

Efficiency and Equity Impacts of Urban Transportation Policies with Equilibrium Sorting

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

February 2022

Abstract

We estimate an equilibrium sorting model of housing location and commuting mode choice with endogenous traffic congestion to evaluate the efficiency and equity impacts of a menu of urban transportation policies. Leveraging fine-scale data from household travel diaries and housing transaction data identifying residents' home and work locations in Beijing, we recover structural estimates with rich preference heterogeneity over both travel mode and residential location decisions. Counterfactual simulations demonstrate that even when different policies reduce congestion to the same degree, their impacts on residential sorting and social welfare differ drastically. First, driving restrictions create large distortions in travel choices and are welfare reducing. Second, distance-based congestion pricing reduces the spatial separation between residences and workplaces and improves welfare for all households when it is accompanied by revenue recycling. Third, sorting undermines the congestion reduction under driving restrictions and subway expansion but strengthens it under congestion pricing. Fourth, the combination of congestion pricing and subway expansion delivers the greatest congestion relief and efficiency gains. It can also be self-financed, with the cost of subway expansion fully covered by congestion pricing revenue. Finally, eliminating preference heterogeneity, household sorting, or endogenous congestion significantly biases the welfare estimates and changes the relative welfare rankings of the policies.

Keywords: equilibrium sorting, housing markets, transportation, urban structure

JEL Classification Codes: L91, R13, R21

*Barwick: Department of Economics, Cornell University and NBER, panle.barwick@cornell.edu. Li: Dyson School of Applied Economics and Management, Cornell University, NBER and RFF, SL2448@cornell.edu. Waxman: LBJ School of Public Affairs, University of Texas – Austin, awaxman@utexas.edu. Wu: Hang Lung Center for Real Estate, Tsinghua University, ireswujing@tsinghua.edu.cn. Xia: Cornell University, tx58@cornell.edu. We thank Kelly Bishop, Jan Brueckner, Shengmao Cao, 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, University of Illinois, University of Maryland, University of Texas – Austin, University of Wisconsin – 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 and 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.

1 Introduction

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 (Davis, 2008; Li, 2018; Akbar et al., 2018; Gu et al., 2020).¹ To address these challenges, local governments around the world have implemented a suite of policies, including driving restrictions, public transit 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 households adjustment of residential locations. This adjustment, 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 their income on transportation. This paper aims to understand the efficiency and equity impacts of urban transportation policies while accounting for multiple adjustment channels and equilibrium effects. To do so, we jointly model residential locations and travel mode choices in an equilibrium sorting framework with endogenous congestion.

The empirical context of our study is Beijing, which has a population of 21.5 million and has routinely been ranked 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 anticongestion policies (driving restrictions and subway expansion) together with a proposed congestion pricing scheme represent three general approaches to regulating the unpriced congestion externality—the first a command-and-control, the second a supply-side, and the third a demand-side approach.

Exploiting this policy-rich context, we first develop a stylized theoretical model based on LeRoy and Sonstelie (1983) and Brueckner (2007) to account for endogenous congestion and heterogeneity in income and commuting technologies. The model highlights the countervailing forces at play in travel mode and housing location choices and illustrates both the efficiency and the distributional consequences of transportation policies. Our reduced-form evidence corroborates the theoretical predictions that driving restrictions steepen the bid-rent curve and lead to a higher premium for properties in desirable locations (e.g., close to subway

¹The 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. The top five cities were Mumbai, Bogota, Lima, New Delhi, and Moscow. Drivers in Mumbai spent 65% more commuting time on average than they would have under the free-flow condition. Los Angeles, the most congested city in the US, was ranked 24th with a congestion index of 41%. Four cities in China (including Beijing) were among the top 30 on the list. Beijing's drivers spend nearly 180 extra hours on the road (or 9% of working hours) per year relative to the travel time under the free-flow speed due to traffic congestion. The full ranking based on the TomTom Traffic Index 2018 is available at https://www.tomtom.com/en_gb/traffic-index/ranking.

stations). Motivated by the theoretical predictions and reduced-form evidence, we then build and estimate an equilibrium model of residential sorting with endogenous congestion that incorporates preference heterogeneity and allows general equilibrium feedback between housing locations and commuting decisions through congestion. In the model, households choose a residence, and conditional on their home location, they also choose travel modes for commutes. A key consideration in a household's choice of a residential property is the associated ease of commute for working members of the household. The utility derived from an easier commute is an equilibrium object that crucially depends on congestion, which varies across locations and is determined by all households' travel choices and residential locations. Once estimated, the model allows us to conduct counterfactual simulations to predict new equilibrium outcomes under different policies in terms of travel mode choices, household locations, the congestion level, housing prices, and welfare distribution.

Our structural analysis leverages data at fine spatial resolution from two unique sources that allow us to jointly model residential locations and commuting choices. The first is the Beijing Household Travel Survey (BHTS) from 2010 and 2014, a large representative survey that records households' home and work locations, trips made in a 24-hour window, and other demographic and transportation-related information. We complement this dataset by constructing all feasible home-to-work commuting choices with historical geographical information system (GIS) maps and the application programming interface (API) from online mapping services. This exercise allows us to compile the commuting route, distance, travel time and pecuniary travel cost for each trip-mode combination (walking, biking, bus, subway, car, or taxi). The second dataset contains housing transactions from a major government-run mortgage program and provides a large representative sample of Beijing home buyers. Critically for our analysis, the housing data report not only the home but also the work locations of both the primary and secondary borrowers. Using this information, we construct over 13 million hypothetical work-commute and travel-mode combinations for all primary and second borrowers with the same GIS and API procedure used for the travel survey data. To our knowledge, these datasets constitute the most comprehensive data on work-commute travel and housing transactions ever used in the context of equilibrium sorting models.

We use a two-step strategy to estimate the equilibrium sorting model. The first step recovers heterogeneous preferences on travel times and monetary costs (and thereby the value of time) using information on the time and pecuniary cost for both the chosen travel mode that is reported in the household travel surveys and alternative modes that we constructed using various GIS and APIs. We then utilize the estimated parameters from this step and the work locations of both the primary and secondary borrower (which primarily correspond to husband and wife in our setting) to construct the ease-of-commute attribute separately for each commuter in the household and for all properties in a household's choice set. These variables are included as household property-specific attributes in the housing demand estimation below.

The second step of our estimation procedure recovers preferences for housing attributes from observed home purchases. The key estimation challenge arises from the potential correlation between unobserved housing attributes and the housing price as well as the ease-of-commute utility, where the latter two variables are equilibrium outcomes. To address this challenge, we construct three sets of instrumental variables (IVs)

in the spirit of [Berry et al. \(1995\)](#) and [Bayer et al. \(2007\)](#). These IVs include the number of houses sold in a two-month window around the sales date within a reasonable distance from a given property. We also use as IVs the average housing and neighborhood attributes for these properties and the time-varying odds of winning a license lottery to purchase a vehicle. The first and second sets of IVs proxy for the extent of competition faced by a given property (the number and attributes of alternative properties on the market). The third set of IVs reflects exogenous policy-induced shifts in demand for houses in premium locations, such as places close to subways or in the city center. We account for a rich set of observed and unobserved preference heterogeneity and estimate parameters through maximum likelihood estimation with a nested contraction mapping combined with the IV approach ([Train and Winston, 2007](#)).

Our results confirm the importance for both the travel mode and residential location decisions of preference heterogeneity, which significantly improves model fit. The average and median value of time (VOT) from our preferred specification is 95.6% and 84.6% of survey respondents' hourly wage, consistent with estimates in recent literature. In addition, the ease-of-commute attributes substantially improve the ability to explain observed housing choices. A noteworthy result of the estimation is that households are willing to pay 18% more for similar reductions in a wife's than in a husband's commuting time. Our structural model provides the first estimates of the income elasticity of housing demand and the income elasticity of the marginal commuting cost in one unified framework. These elasticities are key determinants of the urban spatial patterns of residential locations.

Utilizing these estimates, we then simulate equilibrium residential sorting and transportation outcomes based on three policies of interest in our study: driving restrictions, subway expansion, and congestion pricing, as well as combinations of the three. We decompose the welfare effect along five margins. The first channel measures changes in welfare when households change travel mode in response to increasing commuting costs, holding congestion and residential locations fixed. The second and third channels separately consider partial speed adjustments that are common to the empirical transportation literature. The first adjustment does not clear the transportation sector, while the second is a full equilibrium speed adjustment that clears the transportation sector. The fourth channel incorporates sorting and simulates general equilibrium outcomes with changes in both congestion and residential locations. 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.

Our policy simulations yield four 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. In contrast, distance-based congestion pricing provides strong incentives for commuters in *both* the high- and low-income groups to move closer to their workplaces. In comparison, subway expansion generates the most variable changes in commuting costs across households. The sorting responses from subway expansion are also the opposite of what occurs under congestion pricing. Subway expansion disperses households away from the city center into the suburbs and locations near new

subway stations.

Second, different transportation policies can either exacerbate or alleviate economic inequality ([Waxman, 2017](#); [Tsivanidis, 2019](#); [Akbar, 2020](#)). High-income households' welfare is greater under congestion pricing (in the absence of revenue recycling), and low-income households' welfare is greater under subway expansion. Without revenue recycling, congestion pricing is regressive, creating a significant impediment to its adoption in practice. With appropriate revenue recycling, low-income households can also be better off under congestion pricing than in the no-policy scenario.

Third, residential sorting can either strengthen or undermine the congestion-reduction potential of transportation policies. Sorting enhances the efficacy of congestion pricing for congestion relief 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% for high-income households and 16% for low-income households. On the other hand, sorting in response to subway expansion leads to a further separation between residential and work locations, dampening the congestion-reduction effect and welfare gains from infrastructure investment.

Finally, transportation policies generate different aggregate welfare implications. Beijing's rapid subway expansion increased consumer surplus and aggregate welfare despite the fact that it has achieved only a modest congestion reduction. In contrast, driving restrictions are welfare reducing in spite of their larger associated congestion reduction. Congestion pricing and subway expansion in tandem deliver the largest improvement to traffic speed and net welfare gain—equivalent to 3% of average household 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 strengths 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 several extensions and robustness checks, considering housing supply variation, ring road-specific traffic density, migration, and consumption access, as well as the removal of random coefficients. Housing supply fluctuations in response to housing price changes allow more people to move to desirable locations and magnify the role of sorting. Our results are similar with region-specific traffic density. Shutting down random coefficients grossly underestimates the benefit of subway expansion and overestimates the welfare loss from driving restrictions and congestion pricing. Incorporating migration and consumption access considerations does not change the qualitative results of our analysis.

Our study makes three main contributions. First, it adds to the literature on equilibrium sorting. Sorting models have been used to study consumer preferences for local public goods and urban amenities (e.g., air quality, school quality, and open space) and evaluate policies that address economic, social and environmental challenges ([Epple and Sieg, 1999](#); [Bayer et al., 2007](#); [Kuminoff et al., 2013](#)).² Most existing papers treat both

²See, for example, [Ferreyra \(2007\)](#) and [Epple et al. \(2012\)](#) on school quality; [Sieg et al. \(2004\)](#), [Bayer et al. \(2009\)](#), [Kuminoff \(2009\)](#), and [Bayer et al. \(2016\)](#) on air quality; [Timmis and Murdock \(2007\)](#), [Walsh et al. \(2007\)](#), and [Klaiber and Phaneuf \(2010\)](#) on open space and recreation; [Bajari and Kahn \(2005\)](#), [Bayer et al. \(2007\)](#), [Bayer and McMillan \(2012\)](#), and [Hwang \(2019\)](#) on racial and ethnic composition; [Calder-Wang \(2020\)](#) on the distributional impacts of the sharing economy in the housing market; [Almagro and](#)

the distance to work and the level of congestion as exogenous attributes. An attempt to address this limitation is made in [Kuminoff \(2012\)](#), which models household decisions in both the work and housing markets and endogenizes the commuting distance but keeps congestion exogenous. Our paper is, to our knowledge, the first in the empirical equilibrium sorting literature to explicitly model household residential locations and travel mode choices jointly and evaluate how these choices simultaneously determine both congestion and distance to work in equilibrium. We show that the aggregate welfare implications qualitatively differ based on 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). There has been a limited attempt in this literature to examine the role of preference heterogeneity and endogenous congestion in mediating the welfare effects of transportation policies. Our framework accounts for rich heterogeneity in observed and unobserved preferences by leveraging detailed household-level data instead of relying on more tractable distributional assumptions (i.e., Fréchet) and limited differentiation (such as heterogeneity in educational attainment) as the QSE models do.³ Incorporating such preference heterogeneity matters for evaluating equilibrium responses to transportation policies: removing unobserved preference heterogeneity results in welfare estimates that are both qualitatively and quantitatively different. Some recent QSE studies ([Allen and Arkolakis, 2019](#); [Fajgelbaum and Schaal, 2020](#)) explicitly model endogenous congestion with an increasing marginal external cost ([Anderson, 2014](#)). Unlike these studies, our work estimates the VOT and congestion costs based on individual-level commuting decisions from large travel surveys.

Third, our paper bridges a gap in the literature examining short- and long-run responses to transportation policies.⁴ While most studies in this literature focus on short-run effects on travel choices, traffic congestion, and air pollution, some studies examine longer-run partial equilibrium effects and find substantially different results (e.g., [Duranton and Turner \(2011\)](#)). As pointed out by [Gallego et al. \(2013\)](#), little has been done to understand the set of adjustments that happen in the transition from the short to the medium or long run. Understanding these adjustments is crucial from a policy perspective since municipalities often need to plan for infrastructure provision and address development concerns over the medium run. By characterizing the underlying travel and housing choices, our equilibrium sorting framework provides a microfoundation for linking the results between short- and long-run reduced-form impact evaluation studies. More importantly, the unified framework offers a common yardstick to compare actual and counterfactual policies over a range of outcomes

[Domínguez-Iino \(2020\)](#) and [Couture et al. \(2020\)](#) on endogenous amenities across neighborhoods, and [Ferreira and Wong \(2021\)](#) on recovering neighborhood preferences with imperfect information.

³An important exception is [Couture et al. \(2020\)](#), who examine the welfare effects of US urbanization in a quantitative spatial model that allows significant heterogeneity. However, the model abstracts from the general equilibrium effects in housing and transportation markets at play in this paper.

⁴Various policies have been evaluated. See [Parry and Small \(2005\)](#), [Bento et al. \(2009\)](#), and [Li et al. \(2014\)](#) on gasoline taxes; [Bento et al. \(2005\)](#), [Parry and Small \(2009\)](#), [Duranton and Turner \(2011\)](#), [Anderson \(2014\)](#), [Basso and Silva \(2014\)](#), [Li et al. \(2019\)](#), [Severen \(2019\)](#), and [Gu et al. \(2020\)](#) on public transit subsidies and expansion; [Davis \(2008\)](#), [Viard and Fu \(2015\)](#), [Zhang et al. \(2017\)](#), and [Jerch et al. \(2021\)](#) on driving restrictions; and [Langer and Winston \(2008\)](#), [Anas and Lindsey \(2011\)](#), [Hall \(2018\)](#), [Yang et al. \(2020\)](#), and [Kreindler \(2018\)](#) on congestion pricing.

including congestion reduction, urban spatial structure, social welfare, and distributional consequences.⁵

The paper proceeds as follows. Section 2 describes the data and policy background. Section 3 sets up the theoretical framework and provides reduced-form evidence on how Beijing’s driving restrictions affect the housing market. The reduced-form evidence motivates and grounds the subsequent structural analysis. Section 4 lays out the equilibrium sorting model and the estimation strategy. The estimation results are presented in Section 5. Section 6 explains the counterfactual simulation algorithm. Section 7 examines different transportation policies and compares their welfare consequences. Section 8 concludes.

2 Policy Background and Data Description

2.1 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 a driving restriction scheme, 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.⁶ Initially, half of all vehicles were restricted from driving on a given weekday based on their license plate name. After the Olympics concluded, the restrictions were relaxed to apply to each car on only one weekday per week depending on the last digit of the plate. Acquiring a second vehicle to avoid the restriction is difficult since in 2011, Beijing also put in place a binding quota system that caps the number of new vehicle sales in an attempt to curb the growth in vehicle ownership. Winning a car license in this lottery became increasingly difficult over time: the odds of winning decreased from 1:10 in early 2012 to nearly 1:2,000 in 2018, as the pool of lottery participants increased while the number of licenses fell over time (Xiao et al., 2017; Li, 2018; Liu et al., 2020).⁷ These time-series changes in the odds of winning provide useful exogenous variation for the housing demand analysis, as we discuss below.

Beyond implementing these 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 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 four to over 40 from 2000 to

⁵Another approach in the literature allows feedback effects between the transportation sector and housing market in a calibrated computable general equilibrium framework without estimating the underlying consumer preferences (Anas and Kim, 1996; Langer and Winston, 2008; Parry and Small, 2009; Basso and Silva, 2014). Compared with the approach in these studies, our framework is internally consistent in that the estimation of the structural parameters and policy simulations are based on the same model.

⁶Athens, 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 et al., 2017).

⁷About 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 set considerably lower than the historical vehicle sales in Beijing.

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. (See [Anderson \(2014\)](#); [Yang et al. \(2018\)](#); [Gu et al. \(2020\)](#) for a recent analysis on the impact of subway expansion on traffic congestion.)

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 (7-9 am and 5-7 pm), according to the 2020 Beijing Transportation Report. From a neoclassical microeconomic perspective, the aforementioned policies fail to directly address the root cause of traffic congestion: the mispricing of road capacity.⁸

2.2 Data Description

We rely on two main datasets for our analysis: a) the Beijing Household Travel Surveys from 2010 and 2014 and b) housing mortgage data from 2006-2014 with detailed information on household demographics and the work addresses of home buyers. Appendix Section A provides more detail on the data construction.

Beijing's Geography Beijing's spatial structure is characterized by high population density at the center, with a set of concentric ring roads encircling the city center. The second ring road largely traces the city limits of pre-1980s Beijing, from which the city has subsequently expanded outward. We focus on the geographical areas within the sixth ring road, which approximately separates the city proper from its suburbs. Appendix Figures A2 and A3 map out the city contour and various ring roads, commercial centers (with a greater density of job opportunities), subway lines, districts, and amenities including signature elementary schools and parks.

Beijing is not perfectly monocentric. There are several large work clusters across the city, such as the financial cluster between the second and fourth ring roads on the east side of the city and a high-tech cluster toward the northwest between the third and fifth ring roads. The city has 65 signature schools designated by the municipal government as the key elementary schools. These schools have better resources and better student performance. Signature schools are concentrated within the fourth ring road, while parks are more dispersed across the city.

Beijing has a total of 18 districts, each containing on average eight *jiedao* (neighborhoods). A *jiedao* is an administrative unit, similar to a census tract, where homes share similar observed and unobserved amenities. The average size of a *jiedao* is 15.7 square km. For transportation planning purposes, Beijing is also divided into roughly 2,000 traffic analysis zones (TAZs), which are standardized spatial units based on residential and employment density. TAZs are one square kilometer on average and smaller when they are closer to the center of Beijing. Most of the maps in this paper use TAZs as the spatial unit.

⁸Despite being continuously advocated by economists since [Vickrey \(1963\)](#), adoption of congestion pricing has been limited in practice largely due to technical feasibility and especially political acceptability. During the last 15 years, several European cities (London, Milan, Stockholm, and Gothenburg) have successfully implemented various congestion pricing schemes. After considering several proposals, the New York state legislature recently approved a congestion pricing plan for New York City, which—pending approval of the scheme by the Federal Highway Administration—is set to become the first US city to enact congestion pricing.

Beijing Household Travel Survey We utilize two rounds of the Beijing Household Travel Survey (BHTS) collected in 2010 and 2014 by the Beijing Transportation Research Center (BTRC), an agency of the Beijing municipal government. The survey is designed to inform transportation policies and urban planning. It includes data on individual and household demographics (e.g., income, household size, vehicle ownership, home ownership, age, gender) and occupations, availability of transportation options (vehicles, bikes, etc.), and a travel diary on all trips taken during the preceding 24 hours. This trip information includes the origin and destination, departure and arrival time, trip purpose and travel mode used.

Our analysis focuses on 73,154 work commuting trips (home-to-work and work-to-home). Work trips account for the majority of the total travel distance and time: they constituted 62% and 75% of the total travel distance and 53% and 59% of the weekday trips among working-age respondents in 2010 and 2014, respectively. Table 1 provides summary statistics for variables used in the analysis by survey year. Household income increased dramatically from 2010 to 2014, with the share of the lowest income group (less than ¥50,000 annually) decreasing from 48% to 18%. The proportion of households owning vehicles increased from 44% to 62%. The share of respondents living and the share of those working within the fourth ring road (which proxies for the city center) both decreased by about 10 percentage points from 2010 to 2014, reflecting the increased spatial dispersion of housing and work locations.

To understand commuters' travel mode choices, we construct attributes for all travel modes in their choice set. We focus on six travel modes: *walk*, *bike*, *bus*, *subway*, *car*, and *taxi*, as other modes (motorcycles, company shuttles, and unlicensed taxis) collectively account for less than 4% of all trips. Appendix Figures A4 and A5 illustrate the procedures used to calculate the mode-specific travel time and monetary cost. We use the Baidu API to calculate the travel time and distance for walking, biking, car and taxi trips. Baidu Maps incorporates the predicted congestion level based on the time of day and day of week in its estimated trip duration. We query the Baidu API at the same departure time as that recorded in the travel survey (e.g., 7 am) to capture within-day variation in congestion (i.e., at peak vs. off-peak hours). To account for changes in the average congestion between the survey year and the year that we query the Baidu API, we adjust the predicted driving, taxi, and bus travel times based on the historical traffic congestion index (e.g., a 10% difference in the traffic congestion index is associated with a 10% adjustment of the travel time). The driving speed is the ratio of the travel distance to the travel time.

We use the Gaode Map API to calculate the travel time by bus because Gaode reports the number of transfers and walking time between bus stops and delivers more accurate estimates. To take into account the subway expansion occurring during our sample period, we use historical subway maps and GIS software to reconstruct the historical subway network. The subway travel time is calculated based on the published time schedules of subway lines. Our calculation assumes that commuters use the subway stations closest to their trip origin and destination and incorporates the walking distance and walking time to the subway stations in the total trip distance and duration. We validate these constructed trip-mode attributes (e.g., duration) with information from reported trips in the travel survey.

Figure 1 plots for each travel mode the observed share of commuting trips and the constructed travel time,

cost, and distance. Panel (a) contrasts the travel patterns in 2010 with those in 2014 and presents several notable changes. First, walking accounts for a significant share of all commuting trips: 15.0% and 13.5% in 2010 and 2014, respectively. These trips take 51 and 40 minutes on average with a distance of 4.9 and 3.7 km. Second, 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. Third, 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-income (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 travel modes. As a percentage of the hourly wage, car and taxi trips are much more expensive for low-income than for high-income households. There is little difference in travel distance across the two income groups except for the distance of car trips. This echoes evidence from the housing data below that there is limited income-delineated residential sorting.

Housing Transactions The data on housing transactions come from a major government-sponsored mortgage program in Beijing and cover July 2006 to July 2014. As is reflective of the housing supply in urban China, most housing units are within housing complexes, equivalent to condominiums in the US. 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% lower than commercial mortgage rates. Each transaction in our data corresponds to a mortgage, 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, and street address, the transaction price, and then date when the mortgage was signed.⁹ We also observe household demographic details including income, age, marital status, residency status (*hukou*), and—critically for our analysis—the work addresses of the primary borrower and the coborrower if one is present. We geocode the home and work addresses to their latitude and longitude 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 all housing transactions from a separate dataset by using entropy balancing (Hainmueller, 2012). All of the empirical analyses use the weighted sample. The results estimated with the unweighted sample are robust. Appendix Section A.2 discusses the reweighting procedure in more detail and describes additional data patterns, such as differences in commuting distance by gender.

Table 2 provides summary statistics of the data. Figure 2 shows the spatial pattern of housing and house-

⁹We remove transactions with a missing or zero reported price, a price lower than $\text{¥}5,000/m^2$ (the average price is $\text{¥}19,800/m^2$), buyers with no reported income, and addresses outside the sixth ring road.

hold 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 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 illustrated in the monocentric city model in Section 3. There are also exceptions. For example, the high-tech center in northwestern Beijing outside the fifth ring road has high housing prices and short commutes that are comparable to those corresponding to places in the city center. The northern parts of the city have higher-quality amenities (schools and parks) and more work opportunities and attract high-income households. While household income is generally higher in northern Beijing than in southern neighborhoods, households of different income levels tend to mix together throughout most parts of the city.

Our equilibrium sorting model examines households’ residential choices. In theory, a home buyer’s choice set could potentially consist of *all* properties listed on the market, and researchers need to construct hypothetical commuting attributes of different travel modes for all properties in a buyer’s choice set. However, this is technically infeasible because the number of potential home–work–mode combinations exceeds hundreds of billions. This issue of large choice sets is a common empirical challenge in the housing demand literature. To reduce the computational burden, we follow a choice-based sampling strategy.¹⁰. The choice set for a household is assumed to include the purchased home and a 1% sample of houses randomly chosen from those sold during a two-month window around the purchase date. Beijing’s real estate market was fluid during our sample period. The median number of days on the market was only 8 and 13 in 2013 and 2014, respectively, with the average in the same years being 22 and 38 days. Section 5 conducts robustness analyses using different choice 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 coborrower’s work commute, based on their respective work locations. The construction of the travel mode attributes involves over 13 million route–mode combinations.

Auxiliary Dataset To estimate the relationship between traffic density and speed, we use the same traffic data as [Yang et al. \(2020\)](#). The data contain real-time traffic volume and speed data at 30-minute intervals from over 1,500 remote traffic microwave sensors covering all major roads throughout Beijing for 2014.

3 Theoretical Framework and Reduced-Form Evidence

3.1 Theoretical Framework

We motivate our setup with a graphical presentation of two of the transportation policies examined in our empirical analysis: congestion pricing and driving restrictions. These policies affect both the transportation sector (the primary market) and also the housing market (the secondary market) through sorting.

¹⁰Choice-based sampling for differentiated demand has been demonstrated to yield consistent results in [Wasi and Keane \(2012\)](#) and [Guevara and Ben-Akiva \(2013\)](#)

Figure 3 illustrates the welfare effects in the primary market for vehicle road traffic. The economic cost induced by congestion is the difference between the marginal social cost (*MSC*) and marginal private cost (*MPC*). Both congestion pricing and driving restrictions result in a reduction of the traffic volume from the unregulated level, V^0 , to the socially optimal level, V^* . However, there are crucial differences between the two schemes. Congestion pricing reduces the trips with the lowest marginal benefit and leads to welfare gains (the shaded red area). Driving restrictions, as a command-and-control approach, eliminate trips at all levels of marginal benefit. The blue triangle in Figure 3 represents the welfare loss associated with the quantity restriction. Its size is positively related to the degree of heterogeneity in the marginal benefit of trips. As a result, the welfare impact of the driving restrictions is ambiguous.

Figure 3 is only a partial equilibrium analysis and does not take into account the impact of transportation policies on the housing market. It also abstracts from differences across households. To understand these additional effects, we extend the classical monocentric models of land use in [LeRoy and Sonstelie \(1983\)](#) and [Brueckner \(1987\)](#) by incorporating income heterogeneity, multiple transportation technologies, and endogenous congestion, which are important elements in our empirical model. Households with different incomes sort into different locations in response to transportation policies based on their preferences for housing, other goods and time. Appendix Section B presents the model details and the theoretical predictions.

The model predicts that both the driving restriction and the congestion pricing schemes steepen the bid-rent curve and lead to a higher premium on proximity to the central business district (i.e., workplace) but that subway expansion yields the opposite effect. In addition, driving restrictions are more likely than congestion pricing to push rich households to move closer to their workplaces, as driving restrictions impose higher costs on rich households. Subway expansion disperses both groups to the city suburbs. The model highlights that while the transportation policies' primary effect on commuting costs is straightforward, their secondary effect on the spatial distribution of households and housing price capitalization can be large and depends on relative differences in the marginal cost of commuting, income heterogeneity, and preferences for housing size and travel time. While illustrative, the theoretical model ignores a host of important features, such as the polycentricity of most cities, travel modes beyond car and subway trips, household attributes other than income, and variation in the availability of housing across the city. For these reasons, we turn to an empirical equilibrium sorting model to evaluate different policies.

3.2 Reduced-Form Evidence

Before proceeding to the structural model, we present reduced-form evidence for the capitalization of transportation policies into housing prices. Specifically, we examine the housing market response to the car driving restriction (CDR) policy that began in July 2008. Figure 4 shows scatter plots of housing prices in ¥1,000/ m^2 against the distance to the nearest subway station before and after the CDR.¹¹ The top panel uses raw data,

¹¹We focus on two years before and after the starting date of the program to balance the trade-off between sample size and potential confounding changes in the housing market and transportation sector. We remove observations from July to September 2008. The policy was more aggressive in July–September 2008, when half of all vehicles were subject to the restriction on any given weekday

while the bottom panel shows residualized plots after we include year-by-month and neighborhood fixed effects. The price gradient becomes steeper post-CDR, suggesting that homes close to subways command a higher price premium under the policy: being one kilometer closer to subway increases a house's price premium by about $\text{¥}100/m^2$ (about 1% of the housing price). Consistent with theoretical predictions, the driving restrictions increase the price premium of homes near subway stations. Appendix Table A1 reports regression results for various specifications, suggesting a very similar impact magnitude even after we include richer controls such as complex attributes, city district-by-year fixed effects, and subway density to better account for time-varying amenities.

Appendix Section C provides an event-study analysis and a falsification test, lending additional support to the finding. In addition, we specify a piecewise linear function of price on subway distance to allow the price gradient to vary by distance band. The results suggest a convex price function, where the marginal impact of subway proximity is the strongest for properties within 5 km but tapers off after 10 km. The relationship between subway proximity and housing price is strengthened under the policy, especially for properties between 5 and 10 km away.

The reduced-form analysis above confirms the importance of housing market capitalization and the presence of sorting in response to transportation policies. To compare the impacts of different transportation policies and to quantify the underlying mechanisms and margins of adjustment (changes in commuting modes vs. changes in residential locations), we now turn to an equilibrium sorting model that features preference heterogeneity and endogenous congestion.

4 Empirical Equilibrium Sorting Model

The sorting model characterizes individual commuting choices and residential location decisions. 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 and affect driving demand and hence traffic congestion. On the other hand, traffic congestion affects the attractiveness of different residential locations and consequently housing demand. For example, high congestion levels increase demand for premium locations (places close to subways and the city center). The equilibrium nature of our sorting model allows counterfactual simulations and provides direct comparative statics of congestion levels, residential locations, housing prices, and social welfare across different policies.

The model assumes that work locations are determined ex ante and examines housing choices given work locations. In practice, these decisions could be made either simultaneously or sequentially. Our assumption is motivated by three observations. First, for many households, the choice of work location is likely to be the outcome of a longer-term process of labor supply and migration decisions.¹² Second, employment opportu-

(due to the 2008 Olympics in August), and was relaxed to a restriction on driving during one weekday per week for each car from October 2008.

¹²The annual rate of job switching among formal-sector employees in Beijing was about 4% per year from 2006 to 2014. About 60% of home buyers in our data started a new job within the three years prior to purchasing the home.

nities in the same industry tend to be clustered in Beijing. Hence, switching jobs may not entail meaningful changes in work locations. Third, while the mortgage data provide rich information on housing locations and current employment, they do not report the alternative job opportunities available to each household. In addition, adding the labor market component would significantly complicate our empirical analysis given the rich individual-level observed and unobserved preference heterogeneity incorporated into our model.

Our approach contrasts with the emerging literature that uses QSE models to evaluate transportation policies incorporating the joint processes of work and residential location decisions. Unlike that literature, the current work cannot analyze whether changes in the transportation system translate to higher labor productivity (e.g., through better allocation of time and labor market matching), which is a limitation of our approach.¹³ On the other hand, our approach has several advantages. First, QSE models use observed worker flows and wages to recover iceberg commuting costs via gravity equations and origin–destination-specific (dis)amenities. Recent studies ([Allen and Arkolakis, 2019](#); [Fajgelbaum and Schaal, 2020](#)) allow endogenous congestion but still rely on gravity equations to describe commuting flows. In contrast, we estimate preferences over commuting time and monetary costs and hence the VOT directly through observed individual commuting choices and explicitly model the externality of endogenous congestion. This is important as VOT is the single most important parameter for the welfare effects of transportation policies.¹⁴ Second, in contrast to the flexible and rich observed and unobserved preference heterogeneity in our setting, unobserved preference heterogeneity in QSE models is often limited to Fréchet draws for analytic tractability, with homothetic preferences for ease of aggregation.¹⁵ The ability to flexibly specify heterogeneous preferences (which are nonhomothetic) has important welfare implications, as we document in Section 7 below.

4.1 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 ([Lancaster, 1971](#); [Berry et al., 1995](#)). Our data are longitudinal, but we suppress time t to ease exposition. Variables in bold denote vectors. Conditioning on work locations, the utility for household i choosing housing

¹³In addition, our analysis does not model agglomeration forces. [Diamond \(2016\)](#) is a microfounded study that bridges this gap by incorporating housing and labor markets into the evaluation of the heterogeneous welfare consequences of worker movements between US cities, although it does not endogenize congestion from the transportation sector. We also abstract from dynamic considerations. For recent papers with dynamic models on the housing market, see [Almagro and Domínguez-Iino \(2020\)](#); [Han et al. \(2018\)](#); [Murphy \(2015\)](#); [Wang \(2020\)](#).

¹⁴As Ken Small has written in [Small \(2012\)](#), “It is difficult to name a concept more widely used in transportation analysis than the value of travel time. Its theoretical meaning and its empirical measurement are fundamental to travel demand modeling, social cost analysis, pricing decisions, project evaluation, and the evaluation of many public policies.”

¹⁵Welfare improvements in QSE models usually result from changes in real income due to gains from trade via an increase in market access. This benefit is directly mediated through import elasticity with respect to variable trade costs ([Arkolakis et al., 2012](#)). In the context of urban transportation, this approach seems potentially limiting because spatial mismatch and wasteful commuting due to preexisting distortions, like congestion, may leave open opportunities for Pareto improvements without a change in the level of market access.

unit j is specified as:

$$\max_{\{j \in J_i\}} U_{ij} = \alpha_i p_j + \mathbf{x}_j \beta_i + \sum_k \phi_{ik} EV_{ijk}(v_{ijk}) + \xi_j + \varepsilon_{ij}, \quad (1)$$

where J_i is the choice set for household i , p_j denotes the home price, and \mathbf{x}_j denotes a vector of observed housing attributes such as the size and number of bedrooms. Household members with commuting needs are denoted by $k \in \{\text{Primary borrower, Coborrower}\}$. $EV_{ijk}(v_{ijk})$ is the expected commuting utility for member k in household i derived 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} in member k 's work commute (which is affected by congestion) in addition to the travel time and cost. As shown in Section 4.2, this commuting utility is our key innovation relative to traditional residential sorting models. In our setup, transportation policies can generate differential impacts on the commuting utility of housing units through a variety of channels: households' heterogeneous commuting preferences, changes in households' commuting modes, and endogenous congestion (which affects the driving speed) as a result of the policy. 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 log of household income y_i :

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

Household preferences over housing attributes are denoted as β_i , which consists of a household-specific component and a population average. For each element s in β_i :

$$\beta_{is} = \bar{\beta}_s + \mathbf{z}_i \beta_s,$$

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:

$$\phi_{ik} = \bar{\phi}_k + \phi_k \zeta_{ik}, k \in \{\text{Primary borrower, Co-borrower}\},$$

where ζ_{ik} is i.i.d. normal. In subsequent formulations, we suppress subscript k for ease of exposition and use EV_{ij} to denote the commuting utility for both household members $\sum_k \phi_{ik} EV_{ijk}$.

The probability that household i chooses home j is denoted by:

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

where \mathbf{p} and \mathbf{v} denote prices and driving speed and $\mathbf{EV}(\mathbf{v})$ is a vector of the ease-of-commute utility for different properties given household i 's work locations. The triplet \mathbf{X} , $\boldsymbol{\xi}$, and \mathbf{z}_i denotes observed housing

attributes, unobserved housing quality, and household i 's demographics, respectively.

4.2 Choice of Travel Mode

Utility-maximizing individuals within a household choose from six commuting modes (walk, bike, bus, subway, car, and taxi) based on the trip time and financial cost of each. The travel survey reports travel mode choices for each commuting member of a household.¹⁶ With slight abuse of notation, we use i to denote an individual within 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:

$$\max_{m \in M_{ij}} u_{ijm} = \theta_{im} + \gamma_1 \cdot \text{time}_{ijm}(v_{ij}) + \gamma_2 \cdot \text{cost}_{ijm}/y_i + \mathbf{w}_{ijm}\eta + \varepsilon_{ijm}, \quad (3)$$

where M_{ij} is the set of transportation modes available to individual i commuting from home j . Variable time_{ijm} denotes the commute duration between i 's work location and home j via mode m . The driving time for trips with the commuter's own vehicle or taxis ($\text{time}_{ij,car}$ and $\text{time}_{ij,taxi}$) depends on the driving speed v_{ij} , which is ultimately determined by the congestion level.¹⁷ The monetary cost of the trip is denoted as cost_{ijm} and household income as y_i . The variable \mathbf{w}_{ijm} includes mode–commuter-specific controls (such as the driving dummy interacted with commuter's gender) and time and spatial fixed effects. Finally, ε_{ijm} is the i.i.d. error term with the type I extreme value distribution.

We allow a mode-specific random coefficient, θ_{im} , that has 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 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. The time preference γ_1 follows a chi-squared distribution with three degrees of freedom and mean μ_γ . The chi-squared distribution allows all individuals to have a positive value of time. An individual's sensitivity to the monetary costs of commuting is assumed to decrease in income: γ_2/y_i .

Our utility specification makes it straightforward to calculate the VOT, which is $\frac{\gamma_1}{\gamma_2} \cdot y_i$. This measure scales with income by linking the valuation of time to the hourly wage. The VOT is the most important preference parameter for transportation decisions (Small, 2012). Small et al. (2005) demonstrate the importance of modeling rich preference heterogeneity to recover accurate VOT estimates and evaluate transportation policies. Section 5 illustrates that our estimates accurately reflect the preferences of Beijing commuters.

¹⁶We treat the mode choices of different individuals within a household as independent, as we do not observe whether and how mode choices within households are determined. We also abstract from trip-chaining, which is unlikely to be of first-order importance for home buyers.

¹⁷Road congestion affects the travel time for bus trips in addition to car and taxi trips. However, this effect is more complicated as it depends on the local design of the roadway, structure of bus schedules, and locations of bus stops. For the purpose of our analysis, we treat buses as if they run in dedicated lanes unaffected by congestion, which may result in an overprediction of bus mode shares in simulations with higher congestion levels.

Conditional on home location j , the probability that individual i chooses mode m for his or her work commute is defined as:

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

where \mathbf{cost}_{ij}/y_i and $\mathbf{time}_{ij}(v_{ij})$ denote the vector of travel cost (as a share of individual i 's hourly wage) and travel time from home j to i 's work location for all travel modes, respectively. The 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:

$$EV_{ij}(v_{ij}) = \mathbb{E}_{\varepsilon_{ijm}} \left(\max_{m \in M_{ij}} u_{ijm}(v_{ij}) \right), \quad (5)$$

where the expectation is over the set of i.i.d. draws ε_{ijm} across travel modes.

4.3 Market-Clearing Conditions and the Sorting Equilibrium

The equilibrium market-clearing conditions for the housing market and the transportation sector are interrelated in our model. 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. These two markets 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, shapes the spatial distribution of households. We discuss these market-clearing conditions below.

Housing Market The aggregation of households' choice probabilities P_{ij} gives rise to the housing demand:

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

Housing demand depends on both housing prices \mathbf{p} and the driving speed \mathbf{v} (through the ease-of-commute utility). We consider two scenarios for housing supply. In the first scenario, housing supply is fixed at one for all properties: $S_j(\mathbf{p}) = 1$ (the supply for each property unit is one). In the second scenario, housing supply has a constant elasticity, $\ln(S_j) = c_{j,0} + \epsilon * \ln(p_j)$, and increases $\epsilon\%$ with a 1% increase in the housing price.

Transportation Sector Demand for driving is determined by both housing locations and travel mode choices. Intuitively, mode choices determine the extensive margin (the decision on whether to drive), while housing locations determine the intensive margin (the commuting distance). Total driving demand and hence traffic density in region s is the aggregation of the location and commuting decisions for relevant households, which

ultimately depends on the housing price \mathbf{p} and driving speed \mathbf{v} :

$$D_s^v(\mathbf{p}, \mathbf{v}) \equiv \sum_{i \in s} \sum_j P_{ij}(\mathbf{p}, \mathbf{v}) \cdot \{ [R_{ij,car}(\mathbf{v}) \cdot \text{dist}_{ij,car}] + [R_{ij,taxi}(\mathbf{v}) \cdot \text{dist}_{ij,taxi}] \}, \quad (6)$$

where P_{ij} is the probability that household i chooses location j , $R_{ij,car}$ and $R_{ij,taxi}$ are the probabilities that household i living in location j drives and takes taxi, respectively, and $\text{dist}_{ij,car}$ and $\text{dist}_{ij,taxi}$ are the commuting distance by car and taxi.

The appropriate choice of region s is context specific and defines the geographical scope of congestion's negative externality. Our main analysis assumes that the scope of congestion is city-wide, but we also consider extensions where the traffic density is ring-road specific (i.e., it differs for the regions between the second and third ring roads, between the third and fourth ring roads, etc.).¹⁸

The supply side of the transportation sector describes the relationship between the traffic density S^v (the number of vehicles on the road) and the travel speed v that can be sustained given Beijing's existing road capacity. We assume the density and speed relationship has a constant elasticity:

$$\ln(S^v(\mathbf{v})) = c_0 + \epsilon^v * \ln(\mathbf{v}).$$

For a 1% increase in traffic speed, the traffic density that can be sustained under the existing road capacity goes down by $|\epsilon^v|%$. This supply relationship helps to determine the extent of the congestion externality: other drivers on the road reduce household i 's driving speed. Figure 3 depicts the congestion externality and illustrates how the equilibrium congestion level is determined under driving restrictions and congestion pricing.

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

1. The housing market clears for all properties:

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

2. The transportation sector clears for every region s , where households' aggregate driving demand at speed \mathbf{v}^* is equal to the traffic density that can be sustained under the existing road capacity at speed \mathbf{v}^* :

$$D_s^v(\mathbf{p}^*, \mathbf{v}^*) = S_s^v(\mathbf{v}^*), \forall s. \quad (8)$$

Our model follows the class of equilibrium sorting models with local spillovers studied in [Bayer and](#)

¹⁸The use of a single speed adjustment factor to reflect congestion is consistent with theoretical work modeling congestion spillovers in dense downtown intersections like water rising in a bathtub ([Arnott, 2013](#)).

Timmins (2005) and more closely in Bayer et al. (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 (such as the type I extreme value distribution), then the equation system (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 scalable constant) \mathbf{p}^* solves the system of equations defined by equations (2) and (7), conditional on a set of observed and unobserved housing attributes (\mathbf{X} and $\boldsymbol{\xi}$) and the 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. The fixed point of the equation system (2), (4), (7), and (8) defines the equilibrium housing prices and traffic speed $\{\mathbf{p}^*, \mathbf{v}^*\}$.¹⁹

4.4 Estimation Details

This subsection discusses how we estimate the parameters that characterize housing demand, the travel mode choice, and the traffic density–speed relationship. We estimate the housing demand separately from the travel mode choice using two separate data sets. Appendix Section D includes further details. Following the vast literature on discrete choice models, we assume that the error terms in both the housing demand equation (1) and the commuting mode choice equation (3) have type I extreme value distribution.

Estimation of Travel Mode Choices The parameters of the travel mode choices are estimated via simulated maximum likelihood estimation (MLE) using the household travel surveys. The key parameters of interest are the time and monetary cost preferences. We include mode-specific random coefficients to capture (dis)amenities that do not scale with the time or distance traveled. We also interact mode-specific fixed effects with year fixed effects, district fixed effects, and demographic variables (such as commuters' income and age). These interactions control for a rich set of time-varying and location-specific unobservables by travel mode.

We assume that the error term ε_{ijm} in equation (3) is uncorrelated with commuting trips' time and monetary costs. This assumption would be violated if, for example, the route-specific quality of public transit service (i.e., in terms of congestion, delay, comfort, or safety) is correlated with route-specific monetary costs or travel time. Monetary costs are likely to be exogenous because Beijing's transportation bureau sets bus and subway fares uniformly across all routes. Hence, fares do not vary by the level of congestion or quality of service. Travel time is determined by congestion. We include a rich set of mode, time, and spatial fixed effects to absorb shocks that are common across households and affect both travel speed and the error term ε_{ijm} . The remaining variation in ε_{ijm} reflects idiosyncratic considerations that are unlikely to be correlated

¹⁹The 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 et al. (2004). To address the possibility of multiple equilibria, we simulate our model with 100 different initial starting values. The simulation analyses always converge to the same equilibrium outcomes, providing empirical evidence for uniqueness in our applied setting. Our practice follows the recent literature (such as Couture et al. 2020; Hwang 2019) that incorporates rich preference heterogeneity and multiple margins of adjustment to maintain realism without imposing restrictions to guarantee a unique equilibrium outcome.

with travel time.

Once we have estimated parameters from travel mode choices, we plug them in the housing transaction data and construct the ease-of-commute utility EV_{ij} by equation (5), which takes the usual log-sum formula for errors with the type I extreme value distribution. A key underlying assumption of this ‘plug-in’ approach is that after we account for location and demographic differences, commuters’ preferences as estimated from travel surveys are representative of the commuting preferences of home buyers in the mortgage data. Note that the calculation of EV_{ij} is computationally intensive and 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 2.2.

Choice Set of the Housing Demand Computational and data limitations often require restrictions on the number of alternatives in demand estimation. While it may be logical to restrict households’ choice set to a set of affordable or nearby homes, [Banzhaf and Smith \(2007\)](#) shows that this approach may bias estimation due to unobserved heterogeneity in the choice set definition. Instead of restricting the choice set based on attributes, we rely on choice-based sampling methods, which have been proven to deliver consistent estimates in multinomial logit and mixed logit models by [Wasi and Keane \(2012\)](#) and [Guevara and Ben-Akiva \(2013\)](#). We take a 1% random sample of the houses sold during a two-month window around the purchase date of the chosen home. The average size of a choice set is 27. A robustness check using a 0.5% random sample yields very similar results (Section 5.2).

Estimation of Residential Location The estimated EV_{ij} in the previous stage 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 et al. \(2003\)](#) 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.

The parameters in the housing demand equation are estimated using a two-step procedure: the first step uses simulated MLE with a nested contraction mapping, and the second step uses the linear IV. The two-step strategy follows the approach of [Berry et al. \(1995\)](#) and [Bayer et al. \(2007\)](#) to address the challenge related to the presence of endogenous variables in nonlinear estimations (namely, that unobserved housing attributes ξ_j render the price variable endogenous and bias the price coefficient toward zero). In the first step, we search for preference parameters to maximize simulated MLE while inverting the population-average utilities δ_j (also called the alternative-specific utility; see equation (11) below) by using a nested contraction mapping algorithm. In the second step, we regress the population-average utilities δ_j on prices instrumented by IVs to recover unbiased estimates of the price coefficients.

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):

$$U_{ij} = \mu_{ij}(\theta_1) + \delta_j(\theta_2) + \varepsilon_{ij} \quad (9)$$

$$\mu_{ij}(\theta_1) = \alpha_2 \ln(y_i) p_j + \mathbf{x}_j \mathbf{z}_i \beta + \sum_k \phi_{ik} EV_{ijk} \quad (10)$$

$$\delta_j(\theta_2) = \alpha_1 p_j + \mathbf{x}_j \bar{\beta} + \xi_j. \quad (11)$$

where k denotes each commuting member of household i . Note that we use θ_1 to denote the parameters in equation (10) and θ_2 to denote the parameters in equation (11). Appendix D provides further details about the estimation procedure.

Once θ_1 are estimated and $\{\delta_j\}_j$ inverted from the observed data, we estimate equation (11) using three sets of IVs. The first is the number of properties a) within 3 km of property j , b) outside the same complex, and c) sold within a two-month window of property j 's purchase date.²⁰ This variable is arguably exogenous and correlated with housing price p_j because the availability of desirable properties in close proximity exerts downward pressure on p_j through competition. The second set of IVs is the average attributes of these properties (including their neighborhood attributes), which are also proxies for competition in the housing market. The third set of IVs includes the interaction between the second set of IVs and the odds of winning the license lottery discussed in Section 2.1. These odds decreased dramatically from 9.4% in January 2011 to 0.7% by the end of 2014. The interaction terms capture the likely impact of the license lottery on the nature of housing market competition and price setting. Decreasing winning odds push up demand (and hence prices) for properties in desirable locations, such as places close to the subway or city center.

Estimation of Speed–Density Elasticity We estimate the supply side of the transportation sector with the following equation:

$$\ln(v_{st}) = e^v * \ln(d_{st}) + \mathbf{X}_{st} \beta + \varepsilon_{st}, \quad (12)$$

where the unit of observation is road segment by hour, $\ln(v_{jt})$ is log speed in km/h, $\ln(d_{jt})$ is the log of traffic density, measured by the number of vehicles per lane-km, and e^v is the speed–density elasticity. The vector of X_{jt} includes weather-related variables (temperature, wind speed, etc.) and time and spatial fixed effects (hour-of-day, day-of-week, road segment, etc.). The key regressor $\ln(d_{st})$ could be correlated with the residual due to accidents, road construction, or major events. To address the potential endogeneity of traffic density, we construct IVs based on Beijing's driving restriction policy following Yang et al. (2020). The policy has a preset, rotating schedule that restricts each vehicle from driving one weekday per week based on the last digit of the license plate number. We construct a policy indicator taking 1 for the 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 superstition.

²⁰ Alternative cutoffs, such as 1 or 5 km, deliver similar results.

5 Estimation Results

We discuss the estimation results on travel mode choices, housing demand, and the speed–density relationship, in turn, below.

5.1 Commuting Mode Choice

Table 3 presents parameter estimates for six specifications. The first three specifications do not have random coefficients. The only source of household heterogeneity comes from household income (and the interaction between travel costs and income). The last three specifications include random coefficients on travel time to capture unobserved consumer heterogeneity and random coefficients on travel modes. The VOT is defined by the ratio of the (random) coefficient on travel time and the parameter on travel costs. It is measured as a percentage of a household’s hourly wage.

Column (1) controls for interactions between the year dummies (2010 or 2014) and mode fixed effects (car, taxi, bus, subway, walking, and biking). The implied VOT is 75.7% of the hourly wage. Column (2) adds the interactions between mode fixed effects and trip characteristics, including trip distance bins and trip origin and destination attributes (e.g., whether the trip origin is within the second ring road). These controls account for important features of travel demand and significantly improve the model fit. For example, the transportation literature has documented that drivers value the reliability of travel time (Brownstone and Small, 2005; Small et al., 2005). Uncertainty in travel time likely scales with the distance of a trip 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. Column (3) further includes interactions of mode fixed effects with household demographic variables including age, gender, education, vehicle ownership, number of workers, and household size. These variables help explain different mode choices across demographic groups (e.g., wealthier households’ greater likelihood of driving and using taxis) and further improves 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).²¹ In 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 demographics on mode choices. For example, some commuters choose driving or taxiing not because of a high VOT but rather because of scheduling constraints. Others choose walking or biking for the exercise benefits. The dispersion of these preference parameters is economically large and statistically significant, suggesting significant preference heterogeneity.

The results of our preferred specification are in Column (6). Adding travel time and mode-specific random coefficients leads to a strong sensitivity to travel costs and delivers a much more reasonable estimate of the

²¹Following Petrin (2002), we winsorize the top and bottom 5% of the distribution to minimize the impact of extreme random draws, as the VOT is unlikely to be infinite. The distribution with three degrees of freedom provides the best fit.

VOT. Appendix Figure A6 depicts the VOT estimate histogram. The average and median VOT is 95.6% and 84.6% of the hourly wage, respectively, which is within the range typically found in the recent literature.²²

We construct the commuting utility (EV_{ij}) as defined in equation (5) for both male and female borrowers based on their work locations.²³ These variables are included as part of the (buyer-specific) housing attributes in Section 5.2.

5.2 Housing Location Choice

We now turn to the estimation results of housing demand. We first present the MLE estimates of household-specific preference parameters and then discuss the IV estimates for coefficients in the mean utility.

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.²⁴ 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 the most likely to have school-age children.

We do not observe the household size. To capture preference heterogeneity in home sizes due to variation in the household size, we use the age of the primary borrower as a proxy 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, probably due to the presence of both children and elderly 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 sizeable 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 (or for ¥1 savings in commuting costs), we search for changes in housing prices that would keep households' utility constant. Since we use housing transaction prices instead of rental prices in housing demand (equation (1)), the WTP

²²In the context of travel demand, the VOT estimates typically range between 30% and 100% of hourly income (Small et al., 2007). Using a discrete choice framework similar to ours, Small et al. (2005) estimate the median VOT at 93% of the hourly wage for commuters in Los Angeles. Buchholz et al. (2020) use the trade-off between wait time and price among users on a large ride-hailing platform in Prague and find the average VOT to be roughly 100% of users' wage during work hours. Goldszmidt et al. (2020) find an average (median) VOT of 75% (100%) of the hourly (after-tax) wage based on a large-scale field experiment by Lyft in 13 US cities. The US Department of Transportation recommends using 50% 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 travel time savings for transportation projects (USDOT, 2015). Leveraging the trade-off between vehicle driving speed and gasoline usage, Wolff (2014) estimates the average VOT in eight rural locations in Washington state to be 50% of the hourly wage based on traffic speed data.

²³Around 61% of the primary borrowers are male, with the remaining 39% female. We set $EV_{ijk} = 0$ for unemployed family members and ignore their commuting needs in the house purchase decision.

²⁴The price coefficient is $\alpha_1 + \alpha_2 * \ln(y_i)$. Since α_1 is negative, a positive α_2 means the absolute level of price sensitivity is lower for higher-income households.

estimates reflect the monetized lifetime utility over the property tenure. According to our preferred specification (Columns (5) and (6)), an average household is willing to pay ¥18,500 on a home to shorten the male member's daily work commute by one minute and ¥18,000 to save ¥1 in the male member's commute cost.²⁵ Households are willing to pay 18% more for a similar reduction in their female member's commuting time and monetary cost, suggesting that households prioritize the female member's commute convenience in housing choices. This is consistent with descriptive evidence that women tend to live much closer to their work locations (Appendix Figure A7) and existing literature (Le Barbanchon et al., 2020). There exists significant preference heterogeneity among households: the interquartile willingness to pay for a one-minute shorter commute is between ¥4,510 and 23,400 and ¥5,300 and 33,800 for the male and female member, respectively.

Table 5 reports the coefficient estimates on the population-average utility (equation (11)), conditioning on the specification in Table 4 that allows for random coefficients on the commuting utility (the EV terms). Columns (1) and (2) use OLS, while Columns (3)-(6) are from IV regressions. All regressions include the interaction of the month-of-sample and district fixed effects to capture time-varying changes in market conditions and amenities that could vary across districts in Beijing. Columns (2)-(6) also include 158 neighborhood (*jiedao*) fixed effects to capture unobserved time-invariant neighborhood amenities. We use the three sets of IVs for housing prices that are discussed in Section 4.4: the number of homes within 3 km outside the sample complex sold in a two-month time window of a given home sale; the average attributes of these properties; and the interaction between the second set of IVs and the odds of winning the license lottery. The results of our preferred specification with all instruments appear in Column (6), with a first-stage F statistic of 14.22.

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 average elasticities from the IV estimates vary from -1.34 to -1.94 in Columns (4)-(6). The signs on the other coefficient estimates from the IV regressions in Columns (3)-(6) are all consistent with intuition. Households prefer larger properties and those closer to signature schools but dislike older buildings and places far from parks.²⁶

Incorporating the commuting utility not only improves the model fit but also has implications for other parameter estimates, especially the price coefficient and price elasticity. Appendix Table A2 reports the results for the population-average utility from the regression without the *EV* terms. Both the price coefficient and especially price elasticities are smaller in magnitude, consistent with the downward bias arising from omission of important attributes (*EV* terms). Timmins and Murdock (2007) find a 50% downward bias in the estimation

²⁵Our empirical analysis delivers two estimates of the marginal util per ¥: one based on the marginal utility of the housing price and one based the marginal utility of the commuting cost. Equating these two estimates implies 600 trips per year for the male borrower and 700 trips for the female borrower over a 30-year housing tenure. These implied trip numbers are high though not implausible.

²⁶While the coefficient of distance to key schools is positive for the base group (borrowers under 30), the coefficient for borrowers between 30 and 45 is negative and significant, as the latter are more likely to have school-aged children.

of consumer welfare from recreation sites when on-site congestion is ignored in demand estimation.

To examine the robustness of our results to the choice sampling method, we repeat the analysis with a 0.5% instead of 1% random sample to construct households' choice sets. The results are shown in Appendix Tables A3 and A4. The average price elasticity is -1.64 with the 0.5% random sample and -1.44 with the 1% random sample. The parameter estimates and implied willingness to pay for housing attributes are quite similar across these two samples.

Based on the parameter estimates from our preferred specification (the last set of results in Tables 4 and 5), the average income elasticity of housing size and the income elasticity of marginal driving costs are 0.10 and 0.78, respectively.²⁷ To our knowledge, these are the first estimates of these elasticities for Chinese households. Our estimated elasticity for housing size is somewhat smaller than estimates based on U.S. data, while the elasticity of marginal driving costs is largely consistent with other estimates in the literature. Using the 2003 American Housing Survey, Glaeser et al. (2008) find the elasticity of lot 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.

5.3 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 that cover all major roads throughout Beijing for 2014. We focus on observations with traffic density higher than 35 cars per lane-km (Column 5). The average speed of these observations is 30 km/h, close to the city-wide average speed during peak hours. The peak-hour speed is more relevant since we focus on commuting trips.

To examine potential differences in the speed–density elasticity across regions, we split our sample into four groups based on the location of traffic sensors: between the second and third ring roads, the third and fourth ring roads, the fourth and fifth ring roads, and the fifth and sixth ring roads. Appendix Table A5 reports the IV-estimated speed–density elasticity. The extent of heterogeneity among groups is limited, with the OLS and IV estimates comparable across columns.²⁸ In the counterfactual analysis below, we use the city-wide speed–density elasticity estimate of -1.1.

6 Counterfactual Simulation Algorithm

To evaluate Beijing's transportation policies, we examine five scenarios: driving restrictions, congestion pricing, subway expansion, and combinations of these policies. The first scenario follows the actual driving restriction scheme implemented in Beijing: a vehicle is prohibited from driving on one of the five work-

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

²⁸We 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.

days. Under the congestion pricing scheme, which is hypothetical, we choose a distance-based charge (at ¥1.13/km) to achieve the same level of congestion reduction as that resulting from the driving restriction policy 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, with 8 additional lines in operation. We conduct the entire counterfactual analysis using the 2014 cohort to allow for maximum coverage of the subway expansion. Appendix E explains in detail the simulation algorithm. We provide a brief outline below. Readers not interested in this material can proceed to Section 7.

6.1 Simulating the Counterfactual Equilibrium

The counterfactual equilibrium is defined as new vectors of housing prices and travel speed $\{\mathbf{p}^*, \mathbf{v}^*\}$ that satisfy the market-clearing conditions (equations (7) and (8)). We iterate equations (7) and (8) sequentially to find the unique 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:

$$\frac{\tilde{v}_{ij} - v_{ij}^o}{v_{ij}^o} = e^\nu * \frac{\tilde{d}_s - d_s^o}{d_s^o}, \quad (13)$$

where \tilde{v}_{ij} is the counterfactual driving speed for household i 's work commute from home j , v_{ij}^o is the observed driving speed (see Section 2.2 and Online Appendix A for its construction), and \tilde{d}_s and d_s^o are the counterfactual and observed traffic density for region s , respectively.

The size of region s defines the geographical scope of the congestion externality. We consider two scenarios. In the first scenario, the scope of this externality is assumed to be city-wide. Hence, traffic density is measured by the aggregate driving demand of all Beijing households via equation (6): $d_s = d = \sum_i \sum_j P_{ij} \cdot \{[R_{ij,car} \cdot \text{dist}_{ij,car}] + [R_{ij,taxi} \cdot \text{dist}_{ij,taxi}]\}$. We plug in the estimated speed–density elasticity $e^\nu = -1.1$ to update the counterfactual driving speed. Note that the driving speed \tilde{v}_{ij} differs across households though the traffic density is city-wide. In the second scenario, we assume that the scope of the congestion externality is ring road specific. To accommodate this assumption, we construct the ring road-specific traffic density by allocating each household's drive commute to the appropriate ring road traffic density based on the trip origin and destination.

6.2 Welfare Decomposition

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

$$W_i = \mathbb{E}_{\epsilon_{ij}} \left(\max_{j \in J_i} U_{ij}(\mathbf{p}, \mathbf{v}, \mathbf{t}) \right),$$

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

$$\frac{dW}{dt} = \underbrace{\frac{\partial W}{\partial t} \Big|_{\mathbf{p}=\mathbf{p}_0, \mathbf{v}=\mathbf{v}_0}}_{(1) \text{ direct policy effect}} + \underbrace{\frac{\partial W}{\partial \mathbf{v}'} \frac{\partial \mathbf{v}}{\partial t} \Big|_{\tilde{\mathbf{v}}}}_{(2) \text{ partial speed effect}} + \underbrace{\frac{\partial W}{\partial \mathbf{v}'} \frac{\partial \mathbf{v}}{\partial t} \Big|_{D^v(\mathbf{v}^*)=S^v(\mathbf{v}^*)} - \frac{\partial W}{\partial \mathbf{v}'} \frac{\partial \mathbf{v}}{\partial t} \Big|_{\tilde{\mathbf{v}}}}_{(3) \text{ rebound effect}} \\ + \underbrace{\frac{\partial W}{\partial \mathbf{p}'} \frac{\partial \mathbf{p}}{\partial t} \Big|_{D(\mathbf{p}^*, \mathbf{v}^*)=1}}_{\text{equil. (4) sorting effect}} + \underbrace{\frac{\partial W}{\partial \mathbf{p}'} \frac{\partial \mathbf{p}}{\partial t} \Big|_{D(\mathbf{p}^*, \mathbf{v}^*)=S} - \frac{\partial W}{\partial \mathbf{p}'} \frac{\partial \mathbf{p}}{\partial t} \Big|_{D(\mathbf{p}^*, \mathbf{v}^*)=1}}_{(5) \text{ housing supply effect}}. \quad (14)$$

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 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% of drivers off the road, which leads to an initial 22% 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. In these studies, the partial equilibrium welfare benefit is often the product of implied driving time savings and an estimated value of time ([Anderson, 2014](#); [Hanna et al., 2017](#); [Adler and van Ommeren, 2016](#); [Bauernschuster et al., 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 ([Yang et al., 2020](#); [Bento et al., 2020](#)). The fourth channel, the equilibrium sorting effect, incorporates residential sorting and evaluates changes in welfare when households relocate in response to changes in the commuting utility, with housing supply held fixed. The last channel allows housing supply to adjust in response to housing price changes. In the analysis below, we sometimes refer to the first channel as the direct effect, the third channel as the rebound effect, and the second and third channels together as the equilibrium speed effect.

Before we present the simulation results, we first validate the structural model by comparing its predictions with the reduced-form evidence presented in Section 3.2. To do so, we simulate the market equilibrium under the 2008 subway network with and without the driving restriction and examine changes in the model-predicted housing price gradient with respect to subway access. The results are reported in Appendix Table A6. The model-predicted price gradient change as a result of the driving restrictions is -0.034, consistent with

the reduced-form evidence that the driving restriction steepens the price gradient of subway access.²⁹ This suggests that our structural analysis replicates well the observed pattern of equilibrium price changes under the driving restriction policy.

7 Counterfactual Results

We now evaluate different transportation policies and compare the equilibrium outcomes when households reoptimize both their commuting modes and their residential locations and when both the housing and transportation sectors clear. Sections 7.1 to 7.3 analyze the congestion reduction, sorting patterns, and social welfare effects in the baseline case in which housing supply is held fixed and a city-wide traffic density is used. Section 7.4 considers various extensions, including considering variable housing supply and ring road-specific traffic density, removing random coefficients, and accounting for migration and consumption access.

Table 6 reports the results from our baseline analysis. It considers six different scenarios. The first three columns report the equilibrium outcomes under the 2008 subway network, while the next three illustrate the results under the 2014 subway network. Column (1) presents the scenario with no policies. Columns (2)-(6) describe the differences relative to Column (1). Columns (2), (3), and (4) evaluate the driving restriction, congestion pricing, and subway expansion, respectively. Columns (5) and (6) examine combinations of these policies. All results are shown separately for households with income above or below the median (high vs. low income) to reflect distributional considerations.

7.1 Mode Choice and Congestion Reduction

Driving Restriction Panel A of Table 6 examines changes in the travel mode and congestion.³⁰ The driving restriction policy entails two countervailing forces. On the one hand, it moves households off the road on the 20% of workdays when driving with personal vehicles is restricted, forcing them to switch to slower modes (subway, bus, biking, walking). The increase in commuting time incentivizes households to relocate (move closer to work). Both margins of adjustments reduce congestion and increase the driving speed. On the other hand, the improved travel speed from less congestion induces households to drive more on days when vehicle usage is not restricted, especially among those with a long commute. This rebound effect dampens the congestion reduction from the direct policy effect. On average, the driving restriction increases traffic speed by 18% (from 21.5 km/h to 25.3 km/h).

²⁹The coefficient of -0.034 is smaller in magnitude than the reduced-form analysis because the reduced-form result reflects a short-run response while the structural simulation incorporates long-run equilibrium adjustments (especially the rebound effects). The two samples are also different. The reduced-form analysis uses two years before and after CDR's initial implementation date while the structural analysis uses the 2014 cohort.

³⁰The mode choice shares are slightly different between Table 6 and Figure 1 because the former reports mode choices among home buyers while the latter reports mode choices for all residents in Beijing, including non-home owners, who accounted for 27.8% in 2010.

Congestion Pricing Congestion pricing is levied on a per-kilometer basis and increases the driving cost at both the intensive margin (how far drivers travel) and extensive margin (whether drivers choose to drive). There are three key differences between the congestion pricing and driving restriction schemes. 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. Indeed, low-income households' driving probability is reduced by as much as 25%, as they are more sensitive to congestion charges. Third, even though both policies lead to the same congestion reduction, a larger share of commuters drive under the congestion charge. This is because congestion pricing induces a stronger sorting response (more on this below), with households from both income groups and especially high-income households moving closer to work. In contrast, the commuting distance under the driving restriction barely changes. Thus, there are fewer long commutes but more people on the road under the congestion pricing than under the driving restriction scheme.

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) demonstrates that traffic speed increases by 7%, only 40% of the speed increase under the driving restriction and congestion pricing policies. The reason for this muted response is twofold. First, the reduction in the driving share is smaller under subway expansion than under the driving restriction or congestion pricing schemes. 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 this below and present the spatial sorting responses in Figure 5.

These results, and especially the one on sorting, point to important channels beyond what has been covered in empirical studies that focus on the short-run impact of the subway system on traffic congestion.³¹ Our findings are consistent with the findings in prior literature that a) with sufficient time, induced travel demand increases one-for-one with capacity expansion (Downs, 1962; Duranton and Turner, 2011) and b) subway expansion itself lowers the cost associated with the commuting distance and increases urban sprawl (Gonzalez-Navarro and Turner, 2018; Heblitch et al., 2020).

Nonetheless, the subway expansion dramatically increased subway access: the distance between home and the nearest subway station declined by about 80% for both income groups. Subway ridership increased significantly by 51% and 56% among high- and low-income groups, respectively.

Policy Combinations We now consider an array of policy mixes. We first evaluate Beijing's actual transportation policy in Column (5), which combines the subway expansion with driving restrictions. Then, in

³¹Using a regression discontinuity (in time) approach, Anderson (2014) finds that a 35-day transit strike that shut down subways in Los Angeles resulted in a 47% increase in highway traffic delays during peak hours. Yang et al. (2018) shows, using a 120-day window surrounding the subway opening, that the subway expansion in Beijing from 2009 to 2015 reduced traffic congestion by 15% on average. Using a difference-in-differences framework, Gu et al. (2020) estimate that one new subway line increases traffic speed by 4% during peak hours on nearby roads based on their examination of 45 subway lines opened across 42 Chinese cities during 2016 and 2017.

Column (6), we compare the existing policy to the alternative of subway expansion with congestion pricing. Subway expansion lowers the cost and increases the accessibility of public transit. The driving restriction is a command-and-control policy that leaves households with little discretion. Congestion pricing is distance based and approximates the Pigouvian tax that would internalize the congestion externality. The empirical question at hand is the extent to which congestion pricing (a market-based demand policy) exhibits stronger complementarity in reducing congestion and increasing welfare than the driving restriction (a command-and-control policy) when combined with subway expansion (a supply-side policy).

The improvement to 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 scheme, 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, the driving restriction lead to a 8.5-percentage-point reduction in high-income households' driving probability with the 2014 subway network in comparison to a 7.17-percentage-point reduction with 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.

Our results suggest that both forces are at play. In addition, congestion pricing exhibits stronger complementarity with subway expansion than the driving restriction and is more effective in moving people off the road: the speed improvement is 5.29 km/h in Column (6), higher than 5.08 km/h improvement in Column (5). Congestion pricing affects both the extensive margin (the decision of whether to drive) and the intensive margin (the decision of how far to drive); both effects could be reinforced by subway expansion. In contrast, the driving restriction primarily operates along the extensive margin.

7.2 Sorting and the Housing Price

Sorting and Household Spatial Distribution Panel B of Table 6 examines the differential impact of the three transportation policies on households' spatial distribution. Transportation policies directly affect commuting costs. These direct changes set in motion a set of behavioral responses whereby households substitute across different travel modes and adjust their residential locations. Take **driving restrictions** as an example. Their direct effect pushes households to live closer to their workplaces and thereby offset the increase in commuting time when household members are forced to use a slower travel mode. On the other hand, since people are forced to drive less, congestion is reduced, and driving speed improves. This shortens the commuting time, especially for long-distance trips. For example, the driving time declines by twelve minutes for trips for which both the origin and the destination fall outside the fourth ring road (i.e., longer commuting trips) but only six minutes for other trips. In other words, the speed effect disproportionately benefits long-distance trips. Indeed, the correlation between the driving probability and driving distance increases from

0.28 to 0.33 under the driving restriction. 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 contrast, **congestion pricing** is distance based and causes a much higher increase in commuting costs for longer trips. Hence, both high- and low-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% of high-income households drive to work, in comparison to 21.44% of low-income households. As a result, high-income households are much more affected by congestion pricing (which has an effect close to zero on people who do not drive). Second, high-income households have a higher WTP for commuting convenience, as their value of time is higher. Lastly, properties closer to common 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 associated with congestion pricing. The direct policy effect moves people off the road as they substitute toward subways. The improved driving speed and improved subway system make long-distance commuting by either driving or subway less costly. 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. Ultimately, the sorting response across households is dictated by changes in the commuting utility ΔEV , households' idiosyncratic preferences for commuting convenience (the random coefficient of EV), and households' price sensitivity. Subway expansion creates the strongest sorting responses among the three policies because its effect is local and uneven across households: it primarily affects households experiencing changes in subway access. In contrast, both congestion pricing and driving restrictions affect the commuting costs of all households with drivers. Congestion pricing has a greater impact on households with longer driving commutes, while the effect of driving restrictions is less variable across households. One piece of supporting evidence is that the standard deviation of changes in commuting utility is several times higher under subway expansion than under congestion pricing, while that under driving restrictions is the smallest.

Figure 5 plots changes in the average commuting distances for residents in each TAZ relative to the distances in the no-policy scenario. The driving restriction leads 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 job 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. In contrast, subway expansion increases the commuting distance in most TAZs, especially along the new subway lines, exacerbating “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 et al. \(2020\)](#).

In terms of the distance to subway stations, 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 outcomes under the baseline scenario. This reflects transit-based gentrification, where lower-income households are priced out of premium 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 4.3km to 0.86km for low-income households.

Housing Price Changes in housing prices closely mirror the sorting patterns. Figure 6 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.³² 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, Appendix Figure A8 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/m² per km) than under the 2008 network (-¥700/m² 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.

7.3 Welfare Analysis

Panel C of Table 6 presents the welfare results. In Figure 7, we decompose welfare changes by following equation (14) to illustrate different adjustment margins. Variable housing supply is considered in Section 7.4. Note that consumer surplus reported in this section refers to the discounted lifetime consumer surplus over the property tenure (commonly assumed to be thirty years), as discussed in Section 5.2. This is because our WTP estimates are recovered from housing transaction prices instead of the rental cost of capital.

Key Findings First, despite their effectiveness in congestion reduction, **driving restrictions** generate a welfare loss of ¥129,900 per household over a 30-year period discounted to the present. The annualized loss is roughly 4% of the household income.³³ Driving restrictions force drivers to switch to slower commuting modes and significantly increases the commuting time, especially for households with long commutes. On average, a household spends 16.8 more minutes commuting each day as a result of the driving restrictions.

³²Under congestion pricing, housing prices in northwestern Beijing (near the tech center) would increase by about 2,000 ¥/m², while those in some southeastern areas would decrease by 2,000 ¥/m² from a baseline average price at 24,022 ¥/m². Under subway expansion, home prices increase by as much as 4,000 ¥/m² in the southwest, where the subway expansion is greatest and historical prices have been lowest.

³³According to the Beijing Statistics Bureau, the average household income in Beijing in 2014 was ¥128,000. The average income for high- and low-income households was ¥172,000 and ¥84,000, respectively.

High-income households experience a much steeper reduction than low-income households. As shown in Figure 7, the direct policy effect of a driving restriction is large and negative at ¥226,800 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 the direct policy effect, as shown by the second and third bars in Figure 7. 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 difference between the second and third bar (a loss of ¥97,800 vs. ¥129,200) illustrates the importance of incorporating the rebound effect and allowing the full equilibrium adjustment of the transportation sector. Otherwise, the welfare effect could be overestimated by 32% for driving restrictions and by 30%-49% for other policies. The fourth bar further incorporates the sorting effect. With all four channels incorporated, the welfare loss is at ¥129,900 per household.

Second, before revenue recycling, low-income households experience a greater loss under **congestion pricing** than under driving restrictions. This reflects the fact that low-income households are more responsive to increases in monetary costs from congestion pricing than they are to longer commuting times under driving restrictions. However, when the toll revenue is uniformly recycled across income groups, congestion pricing leads to *welfare gains* for both groups: consumer surplus increases by ¥39,200 and ¥64,300 for high- and low-income households, respectively. Low-income households witness a larger consumer surplus increase because they pay a smaller amount in toll charges but receive 50% of the toll revenue. This highlights the role of properly distributing toll revenues to abate distributional concerns related to congestion pricing.

In terms of the underlying channels, the direct effect of congestion pricing (with revenue recycling) reduces welfare by ¥30,000 per household. The equilibrium speed effect (the partial speed and rebound effect) reverses the welfare loss to yield a welfare gain of ¥41,800 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 ¥51,800 per household. Residential sorting enhances the welfare gain from congestion pricing by 24%, consistent with the result based on 98 US cities in [Langer and Winston \(2008\)](#), who also finds a large benefit of congestion pricing.

Third, while the **subway expansion** from 2008 to 2014 resulted in limited congestion reduction relative to that under the other two policies, it leads to a larger increase in consumer surplus. Much of this increase comes from the greater access to the subway network: the distance to the nearest subway station declined by 80% on average, and subway ridership increased by more than half. Although the substitution from nondriving trips to subway trips does not alleviate traffic congestion, it improves consumer welfare by offering better commuting choices. After taking into consideration the costs of expanding and maintaining the subway, net welfare is almost halved for high-income at ¥117,300 and close to zero for low-income households. The direct effect of subway expansion generates a welfare gain of ¥6,400 per household as a result of improved subway accessibility. The improved driving speed in equilibrium increases consumer welfare by ¥53,000 per household, with the overall welfare gain reaching ¥59,400 per household. Sorting induces households to

move away from their workplaces and the city center, which increases congestion and dampens the welfare gain to ¥57,100 per household.

Fourth, the combination of **congestion pricing and subway expansion** achieves the largest congestion reduction (with a 25% speed improvement) and generates the highest welfare gain at ¥93,400 per household over a 30-year period (about 3% of household income) across all policy scenarios. The revenue from congestion pricing at ¥127,700 per household could fully cover the costs of subway expansion at ¥103,000 per household. We believe this finding has broader applicability for the design of transportation infrastructure outside the context of Beijing.³⁴ 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 greater given the severity of the congestion externality in many cities. Indeed, a broader implication of our analysis is that welfare gains from infrastructure improvements can be mitigated by induced congestion. Pairing these investments with pricing instruments such as congestion pricing is critical to successfully address pre-existing and induced congestion and increase social welfare.

Importance of Sorting and Endogenous Congestion To highlight the role of sorting and endogenous congestion, Panel A of Table 7 reports changes in speed and welfare when we shut down these channels. 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 (i.e., sorting is shut down; Appendix Table A8 reports the full set of results). Sorting amplifies the effectiveness of 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% for high-income households and 16% for low-income households, consistent with Figure 7.

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 competition from high-income households raises prices and hurts low-income households. In contrast, under congestion pricing, both groups are better off with sorting. This is because sorting (moving closer to work under congestion pricing) further reduces congestion and increases welfare. In addition, households' workplaces are not perfectly aligned, and there is room for a Pareto improvement with a better home–work match for everyone. The opposing effect of sorting on congestion pricing and subway expansion and the distributional consequences of sorting highlight the importance of accounting for sorting in the analysis. Otherwise, we risk not only overestimating or underestimating the welfare gains but also getting the signs wrong and making inappropriate policy recommendations.

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

³⁴As 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 make enhancements to affected corridors, including public transit ([GAO, 2012](#)).

(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% and the welfare benefit inflated even more.

7.4 Extensions and Robustness Checks

Our baseline results in Sections 7.1 to 7.3 assume a fixed housing supply and city-wide traffic density. We relax these assumptions in Panel B of Table 7.

Housing Supply Adjustment The first row of Panel B in Table 7 summarizes the speed and welfare changes when the housing supply adjusts with the supply elasticity of 0.53 as in Wang et al. (2012).³⁵ Appendix Table A7 reports the full set of results. To understand how housing supply adjustments affect our previous findings, consider a home whose price appreciates after a transportation policy change. This price appreciation increases housing supply, which mitigates the overall price effect. In addition, the availability of additional housing in desirable locations enhances sorting by allowing more households to move in. In other words, allowing flexibility in the housing supply magnifies the role of sorting. For example, under congestion pricing, the housing price appreciates the most around employment centers. The large housing supply response in these areas allows more people to live closer to work, further alleviating congestion. Indeed, the *reduction* in the commuting distance is amplified from 0.19 km to 0.32 km for high-income households and from 0.07 km to 0.22 km for low-income households with housing supply adjustment. The driving speed improvement increases from 3.83 km/h to 3.97 km/h.

In contrast, subway expansion leads to housing price appreciation and new housing supply in city suburbs, which causes people to live farther away from work. The increase in the commuting distance deteriorates from 0.33 km to 0.76 km for high-income households and from 0.15 km to 0.61 km for low-income households. This attenuates the speed improvement under subway expansion from 1.49 km/h to 1.13 km/h.

Ring-Specific Traffic Density The baseline analysis assumes that the geographic scope of congestion is city-wide and adjusts local speeds with a city-wide traffic density. The second row of Panel B in Table 7 incorporates ring road band-specific traffic densities and adjusts local speeds with the corresponding density in the appropriate ring road bands (Appendix Table A9 reports the full set of results). Both the average speed improvement and welfare effects are very similar to those in the baseline. Perhaps not surprisingly, the simulations with ring-specific densities deliver higher speed improvements for the fourth/fifth and fifth/sixth ring road bands but lower speed improvements toward the city center under the congestion pricing and subway expansions, both of which are more effective at moving long commutes off the road.

³⁵The analysis is based on data for 35 Chinese cities from 1998 to 2009. Baum-Snow and Han (2021) estimate a supply elasticity of 0.3-0.5 for US cities based on data from 2000 to 2010, smaller than the estimates from Saiz (2010) based on data from 1970-2000.

Exclusion of Random Coefficients The sorting model with a rich set of demographic variables and random coefficients predicts intuitive substitution patterns (Table 6). For example, under both the driving restriction and congestion pricing schemes, taxi trips account for a disproportionately higher share of the reduction in driving trips relative to the shares in other models, a pattern that holds for both the high- and low-income groups. Under the subway expansion, taxi and bus trips are affected disproportionately more. The reduction in taxi and driving trips is much larger among low-income households than among high-income households due to the former group’s larger price sensitivity.

To evaluate the importance of incorporating heterogeneous preferences, we re-estimate the entire model excluding the random coefficients. We keep observed heterogeneity (e.g., income, age, gender, education, car ownership) since the model without these demographic variables has a poor fit. We repeat the counterfactual analysis in Appendix Table A10 and summarize the results in the third row of Panel B in Table 7.

While the model without random coefficients can fit the observed travel mode shares and replicate the average housing demand elasticity as in our baseline, its predictions on substitution patterns are often counterintuitive. For example, subway expansion increases ridership only by a modest 14% among high-income households instead of 51% as predicted by the baseline model. This is because the subway’s market share was less than 10% in 2008 and multinomial logit-type models tend to predict “proportionate” changes (and hence a modest increase) in market share. Consequently, the model without random coefficients predicts a negligible speed improvement of 0.16 km/h, which is only 11% of the baseline prediction, and underestimates the welfare gain from subway expansion by as much as 44%. In a similar vein, the model without random coefficients overestimates the value that households attach to driving, especially among high-income households, who drive more than 40% of the time. Not surprisingly, the welfare reductions associated with driving restrictions and congestion pricing are often prohibitive and several times larger than the baseline predictions.

Growing Population The baseline analysis assumes a fixed population. To account for migration, we assume in-migration of 5% under subway expansion and out-migration of 5% under the driving restriction and congestion pricing schemes in the fourth row of Panel B in Table 7. These choices are somewhat arbitrary but serve as upper-bound estimates of policy-induced migration since Beijing’s population grew by 14% during the sample period. The speed improvement and the associated welfare under the driving restriction and congestion pricing are strengthened with out-migration, while the opposite is true for subway expansion with in-migration. Importantly, the qualitative findings remain the same as the baseline results in Table 6.

Consumption Access Our sorting model abstracts from the endogenous availability of consumption services (e.g., restaurants, shops, and theaters) that might respond to improved traffic speeds. Estimates from recent literature (Miyauchi et al., 2021; Rao, 2021) suggest that consumer surplus from consumption access is about one-third that from job access in Tokyo and Beijing. We multiply the baseline consumer surplus by 1.33 and report the welfare changes (which also include toll revenues and subway costs) in the fifth row of Panel B in Table 7. This does not affect the baseline qualitative findings.

Optimal Congestion Pricing Finally, we plot changes in welfare as the toll rate varies in Figure 8. The optimal congestion charge is ¥1.2 per km when we shut down household sorting, ¥1.4 per km with sorting, and approximately the same when we incorporate both sorting and housing supply. At most congestion pricing levels, sorting increases consumer welfare by 20%-30%, and supply adjustment contributes another 10%-20% welfare gain. In addition, changes in consumer surplus are positive for a wide range of congestion charges (<¥2.5/km). This indicates that congestion pricing is likely to be an effective tool even when governments cannot gauge the exact optimal pricing level *a priori*.

8 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 empirically evaluate the efficiency and equity impacts of various urban transportation policies, incorporating 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, the different policies generate drastically different efficiency and equity consequences. 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 recycling of the revenue from the congestion charge. The combination of congestion pricing and subway expansion stands out as the best policy among all scenarios: it delivers the largest congestion reduction and the highest welfare gains. In addition, it is self-financing in that 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**, “Does public transit reduce car travel externalities? Quasi-natural experiments’ evidence from transit strikes,” *Journal of Urban Economics*, 2016, 92, 106–119.
- Akbar, Prottoy**, “Who Benefits from Faster Public Transit?,” 2020. Working Paper.
- Akbar, Prottoy A, Victor Couture, Gilles Duranton, and Adam Storeygard**, “Mobility and Congestion in Urban India,” Technical Report, National Bureau of Economic Research 2018.
- Albouy, David and Aarsh Farahani**, “Valuing Public Goods More Generally: The Case of Infrastructure,” 2017. Upjohn Institute working paper.
- Allen, Treb and Costas Arkolakis**, “The welfare effects of transportation infrastructure improvements,” Technical Report, National Bureau of Economic Research 2019.
- Almagro, Milena and Tomás Domínguez-Iino**, “Location Sorting and Endogenous Amenities: Evidence from Amsterdam,” 2020. Working Paper.
- Anas, Alex and Ikki Kim**, “General Equilibrium Models of Polycentric Urban Land Use with Endogenous Congestion and Job Agglomeration,” *Journal of Urban Economics*, September 1996, 40 (2), 232–256.
- and Robin Lindsey, “Reducing urban road transportation externalities: Road pricing in theory and in practice,” *Review of Environmental Economics and Policy*, 2011.
- Anderson, Michael L**, “Subways, strikes, and slowdowns: The impacts of public transit on traffic congestion,” *The American Economic Review*, 2014, 104 (9), 2763–2796.
- Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare**, “New trade models, same old gains?,” *American Economic Review*, 2012, 102 (1), 94–130.
- Arnott, Richard**, “A bathtub model of downtown traffic congestion,” *Journal of Urban Economics*, 2013, 76, 110–121.
- Arnott, Richard J and Joseph E Stiglitz**, “Aggregate land rents, expenditure on public goods, and optimal city size,” *The Quarterly Journal of Economics*, 1979, 93 (4), 471–500.
- Bajari, Patrick and Matthew E Kahn**, “Estimating Housing Demand With an Application to Explaining Racial Segregation in Cities,” *Journal of Business & Economic Statistics*, 2005, 23 (1), 20–33.
- Banzhaf, H. Spencer and V. Kerry Smith**, “Meta-analysis in model implementation: choice sets and the valuation of air quality improvements,” *Journal of Applied Econometrics*, September 2007, 22 (6), 1013–1031.
- Barbanchon, Thomas Le, Roland Rathelot, and Alexandra Roulet**, “Gender Differences in Job Search: Trading off Commute against Wage,” *The Quarterly Journal of Economics*, 10 2020, 136 (1), 381–426.
- Basso, Leonardo J. and Hugo E. Silva**, “Efficiency and Substitutability of Transit Subsidies and Other Urban Transport Policies,” *American Economic Journal: Economic Policy*, November 2014, 6 (4), 1–33.
- Bauernschuster, Stefan, Timo Hener, and Helmut Rainer**, “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*, 2017, 9 (1), 1–37.
- Baum-Snow, Nathaniel and Lu Han**, “The Microgeography of Housing Supply,” 2021. Working Paper.
- Bayer, Patrick and Christopher Timmins**, “On the equilibrium properties of locational sorting models,” *Journal of Urban Economics*, 2005, 57 (3), 462–477.
- and Robert McMillan, “Tiebout sorting and neighborhood stratification,” *Journal of Public Economics*, 2012, 96 (11), 1129 – 1143.
- , Fernando Ferreira, and Robert McMillan, “A Unified Framework for Measuring Preferences for Schools and Neighborhoods,” *Journal of Political Economy*, August 2007, 115 (4), 588–638.

- , **Nathaniel Keohane, and Christopher Timmins**, “Migration and hedonic valuation: The case of air quality,” *Journal of Environmental Economics and Management*, July 2009, 58 (1), 1–14.
- , **Robert McMillan, Alvin Murphy, and Christopher Timmins**, “A Dynamic Model of Demand for Houses and Neighborhoods,” *Econometrica*, 2016, 84 (3), 893–942.
- , —, and **Kim Rueben**, “An equilibrium model of sorting in an urban housing market,” Technical Report, National Bureau of Economic Research 2004.
- Bento, Antonio, Jonathan Hall, and Kilian Heilmann**, “Estimating Congestion Externalities Using Big Data,” *Technical report*, 2020.
- Bento, Antonio M, Lawrence H Goulder, Mark R Jacobsen, and Roger H Von Haefen**, “Distributional and efficiency impacts of increased US gasoline taxes,” *American Economic Review*, 2009, 99 (3), 667–99.
- Bento, Antonio M., Maureen L. Cropper, Ahmed Mushfiq Mobarak, and Katja Vinha**, “The Effects of Urban Spatial Structure on Travel Demand in the United States,” *The Review of Economics and Statistics*, 2005, 87 (3), 466–478.
- Berry, Steven, James Levinsohn, and Ariel Pakes**, “Automobile prices in market equilibrium,” *Econometrica*, 1995, pp. 841–890.
- Brownstone, David and Kenneth A. Small**, “Valuing time and reliability: assessing the evidence from road pricing demonstrations,” *Transportation Research Part A: Policy and Practice*, 2005, 39 (4), 279 – 293.
- Brueckner, Jan K**, “The structure of urban equilibria: A unified treatment of the Muth-Mills model,” *Handbook of regional and urban economics*, 1987, 2, 821–845.
- , “Urban growth boundaries: An effective second-best remedy for unpriced traffic congestion?,” *Journal of Housing Economics*, 2007, 16 (3-4), 263–273.
- Buchholz, Nicholas, Laura Doval, Jakub Kastl, Filip Matejka, and Tobias Salz**, “The Value of Time: Evidence From Auctioned-Cab Rides,” 2020. Working Paper.
- Calder-Wang, Sophie**, “The Distributional Impact of the Sharing Economy on the Housing Market,” 2020. Working Paper.
- Capps, Cory, David Dranove, and Mark Satterthwaite**, “Competition and market power in option demand markets,” *RAND Journal of Economics*, 2003, pp. 737–763.
- Couture, Victor, Cecile Gaubert, Jessie Handbury, and Erik Hurst**, “Income Growth and the Distributional Effects of Urban Spatial Sorting,” *NBER Working Paper No. 26142*, 2020.
- Davis, Lucas W.**, “The Effect of Driving Restrictions on Air Quality in Mexico City,” *Journal of Political Economy*, 2008, 116, 38–81.
- Diamond, Rebecca**, “The determinants and welfare implications of US workers’ diverging location choices by skill: 1980–2000,” *The American Economic Review*, 2016, 106 (3), 479–524.
- Downs, Anthony**, “The Law of Peak-Hour Expressway Congestion,” *Traffic Quarterly*, 1962, 16 (3), 393–409.
- Duranton, Gilles and Matthew A. Turner**, “Fundamental Law of Road Congestion: Evidence from U.S. Cities.,” *American Economic Review*, 2011, 101(October).
- Eppe, Dennis and Holger Sieg**, “Estimating Equilibrium Models of Local Jurisdictions,” *Journal of Political Economy*, 1999, 107 (4), 645–681.
- , **Richard Romano, and Holger Sieg**, “The intergenerational conflict over the provision of public education,” *Journal of Public Economics*, 2012, 96 (3), 255 – 268.
- Fajgelbaum, Pablo D and Edouard Schaal**, “Optimal transport networks in spatial equilibrium,” *Econometrica*, 2020, 88 (4), 1411–1452.
- Ferreira, Fernando and Maisy Wong**, “Estimating Preferences for Neighborhood Amenities Under Imperfect Information,” *NBER Working Paper 28165*, 2021.

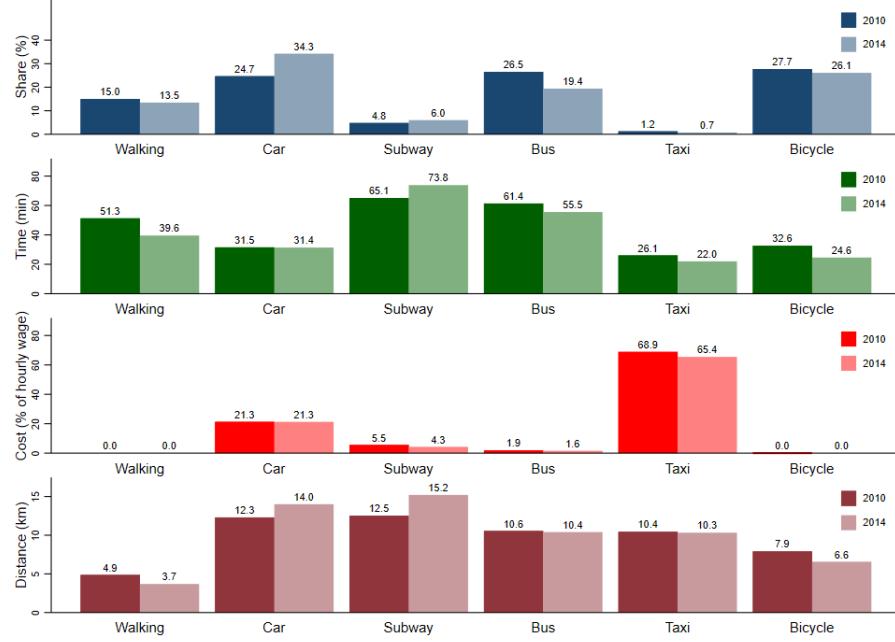
- Ferreyra, Maria Marta**, “Estimating the effects of private school vouchers in multidistrict economies,” *American Economic Review*, 2007, 97 (3), 789–817.
- Gallego, Francisco, Juan-Pablo Montero, and Christian Salas**, “The effect of transport policies on car use: Evidence from Latin American cities,” *Journal of Public Economics*, 2013, 107, 47–62.
- GAO**, “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.
- Glaeser, Edward L, Matthew E Kahn, and Jordan Rappaport**, “Why do the poor live in cities? The role of public transportation,” *Journal of urban Economics*, 2008, 63 (1), 1–24.
- Goldszmidt, Ariel, John A. List, Robert D. Metcalfe, Ian Muir, V. Kerry Smith, and Jenny Wang**, “The Value of Time in the United States: Estimates from a Nationwide Natural Field Experiment,” 2020. Working Paper.
- Gonzalez-Navarro, Marco and Matthew A. Turner**, “Subways and urban growth: Evidence from earth,” *Journal of Urban Economics*, November 2018, 108, 85–106.
- Gu, Yizhen, Jiang Chang, Junfu Zhang, and Ben Zou**, “Subways and Road Congestion,” *American Economic Journal: Applied Economics*, 2020. forthcoming.
- Guevara, C. Angelo and Moshe E. Ben-Akiva**, “Sampling of alternatives in Logit Mixture models,” *Transportation Research Part B: Methodological*, 2013, 58, 185 – 198.
- Hainmueller, Jens**, “Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies,” *Political Analysis*, 2012, 20 (1), 25?46.
- Hall, Jonathan D.**, “Pareto improvements from Lexus Lanes: The effects of pricing a portion of the lanes on congested highways,” *Journal of Public Economics*, 2018, 158, 113 – 125.
- Han, Bing, Lu Han, and Guozhong Zhu**, “Housing Price and Fundamentals in A Transition Economy: The Case of The Beijing Market,” *International Economic Review*, 2018, 59 (3), 1653–1677.
- Hanna, Rema, Gabriel Kreindler, and Benjamin A Olken**, “Citywide effects of high-occupancy vehicle restrictions: Evidence from “three-in-one” in Jakarta,” *Science*, 2017, 357 (6346), 89–93.
- Heblich, Stephan, Stephen J Redding, and Daniel M Sturm**, “The Making of the Modern Metropolis: Evidence from London,” *The Quarterly Journal of Economics*, 05 2020, 135 (4), 2059–2133.
- Hwang, Yujung**, “An Estimable General-Equilibrium Structural Model of Immigrants’ Neighborhood Sorting and Social Integration,” 2019. Working Paper.
- Jerch, Rhiannon, Panle Jia Barwick, Shanjun Li, and Jing Wu**, “Road Rationing Policies and Housing Markets,” 2021. Working Paper.
- Klaiber, Allen H. and Daniel J. Phaneuf**, “Valuing open space in a residential sorting model of the Twin Cities,” *Journal of Environmental Economics and Management*, September 2010, 60 (2), 57–77.
- Kreindler, Gabriel E.**, “The Welfare Effect of Road Congestion Pricing: Experimental Evidence and Equilibrium Implications,” 2018. Working Paper.
- Kuminoff, Nicolai V.**, “Decomposing the structural identification of non-market values,” *Journal of Environmental Economics and Management*, March 2009, 57 (2), 123–139.
- Kuminoff, Nicolai V.**, “Partial identification of preferences in a dual-market sorting equilibrium,” 2012. Working Paper.
- Kuminoff, Nicolai V., V. Kerry Smith, and Christopher Timmins**, “The New Economics of Equilibrium Sorting and Policy Evaluation Using Housing Markets,” *Journal of Economic Literature*, December 2013, 51 (4), 1007–1062.
- Lancaster, Kelvin**, *Consumer Demand: A New Approach*, Columbia University Press, 1971.

- Langer, Ashley and Clifford Winston**, “Toward a Comprehensive Assessment of Road Pricing Accounting for Land Use,” *Brookings-Wharton Papers on Urban Affairs*, 2008, pp. 127–175.
- LeRoy, Stephen F and Jon Sonstelie**, “Paradise lost and regained: Transportation innovation, income, and residential location,” *Journal of Urban Economics*, 1983, 13 (1), 67–89.
- Li, Shanjun**, “Better lucky than rich? Welfare analysis of automobile license allocations in Beijing and Shanghai,” *Review of Economic Studies*, 2018, 85 (4), 2389–2428.
- , **Joshua Linn, and Erich Muehlegger**, “Gasoline taxes and consumer behavior,” *American Economic Journal: Economic Policy*, 2014, 6 (4), 302–42.
- , **Yanyan Liu, Avralt-Od Purevjav, and Lin Yang**, “Does subway expansion improve air quality?,” *Journal of Environmental Economics and Management*, 2019, 96, 213 – 235.
- Liu, Youming, Shanjun Li, and Caixia Shen**, “The Dynamic Efficiency in Resource Allocation: Evidence from Vehicle License Lotteries in Beijing,” Working Paper 26904, National Bureau of Economic Research March 2020.
- Miyauchi, Yuhei, Kentaro Nakajima, and Stephen J Redding**, “Consumption Access and the Spatial Concentration of Economic Activity: Evidence from Smartphone Data,” Working Paper 28497, National Bureau of Economic Research February 2021.
- Mohring, Herbert and Mitchell Harwitz**, *Highway Benefits: An Analytical Framework*, Northwestern University Press, 1962.
- Murphy, Alvin**, “A Dynamic Model of Housing Supply,” Available at SSRN 2200459, 2015.
- Parry, Ian WH and Kenneth A Small**, “Does Britain or the United States have the right gasoline tax?,” *American Economic Review*, 2005, 95 (4), 1276–1289.
- and — , “Should Urban Transit Subsidies Be Reduced?,” *The American Economic Review*, 2009, 99 (3), 700–724.
- Petrin, Amil**, “Quantifying the Benefit of New Products: the Case of Minivan,” *Journal of Political Economy*, 2002, 110 (4), 705–729.
- Phaneuf, Daniel J, V Kerry Smith, Raymond B Palmquist, and Jaren C Pope**, “Integrating Property Value and Local Recreation Models to Value Ecosystem Services in Urban Watersheds,” *Land Economics*, 2008, 84 (3), 361–381.
- Rao, Deyu**, “The Role of Environmental Amenities in the Urban Economy: Evidence from a Spatial General Equilibrium Approach,” 2021. Working Paper.
- Redding, Stephen J. and Esteban Rossi-Hansberg**, “Quantitative Spatial Economics,” *Annual Review of Economics*, 2017, 9 (1), 21–58.
- Saiz, Albert**, “The geographic determinants of housing supply,” *The Quarterly Journal of Economics*, 2010, 125 (3), 1253–1296.
- Severen, Christopher**, “Commuting, Labor, and Housing Market Effects of Mass Transportation: Welfare and Identification,” 2019. Working Paper.
- Sieg, Holger, V Kerry Smith, H Spencer Banzhaf, and Randy Walsh**, “Estimating the general equilibrium benefits of large changes in spatially delineated public goods,” *International Economic Review*, 2004, 45 (4), 1047–1077.
- Small, Kenneth A.**, “Valuation of travel time,” *Economics of Transportation*, 2012, 1 (1-2), 2–14.
- Small, Kenneth A, Clifford Winston, and Jia Yan**, “Uncovering the distribution of motorists’ preferences for travel time and reliability,” *Econometrica*, 2005, 73 (4), 1367–1382.
- , **Erik T Verhoef, and Robin Lindsey**, *The Economics of Urban Transportation*, Routledge, 2007.
- Stiglitz, Joseph E**, “The theory of local public goods,” in “The economics of public services,” Springer, 1977, pp. 274–333.
- Timmins, Christopher and Jennifer Murdock**, “A revealed preference approach to the measurement of congestion in travel cost models,” *Journal of Environmental Economics and management*, 2007, 53 (2), 230–249.
- Train, Kenneth E. and Clifford Winston**, “Vehicle Choice Behavior and the Declining Market Share of U.S. Automakers,” *Inter-*

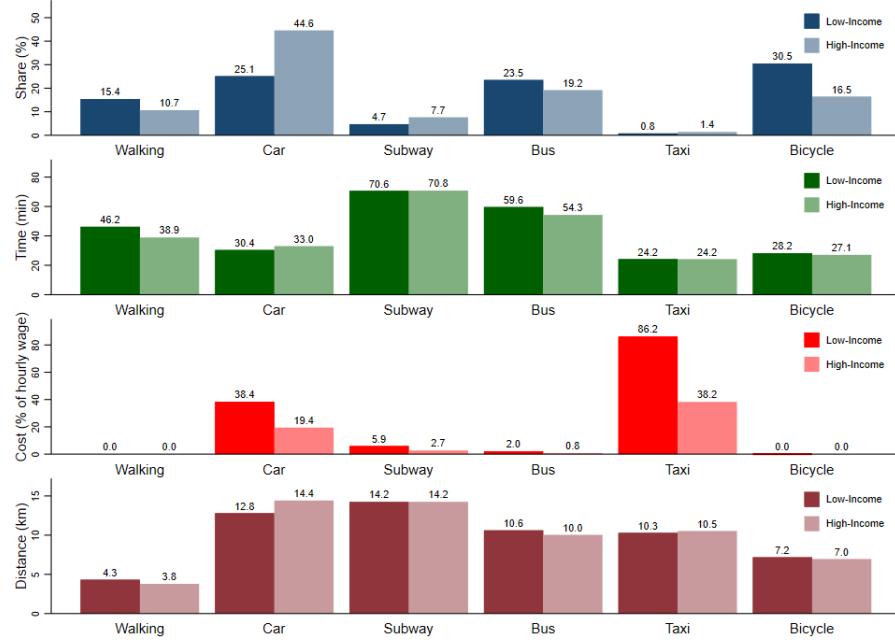
- national Economic Review*, 2007, 48, 1469–1496.
- Tsivianidis, Nick**, “Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogota’s TransMilenio,” 2019. Working Paper.
- USDOT**, “Revised departmental guidance on valuation of travel time in economic analysis,” 2015. US Department of Transportation, Washington, DC.
- Verhoef, Erik T. and Herbert Mohring**, “Self-Financing Roads,” *International Journal of Sustainable Transportation*, 2009, 3 (5-6), 293–311.
- Viard, V Brian and Shihe Fu**, “The effect of Beijing’s driving restrictions on pollution and economic activity,” *Journal of Public Economics*, 2015, 125, 98–115.
- Vickrey, William**, “Pricing in Urban and Suburban Transport,” *American Economic Review*, 1963, 53 (2), 452–465.
- Walsh, Randy et al.**, “Endogenous open space amenities in a locational equilibrium,” *Journal of urban Economics*, 2007, 61 (2), 319–344.
- Wang, Songtao, Su Han Chan, and Bohua Xu**, “The Estimation and Determinants of the Price Elasticity of Housing Supply: Evidence from China,” *Journal of Real Estate Research*, 2012, 33 (3), 311–344.
- Wang, Wen**, “Environmental Gentrification,” *Working Paper*, 2020.
- Wasi, Nada and Michael P. Keane**, “Estimation of Discrete Choice Models with Many Alternatives Using Random Subsets of the Full Choice Set: With an Application to Demand for Frozen Pizza,” *Working Paper*, 2012.
- Waxman, Andrew**, “The Long Road to Work: The Divergent Effects of Transportation Policies by Worker Skill in a Locational Sorting Model,” *working paper*, 2017.
- Winston, Clifford**, “Efficient Transportation Infrastructure Policy,” *The Journal of Economic Perspectives*, 1991, 5 (1), 113–127.
- Wolff, Hendrik**, “Value of time: Speeding behavior and gasoline prices,” *Journal of Environmental Economics and Management*, 2014, 67 (1), 71–88.
- Xiao, Junji, Xiaolan Zhou, and Wei-Min Hu**, “Welfare Analysis of the Vehicle Quota System in China,” *International Economic Review*, 2017, 58 (2), 617–650.
- Yang, Jun, Avralt-Od Purevjav, and Shanjun Li**, “The Marginal Cost of Traffic Congestion and Road Pricing: Evidence from a Natural Experiment in Beijing,” *American Economic Journal: Economic Policy*, 2020, 12 (1), 418–453.
- , **Shuai Chen, Ping Qin, Fangwen Lu, and Antung A Liu**, “The Effect of Subway Expansions on Vehicle Congestion: Evidence from Beijing,” *Journal of Environmental Economics and Management*, 2018, 88, 114–133.
- Zhang, Wei, C-Y Cynthia Lin Lawell, and Victoria I Umanskaya**, “The effects of license plate-based driving restrictions on air quality: Theory and empirical evidence,” *Journal of Environmental Economics and Management*, 2017, 82, 181–220.

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

(a) Year 2010 vs. Year 2014

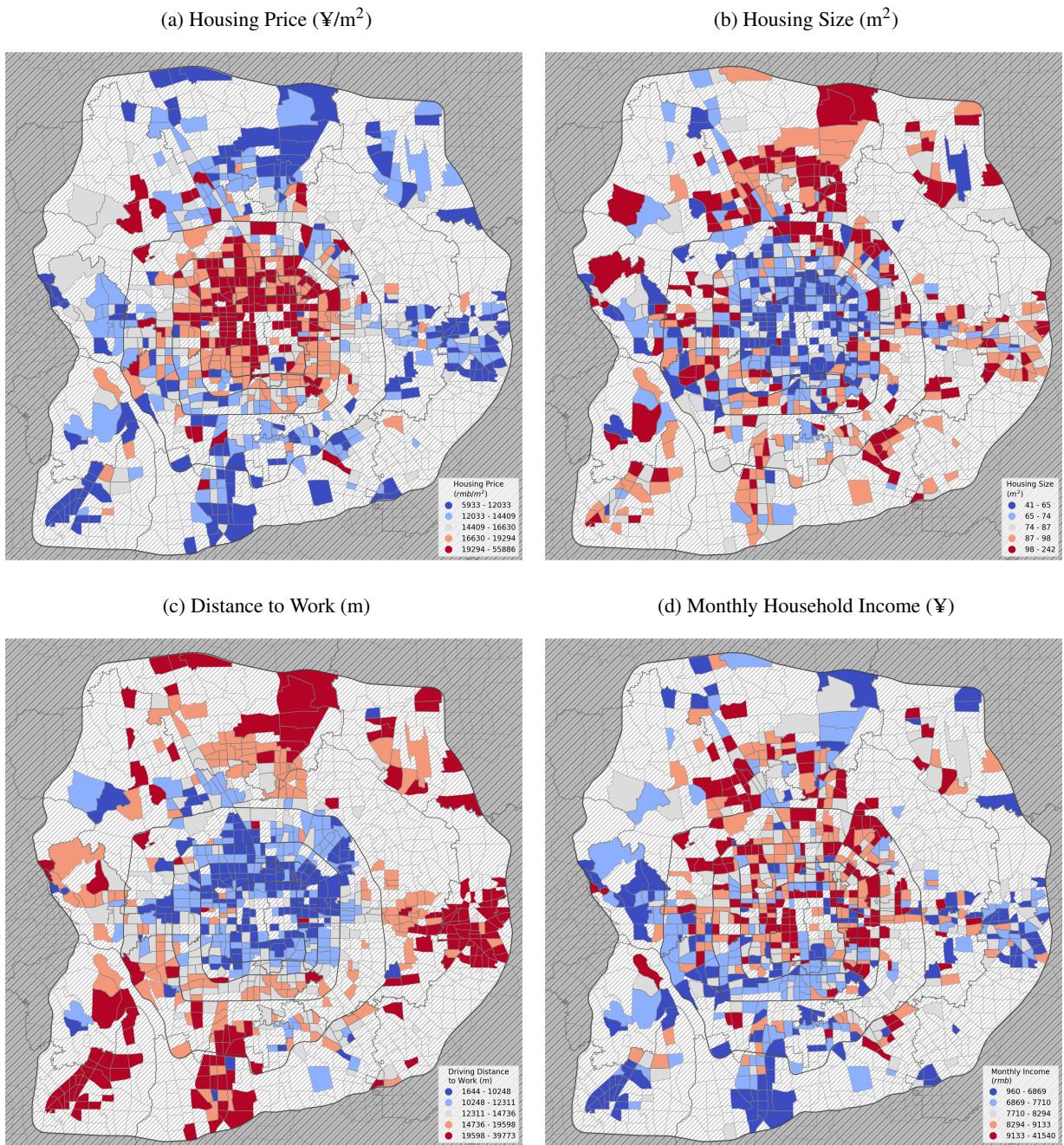


(b) High-income vs. Low-income Households



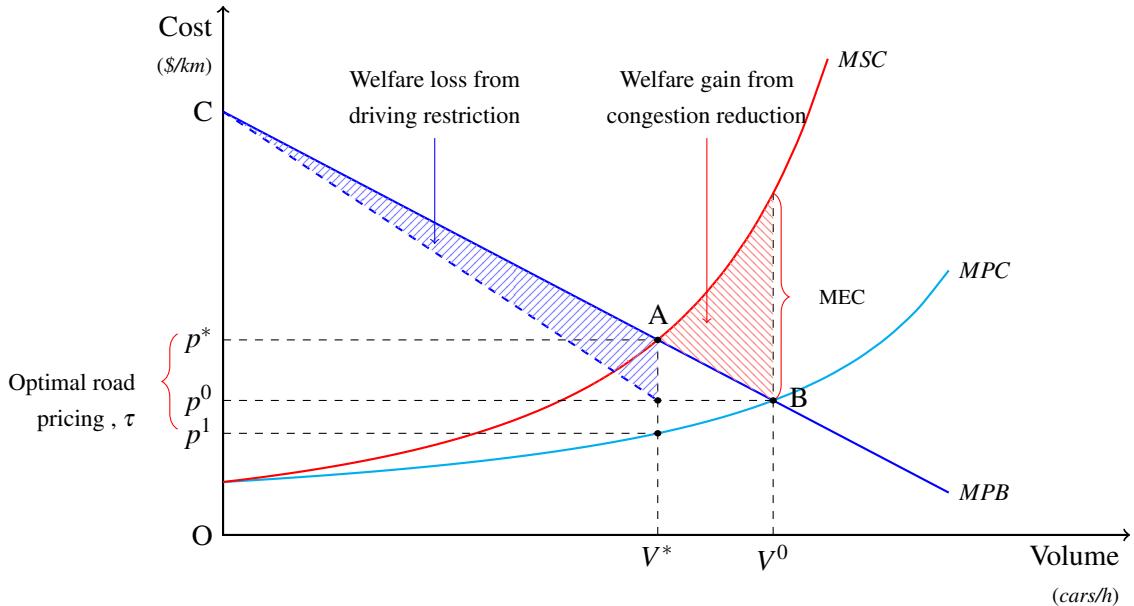
Note: This figure plots the trip share, time, costs, and average distance by travel modes for work commuting trips in the Beijing Household Travel Survey of 2010 and 2014. There are six main trip modes: walk, bike, bus, subway, car, and taxi. Bus and subway trips could include segments with other modes but we characterize them as bus and subway trips. Trips using both bus and subway are rare (less than 3% in the data and are dropped in the analysis.) Travel time, cost (defined as % of hourly wage), and distance are constructed as in Appendix A.1. High-income households are defined as households whose income is greater than the median in the survey year.

Figure 2: Housing and Household Attributes from Mortgage Data



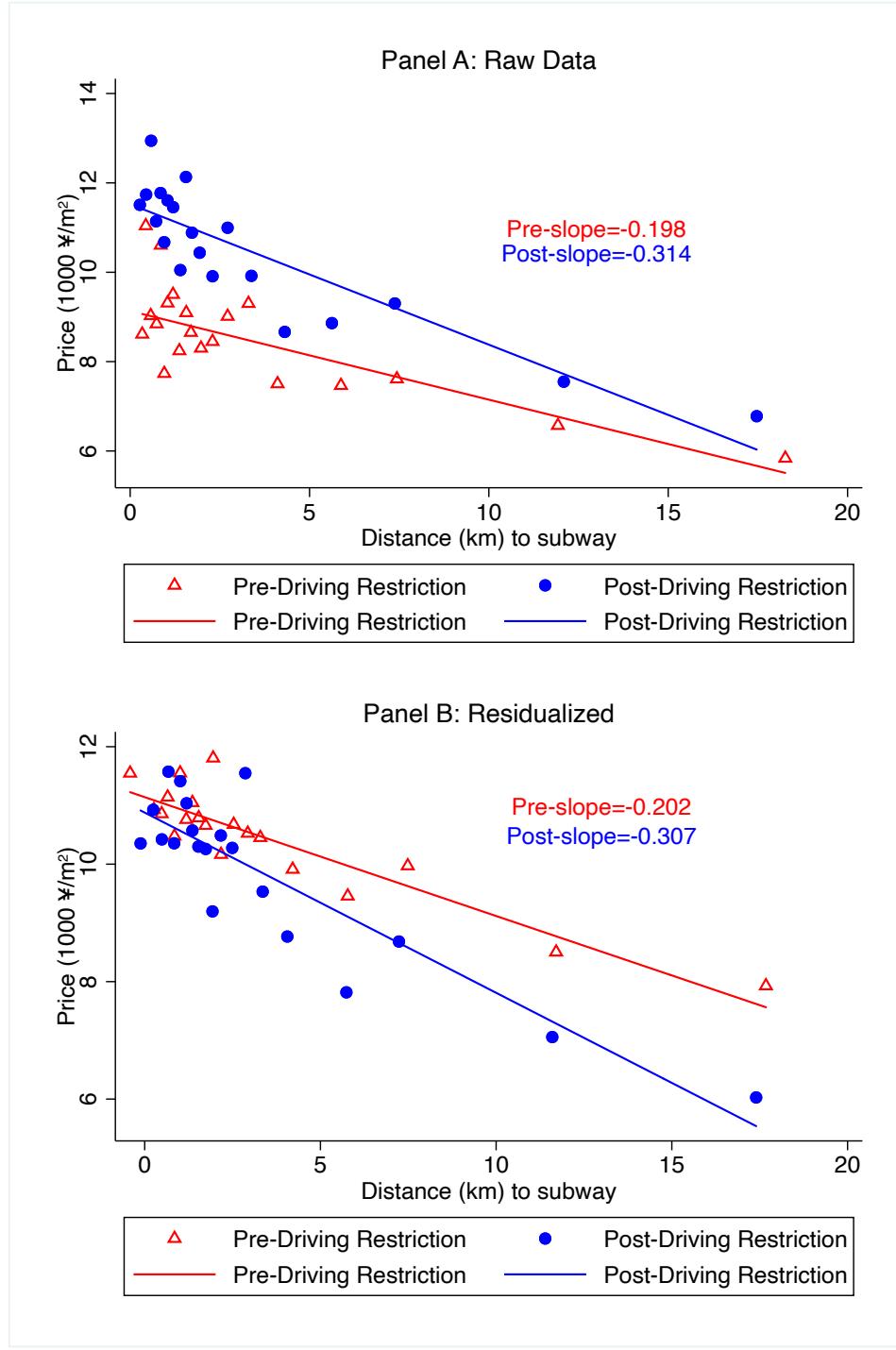
Note: This figure plots the average housing price and size and household commuting distance and monthly income by Traffic Analysis Zones (TAZ) based on the 2006-2014 mortgage data. TAZs are standardized spatial units used by transportation planners. There were 2050 TAZs in Beijing in 2014. Distance to 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.

Figure 3: Welfare under Congestion Pricing and Driving Restriction in Partial Equilibrium



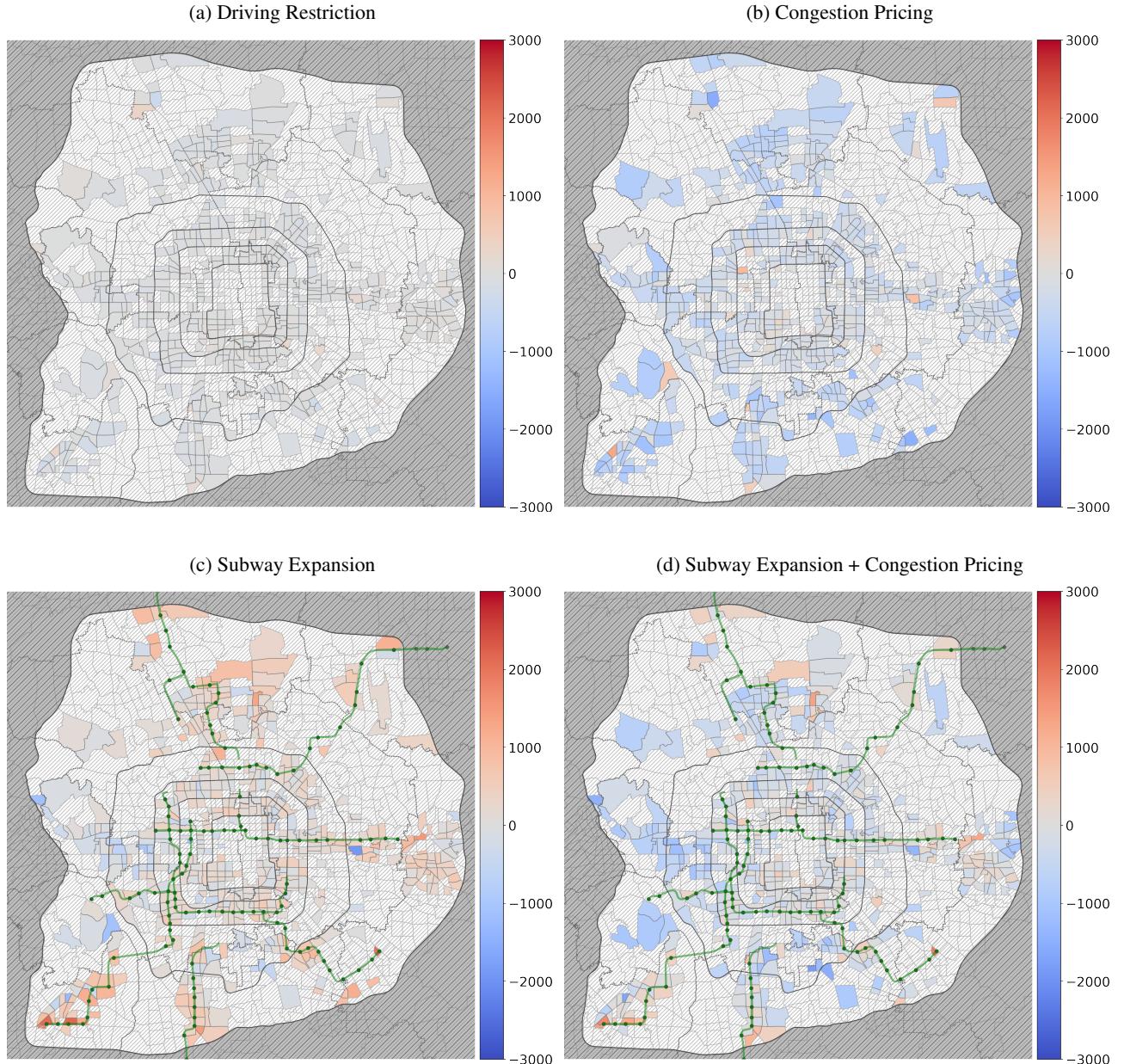
Note: The figure illustrates the welfare impacts of optimal congestion pricing and driving restriction. The x-axis denotes traffic volume (measured by the number of cars per hour passing a location). The marginal private benefit MPB curve represents demand for driving (willingness to pay for driving). The marginal private cost MPC curve reflects the private cost of driving, while the marginal social cost MSC curve reflects changes in the aggregate commuting costs by all drivers when there is an additional driver on the road. The difference between the private cost MPC and the social cost MSC is the congestion externality (or the marginal external cost of congestion, MEC). In the absence of any intervention, equilibrium occurs at V^0 . The shaded red area shows the deadweight loss due to excess congestion. Congestion pricing (i.e., a Pigouvian tax), τ , can be imposed to achieve the socially optimal level of traffic volume V^* . Different from congestion pricing, a driving restriction is a command-and-control approach and eliminates trips at all levels of marginal benefit. Assuming that the driving restriction results in a random reduction of trips from V^0 to V^* , total consumer surplus is reduced in proportionate to $\frac{V^0}{V^*}$ of that from no policy, i.e., from CBV^0O to CAV^*O . Taking into consideration the changes in the aggregated social costs of driving, a driving restriction involves additional welfare losses as shaded in blue compared to congestion pricing. Therefore, the welfare impact of the driving restriction is ambiguous *a priori*. We abstract away from potential income effects that arise from the congestion pricing, which can be offset by recycling the toll revenue.

Figure 4: Reduced-Form Evidence on Housing Price Gradient before and after Driving Restriction



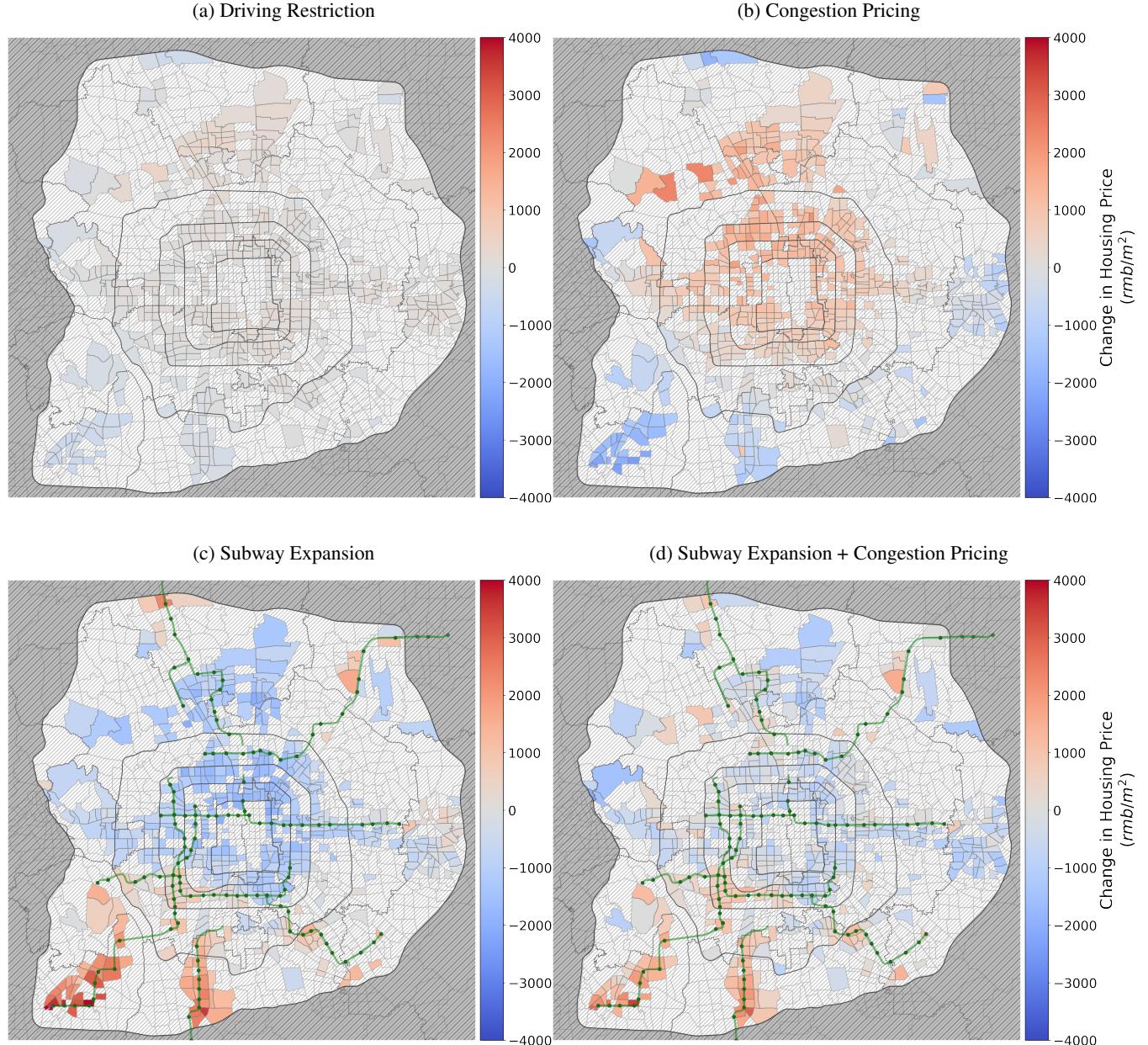
Notes: These binned scatterplots display housing price in ¥1000s per square meter against distance to nearest subway stations before and after the driving restriction. The sample spans 24 months before and after the policy's starting point (July 2008) with 23,917 observations. The top panel is based on raw prices. The bottom panel controls for neighborhood fixed effects, and year by month fixed effects and presents residualized prices. The slopes are based on linear fits.

Figure 5: Changes in Commuting Distances from Counterfactual Simulations (in meters)



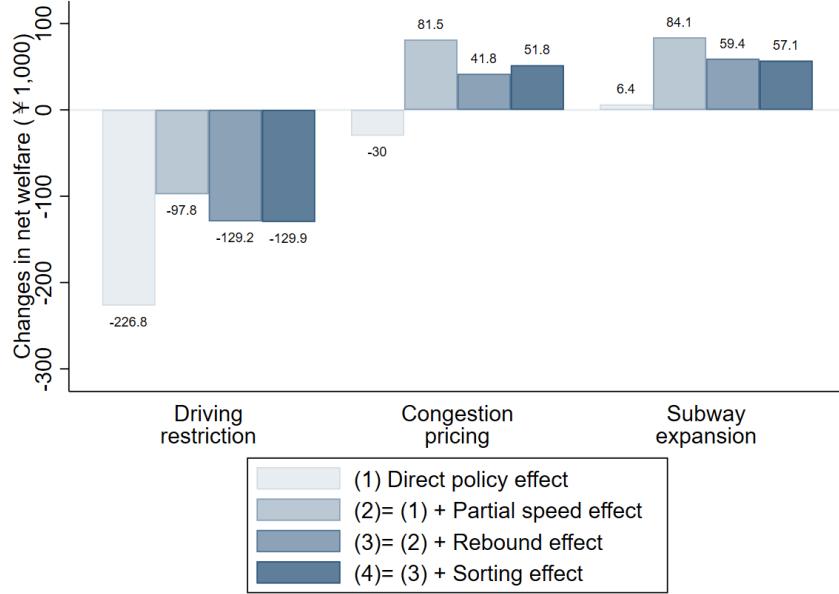
Note: This figure illustrates simulated changes in commuting distances (in meters) across TAZs under different counterfactual policies (relative to the no policy scenario). The results are based on the simulations in Table 6. Warmer colors correspond to increases in commuting distance while colder colors represent decreases. Green lines represent new subway lines built between year 2008 and 2014.

Figure 6: Changes in Housing Prices from Counterfactual Simulations (in $\text{¥}/m^2$)



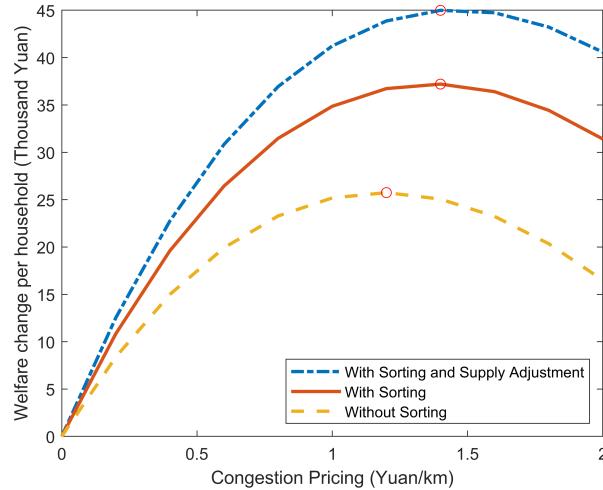
Note: This figure illustrates simulated changes in housing prices (in $\text{¥}/m^2$) across TAZs under different counterfactual policies (relative to the no policy scenario). The results are based on the simulations in Table 6. Warmer colors correspond to increases in commuting distance while colder colors represent decreases. Green lines represent new subway lines built between year 2008 and 2014.

Figure 7: Welfare Decomposition for Different Transportation Policies



Notes: This figure decomposes welfare changes per household along four adjustment margins. For each policy, the bars display the cumulative welfare changes incorporating previous margins. The direct policy effect measures changes in household 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 from 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) and toll revenues (net of capital and operating costs for the congesting pricing system).

Figure 8: Optimal Congestion Pricing under the 2014 Subway Network



Note: The figure plots welfare changes against 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 optimal congestion pricing is ¥1.2/km without sorting and ¥1.4/km with sorting. Sorting increases consumer welfare by 20%-30% and supply adjustment contributes to another 10%-20% for most congestion pricing levels.

Table 1: Summary Statistics of Household Travel Survey

		2010			2014		
	N	Mean	SD	N	Mean	SD	
Respondent characteristics							
Income: <¥ 50k	14780	0.48	0.50	20573	0.18	0.38	
Income: [¥ 50k, ¥ 100k)	14780	0.39	0.49	20573	0.44	0.50	
Income: >=¥ 100k	14780	0.13	0.34	20573	0.38	0.49	
Having a car (=1)	14780	0.44	0.50	20573	0.62	0.49	
Female (=1)	14780	0.44	0.50	20573	0.43	0.50	
Age (in years)	14780	37.59	10.28	20573	38.47	9.84	
College or higher (=1)	14780	0.61	0.49	20573	0.64	0.48	
Home within 4th ring (=1)	14780	0.51	0.50	20573	0.41	0.49	
Workplace within 4th ring (=1)	14780	0.59	0.49	20573	0.50	0.50	
Trip related variables							
Travel time (hour)	30334	0.87	1.06	42820	0.74	0.98	
Travel cost (¥)	30334	2.47	5.55	42820	3.83	6.96	
Distance: <2km	30334	0.25	0.43	42820	0.24	0.43	
Distance: [2, 5km)	30334	0.27	0.45	42820	0.26	0.44	

Note: The table reports survey respondent demographics and trip attributes of all work commuting trips within the 6th ring road from the 2010 and 2014 Beijing Household Travel Survey. Travel time and travel cost are constructed as in Appendix A.1. Trip distance is measured by straight-lines. Distance<2km and Distance within 2-5km flag commuting trips with a short to medium-distance.

Table 2: Summary Statistics of Housing Data

	Mean	SD	Min	Max
Housing attributes				
Transaction year	2011	1.89	2006	2014
Price (¥1000/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 (¥1000)	159.71	103.34	6.24	2556.90
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
Num. of units	1972	1521	24	13031
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

Note: This table reports statistics from the 2006-2014 mortgage dataset. The number of housing transactions is 79,884, all of which are within the 6th 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.

Table 3: Estimation Results for Travel Mode Choices

	Logit			Random coefficient		
	(1)	(2)	(3)	(4)	(5)	(6)
Travel time (γ_1)	-1.194 (0.082)	-0.270 (0.006)	-0.191 (0.006)			
Travel cost/hourly wage (γ_2)	-1.578 (0.324)	-0.788 (0.028)	-0.565 (0.034)	-1.411 (0.041)	-1.424 (0.052)	-2.531 (0.065)
Random coefficients on travel time (μ_γ)						
Travel Time				-0.955 (0.008)	-0.885 (0.008)	-0.931 (0.012)
Random coefficients on mode dummies (σ_m)						
Driving					3.394 (0.049)	3.391 (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

Note: The number of observations are 73,154. The specifications include an increasingly rich set of fixed effects interacting with travel model dummies. Trip related FE includes trip distance bins and ring road dummies for the origin and destination (e.g., if the origin is between the 2nd and 3rd ring roads). Demographics FE includes a respondent's age, gender, education, and car ownership. 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 5th and 95th percentile) with three degrees of freedom to allow for long tails. The estimates of μ_γ are provided in the table. 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). Standard errors are displayed below parameter estimates.

Table 4: Housing Demand - Nonlinear Parameters from Simulated MLE

	(1) No EV		(2) With EV		(3) EV and random coef.	
	Para	SE	Para	SE	Para	SE
Demographic Interactions						
Price (¥mill.) * ln(income)	0.965	0.007	1.005	0.008	1.030	0.016
Age in 30-45 * ln(distance to key school)	-0.329	0.004	-0.391	0.005	-0.420	0.010
Age > 45 * ln(distance to key school)	-0.074	0.009	-0.111	0.010	-0.123	0.021
Age in 30-45 * ln(home size)	1.343	0.014	1.443	0.015	1.486	0.029
Age > 45 * ln(home size)	2.394	0.028	2.665	0.031	2.746	0.061
EV _{Male}			0.709	0.026	0.755	0.006
EV _{Female}			0.833	0.026	0.893	0.006
Random Coefficients						
$\sigma(\text{EV}_{\text{Male}})$					0.379	0.013
$\sigma(\text{EV}_{\text{Female}})$					0.482	0.012
Log-likelihood	-206,829		-170,057		-168,808	

Note: this table reports MLE estimates of housing demand's nonlinear parameters 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 the on EV terms.

Table 5: Housing Demand - Linear Parameters

	OLS (1)	OLS (2)	IV1 (3)	IV2 (4)	IV2+IV3 (5)	All IVs (6)
Price (¥mill.)	-2.240 (0.186)	-2.191 (0.184)	-7.091 (1.640)	-6.283 (0.867)	-6.454 (0.583)	-6.596 (0.534)
Ln(home size)	-3.648 (0.257)	-3.797 (0.261)	4.721 (2.927)	3.331 (1.505)	3.631 (1.022)	3.879 (0.969)
Building age	-0.043 (0.007)	-0.029 (0.006)	-0.144 (0.040)	-0.125 (0.020)	-0.129 (0.014)	-0.132 (0.013)
Floor area ratio	-0.006 (0.034)	-0.009 (0.025)	-0.019 (0.036)	-0.023 (0.032)	-0.023 (0.033)	-0.023 (0.034)
Ln(dist. to park)	0.210 (0.069)	0.074 (0.057)	-0.475 (0.222)	-0.389 (0.117)	-0.408 (0.101)	-0.424 (0.103)
Ln(dist. to key school)	0.950 (0.080)	0.782 (0.137)	0.210 (0.213)	0.323 (0.139)	0.304 (0.121)	0.288 (0.118)
Year-Month-District FE	Y	Y	Y	Y	Y	Y
Neighborhood FE		Y	Y	Y	Y	Y
First-stage Kleinberg-Paap F			9.9	10.5	14.2	14.2
Avg. Housing Demand Price elasticity	2.96	2.96	-1.94	-1.04	-1.34	-1.44

Note: 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. Column (3) use IV1 as price instruments, i.e. the number of homes that are within 3km from a given home, outside the same complex, and sold in a two-month window. Columns (4), (5), and (6) use IV2, i.e. the average attributes of these homes (building size, age, log distance to park, and log distance to key school). Column (5) also includes IV3, the interaction between IV2 and the winning odds of the licence lottery. The winning odds decreased from 9.4% in January 2011 to 0.7% by the end of 2014. Column (6) uses all IVs. Standard errors are clustered at the neighborhood level.

Table 6: Simulation Results with Household Sorting

Income relative to the median	2008 Subway Network						2014 Subway Network					
	(1) No Policy		(2) Driving restriction		(3) Congestion pricing		(4) Subway Expansion		(5) + Driving restriction		(6) + Congestion pricing	
			Δs from (1)		Δs from (1)		Δs from (1)		Δs from (1)		Δs from (1)	
	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	-3.40	-3.48	-5.39	-2.14	-1.66	-8.52	-4.62	-5.20	-6.40
Subway	9.02	10.77	1.29	0.70	0.84	0.96	4.62	6.06	5.79	6.44	5.24	6.83
Bus	22.44	30.47	1.78	0.60	0.57	1.24	-1.54	-2.53	0.31	-1.57	-0.76	-1.03
Bike	15.96	24.01	1.60	0.80	0.77	1.78	-0.80	-1.64	0.52	-0.94	-0.13	-0.13
Taxi	2.20	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
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	4.30	-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	100.0	-14.0	64.0	108.7	28.7
Toll revenue (+)					137.4	137.4					127.7	127.7
Subway cost (-)							103.0	103.0	103.0	103.0	103.0	103.0
Net welfare	-227.1		-32.7	39.2	64.3	117.3	-3.0	-117.0	-39.0	133.4	53.4	

Note: 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. Appendix E explains the simulation procedure. We incorporate household sorting and keep housing supply fixed. 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. Toll revenue, net of the capital and operating costs of the system, is recycled uniformly across households. Subway cost includes the construction and operation costs that are equally distributed among 7.2 million households. Net welfare is consumer surplus plus recycled revenue and minus subway costs. See Appendix Section E for more details.

Table 7: Importance of Sorting and Endogenous Congestion and Various Extensions

Income relative to the median	Driving restriction			Congestion pricing			Subway expansion		
	$\Delta Speed$ (km/h)	$\Delta Welfare$ (¥1,000) High	$\Delta Welfare$ (¥1,000) Low	$\Delta Speed$ (km/h)	$\Delta Welfare$ (¥1,000) High	$\Delta Welfare$ (¥1,000) Low	$\Delta Speed$ (km/h)	$\Delta Welfare$ (¥1,000) High	$\Delta Welfare$ (¥1,000) Low
Panel (A): importance of sorting and endogenous congestion									
With sorting (main results)	3.83	-227.1	-32.7	3.83	39.2	64.3	1.49	117.3	-3.0
Without sorting	3.82	-227.3	-31.0	3.61	28.1	55.6	1.76	104.7	14.0
With sorting but without endogeneous congestion	5.47	-111.6	-8.7	5.62	112.2	75.1	2.36	260.1	107.7
Panel (B): extensions and robustness checks									
With housing supply	3.83	-226.9	-31.6	3.97	51.4	65.7	1.13	84.4	2.3
With location-specific traffic density	3.82	-227.6	-3.3	3.81	31.7	61.1	1.48	113.0	-0.4
Without random coefficients	4.59	-1447.6	-338.0	4.59	-406.9	-3.6	0.16	15.5	-42.5
With migration	4.65	-197.7	-26.9	4.63	72.7	71.3	0.73	80.3	-10.2
With consumption access	3.83	-302.0	-43.5	3.83	52.2	85.5	1.49	190.7	30.3

Note: Panel (A) examines the importance of sorting and endogenous congestion. Each cell reports changes relative to the no-policy scenario. Speed without any policy is 21.49 km/h. The unit of welfare changes is ¥1,000. Congestion pricing is fixed at ¥1.13 per km. The first row summarizes results in Table 6. “Without sorting” holds residential locations 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 equation (13), but do not impose the transportation sector’s equilibrium condition. Panel (B) considers various extensions. “With housing supply” assumes that the housing supply responds to price changes under a constant price elasticity of 0.53. “With location-specific traffic density” incorporates ring road-band specific traffic densities and adjusts travel speeds using the corresponding density in appropriate ring road bands. “Without random coefficients” re-estimate the entire model with no random coefficients and repeat the counterfactual analysis. “With migration” assumes 5% more vehicles (in-migration) under the subway congestion and 5% fewer vehicles (out-migration) under the driving restriction and congestion pricing. “With consumption access” incorporates additional 33% of changes in consumer surplus through the changing availability of consumption services (restaurants, shops, and theaters).

Appendix

A Data Construction

A.1 Household Travel Survey Data

Here we describe the procedures we used to clean the 2010 and 2014 Beijing Household Travel Surveys (BHTS), geocode home and work addresses, and construct commuting routes for trips in the BHTS data and hypothetical trips in the mortgage data. The notation in this appendix follows as closely to that in the main text as possible.

BHTS is designed to be representative using a multistage cluster sampling of households in Beijing. In the first stage, BTRC randomly selects a subset of Traffic Analysis Zones (TAZs) from the entire city. TAZs are one to two square km on average and their size is inversely proportional to the density of trip origins and destinations: the TAZs are smaller closer to the city center. For the first stage of sampling, BTRC selected 642 out of 1,191 TAZs in 2010 and 667 out of 2050 TAZs in 2014, respectively. In the second stage, about 75 and 60 households were randomly selected for in-person interviews for each TAZ.

BHTS includes detailed individual and household demographic (e.g., income, household size, vehicle ownership, home ownership, age, gender) and occupations, availability of transportation options (vehicles, bikes, etc.), and a travel diary on all trips taken during the preceding 24 hours. Household income is reported in bins and we use bin midpoints to measure income.

We focus on six commuting modes: *Walk, Bike, Bus, Subway, Car, and Taxi*. In principle, a traveler could take arbitrary combinations of different travel modes. In our data, single-mode trips account for over 95% of all trips. We therefore eschew multi-mode commuting trips, except for subway and bus trips where we allow commuters to walk to and from subway stations and bus stops.

We use Baidu's API to geocode addresses because its quality of matching Chinese character strings is higher than alternative APIs such as Google Maps. We found that Baidu's Geocoding API performed best for home addresses and its Place API performed best for work addresses. About 36% of 2010 respondents and 44% of 2014 respondents are dropped because their home or work addresses cannot be geocoded.

The 2010 survey contains 46,900 households, 116,142 individuals, and 253,648 trips, while the 2014 survey contains 40,005 households, 101,827 individuals, 205,148 trips. We dropped trips with the origin or destination that could not be geocoded (40%), trips on weekends and holidays (10%), trips of non-working aged respondents (age > 65 or age < 16, 12%), trips using mixed travel modes among subway, bus, and driving (3%), and trips with implausible trip distance and travel time (3%). The remaining sample includes 78,246 trips by 29,770 individuals in the year 2010 and 98,730 trips by 38,829 individuals in the year 2014. The analysis in the main text focuses on work commuting trips, with a total of 73,154 observations.

The size of the choice set varies across commuters and trips. Driving is available for households with personal vehicles on non-restricted days. Car rental is uncommon in Beijing and the mode share of rental cars

is nearly zero in the travel survey. Walk, bike, taxi, and subway modes are available for all trips. We assume that households walk to/from the nearest subway stations if they take subways. Bus availability is determined by the home and work locations. We remove bus from the choice set if Gaode Maps API fails to provide any bus route, indicating a lack of the public bus service in the vicinity.

The monetary cost for walking is zero. For biking, the cost is zero if a household has a bike and the rental price (free for the first hour and then ¥1 per hour with ¥10 as the maximum payment for 24 hours) otherwise. The bus fare is set by the municipality at 0 for senior citizens, ¥0.2 for students, ¥0.4 for people with public transportation cards, and ¥1 for people without public transportation cards. The subway cost per trip is set by the public transport authority at ¥2 and adjusted by the type of public transportation card the traveler holds. Fuel cost is a major component of the monetary cost associated with driving. Based on the average fuel economy reported by vehicle owners in BHTS, we use 0.094 liter/km (10.6 km/liter) for 2010, and 0.118 liter/km (8.5 km/liter) for 2014. Gasoline prices are ¥6.87/liter in 2010 and ¥7.54/liter in 2014. We also assume a tear-and-wear cost which is 0.3 yuan per km. In 2010, the taxi charge is ¥10 for the first 3 km, ¥2 for each additional km, and ¥1 for the gasoline fee. In 2014, the charge increases to ¥13 for the first 3 km, ¥2.3 for each additional km plus ¥1 for the gasoline fee.

The construction of the travel time and distance via API and GIS is illustrated in Appendix Figure A4. To take into consideration differences between peak and off-peak traffic, we queried Baidu and Gaode API at the same departure time on the same weekday as that reported in the survey to obtain driving time predictions. Then we use historic levels of Beijing's Travel Congestion Index (TCI) to adjust the travel time for driving, taxi, and bus to the relevant historical years. We use ring-road specific TCIs for the travel time adjustment.

For subway commuting, we identified the nearest subway stations to home and to work using ArcGIS maps of the historical subway stations and used Baidu's API to calculate walking distances and time from home and work to the nearest subway station. For the BHTS travel survey data, the subway commuting time is calculated using the historical subway system at the time of the survey, including additional time when transferring lines. For hypothetical trips considered by home buyers in the mortgage data, we assume buyers are forward-looking and use the subway network two years after the home purchase date. This is because the subway construction goes through a lengthy process and it takes a few months to a few years from the the public announcement of subway station locations to the actual operation. Households are likely to be aware of new subway stations in the near future and we allow households to consider this in their purchase decisions. We also conduct a robustness check using a one-year projection window in constructing subway time and obtain similar results. Appendix Figure A5 shows travel time and cost of six routes for a particular trip based on the procedure.

The constructed travel distances and reported travel distances of chosen modes in the final dataset are highly correlated (correlation= 0.81). Correlation is highest among walking trips (0.99), followed by bicycle trips (0.98), subway trips (0.94), bus trips (0.88), car trips (0.61), and taxi trips (0.49).

A.2 Mortgage Housing Transaction Data

As part of the social safety net, the mortgage program aims to encourage home ownership by offering prospective homeowners mortgages with a subsidized interest rate. Similar to the retirement benefit, employees and employers are required to contribute a specific percentage of the employee's monthly wage to a mortgage account under this program. The savings contributed to this account can only be used for housing purchases and rental. Workers with formal employment were eligible for this government-backed mortgage program, upon which our data are based.

Although the mortgage data have a good representation of Beijing's middle-class, it under-represents low-income households without employment and high-income households who do not take loans for home purchases. To increase the representativeness of the mortgage data, we re-weight them based on two larger datasets that are more representative of home buyers in Beijing.

The first dataset include sales of new properties that are compiled from home registration records from Beijing Municipal Commission of Housing and Urban-Rural Development, accounting for 90% of all new home sales. It does not include employer-provided/subsidized housing. The second dataset includes 40% of all transactions in Beijing's second-hand market during our data period and is sourced from China's largest real estate brokerage company, Lianjia, that is present at all neighborhoods and across housing segments ([Jerch et al., 2021](#)). Different from the mortgage data, these datasets do not include information on the work location of the owners, therefore preventing us from using it for the main empirical analysis. To ease explanation below, we call these two larger transaction datasets "population dataset".

To improve the representativeness of the mortgage data, we match the distributions of housing price, size, age, and distance to the city center in the mortgage data to those in the population dataset using entropy balancing following [Hainmueller \(2012\)](#). Specifically, we solve the following constrained optimization problem to match sample moments between the mortgage data and the population dataset:³⁶

$$\min_{w_i} H(w) = \sum_i h(w_i) = \sum_i w_i \log(w_i) \quad (\text{A1})$$

subject to balance and normalizing constraints:

$$\frac{1}{N} \sum_{i \in \text{new homes}} w_i (X_{ij}^{\text{mortgage}} - \mu_j^{\text{mortgage}})^r = E_{\text{new homes}}[(X_j - \mu_j)^r]$$

$$\frac{1}{N} \sum_{i \in \text{resales}} w_i (X_{ij}^{\text{mortgage}} - \mu_j^{\text{mortgage}})^r = E_{\text{resales}}[(X_j - \mu_j)^r]$$

$$\sum_i w_i = N = \text{total number of new homes + resales in the mortgage data}$$

$$w_i \geq 0 \text{ for all } i.$$

$$\frac{\sum_{i \in \text{new homes}} w_i}{\sum_{i \in \text{resales}} w_i} = E \left[\frac{\text{new homes}}{\text{resales}} \right]$$

³⁶The objective function, h , is a special case of a Kubelock divergence function, where the base weight, which w_i within the logarithm is divided by, is set to 1.

Here w_i is property i 's weight, $\sum_{i \in \text{new homes}} w_i (X_{ij} - \mu_j)^r$ is the r th order (weighted) moment of matching covariate X_j among new home transactions in the mortgage data, and $E_{\text{new homes}}[(X_j - \mu_j)^r]$ is the r th order moment of covariate X_j among new home transactions in the population dataset. Similarly, $\sum_{i \in \text{resales}} w_i (X_{ij} - \mu_j)^r$ is the r th order moment of covariate X_j among second-hand housing transactions in the mortgage data, and $E(X_j - \mu_j)^r_{\text{resales}}$ is the r th order moments among used properties in the population data.

Matching covariates include housing prices, sizes, building ages, and distances to city center. We match both the mean and variance ($r = 1$ and $r = 2$). The third constraint normalizes the sum of weights to the total number of homes in the mortgage dataset (which is N). The fourth constraint requires weights to be positive. The last constraint requires the ratio of new homes to resales to be the same as the official statistics provided by Beijing Municipal Commission of Housing and Urban-Rural Development. We solve for optimal weights, w_i^* using the entropy package in STATA. The reweighted mortgage data match the larger representative datasets quite well. In all the empirical analysis, we use the reweighted the mortgage data to better represent home buyers in Beijing.

B Theoretical Model

This appendix presents the monocentric city model discussed in Section 3. The city is linear with a fixed population of N_R rich and N_P poor households ($N_R + N_P = N$). All households work at the urban center (CBD) at location 0, where wage income for the rich is larger $y_R > y_P$. The rest of urban space is occupied by homes with lot size normalized to 1 (roads take up no space in this model with a linear city). Land rents are remitted to absentee landlords. Housing consumption (in square meters) is provided by perfectly competitive developers facing constant returns to scale. Beyond the residential area is agricultural land with rental value p_a . The model is a closed-city model with intracity but not intercity migration. Both assumptions could be relaxed without affecting the key predictions of the model.³⁷ Since both the population and land use per household are fixed, the location of the urban boundary \bar{x} which reflects the overall city size is fixed and equal to the lot size multiplied by population.

We begin by assuming that the subway network covers the entire urban area and then relax this assumption when considering the role of public transportation infrastructure. While there could be many urban configurations, we focus on those where the city is segmented by household income and commuting modes, such as the one illustrated in Figure A11.

B.1 A Theoretical Sorting Model

Household Utility Maximization and Housing Demand Households consume two goods: a numeraire good c with a unitary price, and housing $q(x)$ which varies in quantity depending on the distance x from CBD. There are two commuting modes $m \in M = \{C, S\}$: personal vehicles C and subway S . Households maximize

³⁷Brueckner (1987) provides an analysis of a monocentric city model with a perfectly competitive supply side for both a closed and open city.

utility given income y_d , fixed (θ_m) and variable commuting cost ($w_{d,m}$) that vary by mode and household income via differences in the value of time:

$$u_d = \max_{c_{d,m}, q_{d,m}} u(c_{d,m}, q_{d,m}) \quad \text{s.t.} \quad c_d + p(x) \cdot q_d = y_d - \theta_m - w_{d,m}(x), \quad d = R, P; m \in M. \quad (\text{A2})$$

The solution to (A2), $\{c_{d,m}^*, q_{d,m}^*\}_{d=R,P; m \in M}$ determines housing demand conditional on mode choice m . We rewrite the housing demand as the bid-rent equations for each income type $d = R, P$ conditioning on the mode choice $m \in M$:

$$p_{d,m}^*(x; w) \equiv \max_{c_{d,m}, q_{d,m}} \left\{ \frac{y_d - \theta_m - w_{d,m} - c_{d,m}}{q_{d,m}} \right\}. \quad (\text{A3})$$

Travel Mode Choice Commuting is costly: the fixed cost and variable (per-kilometer) cost are θ_C and $w_{d,C}$ for personal vehicles and θ_S and variable cost $w_{d,S}$ for subway. Car commuting has higher fixed costs but lower variable costs (when there is no congestion) than subway commuting. Variable costs include time (monetized by the value of time, VOT) and pecuniary costs. Rich households have a higher VOT: $v_R > v_P$.

We ignore the potential congestion in subway transportation and assume subway commuting's variable cost is linear in distance:

$$w_{d,S}(x) = v_d / \xi \cdot x$$

where v_d is the value of time for income group d and ξ is the average subway speed. Car commuting's variable cost depends on congestion. Congestion at a given location x in turn depends on the flow of vehicles $n_C(x)$ by rich and poor households from the urban boundary to that location:

$$n_C(x) = \int_x^{\bar{x}} \mathbf{1}_R\{m = C\}(s)ds + \int_x^{\bar{x}} \mathbf{1}_P\{m = C\}(s)ds. \quad (\text{A4})$$

The indicator function $\mathbf{1}_d\{m = C\}(x)$ is equal to one if a household of type $d = R, P$ commutes via car (i.e., $m = C$) from location x . The commute cost per unit of travel distance is given by:

$$t_{d,C}(x) = v_d \mathcal{C}(n_C(x)), \quad d = R, P \quad (\text{A5})$$

The congestion function $\mathcal{C}(\cdot)$, which depends on the flow of car commuters $n_C(x)$, is measured in travel time per unit of travel distance with positive first and second derivative. Total variable car commuting costs for a household living at x is:

$$w_{d,C}(x) = \int_0^x t_{d,C}(s)ds. \quad (\text{A6})$$

Market Clearing Conditions and Spatial Equilibrium For simplification, we assume consumption of the numeraire good is fixed, and that housing quantity varies by income group d and travel mode m but is invariant of x conditioning on d, m . This allows us to focus on the effect of congestion on the bid-rent curve across

location x .

A spatial equilibrium is determined by a bid-rent function that is the envelope of willingness to pay across households, keeping the utility for each income type fixed at $\bar{u}_d, d = R, P$:

$$p^*(x) = \max_{d,m} \left\{ p(y_d - \theta_m - w_{d,m}(x), \bar{u}_d) \right\}. \quad (\text{A7})$$

Equations (A2) - (A6) make clear the simultaneous determination of housing locations and traffic congestion across the city. Solving for congestion at any location x requires knowing the distribution of car users at all points in the city. The slope of the bid-rent function for car commuters is equal to the derivative of $\frac{w_{d,C}(x)}{q_{d,m}}$ with respect to x :

$$p'_{d,C}(x) = \frac{t_{d,C}(x)}{q_{d,m}}.$$

Subway commuters do not experience congestion. The slope of their bid-rent function is constant. In residential regions with car commuting, moving toward the city center means adding additional car commuters, further increasing per kilometer commuting time costs and steepening the bid-rent function.

The market clearing conditions for a spatial equilibrium are:

1. The equilibrium bid-rent at the urban boundary $p^*(\bar{x})$ must adjust to equate to the agricultural rent p_a .
2. Given identical preferences, all households in the same income group attain the same level of utility, \bar{u}_d :
$$u(c_d^*, q_d^*) = \bar{u}_d \quad \text{for all } x \in [0, \bar{x}], \quad d = R, P.$$
3. Congestion from car commuting at each location $x \in [0, \bar{x}]$ is determined by equation (A4).
4. Bid-rent curves across income and commuting types intersect at some $x_i \in (0, \bar{x}), i = 1, \dots, 6$:

$$\begin{aligned} p_{R,S}(x_1) &= p_{P,S}(x_1), p_{R,S}(x_2) = p_{R,C}(x_2), p_{R,S}(x_3) = p_{P,C}(x_3) \\ p_{P,S}(x_4) &= p_{R,C}(x_4), p_{P,S}(x_5) = p_{P,C}(x_5), p_{R,C}(x_6) = p_{P,C}(x_6) \end{aligned}$$

5. The market clearing bid-rent function is the envelope of conditional bid-rent functions across all income and commuting groups:

$$p^*(x) = \max_{d,m} \{p_{d,m}(x)\}, \quad d = R, P, \quad m = C, S.$$

This envelope determines the pattern of residential locations as well as commuting choices that clear the housing market at each location $x \in (0, \bar{x})$. The commuting boundaries between each group are defined by the intersections of bid-rent curves across income and commuting types that determine the equilibrium bid-rent function. There must be at least 1 and at most 3 intersections for a spatial equilibrium to exist.

In equilibrium, the bid-rent function $p^*(x)$ and utility level \bar{u} are determined endogenously based on the level of commuting cost t , incomes y_d , and agricultural rent p_a . Many urban configurations are possible, though we focus on a specific one. Given sufficiently high fixed costs for driving relative to subway, high variable costs for subway relative to driving, and large enough differences in the value of time between the rich and poor, a spatial configuration as in Figure A11 may emerge where a mass of rich households live closest to the CBD and commute by subway. Beyond this group, a mass of poor households also commute by subway, followed by a mass of rich households commuting by car who consume more housing than their subway commuting counterparts ($q_{R,S} < q_{R,C}$) to compensate for longer commuting. Finally a mass of car commuting poor households live at the urban boundary given their lower value time, but also consume more housing than their subway commuting counterparts ($q_{P,S} < q_{P,C}$). Bid-rents are steeper for the rich than for the poorer for each respective commuting mode because the rich have a higher value of time.

Effect of Congestion on Bid-Rent Functions The bid-rent functions for each transportation technology evaluated at the CBD ($x = 0$) are:

$$p_{d,m}^0 = \frac{1}{q_{d,m}} [y_d - \theta_m - c_d], \quad d = R, P; m = C, S,$$

where $\frac{y_R}{q_{R,S}}$ is sufficient large compared to $\frac{y_P}{q_{P,S}}$ that $p_{R,S}^0 > p_{P,S}^0$. Similarly, the fixed costs of driving are sufficiently high that $\frac{y_R - \theta_C}{q_{R,C}} > \frac{y_P - \theta_C}{q_{P,C}}$ so $p_{R,C}^0 > p_{P,C}^0$, and both are smaller than those for subway.

The bid-rent function for subway riders is:

$$p_{d,S}(x) = \frac{1}{q_{d,S}} [y_d - \theta_S - c_d - v_d/\xi \cdot x], \quad d = R, P$$

where ξ is the average subway speed. The bid-rent function for drivers is:

$$p_{d,C}(x) = \frac{1}{q_{d,C}} \left[y_d - \theta_C - c_d - v_d \mathcal{C} \left(\int_x^{\bar{x}} \mathbf{1}_R\{m=C\}(s) ds + \int_x^{\bar{x}} \mathbf{1}_P\{m=C\}(s) ds \right) \right], \quad d = R, P,$$

where $\mathcal{C}()$ is an increasing, convex congestion function of car commuting.

Consider the effect of an exogenous increase in vehicle traffic congestion. Increases in congestion make the slope and curvature of the car commuting bid-rent functions adjust based on the extent of car commuting across the city. To make this clear, consider the changes in the slope of $p_{R,C}$ evaluated at x_B , the boundary between poor subway commuting and rich car commuting, when congestion increases from $n_C(x_B)$ to $n_C(x_B) + \varepsilon$:

$$\widetilde{p'_{R,C}}(x_B) = \frac{t_{R,C}(x_B + \varepsilon)}{q_{R,C}} = \frac{v_R}{q_{R,C}} [\mathcal{C}(n_C(x_B + \varepsilon))] > \frac{v_R}{q_{R,C}} [\mathcal{C}(n_C(x_B))] = p'_{R,C}(x_B),$$

where ε also corresponds to the mass of commuters living along that distance given the assumption of fixed lot size at 1 and $q_{R,C}$ does not vary with x as mentioned above. This demonstrates that a change in endogenous

congestion in the model has two effects: it steepens the bid-rent curve overall, and the curve itself gets steeper after passing through residential areas with car commuters as the flow of vehicles onto the roadway builds up.

B.2 Effect of Transportation Policies on Urban Structure and Welfare

Now we consider the effect of transportation policies on the urban structure to motivate our empirical analysis in the main text. While there are many potential equilibrium spatial configurations depending on model parameters, we calibrate the parameters of the model to produce the patterns depicted in Panel (a) of Figure A11.³⁸ LeRoy and Sonstelie (1983) demonstrate that the urban configuration can be explained, in part, by the relation between the income elasticity of marginal commuting costs relative to the income elasticity of housing demand. The spatial pattern in the figure is consistent with the case where the income elasticity of marginal commuting costs is larger than the income elasticity of housing demand. therefore the rich outbid the poor to live close to the city center to save commuting costs.

Congestion Pricing. First we consider a typical first-best approach: a per-kilometer congestion charge. The optimal level would be equal to the marginal external cost of congestion and can be derived from differentiating (A6) with respect to $n_C(x)$ giving the increased cost from one additional car commuter. Multiplied by the number of car commuters, this yields:

$$\tau_C(x) = \left(v_R \frac{n_{R,C}(x)}{n_C(x)} + v_P \frac{n_{P,C}(x)}{n_C(x)} \right) \mathcal{C}'(n_C(x)). \quad (\text{A8})$$

The revenue from congestion pricing can then be recycled lump sum to each resident of the city. Panel (b) of Figure A11 shows the effect of the congestion pricing on spatial equilibrium in this hypothetical city. Due to the recycling of the revenue, the bid-rent curves shift up for subway users. Under the uniform recycling of the revenue, the shift is larger for the poor than for the rich due to the smaller home size among the poor (the denominator of the intercept). The intercept of the bid-rent curves for car users moves up as well. For poor drivers, the slope of the bid-rent curve steepens as the congestion toll defined above would be larger than the savings from improved speed (due to their low VOT), hence leading to a higher travel cost per unit of distance (the numerator of the slope). For richer drivers, the bid recent curve would be flatter as the congestion toll would be smaller than the savings from improved speed (due to high VOT). However, the curve to the right of x'_B becomes steeper as it moves to x'_B from the right. This is due to the fact that congestion worsens as it is closer to x'_B from the right, leading to an increasing unit travel cost.

There are two competing forces at work that affect the spatial pattern of residential locations. First, congestion pricing increases the unit travel cost and incentivizes residents to move closer, hence bidding up home prices near the city center. Second, the reduction in congestion leads to time savings and reduces the

³⁸The bid-rent curve is convex in the standard monocentric city model where the travel cost is assumed to linear in distance: prices do not need to fall as fast as the increase in travel cost to keep residents indifferent since they are compensated by living in a larger home the further they live away from the city center. The bid-rent curve becomes concave only if the travel cost is sufficiently convex in distance. To ease exposition, we draw linear bid-rent curves in the figures.

travel cost. However, due to the differences in VOT, the time saving is more valuable to the rich than to the poor. That is, the second force is relatively stronger compared to the first force for the rich than for the poor. Given the initial spatial configuration, congestion pricing results in some poor residents shifting away from driving to subway while moving their residence from the outer ring to the inner ring. At the same time, some rich residents move out of the inner ring and switch to driving while living in larger homes. If congestion pricing or the cost of driving increases enough, the poor may occupy the city center as in the case when driving is prohibitively expensive for the poor ([LeRoy and Sonstelie, 1983](#)).

Subway Expansion. Panel (a) of Figure [A12](#) examines subway expansion by equivalently considering the effect of constraining transit to \bar{x}_B in panel (a) of Figure [A12](#). Given the fixed supply of housing, the constraint of public transit from x_B in the baseline to \bar{x}_B shifts the mass of poor subway commuters $\bar{x}_B - x_B$ to the poor car commuting region. This increases congestion to the left of x_B for all car commuters, steepening the slope of the bid-rent curves to the left of x_B more for the rich than for the poor due to the high VOT among the rich. In the extreme case of removing all subway from the city, the bid-rent curve for the rich would become even steeper as it gets closer to the city as the congestion worsens. Only the rich will occupy the city center. The results are consistent with [Glaeser et al. \(2008\)](#) which calibrates this class of models to corroborate the narrative that better public transportation leads more poor to live in the city center.

Driving Restriction. Finally, we consider the effect of raising the cost of driving via a driving restriction. In practice, the driving restriction only bans a portion of the cars from driving each day. We assume that during those days, residents need to take the subway to work. This would imply that the car commuters would need to pay the fixed cost of two modes, while the variable cost would be a weighted sum of the two modes. Panel (b) of Figure [A12](#) shows the effect of the driving restriction. Due to the increase in the fixed cost for car commuters, the intercept of the bid-rent curves shifts down, more so for the poor than for the rich (the denominator being larger for the rich). The change in the slope for the car commuters is subjects to two countervailing forces. On the one hand, the added variable cost (due to the higher variable cost when using the subway) will increase the slope. On the other hand, congestion reduction to the left of x'_B will reduce the slope, more so for the rich than for the poor. The first force likely dominants. The impact of the policy is that the rich reduce car commuting ($x'_A - x_A$) by more than the poor ($x'_C - x_C$).

Welfare. Given the stylized nature of this model, the welfare implications from this stylized example are probably less informative to real world policy applications than the empirical exercises performed in Section 7 in the main text. That said, Figures [A11](#) and [A12](#) yield an important observation: key welfare effects follow the movement up and down of equilibrium bid-rent functions from capitalization of changes in the transportation system. Put simply, an approximation to total welfare is the sum of equilibrium rents paid.³⁹

³⁹To be precise, this provides the sum of willingness to pay as reflected by equilibrium rents. As we consider the effects of different transportation policies on rents, their differences can be interpreted as an approximation to differences in welfare. The rationale for

$$\int_0^{\bar{x}} p^*(s)ds.$$

This can be visualized as the area under the equilibrium bid-rent envelope in the figures: when the envelope is lower than the baseline under a given policy scenario, the aggregate rent (and thus welfare) is lower, and visa versa. Considering Figure A12 panel (b), the effect of the driving restriction seems to lower the sum of rents by more than it is increased as the envelope is lower beyond x'_B . This accounts for a larger area than the increase in the level of the equilibrium bid-rent to the left of x'_B . In contrast, the overall effect of congestion pricing in Figure A11 panel (b) is to increase the equilibrium bid-rents envelope at nearly all points in the city because congestion reduction flattens bid-rents and the redistribution of the toll revenues increases their value for subway commuters. This points to an important general equilibrium effect of transportation policies in cities, where their benefits are capitalized into the housing market and provide additional welfare gain relative to a driving restriction beyond a partial equilibrium framework on the transportation sector as shown in Figure 3 in the main text. Our simulations suggest the welfare impacts of these capitalization effects (and subsequent sorting) may be larger than the direct effects on transportation choice itself.

C Reduced-form Evidence

This section presents reduced-form evidence of the impact of the car driving restriction policy (CDR) on the housing market. We examine the price gradient with respect to subway proximity as well as household sorting behavior.

Price Gradient w.r.t. Subway Proximity A number of confounding factors could undermine identification of the causal relationship between housing price gradients and the driving restriction policy. For example, if amenities improve over time in locations near subway stations more than in locations that are farther away from subway stations, this would result in a larger price increase for homes close to subway stations, leading to an overestimation of the true impact of the policy. On the other hand, if changes in dis-amenities such as congestion or noise follow the aforementioned pattern, we would underestimate the impact of the policy. Causal identification requires an assumption that the housing price gradient with respect to the distance to the nearest subway station would be unchanged in the absence of the driving restriction policy. We test the plausibility of this assumption by examining trends in price gradients in the periods leading up to the policy based on an event study framework:

$$\text{Price}_{jt} = \sum_{k=-24}^{24} \beta_k \times \text{Dist}_{jt} \times \mathbb{1}(t = k) + \mathbf{x}'_{jt} \gamma + \varepsilon_{jt} \quad (\text{A9})$$

this approach follows the Henry George Theorem since transportation system investments can be thought of as spatially heterogeneous public goods (Stiglitz, 1977; Arnott and Stiglitz, 1979). If transportation policies create or remove welfare reducing distortions, this will be reflected in relative changes in the bid-rent curves. Albouy and Farahani (2017) show that the value of infrastructure could be underestimated using this approach and argue for a more broad method to incorporate imperfect mobility, federal taxes, and non-traded production.

where j denotes a home and t denotes a month. The outcome variable is the unit price (¥1000 per m²). We allow the slope of the price gradient β 's to vary over time. The regression includes a flexible set of controls (\mathbf{x}_{jt}) that include neighborhood fixed effects, year by month fixed effects, city district by year fixed effects. We also control for complex-level attributes such as complex age, the floor area ratio and green space ratio, the complex land area, the number of units and buildings in the complex, and home management fee (HOA fee). Standard errors are clustered at the neighborhood level to allow for correlations among the homes in the same neighborhood (e.g., due to unobservables).

Panel (a) of Appendix Figure A9 shows the coefficient estimates of β_k that vary by quarter. There does not appear to be a pre-existing trend before the policy, alleviating the concern of the time-varying and location-specific unobservables. While there is not a clear relationship between subway proximity and housing price before the policy, there is a clear downward shift in the slope of price gradient. Moreover, the negative relationship between subway distance and housing price becomes stronger over time. The increasingly larger impact over time after the policy could be driven by the fact that the policy uncertainty is reduced over time and enforcement is tightened gradually.

One additional concern for identification is that subway expansion may result in network externalities so that the benefit of a new station occurs not just to those living or working nearby but to all who use that station. Since subway construction takes a long time with a significant lead time of public announcements, there should have been a steepening of the slope *before* the driving restriction if our results are driven by the subway expansion. To further address this issue, we include in the regressions a measure of subway density that is constructed as the inverse distance weighted number of subway stations from a given location following [Li et al. \(2019\)](#). This measure can be considered as the number of subway stations per unit area centered around a given housing unit and it increases as the subway network expands.

Appendix Table A1 provides the regression results for seven specifications. The parameter of interest is the interaction between subway distance and the policy dummy. Adding neighborhood fixed effects to control for neighborhood amenities significantly changes the main coefficient on subway distance and the interaction coefficient from Column (1) to Column (2). Further including district by year fixed effects and the rich set of complex-level variables in Columns (3) and (4) barely changes the results. Column (4) corresponds to the specification for the event study in Panel (a) of Appendix Figure A9. Column (5) is a weighted regression to make the sample more representative of the universe housing transactions in Beijing, where the weighting procedure is described in Section A.2. The last two columns include the subway density measure. The results from different specifications are qualitatively the same: the driving restriction increases the price premium for homes that are closer to subway.

Panel (b) of Appendix Figure A9 provides a falsification test by randomizing the treatment status (before or after the driving restriction policy) of each transaction while keeping the share of post-policy transactions fixed. The figure shows the histogram of the coefficient estimates of the interaction between distance and treatment dummy from 500 iterations. The estimate from the true sample (-0.075) lies far away from the histogram. This alleviates the concern that the estimated impact might be driven by unobservables.

Our analysis so far assumes a linear relationship between housing price and subway proximity. To relax this assumption, we specify a piece-wise linear relationship between housing pricing and subway distance. Appendix Figure A10 presents the price gradient estimates for three distance bands for the pre- and post-policy periods separately. The marginal impact of subway proximity is the strongest for properties within five km but tapers off after ten km, implying a convex price function. In addition, consistent with our previous analysis, the relationship between subway proximity and housing price is strengthened after the policy for all distance bands but especially for properties within five to ten km.

D Estimation Details

Estimating Travel Mode Choices To recover preference parameters in travel model choices, we use simulated maximum likelihood. The notation follows that in Section 4.2 of the main text. The log likelihood function is defined as:

$$\begin{aligned} \ln \mathcal{L}(\gamma, \eta, \theta) &= \sum_i \sum_{m=1}^{M_i} \mathbb{I}_j^i \ln R_{ijm}(\gamma, \eta, \theta) \text{ where} \\ R_{ijm}(\gamma, \eta, \theta) &= \frac{1}{H} \sum_{h=1}^H \frac{\exp(\bar{u}_{ijmh}(\gamma, \eta, \theta))}{\sum_k \exp(\bar{u}_{ijkh}(\gamma, \eta, \theta))} \end{aligned}$$

where i denotes survey respondents, j denotes home locations, m is the travel mode, and M_i includes modes available to commuter i . The indicator function \mathbb{I}_j^i takes value one if commuter i chooses travel mode m . The mode choice probability is denoted by R_{ijm} and calculated by averaging over $H = 100$ vectors of Halton simulation draws. Utility $\bar{u}_{ijmh}(\gamma, \eta, \theta)$ is similar to equation (3) but without the error term: $\bar{u}_{ijmh} = \theta_{imh} + \gamma_{ih} \cdot \text{time}_{ijm}(v_{ij}) + \gamma_2 \cdot \text{cost}_{ijm}/y_i + \mathbf{w}_{ijm}\eta$, where h (and $\{\theta_{imh}, \gamma_{ih}\}$) refers to the h th Halton draw. In the estimation, we leverage the fact that we observe the round trips for work commute and use the same simulation draw for the two trips by the same commuter to capture the mode-specific preference for each individual.

The parameter estimates $\hat{\gamma}, \hat{\eta}, \hat{\theta}$ maximize the simulated log-likelihood defined above. Once we have these estimates for the commuting preference, we plug them in the housing transaction data and calculate ease-of-commuting EV for each i and j pair based on the observed housing and job locations via:

$$EV_{ij}^t = \frac{1}{H} \sum_{h=1}^H \log \left(\sum_m \exp(\bar{u}_{ijmh}(\hat{\gamma}, \hat{\eta}, \hat{\theta})) \right)$$

and include it as a housing attribute in the estimation of housing demand below. Importantly, we construct the EV term separately for husband and wife and include both in the housing demand.

Estimating Housing Demand The housing demand model is estimated using a two-step procedure. The notation follows Section 4.4 of the main text. Nonlinear parameters θ_1 is estimated via simulated Maximum

Likelihood with a nested contraction mapping in the first step, and linear parameters θ_2 are estimated in the second step via linear IV/GMM. The log likelihood function is defined as:

$$\begin{aligned} \ln \mathcal{L}(\theta_1, \delta_j) &= \sum_i \sum_j \mathbb{I}_j^i w_i \ln P_{ij}(\theta_1, \delta_j), \text{ where} \\ P_{ij}(\theta_1, \delta_j) &= \frac{1}{Q} \sum_{q=1}^Q \frac{\exp[\mu_{ijq}(\theta_1) + \delta_j]}{\sum_k \exp[\mu_{ikq}(\theta_1) + \delta_k]}, \end{aligned}$$

The indicator function \mathbb{I}_j^i takes value one if household i chooses housing j and w_i is the weight of household i (obtained from the entropy balancing in Section A.2 to make the mortgage data more representative of home buyers in Beijing). The housing demand choice probability is denoted as P_{ij} and calculated by averaging over $Q = 200$ vectors of Halton simulation draws. Parameters θ_1 are non-linear parameters (since the log-likelihood function is a nonlinear function of these parameters) that characterize household preference heterogeneity.

We search for θ_1 and population average utilities δ_j to maximize the log-likelihood function, subject to the constraint that the model predicted housing demand based on δ_j can replicate observed housing demand (as in [Berry et al. \(1995\)](#)):

$$\begin{aligned} \max_{\theta_1, \{\delta_j\}_j} \ln \mathcal{L}(\theta_1, \delta_j) &\quad \text{s.t.} \\ \sum_{i \in C^{-1}(j)} \frac{1}{Q} \sum_{q=1}^Q w_i \frac{\exp[\mu_{ijq}(\theta_1) + \delta_j]}{\sum_k \exp[\mu_{ikq}(\theta_1) + \delta_k]} &= w_j, \forall j \in J \end{aligned} \quad (\text{A10})$$

The left-hand-side of constraint (A10) is model predicted housing demand for property j . The first summation $\sum_{i \in C^{-1}(j)}$ aggregates simulated choice probabilities over all households whose choice set contains property j . The second summation averages over $Q = 200$ vectors of Halton simulations draws to simulate household i 's probability of choosing property j . The right-hand-side is the observed housing demand for property j (which is 1 weighted by the entropy weight w_j).

We follow the literature and use the following contraction mapping to solve for $\{\delta_j\}_j$ that satisfies constraint (A10):

$$\begin{aligned} \delta_j^{d+1} &= \delta_j^d + \ln(w_j) - \ln D_j(\theta_1, \delta_j^d), \text{ where} \\ D_j(\theta_1, \delta_j^d) &= \sum_{i \in C^{-1}(j)} \frac{1}{Q} \sum_{q=1}^Q w_i \frac{\exp[\mu_{ijq}(\theta_1) + \delta_j^d]}{\sum_k \exp[\mu_{ikq}(\theta_1) + \delta_k^d]} \end{aligned}$$

where $d + 1$ is the $d + 1$ -th iteration, w_j is observed property j 's demand, and $D_j(\theta_1, \delta_j^d)$ is model-predicted demand in iteration d . Our model assumes a closed city – all households choose a place to live in Beijing – and there is no outside option, hence we normalize the population-average utility of the home with the lowest unit price to zero (results are the same) regardless of which home we use for the normalization). As the

unobserved housing attributes ξ_j that are correlated with price and EV terms are absorbed by property fixed effects δ_j , the simulated MLE produces consistent estimates on household specific parameters θ_1 .

After obtaining $\hat{\theta}_1$ and $\{\hat{\delta}_j\}_j$, we recover the linear parameters θ_2 via IV:

$$\hat{\delta}_j(\theta_2) = \alpha_1 p_j + \mathbf{x}_j \bar{\beta} + \xi_j$$

where the IVs are discussed in Section 4.4 of the main text.

E The Simulation Algorithm

Simulation Algorithm The notation in this section follows that in the main text. Household-trip characteristics $\{\mathbf{w}\}$, housing attributes $\{\mathbf{X}\}$, and commuting distance are fixed at the observed level. Demand parameters are denoted by: $\{\gamma_1, \gamma_2, \eta, \beta, \alpha, \phi, \theta, \xi\}$. We fix the random Halton draws throughout the simulation.

We now describe the algorithm for counterfactual simulations. The goal of each simulation is to find the equilibrium vector of traffic density and housing prices. The simulation algorithm has both the outer loop and the inner loop. The outer loop searches for driving speeds and traffic density that clear the traffic sector, while the inner loop searches for housing prices that clear the housing market. In counterfactual analyses that shut down sorting, there is no inner loop.

The algorithm starts with an initial vector of housing price \mathbf{p}^0 and driving speed \mathbf{v}^0 (and traffic density d^0). We use the observed levels as the starting point. Repeat the following steps until convergence:

1. Based on density level d^t and price vector \mathbf{p}^t for iteration t (d^0 and \mathbf{p}^0 for the first iteration):
 - (a) Update the driving speed for every household:

$$\frac{v_{ij}^t - v_{ij}^0}{v_{ij}^0} = e^\nu * \frac{d^t - d^0}{d^0},$$

where v_{ij}^t is updated driving speed for household i 's work commute from home j and e^ν is the speed-density elasticity -1.1.

- (b) Use v_{ij}^t from step (a) to revise each commuter's driving time: $time_{ijk,drive}^t = \frac{dist_{ijk,drive}}{v_{ijk}^t}$ (as well as the commuting time via taxi: $time_{ijk,taxi}^t$), where k denotes member k in household i ;

- (c) Update the ease-of-commuting measure EV for every household member:

$$EV_{ijk}^t = \frac{1}{H} \sum_{h=1}^H \log \left(\sum_m \exp \left(\theta_{hm} + \gamma_{1ih} time_{ijk,m}^t + \gamma_2 \frac{cost_{ijk,m}}{\text{hourly wage}_{ik}} + \mathbf{w}_{ijk,m} \eta \right) \right)$$

where h is a Halton simulation draw for the random coefficient of travel time and the random coefficients of each travel mode, m stands for travel mode, and θ_{hm} and γ_{1ih} are simulated random coefficients for travel time and mode dummies.

In a similar manner, update member k 's driving probability:

$$R_{ijk,driving}^t = \frac{1}{H} \sum_{h=1}^H \frac{\exp\left(\theta_{h,drive} + \gamma_{1ih} time_{ijk,drive}^t + \gamma_2 \frac{\text{cost}_{ijk,drive}}{\text{hourly wage}_{ik}} + \mathbf{w}_{ijk,drive} \eta\right)}{\sum_m \exp\left(\theta_{hm} + \gamma_{1ih} time_{ijk,m}^t + \gamma_2 \frac{\text{cost}_{ijk,m}}{\text{hourly wage}_{ik}} + \mathbf{w}_{ijk,m} \eta\right)}$$

If the counterfactual analysis incorporates sorting, continue with steps (d)-(e). Otherwise skip them and move to step (f);

(d) Given the updated EV term, search for a new housing price vector \mathbf{p}^t that clears the housing market with housing demand equal to housing supply:

$$\sum_{i \in C^{-1}(j)} \frac{1}{Q} \sum_{q=1}^Q w_i \frac{\exp\left(\alpha_i p_j^t + \mathbf{X}_j \beta_i + \sum_k \phi_{kq} EV_{ijk}^t + \xi_j\right)}{\sum_{s \in C(i)} \exp\left(\alpha_i p_s^t + \mathbf{X}_s \beta_i + \sum_k \phi_{kq} EV_{isk}^t + \xi_s\right)} = S_j, \forall j \in J \quad (\text{A11})$$

where the left-hand-side (LHS) is the simulated demand for property j and the right-hand-side is the housing supply S_j (under fixed housing supply, S_j is fixed at w_j , the weight for property j). The first summation of the LHS is over households whose choice set includes property j , denoted as $C^{-1}(j)$. The second summation aggregates over $Q = 200$ Halton draws of random coefficients. Each household has weight w_i to make the sample more representative of home buyers in Beijing. The coefficients ϕ_{kq} are member k 's simulated random coefficients for the EV term.

As the model is a closed city with no outside options, the market clearing condition pins down the housing price vector \mathbf{p}^t up to a constant. We normalize the average housing price (the mean of vector \mathbf{p}^t) to be the same as the average price observed in the sample.

For counterfactual analyses that allow the housing supply to respond to housing price changes, we use a constant supply-price elasticity at 0.53 following [Wang et al. \(2012\)](#):

$$\frac{S_j^t - w_j}{w_j} = 0.53 * \left(\frac{p_j^t - p_j^0}{p_j^0} \right)$$

Solving equilibrium price vector \mathbf{p}^t requires iterating the market clearing condition (A11) many times. This is the inner loop as illustrated below.

- (i) Suppose that the price vector is \mathbf{p}^l in iteration l ($l = 1$ for the first iteration);
- (ii) Update the price vector:

$$\mathbf{p}_j^{l+1} = \mathbf{p}_j^l + [\log(S_j) - \log(D_j^l(\mathbf{p}_j^l))] / k$$

where \mathbf{p}^{l+1} is the updated price vector, S_j is the observed housing supply, $D_j^l(\mathbf{p}_j^l)$ is the model predicted demand in iteration l (the LHS of equation (A11)), and k is a pre-set constant that controls the step size of each iteration.

- (iii) If $\|\mathbf{p}^{l+1} - \mathbf{p}^l\| < \varepsilon_{tol}^{\text{inner}}$ where ε_{tol} is a pre-set tolerance level, stop. Otherwise, return to step (ii). In

our simulation, this algorithm always converges. Let $\mathbf{p}^t = \mathbf{p}^{t+1}$, the fixed point from the inner loop.

(e) At the new equilibrium housing price \mathbf{p}^t , revise the housing demand choice probability:

$$P_{ij}^t = \frac{1}{Q} \sum_{q=1}^Q \frac{\exp(\alpha_i p_j^t + \mathbf{X}_j \beta_i + \sum_k \phi_{kq} E V_{ijk}^t + \xi_j)}{\sum_{s \in C(i)} \exp(\alpha_i p_s^t + \mathbf{X}_s \beta_i + \sum_k \phi_{kq} E V_{isk}^t + \xi_s)}$$

(f) Update the traffic density using the revised probability to drive and take taxis (and the revised probability that household i chooses property j in the case of sorting):

$$\tilde{d} = \sum_i w_i \sum_j P_{ij}^t \left[\sum_k (R_{ijk,drive}^t \times dist_{ijk,drive} + R_{ijk,taxi}^t \times dist_{ijk,taxi}) \right]$$

where the summation inside the large brackets $[\cdot]$ aggregates over commuting member k within each household. If sorting is shut down, the housing demand choice probability P_{ij}^t is fixed at initial levels.

2. If $\|\tilde{d} - d^t\| < \varepsilon_{tol}^{\text{outer}}$ where $\varepsilon_{tol}^{\text{outer}}$ is a pre-set tolerance level, stop. Otherwise, set $d^{t+1} = \varphi d^t + (1 - \varphi) \tilde{d}$ for some $\varphi \in (0, 1)$ and return to step 1.

Ring-specific Density The discussion above assumes that the traffic density is city-wide. To accommodate ring-specific traffic density, we start with apportioning each trip to the relevant ring road segment according to the radius of each ring. For example, if the radius of ring road segment between the k th and $k+1$ th ring is $r_{k,k+1}$, a trip that originates in ring 5 and ends in ring 3 will contribute $\frac{r_{3,4}}{r_{3,4} + r_{4,5}}$ to the 3rd-4th ring road traffic density and $\frac{r_{4,5}}{r_{3,4} + r_{4,5}}$ to the 4th-5th ring road traffic density. Accordingly, we break down the driving distance of each trip $dist_{ijk,drive}$ to the ring-specific driving distance $\{dist_{ijk,s,drive}\}_s$, where s denotes a ring-road segment.

The simulation algorithm is the same as above, except that we now have ring-road specific traffic speed v_{ijs}^t and ring-road specific traffic density d_s . Specifically, in step (a) above, we update the ring-road specific speed v_{ijs}^t via:

$$\frac{v_{ijs}^t - v_{ijs}^0}{v_{ijs}^0} = e^\nu * \frac{d_s^t - d_s^0}{d_s^0}.$$

In step (b) above, Use v_{ijs}^t from step (a) to revise each member k 's driving time spent in ring-road segment s . The total driving time is the sum of the ring-road specific driving time: $time_{ijk,drive}^t = \sum_s \frac{dist_{ijk,s,drive}}{v_{ijs}^t}$ (as well as the commuting time via taxi: $time_{ijk,taxi}^t$). In step (f) above, we aggregate traffic demand to ring-road specific density measures as follows,

$$\tilde{d}_s = \sum_i \sum_j P_{ij}^t \left[\sum_k (R_{ijk,drive}^t \times dist_{ijk,s,drive} + R_{ijk,taxi}^t \times dist_{ijk,s,taxi}) \right]$$

The rest of the simulation steps remain the same.

Assumptions for the Welfare Analysis We make some assumptions in the welfare analysis. First, we use a 30-year time horizon for our welfare analysis. The assumption on the time horizon affects the welfare magnitude (a shorter period is associated with lower welfare gains), but does not affect the comparison across different transportation policies. Second, to be conservative, we only include commuting trips (which account for 60% of all trips and 75% of total travel distance in 2014) and ignore non-commuting trips in calculating the benefits of subway expansion. Lastly, we abstract away from distortionary taxes and assume instead that subways' construction costs are financed via a non-distortionary uniform head tax. Similarly, when congestion toll revenues are recycled, they are distributed evenly across households in a lump sum manner.

To facilitate comparison, we calibrate the congestion charge at ¥1.13/km to achieve the same level of congestion reduction as under driving restriction with the 2008 subway network. We include (lifetime) subway and bus fares as part of the government revenue in the counterfactual analyses, though they account for a negligible fraction of the total welfare. For example, the subway fare is ¥2 per ride. Total lifetime subway fares paid by a household is roughly ¥2,400 ($\text{¥2 per ride} * 8\% \text{ likelihood to take subway} * 500 \text{ rides/year} * 30 \text{ years}$), much smaller than the net welfare per household as reported in the main text. Lastly, consumer surplus reported in the counterfactual analyses refers to the discounted lifetime consumer surplus over the property tenure, as discussed in Section 5.2.

We account for both the capital and operating costs of the congestion pricing system based on the system in Singapore. Singapore's electronic road pricing scheme launched in 1998 costed \$110 million to set up with an annual operating cost of \$18.5 million.⁴⁰ Singapore is upgrading the system to be satellite-based and the setup cost would be about \$370 million but the operating cost is expected to be lower than the current system. We assume that the congestion pricing system in Beijing would be satellite-based and we scale both the setup and operating costs up by the population of Beijing relative to that of Singapore. To facilitate policy comparison, we calculate the 30-year total discounted cost per household (assuming 7.2 mil. households in Beijing). This amounts to ¥3000, about 2.5% of the total toll revenue per household during the same period.

We include the construction and operating costs of the new subway lines built between 2008 and 2014. The subway construction cost is ¥245.23 billion (Li et al., 2019), implying about ¥34,000 per household. Based on an annual operating cost per household per year of ¥1246⁴¹, the total operating cost during a 30-year period is ¥69,000 per household. Together, the discounted construction and operating cost amounts to ¥103,000 per household during a 30-year period in the welfare analyses.

⁴⁰<https://www.zdnet.com/article/singapore-readies-satellite-road-toll-system-for-2021-rollout/> and https://nyc.streetsblog.org/wp-content/uploads/2018/01/TSTC_A_Way_Forward_CReport_1.4.18_medium.pdf.

⁴¹See <http://fgw.beijing.gov.cn/zb/ggjt/gg/201410/P020141013443109115783.pdf>.

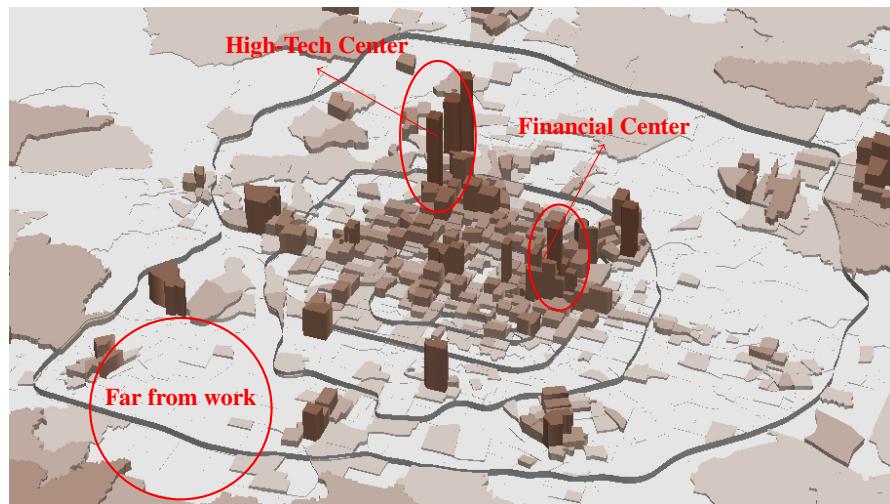
F Figures and Tables

Figure A1: Subway Network Expansion in Beijing



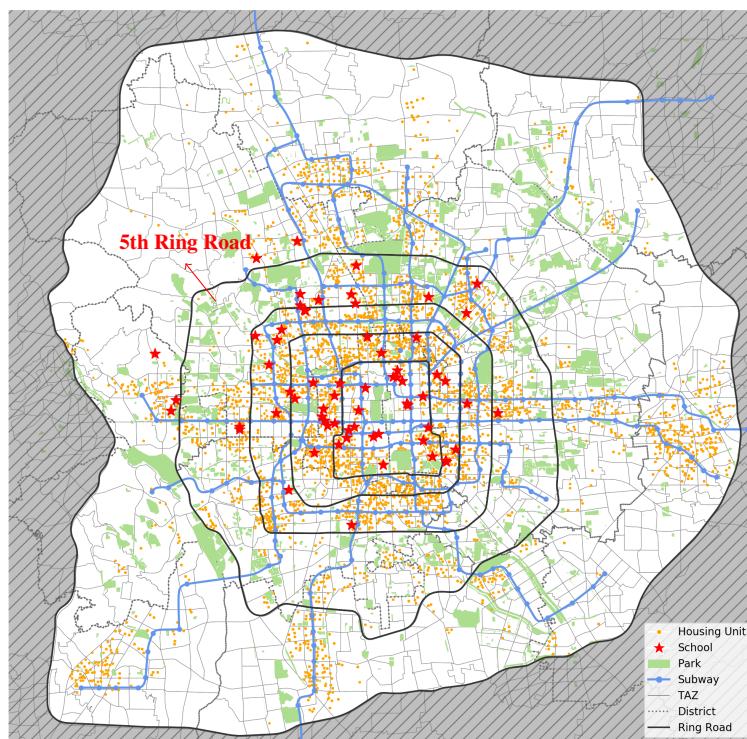
Note: Subway expansion from 1999 to 2019 in Beijing expanded from 2 lines to 22 lines. From 2007 to 2018, 16 new subway lines were built with a combined length of over 500km. By the end of 2019, the Beijing Subway is the world's longest and busiest subway system with a total length of nearly 700km and daily ridership over 10 million.

Figure A2: Job Density



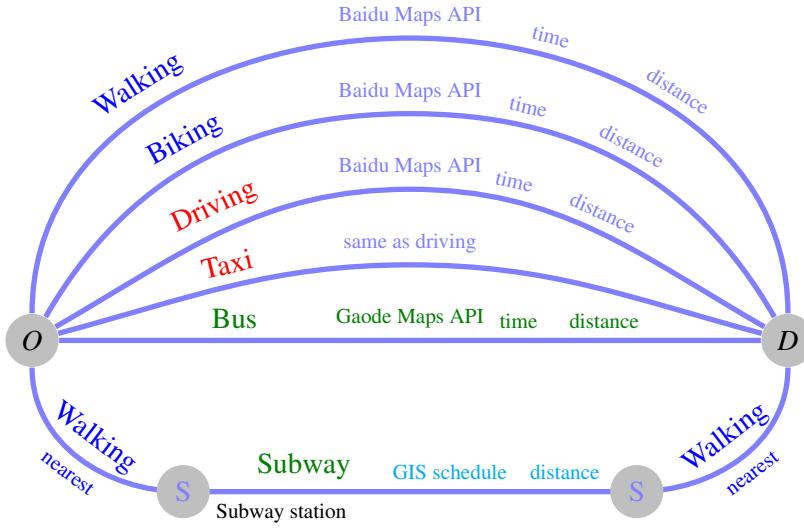
Note: This figure plots work density by TAZ based on work locations from the mortgage data. Darker colors/taller shapes indicate greater work density.

Figure A3: Housing, Amenities, and Transportation Network



Note: The figure shows the home locations in the mortgage data overlaid with ring roads (black lines), subway lines in blue (as of 2015), government-designated key schools (red stars), and government-designated parks (green area). The outermost black line traces out the 6th ring road.

Figure A4: Construction of the Travel Choice Set



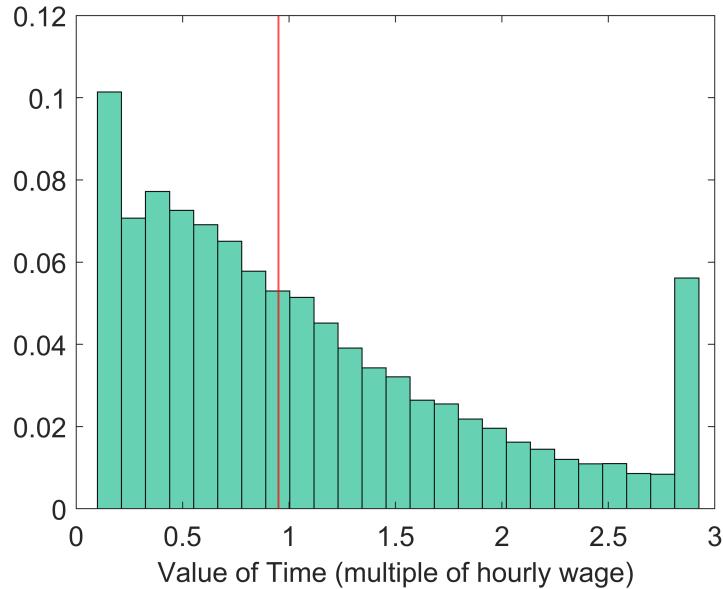
Note: Travel time and distance using walk, bike, car, and taxi are constructed using the Baidu Maps API based on the survey reported departure time and day of the week. Bus travel time and distance are constructed using Gaode Maps API because it provides the number of transfers and walking time between bus stops. The travel time for bus, car and taxi are adjusted based on the historical traffic congestion condition in the survey month-year. For trips via subway, the walking time and distance to and from subway stations are provided by Baidu Maps API. Subway transit distance and time from the origin subway station to the destination station are calculated using GIS based on the historical subway network and subway time tables in 2010 and 2014.

Figure A5: An Example of Travel Routes



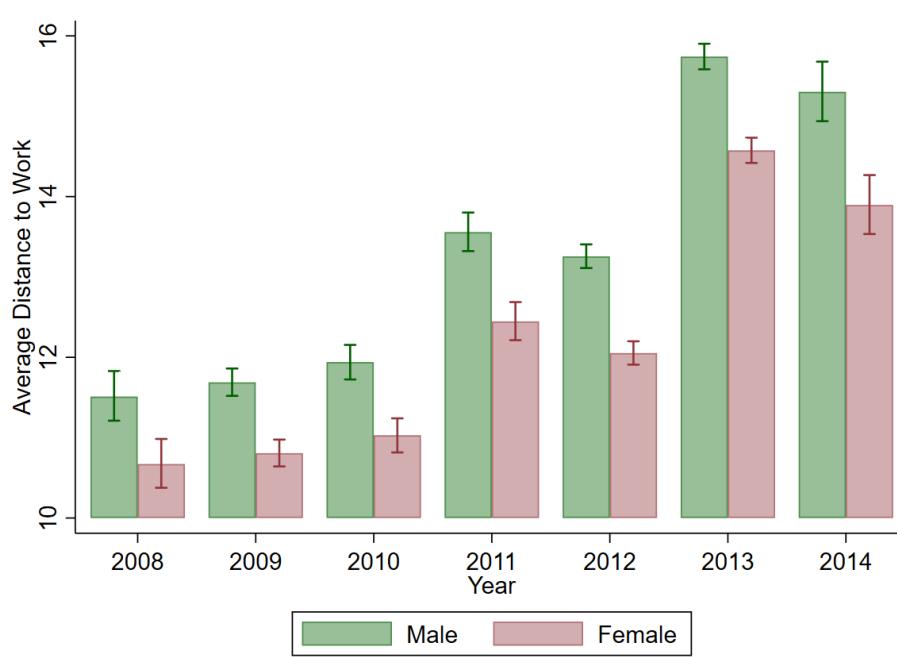
Note: the figure shows travel time and cost of six routes for a particular trip that started at 7:09am on 9/12/2010. The chosen mode was subway. The left panel shows the the straight-line direction of travel, while the right panel shows the time, monetary cost, and distance for each travel mode and the corresponding route constructed by Baidu API, Gaode API and GIS.

Figure A6: Implied Value of Time Distribution from the Mode Choice Estimation



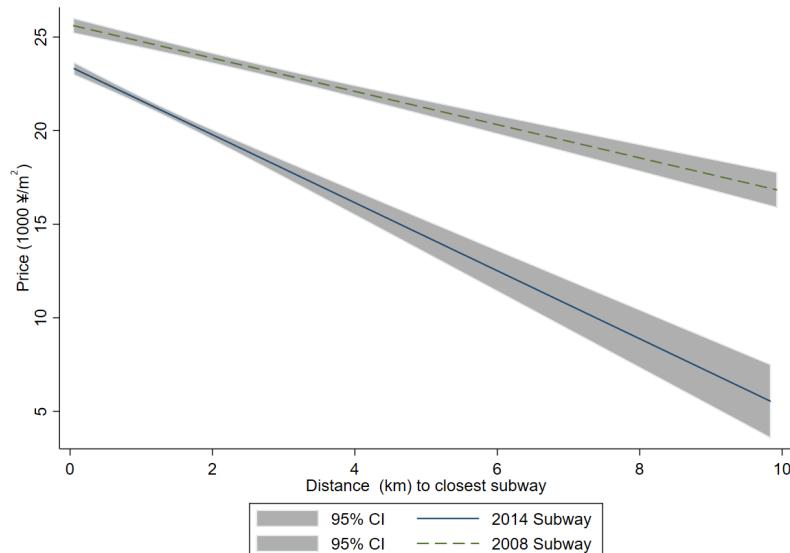
Note: The figure plots the estimated distribution of value of time (VOT) in terms of hourly wage that is based on Column (6) of Table 3. VOT is measured by the ratio of the preference on travel time over the preference on monetary travel cost. The former one has a winsorized (at the 5th and 95th percentile) chi-square distribution with three degrees of freedom, while the latter one is inversely related to income. The red line shows the average VOT (95.6% of the hourly wage). The median VOT is 84.6% of the hourly wage.

Figure A7: Distance to Work by Gender in km



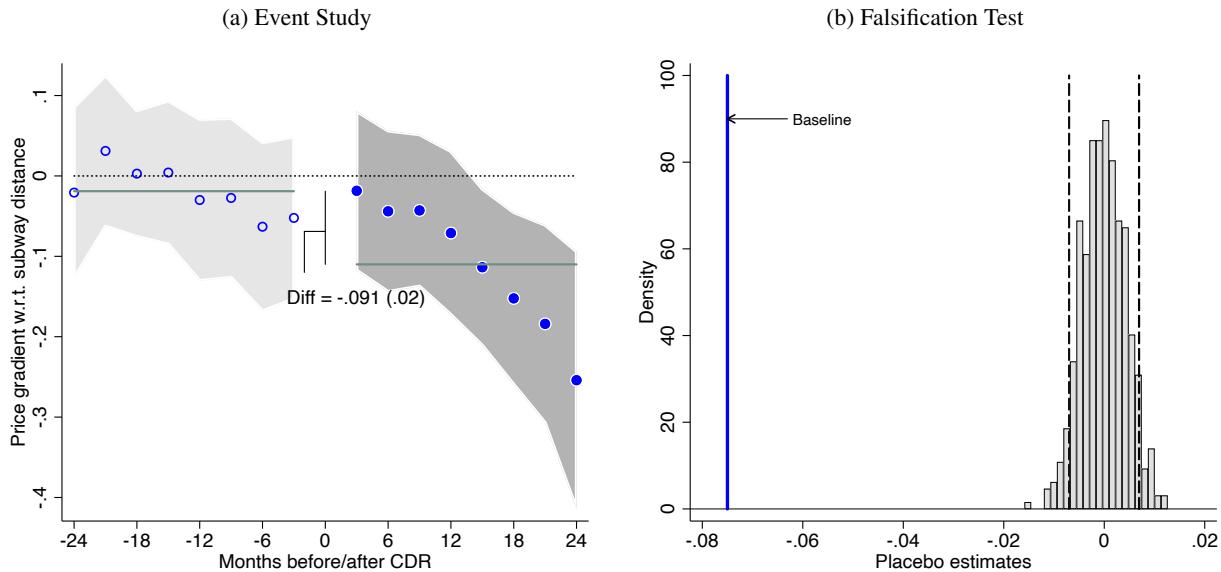
Note: the figure displays the average distance to work by year for male (green bars) and female (red bars) household members, separately. The whiskers denote 95% intervals. Males have longer commutes than females. The increasing commuting distance over time reflects the expansion of the city and its transportation infrastructure.

Figure A8: Price Gradient under the 2008 and 2014 Subway Network



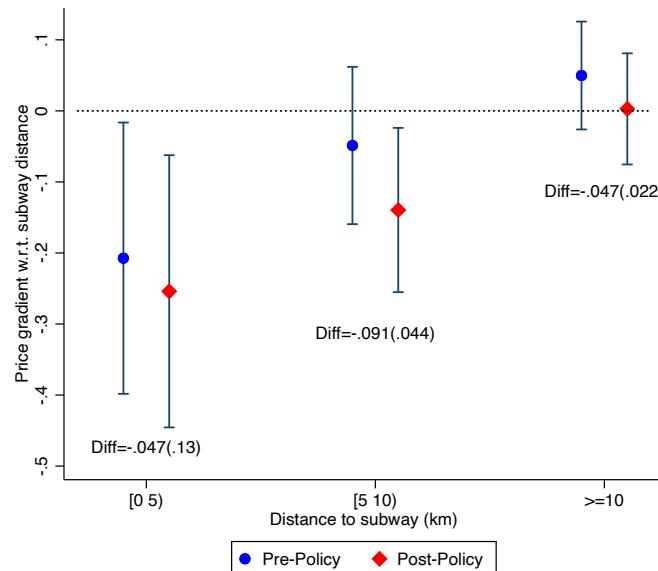
Note: This plot shows the simulated bid-rent curve with respect to subway distance under the 2008 and 2014 network, respectively. The results corresponds to Columns (1) and (4) of Table 6. The gradient of the bid-rent curve under the 2014 subway system (-¥1900/m² per km) is steeper than the 2008 subway system (-¥700/m² per km), reflecting households' higher WTP for proximity to subway stations when the subway system is more desirable. The bid-rent shifts down under the 2014 subway system that reaches to cheaper homes farther away from the city center.

Figure A9: Event Study and Falsification Test



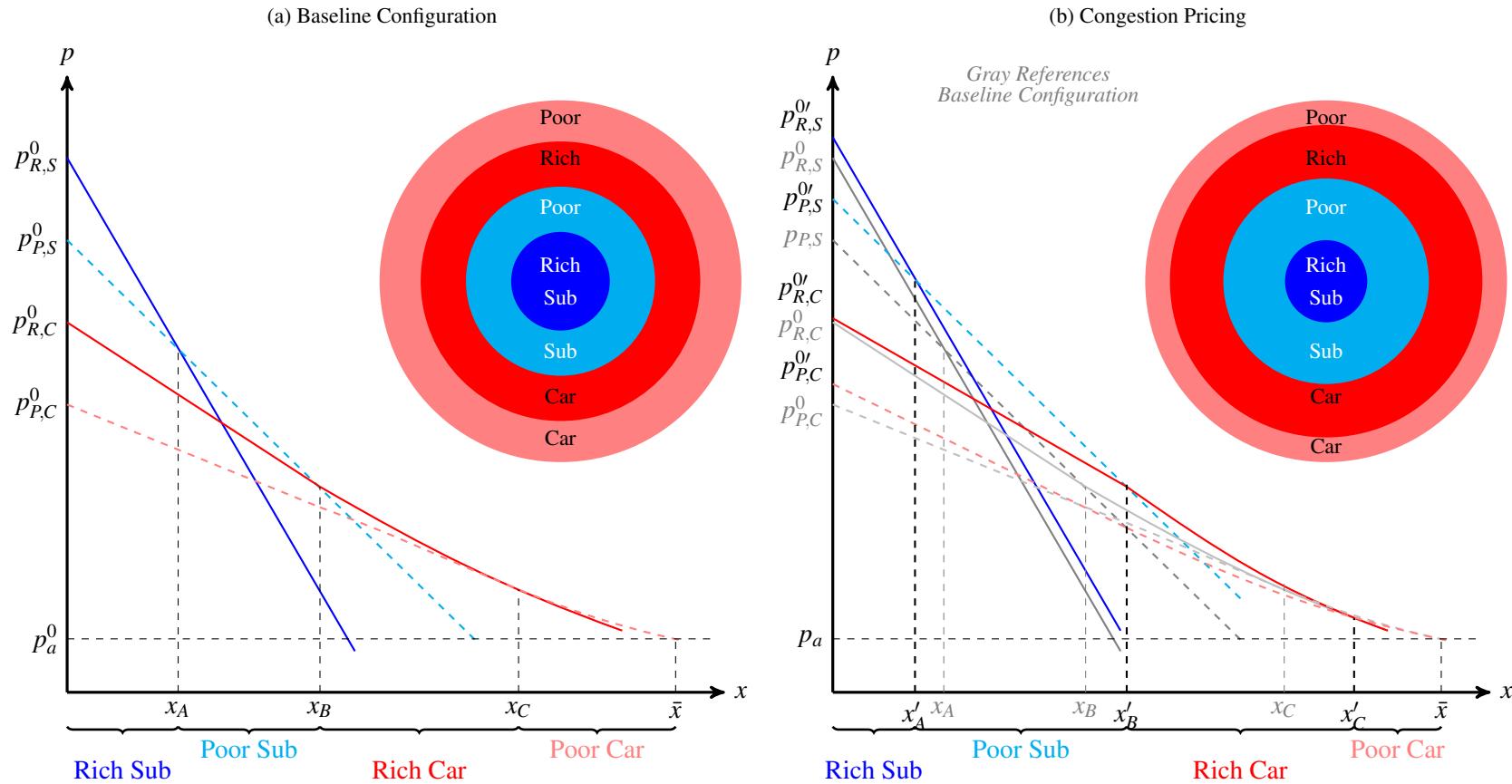
Note: Panel (a) shows estimates from the event study of Beijing's driving restriction on the price gradient w.r.t the subway distance. The blue dots report the estimates by quarter relative to the start of the driving restriction in July 2008. Panel (b) reports the distribution of 500 coefficient estimates from placebo tests with randomized event time. Dashed vertical lines indicate the 95% confidence interval, while the solid vertical line indicates the baseline estimate from Column (4) in Table A1. The regressions in both panels include neighborhood fixed effects, year by month fixed effects, city district by year fixed effects, and complex characteristics.

Figure A10: Price Gradient by Distance Bands



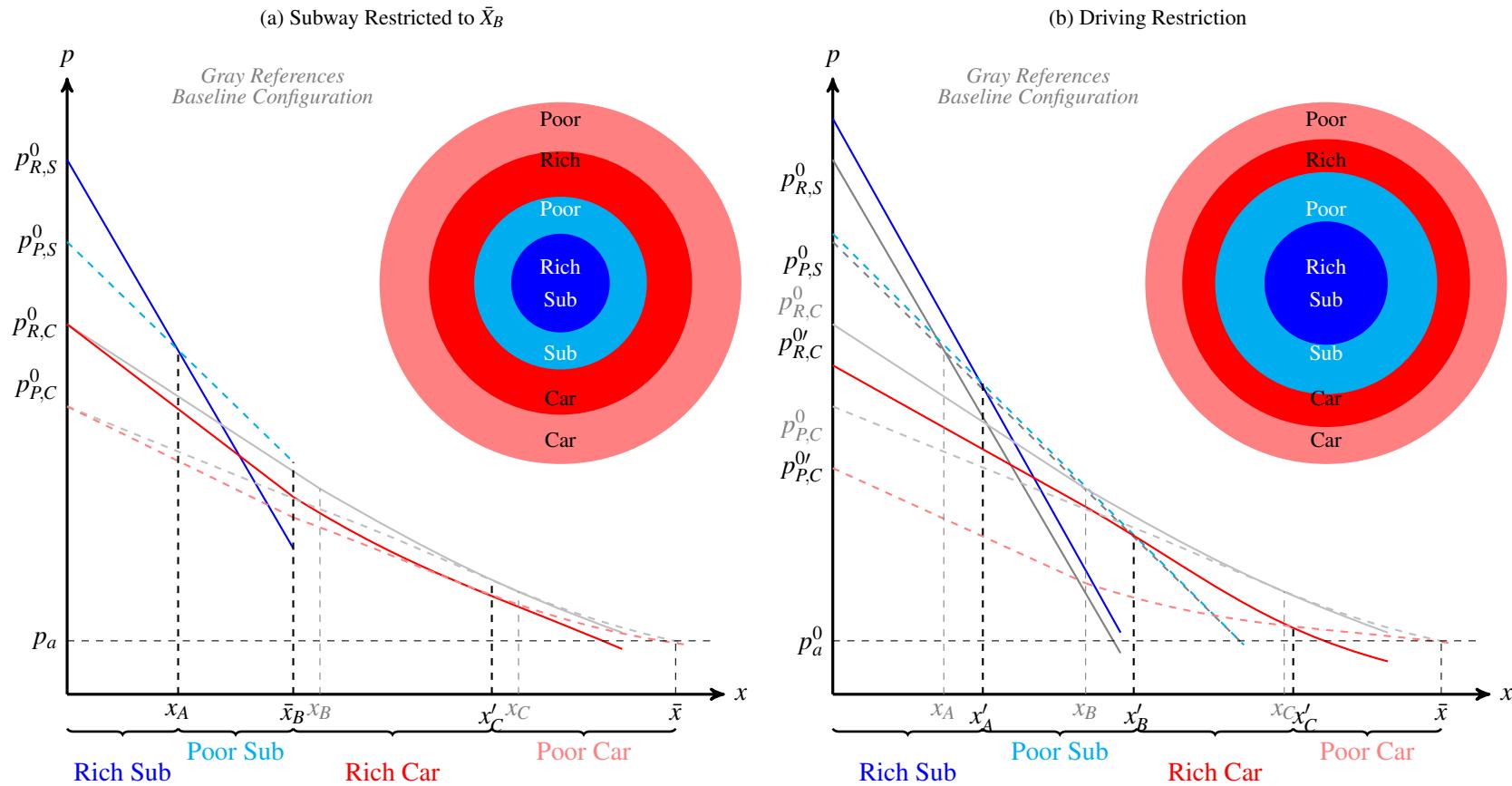
Note: This figure shows the price gradient estimates and their 95% confidence intervals by subway distance bands for the pre-policy (blue dots) and post-policy (red diamond) periods, respectively. It also displays the differences between the pre- and post-policy price gradient estimates and the standard errors of the differences. The controls include neighborhood fixed effects, year by month fixed effects, city district by year fixed effects, and complex characteristics. The standard errors are clustered at the neighborhood level.

Figure A11: Spatial Equilibrium with Two Modes & Income Heterogeneity: Baseline & Congestion Pricing



Note: The x-axis denotes the distance to city center and the y-axis is the rental price. Concentric circles show the equilibrium pattern of housing location where the linear city is drawn as radially symmetric. Commuters are in two income groups (Rich, R , denoted by solid lines, and Poor P , denoted by dashed lines) and choose from two modes: car (C , in red) and subway (S , in blue). The model is calibrated to generate an urban configuration as indicated. Bid-rent curves are steepest for the rich because of their higher value of time. Bid-rent curves are flatter for car commuting because the variable time cost is assumed to be smaller. Values indicated by $p_{d,m}^0$ are the value of bid-rent curves at the origin, while those with a $'$. Gray curves, dashed lines and text in panel (b) reference the baseline configuration in panel (a). Panel (b) shows the effect of congestion pricing in a per-kilometer basis. It increases the y-intercept through lump-sum remittance of toll revenue back to all households. For car commuting, there are offsetting effects: congestion is lower with the toll because there are fewer car commuters, but longer commutes face larger total tolls, making bid-rent curves steeper.

Figure A12: Spatial Equilibrium with Two Modes & Income Heterogeneity: Subway & Driving Restriction



Note: The x-axis denotes the distance to city center and the y-axis is the rental price. Concentric circles show the equilibrium pattern of housing location where the linear city is drawn as radially symmetric. Commuters are in two income groups (Rich, R , indicated with solid lines, and Poor P , indicated with dashed lines) and choose from two modes: car (C , in red) and subway (S , in blue). The model is calibrated to generate an urban configuration as indicated. Bid-rent curves are steepest for the rich because of their higher value of time. Bid-rent curves are flatter for car commuting because the variable time cost is assumed to be smaller. Values indicated by $p_{d,m}^0$ are the value of bid-rent curves at the origin, while $p_{d,m}^{0'}$ are the new values at the origin after the policy change. Gray curves reference the baseline configuration in panel (a) of Figure A11. Panel (a) shows the effect of restricting the subway network to \bar{x} . Comparing this panel to panel (a) of Figure A11 shows the effect of expanding the subway network from \bar{x}_B to x_B . There is no effect of expansion on the bid-rents for subway except that they do not extend beyond \bar{x}_B in panel (a) of Figure A11 because the subway has not been built that far. Because expansion allows a larger share of commuters to use the subway, here only from the poor, it induces lower congestion, flattening out the bid-rent curves for driving, for the rich more than the poor because of value of time differences. Panel (b) considers a driving restriction policy, which induces increases in both the fixed and variable costs of driving not remitted to households (unlike congestion pricing). Car commuters will need to incur the fixed cost of both driving and using subway, and when they cannot drive, they will have to use subway which has a higher variable cost due to time. The bid-rent curves for car commuters move down and become steeper because of the change in fixed cost and variable cost of commuting, respectively. The increase in commuting costs is larger for the rich due to their high VOT than for the poor, leading to a larger movement of the rich away from driving to subway.

Table A1: Effect of Driving Restriction on Price Gradient w.r.t. Subway Proximity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Subway Distance (km)	-0.199 (0.021)	0.004 (0.052)	0.003 (0.050)	0.001 (0.048)	0.005 (0.076)	0.021 (0.047)	0.017 (0.078)
Subway Distance (km) \times CDR	-0.125 (0.030)	-0.080 (0.019)	-0.081 (0.019)	-0.075 (0.019)	-0.126 (0.033)	-0.063 (0.019)	-0.118 (0.033)
Subway Density						0.019 (0.007)	0.011 (0.012)
Year x Month FE	Y	Y	Y	Y	Y	Y	Y
Neighborhood FE	N	Y	Y	Y	Y	Y	Y
District x Year FE	N	N	Y	Y	Y	Y	Y
Complex-level Controls	N	N	N	Y	Y	Y	Y
Weighted	N	N	N	N	Y	N	Y
Adjusted R^2	0.248	0.495	0.503	0.520	0.773	0.522	0.774

Note: The dependent variable is housing price (¥1000/m²), with a sample mean of ¥10,078/m². The sample spans 24 months before and after the start of car driving restriction policy (CDR) in July 2008. We remove observations from July to September 2008. The policy was more aggressive during this period whereby half of the vehicles were restricted on a given weekday (due to the 2008 Olympics in August), and changed to one weekday per week from October 2008 onward. The number of observations is 23,917. Subway distance is the distance (in km) from the housing unit to the nearest subway station. Subway density is constructed at the TAZ level as the inverse distance weighted number of subway stations from the centroid of an TAZ. There are 150 neighborhoods, 16 district, and 2804 complex in the sample. The complex-level attributes include complex age, the floor area ratio, the green space ratio, the land area of the complex, home management fee (HOA fee), and the number of building and apartment units in the complex. Weights are constructed via entropy balancing in Appendix Section A.2.

Table A2: Housing Demand Without EV Terms - Linear Parameters

Variables	OLS (1)	OLS (2)	IV1 (3)	IV2 (4)	IV2+IV3 (5)	All (6)
Price (¥mill.)	-2.073 (0.180)	-2.062 (0.176)	-7.101 (1.646)	-4.224 (0.535)	-5.031 (0.432)	-5.356 (0.418)
Ln(home size)	-3.590 (0.248)	-3.657 (0.251)	5.102 (2.937)	0.104 (0.932)	1.512 (0.761)	2.079 (0.765)
Building age	-0.032 (0.006)	-0.026 (0.006)	-0.144 (0.040)	-0.076 (0.012)	-0.095 (0.010)	-0.103 (0.010)
Floor area ratio	0.017 (0.031)	0.001 (0.023)	-0.009 (0.036)	-0.007 (0.021)	-0.009 (0.025)	-0.009 (0.027)
Ln(dist. to park)	0.167 (0.059)	0.052 (0.054)	-0.513 (0.225)	-0.196 (0.073)	-0.285 (0.075)	-0.321 (0.081)
Ln(dist. to key school)	0.631 (0.060)	0.555 (0.091)	-0.034 (0.213)	0.312 (0.086)	0.223 (0.089)	0.187 (0.091)
Year-Month-District FE	Y	Y	Y	Y	Y	Y
Neighborhood FE		Y	Y	Y	Y	Y
First-stage Kleinberg-Paap F			9.9	10.5	14.2	14.2
Avg. Housing Demand Price elasticity	3.09	3.10	-1.94	0.94	0.13	-0.19

Note: The number of observations is 79,894. The dependent variable is the recovered population-average utilities $\{\delta_j\}_j$ when EV is excluded from housing attributes. The first two columns and the last four present OLS and IV estimates, respectively. The floor area ratio 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. Column (3) use IV1 as price instruments, i.e. the number of homes that are within 3km from a given home, outside the same complex, and sold in a two-month window. Columns (4), (5), and (6) use IV2, i.e. the average attributes of these homes (building size, age, log distance to park, and log distance to key school). Column (5) also includes IV3, the interaction between IV2 and the winning odds of the licence lottery. The winning odds decreased from 9.4% in January 2011 to 0.7% by the end of 2014. Column (6) uses all IVs. Standard errors are clustered at the neighborhood level.

Table A3: Housing Demand - Nonlinear Parameters with Alternative Sampling

	0.5% Sample		1% Sample	
	(1) Para	SE	(2) Para	SE
Demographic Interactions				
Price (¥mill.) * ln(income)	1.153	0.018	1.030	0.016
Age in 30-45 * ln(distance to key school)	-0.459	0.011	-0.420	0.010
Age > 45 * ln(distance to key school)	-0.122	0.024	-0.123	0.021
Age in 30-45 * ln(home size)	1.681	0.034	1.486	0.029
Age > 45*ln(home size)	3.011	0.070	2.746	0.061
EV _{Male}	0.831	0.064	0.755	0.006
EV _{Female}	0.976	0.069	0.893	0.006
Random Coefficients				
$\sigma(EV_{Male})$	0.333	0.154	0.379	0.013
$\sigma(EV_{Female})$	0.408	0.142	0.482	0.012
Log-likelihood	-128,976		-168,808	

Note: The table replicates Table 4 and reports MLE estimates of housing demand's non-linear parameters using mortgage data from 2006-2014 with 77,696 observations. Column (1) constructs households' choice set using a 0.5% random sample of all houses sold during a two-month window around the purchase date of the chosen home. Column (2) reproduces our preferred specification in Column (3) of Table 4 that uses a 1% random sample.

Table A4: Housing Demand - Linear Parameters with Alternative Sampling

Variables	0.5% Sample	1% Sample
	(1)	(2)
Price (in 1 million RMB)	-7.417 (0.590)	-6.596 (0.534)
Ln(property size)	4.355 (1.073)	3.879 (0.969)
Building age	-0.139 (0.014)	-0.132 (0.013)
Complex FAR	-0.019 (0.037)	-0.023 (0.034)
Ln(dist. to park)	-0.442 (0.114)	-0.424 (0.103)
Ln(dist. to key school)	0.321 (0.128)	0.288 (0.118)
Year-Month-District FE	Y	Y
Neighborhood FE	Y	Y
First-stage Kleinberg-Paap F	14.22	14.22
Avg. Price elasticity	-1.64	-1.44
Avg. Price elasticity CI	[-2.83, -0.44]	[-2.52, -0.35]

Note: this table replicates Table 5 and reports IV estimates of linear parameters in housing demand. Column (1) constructs households' choice set using a 0.5% random sample of all houses sold during a two-month window around the purchase date of the chosen home. It uses all three sets of price IVs. Column (2) reproduces our preferred specification in Column (6) of Table 5 that uses a 1% random sample. Standard errors clustered at the neighborhood level.

Table A5: The Speed Traffic Density Elasticity Estimate

Region	(1) 2-3 Ring Roads	(2) 3-4 Ring Roads	(3) 4-5 Ring Roads	(4) 5-6 Ring Roads	(5) All
Log of Density (IV)	-1.250 (0.148)	-1.185 (0.111)	-1.287 (0.417)	NA	-1.099 (0.089)
Log of Density (OLS)	-0.583 (0.065)	-0.645 (0.046)	-0.362 (0.043)	-0.542 (0.048)	-0.554 (0.027)
Observations	45,152	49,351	29,241	32,926	156,670
Average speed (km/h)	28.00	30.39	32.86	31.20	30.3

Note: This table presents 2SLS results on the speed-density relationship by ring-road segments (e.g., between the 2nd and 3rd ring roads). The segment within the 2nd ring road is omitted due to the lack of observation. The dependent variable is $\ln(\text{speed in km/h})$ and the key explanatory variable is $\log(\text{traffic density in the number of cars/lane-km})$. The IVs are based on the driving restriction policy which has a preset rotation schedule using the last digit of the license plate number. They include a policy indicator for days when vehicles with a license number ending 4 or 9 are restricted from driving and interactions between this variable and hour-of-day dummies. Our sample consists of road segment by hour during peak hours within the 6th ring road in 2014. We focus on the top quintile observations with traffic density larger than 35 cars per lane-km. The average speed for these observations is 30km/h, close to the city-wide average speed during peak hours and more relevant for our analysis on commuting trips. The control variables include temperature (C°), wind speed (km/h), visibility (km), dummies for wind directions and sky coverage at the hourly level. The time and spatial fixed effects include day-of-week, month-of-year, hour-of-day, holiday, and monitoring stations fixed effects. Parentheses contain standard errors clustered by road segments. Significance: * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

Table A6: Model Prediction on Changes in Housing Price Gradient due to Driving Restriction

	(1)	(2)	(3)
Subway Distance	-0.725 (0.066)	-0.302 (0.159)	
Subway Distance \times CDR	-0.034 (0.001)	-0.034 (0.001)	-0.034 (0.001)
Neighborhood FE	N	Y	N
home FE	N	N	Y
Adjusted R^2	0.329	0.400	0.999

Note: The analysis is based on the mortgage data's 2014 cohort with 7,136 observations. Using the parameter estimates from the sorting model, we simulate the equilibrium housing prices under the 2008 network for two scenarios: with and without the car driving restriction (CDR). We then regress the simulated housing prices in ¥1,000/ m^2 on subway distance (in km) as in Table A1. Subway distance is the observed distance from the housing unit to the nearest subway station based on the 2008 subway network. The driving restriction steepens the price gradient with respect to subway access, consistent with results in Table A1 based on observed data. The magnitude is somewhat smaller, because the reduced-form result reflects a short-run response while the structural simulation incorporates long-run equilibrium adjustments (especially the rebound effects). The two samples are also different. The reduced-form analysis uses two years before and after CDR's initial implementation date while the structural analysis uses the 2014 cohort. Standard errors clustered at the neighborhood level.

Table A7: Simulation Results with Household Sorting and Variable Housing-Supply

Income relative to the median	2008 Subway Network						2014 Subway Network					
	(1) No Policy		(2) Driving restriction		(3) Congestion pricing		(4) Subway Expansion		(5) + Driving restriction		(6) + Congestion pricing	
			Δs from (1)		Δs from (1)		Δs from (1)		Δs from (1)		Δs from (1)	
	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	-7.19	-3.41	-3.49	-5.37	-2.32	-1.80	-8.61	-4.69	-5.32	-6.53
Subway	9.02	10.77	1.31	0.73	0.87	1.02	5.00	6.57	6.17	6.97	5.52	7.25
Bus	22.44	30.47	1.79	0.61	0.52	1.13	-1.52	-2.48	0.31	-1.53	-0.76	-1.00
Bike	15.96	24.01	1.59	0.79	0.76	1.74	-0.86	-1.78	0.42	-1.12	-0.19	-0.25
Taxi	2.20	1.32	1.19	0.55	0.63	0.58	-0.20	-0.15	0.86	0.31	0.38	0.33
Walk	8.74	11.99	1.31	0.73	0.71	0.90	-0.10	-0.36	0.86	0.05	0.37	0.19
Speed (km/h)	21.49		3.83		3.97		1.13		4.68		5.09	
Panel B: sorting outcomes												
Distance to work (km)	18.56	15.66	0.02	0.02	-0.29	-0.19	0.76	0.61	0.85	0.62	0.43	0.44
Distance to subway (km)	5.33	4.30	-0.05	0.01	-0.11	-0.07	-4.13	-3.45	-4.14	-3.44	-4.14	-3.45
Panel C: welfare changes per household (thousand ¥)												
Consumer surplus (+)	-226.9		-31.6		-85.1		-70.7		187.4		105.3	
Toll revenue (+)					136.4		136.4					
Subway cost (-)							103.0		103.0		103.0	
Net welfare	-226.9		-31.6		51.4		65.7		84.4		2.3	
									-145.7		-31.2	
											112.7	
											60.3	

Note: the table replicates Table 6 with sorting but allows housing supply to adjust with a price elasticity of 0.53. The 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. Appendix E explains the simulation procedure. 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 fixed at ¥1.13 per km as in Table 6. High-income household are those with income above the median. Toll revenue is recycled uniformly across households. Subway cost includes the construction and operation costs that are equally distributed among 7.2 million households. Net welfare is consumer surplus plus recycled revenue and minus subway costs.

Table A8: Simulation Results without Household Sorting

Income relative to the median	2008 Subway Network						2014 Subway Network					
	(1) No Policy		(2) Driving restriction		(3) Congestion pricing		(4) No Policy		(5) Driving restriction		(6) Congestion pricing	
	Baseline levels		Δs from (1)		Δs from (1)		Δs from (1)		Δs from (1)		Δs from (1)	
	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low
Panel A: travel mode shares in percentage points and average speed												
Drive	41.02	21.02	-6.49	-2.99	-2.87	-5.01	-1.26	-1.07	-7.68	-4.08	-4.69	-6.11
Subway	9.43	11.24	0.83	0.26	0.45	0.54	3.61	5.08	4.73	5.47	5.11	6.77
Bus	22.31	30.07	1.88	0.98	0.77	1.65	-1.35	-2.03	0.51	-1.05	-0.63	-0.60
Bike	16.08	24.08	1.48	0.73	0.64	1.72	-0.81	-1.57	0.55	-0.85	-0.32	-0.31
Taxi	2.17	1.32	1.24	0.55	0.65	0.57	-0.08	-0.07	0.99	0.41	0.40	0.34
Walk	8.98	12.26	1.06	0.47	0.37	0.53	-0.10	-0.34	0.91	0.09	0.13	-0.08
Speed (km/h)	21.49		3.82		3.61		1.76		5.39		5.09	
Panel B: sorting outcomes												
Distance to work (km)	18.56	15.66										
Distance to subway (km)	5.33	4.30					-4.13	-3.45	-4.13	-3.45	-4.13	-3.45
Panel C: welfare changes per household (thousand ¥)												
Consumer surplus (+)			-227.3	-31.0	-110.7	-83.2	207.7	117.0	-32.2	82.1	83.2	37.2
Toll revenue (+)					138.8	138.8					127.9	127.9
Subway cost (-)							103.0	103.0	103.0	103.0	103.0	103.0
Net welfare			-227.3	-31.0	28.1	55.6	104.7	14.0	-135.2	-20.9	108.1	62.1

Note: the table replicates Table 6 but shuts down household sorting. In other words, travel mode choices adjust and clear the traffic sector but households do not change residential locations. 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 as in Table 6. High-income household are those with income above the median. Toll revenue is recycled uniformly across households. Subway cost includes the construction and operation costs that are equally distributed among 7.2 million households. Net welfare is consumer surplus plus recycled revenue and minus subway costs.

Table A9: Simulation Results with Ring-Road Specific Density Measures

Income relative to the median	2008 Subway Network						2014 Subway Network					
	(1) No Policy		(2) Driving restriction		(3) Congestion pricing		(4) Subway Expansion		(5) + Driving restriction		(6) + Congestion pricing	
	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low
Panel A: travel mode shares in percentage points and average speed												
Drive	41.66	21.44	-7.18	-3.40	-3.52	-5.39	-2.17	-1.66	-8.53	-4.62	-5.25	-6.41
Subway	9.02	10.77	1.29	0.70	0.86	0.97	4.62	6.05	5.79	6.44	5.25	6.81
Bus	22.44	30.47	1.78	0.60	0.58	1.23	-1.54	-2.53	0.31	-1.57	-0.74	-1.04
Bike	15.95	24.01	1.60	0.80	0.78	1.78	-0.78	-1.63	0.53	-0.93	-0.10	-0.11
Taxi	2.20	1.32	1.19	0.55	0.62	0.57	-0.17	-0.11	0.89	0.36	0.38	0.35
Walk	8.74	11.99	1.31	0.74	0.68	0.85	0.03	-0.12	1.02	0.32	0.47	0.39
Speed (<i>km/h</i>)	21.50		3.83		3.81		1.48		5.07		5.29	
Panel B: sorting outcomes												
Distance to work (km)	18.56	15.66	0.01	0.01	-0.16	-0.07	0.37	0.17	0.42	0.16	0.17	0.10
Distance to subway (km)	5.33	4.30	-0.03	0.03	-0.02	0.02	-4.13	-3.44	-4.13	-3.44	-4.13	-3.44
Panel C: welfare changes per household (thousand ¥)												
Consumer surplus (+)	-227.6		-32.8		-104.0		-74.6		215.6		98.9	
Toll revenue (+)					135.7		135.7					
Subway cost (-)							103.0		103.0		103.0	
Net welfare	-227.6		-32.8		31.7		61.1		112.6		-4.1	
									-119.5		-39.6	
									121.4		48.6	

Note: this table replicates Table 6 but incorporates ring-road specific traffic densities: density within the 3rd ring road, between 3rd and 4th ring roads, between 4th and 5th ring roads, and between 5th and 6th ring roads. 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 as in Table 6. High-income household are those with income above the median. Toll revenue is recycled uniformly across households. Subway cost includes the construction and operation costs that are equally distributed among 7.2 million households. Net welfare is consumer surplus plus recycled revenue and minus subway costs.

Table A10: Simulation Results using Preferred Specifications but without Random Coefficients

Income relative to the median	2008 Subway Network						2014 Subway Network					
	(1) No Policy		(2) Driving restriction		(3) Congestion pricing		(4) Subway Expansion		(5) + Driving restriction		(6) + Congestion pricing	
	Baseline levels		Δs from (1)		Δs from (1)		Δs from (1)		Δs from (1)		Δs from (1)	
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	47.17	28.26	-9.05	-5.38	-2.38	-3.25	-0.43	-0.35	-9.40	-5.67	-2.81	-3.56
Subway	6.38	7.99	1.22	0.66	0.30	0.35	0.92	1.32	2.32	2.11	1.25	1.83
Bus	19.67	27.85	3.45	2.09	0.89	1.39	-0.24	-0.51	3.08	1.47	0.63	0.77
Bike	16.70	23.54	2.68	1.66	0.68	1.02	-0.14	-0.31	2.46	1.29	0.52	0.64
Taxi	1.10	0.96	0.26	0.12	0.08	0.06	-0.02	-0.02	0.23	0.09	0.06	0.04
Walk	8.98	11.40	1.44	0.85	0.42	0.44	-0.08	-0.13	1.31	0.71	0.34	0.29
Speed (km/h)	22.36		4.59		4.59		0.16		4.67		4.77	
Panel B: sorting outcomes												
Distance to work (km)	18.66	15.66	-0.04	0.08	-0.59	-0.13	0.15	0.05	0.17	0.11	-0.46	-0.08
Distance to subway (km)	5.43	4.36	-0.06	0.06	-0.12	0.12	-4.24	-3.51	-4.24	-3.51	-4.25	-3.50
Panel C: welfare changes per household (thousand ¥)												
Consumer surplus (+)	-1447.6		-338.0	-797.6	-394.2	118.5	60.5	-1308.5	-274.0	-669.9	-327.4	
Toll revenue (+)			390.7		390.7				386.0		386.0	
Subway cost (-)					103.0		103.0		103.0		103.0	
Net welfare	-1447.6		-338.0	-406.9	-3.6	15.5	-42.5	-1411.5	-377.0	-386.9	-44.4	

Note: the table replicates Table 6 but removes random coefficients in both travel and housing choices from the preferred specification. The model without random coefficients produces low VOT, counter-intuitive substitution patterns, and very different welfare predictions from Table 6. 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 as in Table 6. High-income household are those with income above the median. Toll revenue is recycled uniformly across households. Subway cost includes the construction and operation costs that are equally distributed among 7.2 million households. Net welfare is consumer surplus plus recycled revenue and minus subway costs.