Raymond B. Palmquist, and Jaren C. Pope. 2008. "Integrating Property Value and Local Recreation Models to Value Ecosystem Services in Urban Watersheds." *Land Economics* 84 (3): 361–81.
Redding, Stephen J., and Esteban Rossi-Hansberg. 2017. "Quantitative Spatial Economics." *Annual Review of Economics* 9 (1): 21–58.
Small, Kenneth A. 1982. "The Scheduling of Consumer Activities: Work Trips." *American Economic Review* 72 (3): 467–479.
Small, Kenneth A. 2012. "Valuation of Travel Time." *Economics of Transportation* 1 (1-2): 2–14.
Small, Kenneth A., Erik T. Verhoef, and Robin Lindsey. 2007**.** *Economics of Urban Transportation***.** Oxfordshire, UK: Routledge.
Small, Kenneth A., Clifford Winston, and Jia Yan. 2005. "Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability." *Econometrica* 73 (4): 1367–82.
Timmins, Christopher, and Jennifer Murdock. 2007. "A Revealed Preference Approach to the Measurement of Congestion in Travel Cost Models." *Journal of Environmental Economics and Management* 53 (2): 230–49.
Tsivanidis, Nick. 2023. "Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogota's TransMilenio." Unpublished.
USDOT. 2015. "Revised Departmental Guidance on Valuation of Travel Time in Economic Analysis." US Department of Transportation, Washington, DC.
Verhoef, Erik T., and Herbert Mohring. 2009. "Self-Financing Roads." *International Journal of Sustainable Transportation* 3 (5-6): 293–311.
Viard, V. Brian, and Shihe Fu. 2015. "The Effect of Beijing's Driving Restrictions on Pollution and Economic Activity." *Journal of Public Economics* 125: 98–115.

Vickrey, William. 1963. “Pricing in Urban and Suburban Transport.” *American Economic Review* 53 (2): 452–65.

Vickrey, William. 1965. “Pricing as a Tool in Coordination off Local Transportation.” In *Transportation Economics,* 275–96. Cambridge, MA: National Bureau of Economic Research.

Wasi, Nada, and Michael P. Keane. 2012. “Estimation of Discrete Choice Models with Many Alternatives Using Random Subsets of the Full Choice Set: With an Application to Demand for Frozen Pizza.” Unpublished.

Winston, Clifford. 1991. “Efficient Transportation Infrastructure Policy.” *Journal of Economic Perspectives* 5 (1): 113–27.

World Bank Open Data. 1987-2023. “FP.CPI.TOTL.ZG Inflation, Consumer Prices” Annual Statistics https://data.worldbank.org/ indicator/FP.CPI.TOTL.ZG?locations=CN. Accessed 2024-04-14.

Xiao, Junji, Xiaolan Zhou, and Wei-Min Hu. 2017. “Welfare Analysis of the Vehicle Quota System in China.” *International Economic Review* 58 (2): 617–50.

Yang, Jun, Avralt-Od Purevjav, and Shanjun Li. 2020a. “The Marginal Cost of Traffic Congestion and Road Pricing: Evidence from a Natural Experiment in Beijing.” *American Economic Journal: Economic Policy* 12 (1): 418–53.

Yang, Jun, Avralt-Od Purevjav, and Shanjun Li. 2020b. “Replication package for: The Marginal Cost of Traffic Congestion and Road Pricing: Evidence from a Natural Experiment in Beijing.” American Economic Association [publisher] https://www.aeaweb.org/journals/dataset?id=10.1257/pol.20170195.

Yinger, John. 1993. “Bumper to Bumper: A New Approach to Congestion in An Urban Model.” *Journal of Urban Economics* 34 (2): 249–74.

Zhang, Wei, C-Y Cynthia Lin Lawell, and Victoria I. Umanskaya. 2017. “The Effects of License Plate-Based Driving Restrictions on Air Quality: Theory and Empirical Evidence.” *Journal of Environmental Economics and Management* 82: 181–220.