

Efficiency and Equity Impacts of Urban Transportation Policies with Equilibrium Sorting

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

July 2023

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 granular 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 that incorporates endogenous and spatially varying congestion. 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 reduce welfare. 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 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. The cost of subway expansion can also be fully financed 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, UW-Madison and NBER, pbarwick@wisc.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: UW-Madison, txia46@wisc.edu. We thank the editor and four anonymous referees whose comments have greatly improved the paper. We also thank Kelly Bishop, Jan Brueckner, Shengmao Cao, Jack Collison, Gilles Duranton, Fernando Ferreira, Edward Glaeser, Matthew Kahn, Robert Metcalfe, Alvin Murphy, Matías Navarro, Stuart Rosenthal, Kerry Smith, Christopher Timmins, and seminar participants at Boston University, Brazilian School of Economics and Finance, Cornell, Duke, Jinan University, MIT, Peking University, Shanghai University of Finance and Economics, Stanford, Syracuse, University of California Berkeley, University of Illinois, University of Maryland, University of Texas – Austin, 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. Barwick acknowledges generous support from the National University of Singapore during her sabbatical visit.

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 household 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 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 anti-congestion 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.

Both the theoretical literature (LeRoy and Sonstelie, 1983; Brueckner, 2007) and empirical analyses (Baum-Snow and Kahn, 2000; McMillen and McDonald, 2004; Jerch et al., 2021) have demonstrated a close connection between transportation policies and the housing market. Motivated by these papers, we develop 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. A key aspect of our model is that congestion, which varies spatially and spills over to other regions like “water in a bathtub”, is determined by the commuting mode choices and residen-

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. Four cities in China (including Beijing) were among the top 30 list. Beijing's drivers spend nearly 180 extra hours on the road (or 9% of working hours) per year relative to the travel time under the free-flow speed. See https://www.tomtom.com/en_gb/traffic-index/ranking. Commuting is one of the most unpleasant uses of time according to subjective well-being assessments (Kahneman et al., 2004).

tial locations of all households. At the same time, congestion directly affects households' location choices through the ease-of-commuting. Once estimated, the model allows us to conduct counterfactual simulations to predict new equilibrium outcomes under different transportation policies in terms of travel mode choices, household locations, congestion level, housing prices, and welfare distribution.

Our analysis leverages two unique data sources with fine spatial resolution (street addresses) that allow us to jointly model residential locations and commuting choices. The first dataset is the Beijing Household Travel Survey (BHTS) from 2010 and 2014, a large representative survey that records households' home and work addresses, trips made in a 24-hour window, and other demographic and transportation-related information. We complement this dataset by constructing the historical commuting route, distance, travel time, and pecuniary travel cost for all travel modes (walking, biking, bus, subway, car, or taxi) for each observed trip. The second data source contains housing transactions from a major government-run mortgage program for Beijing residents. Critically for our analysis, the housing data report not only the home addresses but also the work addresses of both the primary and secondary borrowers. Using these home and work addresses, we put together over 13 million hypothetical work-commute and travel-mode combinations for all primary and secondary borrowers. To our knowledge, these datasets constitute the most comprehensive data that link work commutes and housing transactions in the context of equilibrium sorting models.

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

The average and median 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 ([Small et al., 2007](#)). Households are willing to pay 18% more in terms of the home price for an equivalent reduction in a wife's than in a husband's commuting time. We also report the income elasticity of housing demand and the income elasticity of the marginal commuting cost, both of which are estimated within a unified framework. These elasticities are key determinants of the urban spatial patterns of residential locations.

We next simulate equilibrium residential sorting and transportation outcomes based on three policies of interest: driving restrictions, congestion pricing, and subway expansion, as well as combinations of the three. We decompose the welfare effect along five margins. The first margin measures the change in welfare when households adjust travel mode in response to increasing commuting costs, holding congestion and residential locations fixed. The second and third margins separately consider partial speed adjustments that are common to the empirical transportation literature. The second margin does not clear the transportation

sector, while the third is a full equilibrium speed adjustment that clears the transportation sector. The fourth margin 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. Conversely, 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 and triggers the strongest sorting responses, though in the opposite direction of congestion pricing. Subway expansion disperses households away from the city center and workplaces into locations near new subway stations in the suburbs.

Second, different transportation policies can either exacerbate or alleviate economic inequality ([Waxman, 2017](#); [Tsivanidis, 2019](#); [Akbar, 2020](#)). 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 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 reduction in congestion. 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 robustness checks to examine choices of housing demand IVs, measurement errors, and speed endogeneity. In several extensions, we consider housing supply responses, more granular congestion measures, removing random coefficients, and incorporating migration and consumption access. Housing

supply fluctuations in response to price changes allow more people to move to desirable locations and magnify the role of sorting. Results with more granular congestion measures are similar to the baseline with city-wide congestion as most of Beijing’s urban core is severely congested during rush hour. Shutting down random coefficients produces unrealistic substitution patterns between commuting modes, grossly overestimates the welfare losses from driving restrictions and congestion pricing, and underestimates the benefits of subway expansion. Incorporating migration and consumption access does not change the qualitative results of our analysis.

Our study makes three main contributions (see Section 2.1 for a detailed discussion of related papers). It is related to the large literature on equilibrium sorting (see [Kuminoff et al. \(2013a\)](#) for a review). To our knowledge, our paper is the first in the empirical equilibrium sorting literature to jointly model residential locations and travel mode choices and evaluate how these choices simultaneously determine both congestion and distance to work in equilibrium. The aggregate welfare implications differ qualitatively whether we account for or abstract from endogenous changes in congestion.

Second, our study relates to the recent advances using quantitative spatial equilibrium (QSE) models to explore the role of transportation in urban systems (see [Redding and Rossi-Hansberg \(2017\)](#) for a review). Our framework accounts for rich heterogeneity in observed and unobserved preferences by leveraging detailed household-level data instead of relying on tractable distributional assumptions (i.e., Fréchet) as many QSE models do. As shown above, removing unobserved preference heterogeneity results in welfare estimates that are both qualitatively and quantitatively different. In addition, we estimate 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. As pointed out by [Gallego et al. \(2013\)](#), little has been done to understand 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 to long run. By characterizing the underlying travel and housing choices, our equilibrium sorting framework provides a micro-foundation for linking the results between short- and long-run impact evaluation studies. For example, we illustrate that the equilibrium adjustments in the transportation sector and housing market offset more than half of the initial congestion relief of Beijing’s subway expansion. More importantly, the unified framework offers a common yardstick to compare actual and counterfactual policies over a range of outcomes including congestion reduction, urban spatial structure, social welfare, and distributional consequences.

The paper proceeds as follows. Section 2 discusses the related literature and describes the data and policy background. Section 3 lays out the equilibrium sorting model and the estimation strategy. The estimation results are presented in Section 4. Section 5 explains the counterfactual simulation algorithm. Section 6 examines different transportation policies and compares their welfare consequences. Section 7 concludes.

2 Literature Review, Policy Background, and Data Description

2.1 Literature Review

Our study is closely related 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., 2013b). Relevant papers include 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; Timmins and Murdock (2007), Walsh (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 sharing economy; Almagro and Domínguez-Iino (2020) and Couture et al. (2020) on endogenous amenities across neighborhoods; Murphy (2015); Bayer et al. (2016); Han et al. (2018); Wang (2020) that incorporate housing market dynamics and Ferreira and Wong (2021) on recovering neighborhood preferences with imperfect information. Diamond (2016) considers national-level household sorting accounting for agglomeration and inequality and models housing supply adjustments similar to our approach.

Most sorting papers treat both the distance to work and the level of congestion as exogenous attributes. An exception is Kuminoff (2012) which models household decisions in both the work and housing markets and endogenizes the commuting distance, but keeps congestion exogenous. In contrast, we treat both congestion and distance to work as endogenous objects that are determined simultaneously in equilibrium. Suri (2022) estimates Mumbai residents' housing and mode choices, but not in a unified framework as ours.

Another related literature is recent quantitative spatial equilibrium models (QSE) that explore the role of transportation in urban systems (see Redding and Rossi-Hansberg (2017) for a review). Compared to these papers, our approach has both limitations and advantages. One limitation is that we cannot analyze whether changes in the transportation system translate to higher labor productivity, for example through better allocation of time and labor market matching. Tsivanidis (2018) shows that these labor market implications are important in the context of the Transmilenio project in Bogotá. In addition, our analysis abstracts away from agglomeration forces often modeled in QSE papers (Ahlfeldt et al., 2015).

On the other hand, our modeling of endogenous congestion and endogenous commuting costs is more realistic and flexible. 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 from large travel surveys. This is essential as VOT is the single most important parameter for the welfare effects of transportation policies. As Small (2012) put it, "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 model-

ing, social cost analysis, pricing decisions, project evaluation, and the evaluation of many public policies.” In addition, the spatial feature of our framework accommodates the spillover effects of endogenous congestion and allows congestion in different areas to be interdependent.

Our framework accounts for rich heterogeneity in observed and unobserved preferences by leveraging detailed household-level data. Incorporating nonhomothetic preference heterogeneity matters for equilibrium responses and leads to meaningful changes in welfare evaluations. In contrast, unobserved preference heterogeneity in QSE models is often limited to Fréchet draws for analytic tractability with homothetic preferences for ease of aggregation. An important exception is [Couture et al. \(2020\)](#) which examines the welfare effects of US urbanization in a quantitative spatial model that allows significant heterogeneity. However, it abstracts from the general equilibrium feedback effects between housing and transportation markets at play in this paper.

Welfare improvements in QSE models usually result from changes in real income due to gains from trade via an increase in market access ([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 such as congestion may leave open Pareto improvements without changes in market access.

Third, our paper is related to the literature that evaluates different transportation policies. Related papers include [Bento et al. \(2005\)](#), [Parry and Small \(2009\)](#), [Duranton and Turner \(2011\)](#), [Anderson \(2014a\)](#), [Basso and Silva \(2014\)](#), [Yang et al. \(2018\)](#), [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; [Langer and Winston \(2008\)](#), [Anas and Lindsey \(2011\)](#), [Hall \(2018\)](#), [Yang et al. \(2020\)](#), and [Kreindler \(2018\)](#) on congestion pricing; and [Parry and Small \(2005\)](#), [Bento et al. \(2009\)](#), and [Li et al. \(2014\)](#) on gasoline taxes.

Studies in this literature either focus on short-run effects on travel choices, traffic congestion, and air pollution or examine longer-run partial equilibrium effects, often finding results that are different or with opposite signs (e.g., [Duranton and Turner \(2011\)](#)). Our framework on commuting and residential choices provides a micro-foundation that rationalizes differences between short- and long-run impact evaluation studies.

Another related literature allows feedback effects between the transportation sector and housing market in a calibrated computable general equilibrium framework without estimating the underlying consumer preferences ([Yinger, 1993](#); [Anas and Kim, 1996](#); [Langer and Winston, 2008](#); [Parry and Small, 2009](#); [Basso and Silva, 2014](#)). Our work complements but differs from these studies in that the estimation of the preference parameters and simulation of policy counterfactuals are internally consistent and based on the same model.

2.2 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.² 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 hard 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](#)).³ These time-series changes in the winning odds provide useful exogenous variation for the housing demand analysis, as we discuss below.

Beyond implementing demand-side policies, the Beijing municipal government also invested heavily in public transportation infrastructure. From 2007 to 2018, 16 new subway lines were built with a combined length of over 500 km (See 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 2019, and the total urban rail network reached over 6,700 km by the end of 2019. These expansions were designed, in part, to slow the growth of personal vehicle use by making public transportation more accessible.

Despite these policy efforts, traffic congestion continues to be a pressing issue: the average traffic speed in 2019 was 24.6 km/h during peak hours (7-9 am and 5-7 pm), according to the 2020 Beijing Transportation Report. From a neoclassical microeconomic perspective, the aforementioned policies fail to address the root cause of traffic congestion: the mispricing of road capacity.⁴ The Beijing municipal government recently announced a plan to introduce road pricing in the future while soliciting feedback from experts and the general public ([Yang et al., 2020](#)).

2.3 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. We believe the data compiled in this paper are state-of-the-art in

² Athens, Greece, implemented the first driving restrictions in 1982. Since then, a dozen other large cities in the world, including Bogotá, Mexico City, and New Delhi, have adopted similar policies. The impacts of these policies on congestion and air pollution have been mixed ([Davis, 2008](#); [Viard and Fu, 2015](#); [Zhang et al., 2017](#)).

³ About 20,000 new licenses were distributed each month through nontransferable lotteries from 2011 to 2013. The monthly quota was reduced to 12,000 after 2013. These quotas were considerably lower than the historical vehicle sales in Beijing.

⁴ Despite being continuously advocated by economists since [Vickrey \(1963\)](#), adoption of congestion pricing is limited in practice due to technical feasibility and especially political acceptability. Several European cities (London, Milan, Stockholm, and Gothenburg) have implemented various congestion pricing schemes over the past couple of decades. The New York state legislature recently approved a congestion pricing plan for New York City, which—pending approval by the Federal Highway Administration—is set to become the first US city to enact congestion pricing.

that they combine commuter- and household-level data from multiple sources and provide one of the most comprehensive descriptions of commuting patterns and residential locations in any city to date. Appendix Section A provides more detail on the data construction and how our data compare to other data sets used in recent empirical studies of transportation and urban issues.

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 urban core from the 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.

In addition to a vibrant downtown with many job opportunities, Beijing has 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 key 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 nine *jiedao* (neighborhoods). A *jiedao* is an administrative unit, similar to but larger than a census tract. 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.

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). 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, education) and a travel diary on all trips taken during the preceding 24 hours. The survey reports detailed information for each trip by each commuting member of a household, including the origin and destination, departure and arrival time, trip purpose, and travel mode used.

Our analysis focuses on 73,154 work commuting trips (home-to-work and work-to-home). Work trips constitute 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.

We construct attributes for all travel modes in commuters' choice set. We focus on six travel modes: *walk*, *bike*, *bus*, *subway*, *car*, and *taxi*, as other modes (e.g., motorcycles, company shuttles, carpooling, multi-modal) collectively account for less than 4% 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 it reports the number of transfers and walking time between bus stops and delivers more accurate estimates than Baidu. 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. We assume commuters use the subway stations closest to their trip origin and destination and incorporate the walking distance and time to the nearest 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) presents travel patterns in 2010 and 2014. 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. From 2010 to 2014, the shares of walk, bike, and especially bus trips decreased while the shares of car (i.e., driving) and subway trips increased, reflecting rising vehicle ownership and expansion of the subway network. Walking and subway trips are the longest in duration, while subway and car trips are the longest in distance. Car trips have a slightly longer duration and distance than taxi trips but are cheaper. Overall, the trade-off between time and cost is clear: walking trips are the slowest but also the cheapest. Car and taxi trips are faster but more expensive than other trip types.

Panel (b) of Figure 1 contrasts travel patterns between high- and low-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 are limited differences in travel distance across the two income groups except for the distance of car trips.

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 application, and there are no refinancing loans in the sample.

The final dataset includes 77,696 mortgage transactions, with detailed information on housing attributes such as the property size, age, street address, transaction price, and date when the mortgage was signed.⁵ We observe household demographics 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 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 re-weight the mortgage data to match the distribution of housing price, size, age, and distance to the city center among more representative housing transactions (obtained from separate datasets) by using entropy balancing (Hainmueller, 2012). All of the empirical analyses as well as the counterfactual analyses use the weighted sample. Results using the unweighted sample are similar. Appendix Section A.3 discusses the re-weighting procedure in more detail and describes additional data patterns, such as differences in the commuting distance by gender.

Table 2 provides summary statistics of the data. Figure 2 shows the spatial pattern of housing and household attributes based on mortgage transactions from 2006 to 2014, with a warmer color representing a higher value. Housing prices tend to be higher and the distance to work shorter near the city center. The outskirts of Beijing have larger homes with a lower unit price, reflecting the classic distance–housing size trade-off. There are some exceptions. 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 places in the city center. The northern parts of the city have better amenities (schools and parks) and more work opportunities and attract high-income households. While household income is generally higher in northern Beijing than in southern neighborhoods, households of different income levels tend to mix together throughout the city.

Our equilibrium sorting model requires us to construct home buyers’ choice sets. In theory, these choice sets could consist of *all* properties listed on the market. Researchers need to construct hypothetical commuting attributes of different travel modes for all properties in these choice sets. However, this is technically infeasible because the number of potential home–work–mode combinations exceeds hundreds of billions. Large choice sets are a common empirical challenge in the housing demand literature. To reduce the computational burden, we follow a choice-based sampling strategy as in Bayer et al. (2009); Tra (2010); Klaiber and Phaneuf (2010). The choice set for a household is assumed to include the purchased home and a 1% sample of

⁵We remove transactions with a missing or zero reported price, a price lower than ¥5,000/m² (the average price is ¥19,800/m²), buyers with no reported income, and addresses outside the sixth ring road.

houses randomly chosen from those sold during a two-month window around the purchase date.⁶ Section 4.2 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 addresses. 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 two-minute intervals (which we aggregate to one-hour intervals) from over 1,500 remote traffic microwave sensors covering all major roads throughout Beijing for 2014.

2.4 Commuting Route, Speed, and Congestion

Section 2.3 and Appendix Section A discuss how we construct households’ commuting routes and measure travel speeds. A few points are worth emphasizing. First, the travel survey does not report the actual commuting routes chosen by individuals. We assume households follow the routes recommended by the Baidu and Gaode APIs, which may differ between years (e.g., 2010 and 2014). In counterfactual analyses, commuting routes are held fixed under driving restrictions and congestion pricing and only re-optimized with subway expansions. In other words, we do not endogenize commuting routes in the model, which is computationally infeasible with a complex road network.

Driving speeds that we construct using Baidu and Gaode vary by individual commuting routes. There is significant variation in driving speeds among households living in the same neighborhood but traveling to different workplaces.

Congestion (also called density) is measured by the mileage-weighted number of vehicles on the road (see Section 3). One way to conceptualize the spatial scope of congestion is through the analogy of a bathtub, as described by [Arnott \(2013\)](#). Congestion builds in highly trafficked areas of a city and then spreads outward, like a bathtub filling with water. This concept emphasizes the fluid nature of congestion and the spillover effects experienced across different areas. Determining the appropriate level of spatial granularity for measuring congestion is a complex empirical matter. To account for spatial variation and incorporate spillover effects, we adopt three measures: city-wide congestion, congestion within each ring-road band, and congestion within each ring-road quadrant. There are fifteen ring-road-quadrant areas, based on the intersection of ring roads and quadrants (NW, SW, NE, SE, see Figure A6). City-wide congestion allows for the maximum spillover but may be too coarse. The ring-road-quadrant congestion measure strikes a reasonable balance between the need for localized measures and the complexities associated with the network effects of congestion.

⁶Choice-based sampling for differentiated demand has been demonstrated to yield consistent results in [Wasi and Keane \(2012\)](#) and [Guevara and Ben-Akiva \(2013\)](#). We choose a two-month window because Beijing’s real estate market was fluid during our sample period. The median 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.

To construct region-level congestion, commuting routes are divided into the corresponding segments by region. Congestion within a region equals the sum of mileage-weighted driving demand of individuals from different neighborhoods whose commuting paths cross the focal area. Correspondingly, a driver's commuting time is the sum of driving time for each trip segment within the relevant area. This approach provides a precise representation of how congestion affects commuting time and allows congestion in one area to spill over to other areas with a complex road network.

The effect of congestion on speed is governed by the speed-density elasticity, which states the percentage improvement in speed when congestion decreases by one percent. This speed-density relationship embodies Beijing's existing transportation technology. We estimate this relationship in Section 4.3 below, which exhibits limited heterogeneity across regions during rush hour.

2.5 Reduced-form Evidence

Due to space constraints, we do not report reduced-form evidence and refer readers to the event studies in a concurrent study [Jerch et al. \(2021\)](#) that demonstrate the impact of Beijing's driving restriction policy on the housing market. Appendix Section B.1 presents additional evidence on sorting where neighborhoods that gain access to new subway lines account for a greater fraction of aggregate property transactions after the expansion of the subway network. These findings indicate that Beijing households actively sort across residential areas in response to transportation policies. We now turn our discussion to a structural model integrating the transportation sector with the housing market.

3 Empirical Equilibrium Sorting Model

The sorting model characterizes how household members choose commuting modes and residential locations. 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. Once estimated, the model allows counterfactual simulations and provides direct comparative statics of congestion levels, residential locations, housing prices, and social welfare across different policies.

The empirical framework can accommodate both a fixed and a variable housing supply. The counterfactual analyses in Section 6.1 to 6.3 assume a fixed housing supply. Section 6.4 allows local housing supply to adjust in response to changes in neighborhood housing prices.

Work locations are fixed ex-ante and do not change in this model (See Appendix Figure A2 for a map). This assumption is motivated by three observations. First, event studies in Appendix Section B.2 suggest that job changes tend to predate home purchases, not the other way around.⁷ Second, employment opportuni-

⁷About 60% of home buyers in our data changed jobs within three years prior to purchasing the home.

ties 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.

A location refers to a specific address. Households are identified by i and consist of a borrower and co-borrower, each denoted by k , who commute from home to the workplace. All households reside in a house that is indexed by j . They derive utility both from property and neighborhood attributes and from the convenience of commuting from home to work. Appendix Table A1 tabulates all mathematical notation used in the paper in the order of appearance.

3.1 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 unit j is specified as:

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

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 property size and age. Commuting members within household i are denoted by $k \in \{\text{Primary borrower, Coborrower}\}$.⁸ 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 third term on the right, $EV_{ijk}(v_{ijk})$, 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} of member k 's work commute (which is affected by congestion) in addition to the travel cost. As shown in Section 3.2 below, this commuting utility is our key innovation relative to traditional residential sorting models. Transportation policies generate different impacts on the commuting utility. They can affect either commuting time (such as driving restrictions and subway expansion) or costs (such as congestion pricing), which interact with households' heterogeneous commuting preferences. These changes then influence individuals' travel modes, which collectively determine congestion and further affect their commuting utility.

⁸ As is standard in the sorting literature (Tra, 2010; Bayer et al., 2009; Klaiber and Phaneuf, 2010), we do not consider an outside option and assume all households live in a property.

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 ℓ in β_i :

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

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

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

where ζ_{ik} is independently and identically distributed (i.i.d.) normal.

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

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

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

Equation (1) does not control for neighborhood composition, such as race and ethnicity (Bayer et al., 2007; Shertzer and Walsh, 2019). There is not a close analogue to race or ethnicity in Beijing where 96% of the population is Han Chinese according to the 2010 Census. In addition, as Figure 2 illustrates, Beijing is characterized by co-mingling among households with different socioeconomic statuses and exhibits relatively high income-heterogeneity within small neighborhoods. This is a reflection of the city's historical development and the government's housing policies that promoted neighborhood diversity (Wu et al., 2013).

3.2 Choice of Travel Mode

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

$$\max_{m \in M_{ij}} u_{ijm} = \theta_{im} + \gamma_1 \cdot \text{time}_{ijm}(v_{ij}) + \gamma_2 \cdot \text{cost}_{ijm}/y_i + \mathbf{w}_{ijm} \boldsymbol{\eta} + \varepsilon_{ijm}, \quad m = 1, \dots, M_{ij} \quad (3)$$

where M_{ij} is the set of transportation modes available to individual i commuting from home j .⁹ 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.¹⁰ The monetary cost of the trip is denoted as cost_{ijm} and household income as y_i . The variable \mathbf{w}_{ijm} includes a rich set of interactions between mode dummies and year fixed effects, trip attributes, and commuter demographics to control for time-varying and location-specific factors by travel mode (such as changes in public transportation) as well as observed preference heterogeneity. Finally, ε_{ijm} is the i.i.d. error term with the type I extreme value distribution.

The mode-specific random coefficients, θ_{im} , 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 unobserved heterogeneous preferences that vary across individuals, such as the enjoyment of driving a car, the perceived environmental friendliness of using public transportation, scheduling or inconvenience costs that vary across individuals but do not scale with the time or distance traveled, and the health benefits of biking and walking.

The key parameters in the travel demand analysis are γ_1 and γ_2 . The time preference γ_1 follows a chi-squared distribution with mean μ_γ as in Petrin (2002).¹¹ An individual's sensitivity to the monetary costs of commuting is assumed to decrease in income: γ_2/y_i . VOT, the most important preference parameter for transportation decisions (Small, 2012), is measured by $\frac{\gamma_1}{\gamma_2} \cdot y_i$ and directly linked with the hourly wage.

Our specification of VOT follows the standard approach in the transportation literature that estimates the VOT through the ratio of coefficients on travel time and travel cost, where travel cost is measured as a fraction of the hourly wage (Jara-Díaz and Videla, 1989). Its theoretical foundation is the time allocation models of Becker (1965) and Small (1982), where the value of time spent on commuting is a function of the lost wage income and scheduling preferences. Note that demographic interactions with travel time or costs do not have clear interpretations, hence excluded.

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

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

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

⁹ As is typical in transportation studies (McFadden, 1974; Train and McFadden, 1978; Jara-Díaz and Videla, 1989), we do not consider the outside option of not commuting, such as telecommuting or exiting the labor market. In addition, the mode choices of different individuals within a household are assumed independent, as we do not observe whether and how mode choices within households are determined. We abstract from trip-chaining, which is unlikely to be of first-order importance for home buyers.

¹⁰Road congestion affects the travel time for bus trips in addition to car and taxi trips. We treat buses as if they run in dedicated lanes unaffected by congestion. We also abstract away from capacity constraints for buses and subways. Hence, we may overpredict bus and subway mode shares in simulations with high congestion levels.

¹¹A chi-squared distribution ensures all individuals have a positive value of time, accommodates richer heterogeneity than normal, and is computationally tractable.

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

$$\begin{aligned} EV_{ij}(v_{ij}) &= \mathbb{E}_{e_{ijm}} \left(\max_{m \in M_{ij}} u_{ijm}(v_{ij}) \right) \\ &= \log \left(\sum_{m \in M_{ij}} \exp [\theta_{im} + \gamma_1 time_{ijm}(v_{ij}) + \gamma_2 cost_{ijm}/y_i + w_{ijm}\eta] \right), \end{aligned} \quad (5)$$

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

3.3 Market-Clearing Conditions and the Sorting Equilibrium

The equilibrium market-clearing conditions for the housing market and the transportation sector are inter-related 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, determines households' sorting decisions and shapes their spatial distribution. We discuss these market-clearing conditions below.

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

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

which depends on the vector of housing prices \mathbf{p} as well as the vector of the driving speeds \mathbf{v} (through the ease-of-commute utility). We consider two scenarios for housing supply. In the first 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, we assume that the housing supply adjusts at the neighborhood level in response to changes in the average neighborhood price. This assumption mimics developers' considerations to build more properties in desirable neighborhoods. Let S_n denote the housing supply in neighborhood n . We assume that the housing supply has a constant elasticity, $\ln(S_n) = c_{n,0} + \epsilon * \ln(p_n)$, and increases $\epsilon\%$ with a 1% increase in the average neighborhood housing price p_n .

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). Traffic density, or congestion, is the aggregation over all households' driving demand (the notation makes it explicit that traffic density depends on

the vector of housing prices \mathbf{p} and driving speeds \mathbf{v}):

$$D_{T,r}(\mathbf{p}, \mathbf{v}) \equiv \sum_i \sum_j P_{ij}(\mathbf{p}, \mathbf{v}) \cdot \{ [R_{ij,car}(\mathbf{v}) \cdot \text{dist}_{ijr,car}] + [R_{ij,taxi}(\mathbf{v}) \cdot \text{dist}_{ijr,taxi}] \}. \quad (6)$$

The subscript T denotes the transportation sector, r defines the spatial granularity of the traffic density measure, P_{ij} is the probability that household i chooses property j , $R_{ij,car}$ and $R_{ij,taxi}$ are the probabilities that household i living in property j drives and takes taxi, and $\text{dist}_{ijr,car}$ and $\text{dist}_{ijr,taxi}$ are the distance travelled by car and taxi within region r .

As discussed in Section 2.4, the appropriate level of spatial granularity for congestion r is a complex empirical matter. We consider three levels of congestion: city-wide congestion, congestion at the ring-road-band level, and congestion at the ring-road-quadrant level. Calculating ring-road or ring-road-quadrant congestion requires us to divide commuting trips into the appropriate ring-road or ring-road-quadrant segments and aggregate over all individuals whose commuting paths cross the focal area. If a household member's commuting route passes through multiple regions, then his trip contributes to congestion in all these relevant regions through $\text{dist}_{ijr,car}$ and $\text{dist}_{ijr,taxi}$.¹²

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

$$\ln(S_{T,r}(\mathbf{v})) = c_0 + \epsilon_{T,r} * \ln(\mathbf{v}).$$

For a 1% increase in traffic speed, the traffic density that can be sustained under the existing road capacity goes down by $|\epsilon_{T,r}|%$. This supply relationship characterizes the nature of congestion externality: drivers on the road reduce other drivers' speed.

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

1. The housing market clears for all properties:

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

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

2. The transportation sector clears for every region r , where households' aggregate driving demand at

¹²Correspondingly, a driver's commuting time is the sum of driving time for each trip segment within the relevant ring-road-band or ring-road-quadrant areas.

speeds \mathbf{v}^* is equal to the traffic density that can be sustained under the existing road capacity at speed \mathbf{v}^* :

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

Our model follows the class of equilibrium sorting models with local spillovers studied in [Bayer and 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 system of Equations (2), (4), (7), and (8) is continuous. The existence of a sorting equilibrium follows Brouwer's fixed point theorem. Intuitively, a unique vector of housing prices (up to a 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 system of Equations (2), (4), (7), and (8) defines the equilibrium housing prices and traffic speeds $\{\mathbf{p}^*, \mathbf{v}^*\}$.¹³

4 Estimation Results

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

4.1 Commuting Mode Choice

The parameters of the travel mode choices are estimated via simulated maximum likelihood estimation (MLE) using household travel surveys. The key parameters of interest are time and monetary cost preferences.

We assume that the error term ε_{ijm} in Equation (3) is uncorrelated with commuting trips' time and monetary costs. Monetary costs are likely to be exogenous because Beijing's transportation bureau sets bus and subway fares uniformly across all routes. Gas prices are determined by the National Development and Reform Council and adjusted periodically and taxi fares are regulated by Beijing government.¹⁴ Hence, monetary costs do not vary by the level of congestion or quality of service. Travel time is determined by congestion. We include a rich set of interactions of travel mode with demographics, time, and spatial fixed effects to absorb

¹³The 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.

¹⁴For gas price regulations, see page 8 of <https://www.globalpetrolprices.com/articles/43/>. For taxi fare regulations, see http://fgw.beijing.gov.cn/bmcx/djcx/cxldj/202003/t20200331_1752789.htm.

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 at the individual home-workplace pair level that are unlikely to be correlated with travel time.

Table 3 presents parameter estimates for six specifications. The first three specifications control for demographics but do not have random coefficients. The last three specifications include random coefficients on travel time and travel mode dummies.

Column (1) controls for interactions between the year dummies (2010 or 2014) and mode fixed effects (car, taxi, bus, subway, walking, and biking) to capture changes in public transportation service over time. The implied VOT is 75.7% of the hourly wage. Column (2) adds the interactions between mode fixed effects and trip characteristics, which are trip distance bins (shorter than two km, between 2-5 km, longer than 5km), whether the origin is within the 4th ring road (a proxy for city center), and whether the destination is within the 4th 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 trip distance and is partially absorbed by the mode and trip-distance bin fixed effects. Ring road dummies for trip origins and destinations capture differences in the frequency and quality of public transit services as well as congestion. Column (3) further includes interactions of mode fixed effects with household demographic variables which are age, age squared, gender, education, vehicle ownership, and number of workers. These variables help explain different mode choices across demographic groups (e.g., wealthier households' greater likelihood of driving and using taxis) and improve the model fit.

Columns (4) to (6) use a chi-squared distribution with three degrees of freedom to approximate heterogeneous travel time preferences following Petrin (2002).¹⁵ 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 preferences on mode choices. For example, some commuters choose driving or taxi not because of a high VOT but because of scheduling constraints. Others choose walking or biking for exercise benefits. The dispersion of these preference parameters is economically large and statistically significant, suggesting significant preference heterogeneity.

Column (6) is our preferred specification. Adding travel time and mode-specific random coefficients leads to a stronger sensitivity to travel costs and delivers a much more reasonable estimate of the VOT. Appendix Figure A7 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.¹⁶

¹⁵ As in Petrin (2002), we winsorize the top and bottom 5% of the distribution to minimize the impact of extreme random draws, as VOT is unlikely to be infinite. The distribution with three degrees of freedom provides the best fit, though results are similar with two or four degrees of freedom.

¹⁶ In 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 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

Once we have estimated parameters from travel mode choices, we construct the commuting utility EV_{ij} as defined in Equation (5) for both male and female borrowers based on their work locations.¹⁷ These variables are included as part of the (buyer-specific) housing attributes.

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

One common issue encountered in the estimation of travel mode choices pertains to the simultaneous relationship between equilibrium mode choices and travel times. If households in a neighborhood share similar preferences for driving, it would result in a high driving share and at the same time low driving speed due to congestion. Table 3 incorporates a rich set of interaction terms between modes, trip attributes, and household characteristics to address this.

To further investigate the issue of simultaneity, we include the interaction of driving with fine spatial controls in Appendix Table A2. These regressors absorb correlated preferences in local areas and alleviate endogeneity concerns. Column (1) replicates Column (6) of Table 3. Column (2) adds driving and district fixed effect interactions, while Column (3) further includes driving and neighborhood fixed effect interactions. Estimating these specifications, particularly Column (3) with hundreds of additional fixed effects within a nonlinear framework, takes much longer. Reassuringly, the resulting estimates closely resemble those of the baseline. Parameters for travel time and costs and VOT remain essentially unchanged even with these fine spatial controls. The model's overall fit only improves marginally. These results suggest that the extensive set of controls included in our baseline model adequately addresses potential endogeneity concerns.

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

(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.

¹⁷ Around 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. Note that the calculation of EV_{ij} requires us to construct the travel time and cost for all available travel modes for every property in households' choice sets, as described in Section 2.3.

To address measurement errors in travel times, we conducted several robustness checks beyond controlling for a rich set of trip-related fixed effects. First, we construct alternative travel time variables based on the average speed either at the ring-road-band level or at the ring-road-quadrant level. The average local speed might better reflect households' expectations. Both parameter estimates and VOT are comparable to the baseline (Appendix Table A3).

Second, we repeat the exercise using self-reported travel times in Appendix Table A4.¹⁸ Results are qualitatively similar. The value of time (at 124% of the hourly wage) is higher than the baseline estimate (at 96% of the hourly wage). This is driven by the fact that self-reported values tend to underestimate the actual travel times, the so-called recall biases as shown in Appendix Section A.4, and therefore inflate the implied value of time. Overall, our findings remain robust to measurement errors.

4.2 Housing Location Choice

We now turn to the estimation of housing demand using the mortgage data. The ease-of-commute variable EV_{ij} that is derived from travel mode choices enters the housing demand Equation (1) as an observed housing attribute. Similar approaches that nest the expected utility as a choice attribute have been used by Capps 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. Because \widehat{EV}_{ij} is estimated separately, we bootstrap the standard errors for housing demand parameters.

Nonlinear Parameters Housing demand is estimated using a two-step procedure: the first step uses simulated MLE with a nested contraction mapping to estimate household-specific preference parameters (nonlinear parameters), and the second step uses linear IV for coefficients in the mean utility (linear parameters). The two-step strategy follows the approach of Berry et al. (1995) and Bayer et al. (2007). Specifically, we reorganize household i 's utility of purchasing property j into a sum of household-specific utility μ_{ij} and population-average utility δ_j , which absorbs the unobserved housing attribute ξ_j (we suppress time subscript t to ease exposition):

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

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

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

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

¹⁸Travel time for non-chosen modes uses the constructed time from Baidu/Gaode.

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.¹⁹ 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.

Household size is not reported in the mortgage data. Instead, we use the age of the primary borrower as a proxy for household size and interact age group dummies with the property size. Older households have a stronger preference for large houses. The group over 45 has the strongest large-house preference, 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, we search for changes in housing prices that would keep households' utility constant. According to our preferred specification in Column (3), an average household is willing to pay an additional ¥18,525 for a home that shortens the male member's daily work commute by one minute and ¥21,885 for a similar reduction in the female member's commute time.²⁰ This gap suggests that households prioritize the convenience of female members' commute in housing choices. This is consistent with descriptive evidence that women tend to live closer to their work locations (Appendix Figure A8) and existing literature (Le Barbanchon et al., 2020).

The VOT estimates from the travel survey suggest that each minute of a shorter commute is valued at approximately ¥1.1 (VOT is 96% of the hourly wage at ¥67.6 on average). In order for the WTP estimates based on housing demand to align with these VOT estimates, households would need to expect an average of 500 commuting trips per year for a duration of thirty to forty years.

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

¹⁹The 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.

²⁰This is derived by $\frac{\phi_{ik} \partial EV_{ijk}}{\partial \text{travel time}_{ijm}} * \frac{1}{\alpha_i} * 10^6$, as housing price is measured in millions of ¥.

Linear Parameters Table 5 reports the coefficient estimates on the population-average utility in Equation (11), conditioning on the specification in Column (3) of Table 4. 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 local market conditions and amenities that could vary across districts in Beijing. Columns (2)-(6) also include neighborhood fixed effects to capture unobserved time-invariant neighborhood amenities.

We use three sets of IVs for housing prices: the number of properties that are located in a different complex and within 3 km of unit j and sold within a two-month window around property j 's sale; the average attributes of these properties; and the interaction between the average attributes and the odds of winning the license plate lottery (see Section 2.2 for its policy background). The first two sets of IVs are sometimes called “donut instruments” (Bayer et al., 2007) in the housing literature, because the instruments are constructed from properties that are located between concentric circles around a given house. Our preferred specification is Column (6), with a first-stage F-statistic of 14.2.

The price coefficient estimate is negative and statistically significant across all columns. The IV estimates are larger in magnitude than the OLS estimates, consistent with the finding in the demand literature that unobserved product attributes bias OLS estimates toward zero. The average price elasticities derived from the OLS estimates are of the wrong sign, as the population average price coefficient is not negative enough to offset the positive coefficient of the income–price interaction. The signs on the other coefficient estimates from the IV regressions in Columns (3)-(6) are as expected. Households prefer larger properties and those closer to key schools but dislike older buildings and places far from parks.²¹

Incorporating the commuting utility EV not only improves the model fit but also has implications for other parameter estimates, especially the price coefficient and price elasticity. Appendix Table A5 reports the linear parameter estimates without the EV terms. Both the price coefficient and price elasticities are smaller in magnitude, consistent with the downward bias arising from omission of important attributes (EV terms). Timmins and Murdock (2007) find a 50% downward bias in the estimation of consumer welfare from recreation sites when on-site congestion is ignored in demand estimation.

Based on the parameter estimates from our preferred specification (the last set of results in Tables 4 and 5), the income elasticity of marginal driving costs and income elasticity of housing size is 0.78 and 0.10, respectively.²² To our knowledge, these are the first such estimates for Chinese households. The elasticity of marginal driving costs is largely consistent with other estimates in the literature (LeRoy and Sonstelie, 1983; Glaeser et al., 2008), while the elasticity for housing size is somewhat smaller than estimates based on U.S. data. 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,

²¹While the coefficient of distance to key schools is positive for the base group (borrowers under 30), the coefficients for borrowers in other age groups are negative and significant, as they are more likely to have school-aged children.

²²To calculate these elasticities, we increase household income, re-solve the equilibrium for both the transportation sector and the housing market (holding housing supply fixed), and calculate the changes in driving costs and housing size.

which might explain the lower values.

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

We do not model potential policy-induced changes in amenities as we lack appropriate measures such as retail shops, restaurants, and entertainment facilities. It is worth pointing out that any time-varying amenities are absorbed by the population-average utility for property j , δ_{jt} , and have no effect on nonlinear parameters that govern household preference heterogeneity, such as the random coefficient for ease-of-commute. We use neighborhood fixed effects and district-month-of-sample fixed effects to control for policy-induced amenities in estimating linear parameters. Appendix Section C.5 provides suggestive evidence that amenities might have improved after the subway expansion, and Section 6.4 discusses the implications of improved amenities.

For linear parameters, the first IV on the number of nearby properties and the second set of IVs on these properties’ attributes are often called the “BLP instrument” (Berry et al., 1995). The third set of IVs exploits city-wide shocks induced by exogenous policy changes (time-varying winning odds of license lotteries) that make areas close to city centers and job clusters more attractive.²³

To examine robustness to the choice of IVs, Appendix Table A7 presents parameter estimates for all combinations of IVs, the F-statistics and the Sargan-Hansen J-statistics, as well as the average housing demand price elasticities. The last column that uses all three sets of IVs is our preferred specification Column (6) in Table 5. The parameter estimates are robust across all columns, with the same sign, significance, and similar magnitudes for all coefficients. Results from the weak-IV tests and the Sargan-Hansen J-tests confirm that the choice of instruments in our preferred specification is valid and that the parameter estimates are robust to instruments.

Lastly, to examine the robustness of our results to the choice sampling method (as described in Section 2.3), we repeat the housing demand estimation with a 0.5% instead of 1% random sample to construct households’ choice sets (Appendix Tables A9 and A10). The parameter estimates, implied willingness to pay for housing attributes, and housing demand elasticity are quite similar across these two samples.

²³The effect of Beijing’s license lottery on the housing market has been demonstrated in previous studies (Lyu, 2022), which illustrates the validity of the lottery winning odds as an IV. Similar findings have also been reported for Singapore (Huang et al., 2018).

4.3 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 during peak hours with traffic density higher than 35 cars per lane-km which are more relevant since we focus on commuting trips.

We estimate the supply side of the transportation sector with the following equation:

$$\ln(v_{st}) = e_{T,r} * \ln(\text{Traffic Density}_{st}) + \mathbf{X}_{st}\beta_T + \varepsilon_{st}, \quad (12)$$

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

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

To examine potential differences in the speed–density elasticity across regions, we split our sample into four groups: between the second and third ring roads, the third and fourth ring road band, the fourth and fifth ring road band, and the fifth and sixth ring road band. Appendix Table A12 reports the OLS and IV estimates. The extent of heterogeneity across regions is limited, with the OLS and IV estimates comparable across columns.²⁴ In the counterfactual analysis below, we use the city-wide speed–density elasticity estimate of -1.1 in counterfactual analyses with city-wide congestion and ring-road level speed-density elasticity for ring-road congestion and ring-road-quadrant congestion, though results do not change much regardless of which elasticity is used.

5 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 policy implemented in Beijing: a vehicle is prohibited from driving on one of the five workdays. 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 facili-

²⁴We 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.

tate 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. We first simulate equilibrium outcomes without any transportation policy in place and gradually introduce different policies. Appendix D explains in detail the simulation algorithm. We provide a brief outline below.

5.1 Simulating the Counterfactual Equilibrium

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

$$\frac{\tilde{v}_{r,ij} - v_{r,ij}^o}{v_{r,ij}^o} = e_{T,r} * \frac{\tilde{D}_{T,r} - D_{T,r}^o}{D_{T,r}^o}, \quad (13)$$

where r denotes the spatial scope of congestion, either city-wide, ring-road specific, or at the ring-road-quadrant level. $\tilde{v}_{r,ij}$ is the counterfactual driving speed for the segment of household i 's work commute in region r , $v_{r,ij}^o$ is the observed driving speed (see Section 2.3 for its construction), $e_{T,r}$ is the speed-density elasticity estimate for region r from Equation (12). $\tilde{D}_{T,r}$ and $D_{T,r}^o$ are the counterfactual and observed traffic density for region r , respectively.

We repeatedly update speed, choices of travel mode, ease-of-commute, residential locations, and traffic density through Equations (13), (4), (5), (2) and (6) until we find the unique fixed point $\{\mathbf{p}^*, \mathbf{v}^*\}$.

Environmental Considerations Our housing demand model does not include local pollution as a neighborhood attribute. Air pollution in Beijing varies less across locations within the urban core than it does from day-to-day (Chen et al., 2015). In addition, past work has shown that public awareness of air pollution was limited before the installation of air quality monitors in 2013 (Barwick et al., 2023). We control for pollution indirectly using district-and-month-of-sample fixed effects as well as neighborhood fixed effects.

Given the importance of air pollution as a motivating factor in anti-congestion policies, we report welfare benefits from reduced air pollution as a result of lower congestion. Specifically, the expected air pollution

damage caused by household i in counterfactual simulations can be measured by:

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

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

where B_{ij} is the pollution damage if household i resides in property j . It consists of three terms: EF_{ijk} is the emissions factor that converts the kilometers driven by household i into grams of pollutant k (such as CO₂, NOx, and PM_{2.5}), VKT_{ij} denotes the commuting distance, and MD_k indicates the marginal damage per gram of pollutant k , which are derived using an intake fraction approach following the air pollution literature (Apte et al., 2012). We calculate the decrease in environmental damages resulting from reduced household driving under various transportation policies, aggregate the impacts across households, and divide them by the number of Beijing households to obtain welfare benefits per household. See Appendix D.4 for more details.

Fiscal Balance For subway construction and congestion pricing, we account for capital and operating costs as well as toll revenues. Further details about the source of these numbers can be found in Appendix D. We assume that taxes to cover subway costs are raised via a uniform head tax and that net congestion price revenues are redistributed uniformly by a lump sum, following the standard practice of non-distortionary distributions in the literature.

5.2 Welfare Decomposition

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

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

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

$$\begin{aligned} \frac{dW}{dcost} &= \underbrace{\frac{\partial W}{\partial cost} \Big|_{\mathbf{p}=\mathbf{p}_0, \mathbf{v}=\mathbf{v}_0}}_{(1) \text{ direct policy effect}} + \underbrace{\frac{\partial W}{\partial \mathbf{v}'} \frac{\partial \mathbf{v}}{\partial cost} \Big|_{\tilde{\mathbf{v}}}}_{(2) \text{ partial speed effect}} + \underbrace{\frac{\partial W}{\partial \mathbf{v}'} \frac{\partial \mathbf{v}}{\partial cost} \Big|_{D_T(\mathbf{v}^*)=S_T(\mathbf{v}^*)} - \frac{\partial W}{\partial \mathbf{v}'} \frac{\partial \mathbf{v}}{\partial cost} \Big|_{\tilde{\mathbf{v}}}}_{(3) \text{ rebound effect}} \\ &\quad + \underbrace{\frac{\partial W}{\partial \mathbf{p}'} \frac{\partial \mathbf{p}}{\partial cost} \Big|_{D(\mathbf{p}^*, \mathbf{v}^*)=1}}_{(4) \text{ equil. sorting effect}} + \underbrace{\frac{\partial W}{\partial \mathbf{p}'} \frac{\partial \mathbf{p}}{\partial cost} \Big|_{D(\mathbf{p}^*, \mathbf{v}^*)=S} - \frac{\partial W}{\partial \mathbf{p}'} \frac{\partial \mathbf{p}}{\partial cost} \Big|_{D(\mathbf{p}^*, \mathbf{v}^*)=1}}_{(5) \text{ housing supply effect}}. \end{aligned} \quad (14)$$

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 v^0 to \tilde{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, 2014a; 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). 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. 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 neighborhood housing price changes.

Before we present the simulation results, we first validate the structural model by comparing its predictions with results in the literature. 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 A13. The model-predicted change in the price gradient as a result of the driving restrictions is -0.034, consistent with the reduced-form evidence in Jerch et al. (2021).²⁵ This suggests that our structural analysis replicates well the documented pattern of equilibrium price changes under the driving restriction policy.

6 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 6.1 to 6.3 analyze the congestion reduction, sorting patterns, and social welfare effects in the baseline case where we fix housing supply and use city-wide congestion. Section 6.4

²⁵The coefficient of -0.034 is somewhat smaller in absolute magnitude than the reduced-form analysis in Jerch et al. (2021). This is partly because the reduced-form result reflects a short-run response while the structural simulation incorporates long-run equilibrium adjustments (especially the rebound effects) and partly because the data sources and periods are different.

considers various extensions, including variable housing supply, more granular measures of congestion, removing random coefficients, and accounting for migration and consumption access. Appendix E provides additional simulation results on speed variability across regions and differences between models with and without unobserved heterogeneity.

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. The average across the two groups delivers the welfare effect per household.

6.1 Mode Choice and Congestion Reduction

Driving Restriction Panel A of Table 6 examines changes in the travel mode and congestion.²⁶ 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 (i.e., subway, bus, biking, and walking). This reduces congestion and increases 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.

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 congestion pricing and driving restrictions. First, congestion pricing imposes a higher monetary cost of driving that scales with the distance traveled, while driving restrictions lead to a longer travel time. Second, congestion pricing reduces driving much more than driving restrictions among low-income households and much less among high-income households. Indeed, low-income households' driving probability is reduced by 26%, 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

²⁶The mode choice shares are slightly different between Column (1) of Table 6 and Figure 1 because the former reports mode choices among home buyers in the no policy scenario while the latter reports mode choices for all residents in Beijing, including non-homeowners, who accounted for 27.8% residents in 2010.

commutes but more people on the road under congestion pricing than under the driving restriction.

Subway Expansion Despite the immense scale of Beijing’s subway expansion over 2008-2014, it leads to the smallest congestion reduction among the three policies. Column (4) 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. 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 3.

These results, and especially the one on sorting, point to important channels beyond what has been examined in empirical studies that focus on the *short-run* impact of the subway system on traffic congestion.²⁷ Our findings are consistent with the 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; Heblich et al., 2020).²⁸

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 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 with subway expansion (a supply-side policy) in reducing congestion and increasing welfare than the driving restriction (a command-and-control policy).

The improvement in driving speed under the policy combinations is close to the sum of the speed improvements under the individual policies. For example, the speed improvement is 3.83 km/h under the driving restriction scheme, 1.49 km/h under the subway expansion, and 5.08 km/h under the combination of both

²⁷Using a regression discontinuity (in time) approach, Anderson (2014a) 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.

²⁸Anas (2022) provides some theoretical evidence indicating that road capacity enhancements may increase road usage (VKT) but still lower congestion even in the long-run.

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 leads to a 8.52-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 the 5.08 km/h improvement in Column (5). Congestion pricing affects both the extensive margin (whether to drive) and the intensive margin (how far to drive); both effects could be reinforced by subway expansion. In contrast, the driving restriction primarily operates along the extensive margin.

6.2 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 series of behavioral responses whereby households substitute across different travel modes and adjust their residential locations. We report the average distance to work and to the subway for both the rich and the poor.²⁹ Note that with a fixed housing supply, the average distance to the subway system across all households remains fixed under both driving restrictions and congestion pricing. If rich households move closer to subway stations, poor households would be displaced and move farther away from the subway by construction. We report the average distance to subway separately for the rich and poor to illustrate this displacement effect.

Transportation policies lead to different sorting patterns. Take **driving restrictions** as an example. The direct effect pushes households to live closer to their workplaces and thereby offsets 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 commuting time, especially for long-distance trips. For example, the driving time declines by twelve minutes for trips whose origin and destination both 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.

²⁹Because work locations are fixed, distances to work and the subway are indications of residential sorting. We report straight-line distances, which are not affected by travel mode changes.

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 3 plots changes in the average commuting distances for residents in each TAZ relative to the distances in the no-policy scenario. Driving restrictions lead to modest commuting distance changes that are often in opposite directions across neighborhoods. The commuting distance is reduced in almost all TAZs under congestion pricing, suggesting better spatial matches between work and housing locations. The reduction in commuting distance is most pronounced for TAZs outside the fourth ring road, where the average commuting distance is 21.9 km, versus 11.9 km for households living inside the fourth ring road. 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 baseline scenario. This reflects transit-based gentrification, where lower-income households are priced out of premium locations closer to the subway (Appendix Figure A9). 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 4 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.³⁰ 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 A10 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.

6.3 Welfare Analysis

Panel C of Table 6 presents the welfare results, where we report changes in consumer surplus, toll revenue, costs of subway expansion, and environmental benefits. Figure 5 decomposes changes in consumer surplus following Equation (14) to illustrate different adjustment margins. Note that the consumer surplus estimates are recovered from housing transaction prices. That is, we search for changes in housing prices that would maintain households' utility before the transportation policy. As a result, they should be interpreted as total consumer surplus over a property's life span.³¹ Housing supply is fixed in this subsection. We consider variable housing supply and other extensions in Section 6.4.

³⁰Under 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.

³¹We assume that a property lasts for thirty years, which is consistent with the revealed WTP estimates for shorter commutes documented in Section 4.2. A household is willing to pay ¥18,525 for a home that shortens the male member's daily work commute by one minute. VOT of a minute-shorter commute is approximately 1.1 RMB. In order for the WTP estimate based on the housing transactions to align with the VOT estimate from the travel survey, households would need to expect an average of 500 commuting trips per year for a duration of thirty to forty years. The thirty-year assumption is also consistent with the U.S. Internal Revenue Service's rules that residential properties depreciate 3.6% per year and last 27.5 years. We use the U.S. reference as China does not collect property taxes and we could not find relevant depreciation rates for Beijing. We also assume that transportation policies last over a property's life span, namely thirty years. The thirty-year assumption does not affect estimates of consumer surplus. It is only relevant for the calculation of aggregate congestion pricing revenue, the expenses of collecting tolls, and the costs of operating the subway system, all of which are discounted over thirty years at a discount rate of 0.98.

Key Findings First, despite their effectiveness in congestion reduction, **driving restrictions** generate a total welfare loss of ¥125,700 per household, though high-income households experience a much steeper reduction than low-income households.³² The annualized loss is roughly 4% of the household income.³³ Driving restrictions force drivers to switch to slower commuting modes and significantly increase 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.

Figure 5 decomposes the welfare loss along different adjustment margins, and we present the same analysis for speed adjustments in Appendix Figure A11. The direct policy effect of a driving restriction is large and negative at ¥223,200 per household since it distorts commuting choices and forces households to substitute toward inferior travel modes. As commuters switch to non-driving 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 5. 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 is a loss of ¥56,000 versus ¥124,900, and illustrates the importance of incorporating the rebound effect and allowing the full equilibrium adjustment of the transportation sector. Otherwise, the welfare losses could be underestimated by 55% for driving restrictions and welfare gains overstated by 36%-116% for other policies. The fourth bar further incorporates the sorting effect. With all four channels incorporated, the welfare loss is at ¥125,700 per household.

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 affected by increases in monetary costs from congestion pricing than they are by longer commuting times under driving restrictions. However, when the toll revenue is uniformly recycled across income groups, congestion pricing leads to *welfare gains* for both groups: consumer surplus increases by ¥43,500 and ¥68,600 for high- and low-income households, respectively. Low-income households witness a larger consumer surplus increase than the high-income group partly because they pay a smaller amount in 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 ¥55,500 per household. The equilibrium speed effect (the partial speed and rebound effect) reverses the welfare loss to yield a welfare gain of ¥46,000 per household. As sorting works in the same direction as the speed effect and moves households closer to their places of work, the welfare gain further increases to ¥56,000 per household. Residential sorting enhances the welfare gain from congestion pricing by 22%, consistent with the result based on 98 US cities in [Langer and Winston \(2008\)](#) that also finds a large

³²The welfare effect per household is the average of the numbers reported for high-income and low-income households.

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. Assuming a thirty-year period, the annualized loss is 4%.

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 non-driving 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 ¥119,000 and close to zero for low-income households.

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

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 ¥99,400 per household across all policy scenarios. The annualized welfare gains are equivalent to 3% of household annual income. The revenue from congestion pricing at ¥127,700 per household could fully cover the costs of subway expansion at ¥103,000 per household. 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 equally relevant. Our analyses suggest that welfare gains from infrastructure improvements could be mitigated by induced congestion. Pairing these investments with pricing instruments such as congestion pricing is critical to successfully address pre-existing and induced congestion and finance the cost of infrastructure investment to increase social welfare.

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

Comparison to the Literature The magnitudes discussed above, such as the annualized welfare gains at 3% of household annual income for the combined policy of subway expansion and congestion pricing, are within the spectrum of welfare effects for large-scale transportation policies reported in the literature. [Tsivanidis \(2018\)](#) finds that GDP in Bogota increased by 2.36% as a result of the expansion of the Transmilenio bus rapid

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 enhance affected corridors, including public transit ([GAO, 2012](#)).

transit network without migration and 12.09% with migration in a general equilibrium framework. [Kreindler and Miyauchi \(2019\)](#) find that a peak-hour pricing experiment in Bangalore, India increased welfare by 2.3% through a partial equilibrium analysis.

[Khanna et al. \(2020\)](#) studied the new tram and cable car systems in Medellin and reported welfare effects at 1.6% and 1% of GDP for the tram line and new cable line, respectively. [Warnes \(2020\)](#) studied the new Bus Rapid Transit System (BRT) in Buenos Aires, Argentina and reported welfare gains for residents near a BRT equivalent to a 0.4% increase in the 2010 GDP. Both [Khanna et al. \(2020\)](#) and [Warnes \(2020\)](#) use a general equilibrium framework.

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 by shutting down sorting (See Appendix Table A14 for 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 5.

Sorting also has important distributional implications. This is most evident under subway expansion, where sorting improves the welfare of high-income households at the cost of low-income households. This reflects transit-based gentrification: both high- and low-income households prefer places near the subway, but higher WTP from high-income households raises prices and displaces low-income households. In contrast, under congestion pricing, both groups are better off with sorting. This is because sorting, in the form of 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 under congestion pricing and subway expansion and the unequal distributional consequences 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 (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.

6.4 Extensions and Robustness Checks

Our baseline results in Sections 6.1 to 6.3 assume a fixed housing supply and city-wide traffic density (i.e., congestion). We relax these assumptions in Panel B of Table 7 and describe the results below.

Housing Supply Adjustment The first row of Panel B in Table 7 summarizes the speed and welfare changes when the housing supply responds to changes in neighborhood average housing prices and adjusts at a price elasticity of 0.53 as in Wang et al. (2012).³⁵ Appendix Table A15 reports the full set of results. To understand how housing supply adjustments affect our previous findings, consider a neighborhood where average home prices appreciate after a policy change. This price appreciation increases neighborhood 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 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 0.17 km to 0.35 km for high-income households and from 0.06 km to 0.24 km for low-income households with housing supply adjustment. The driving speed improvement increases from 3.83 km/h to 4.02 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. By allowing for housing supply responses, the increase in commuting distance expansion goes from 0.36 km to 0.86 km for high-income households and from 0.18 km to 0.72 km for low-income households. This attenuates the speed improvement under subway expansion from 1.49 km/h to 0.95 km/h.

Region-Specific Congestion The baseline analysis assumes that the geographic scope of congestion is city-wide and adjusts travel speeds with a city-wide traffic density. The second row of Panel B in Table 7 allows congestion to vary across ring-road-quadrants and adjusts local speeds with the congestion measure in the corresponding regions (Appendix Table A16 reports the full set of results). Appendix Table A17 illustrates changes in travel speeds across these fifteen regions, with the biggest effect in the outer part of southwestern Beijing, which had no subway coverage prior to 2014 and was subsequently connected. Overall, the policies have larger effects within the fifth ring road, as traffic is closer to road capacity in these regions. Despite these regional differences, both the average speed improvement and welfare effects are very similar to those in the baseline. Results using ring-road-band congestion measures are also similar. These patterns should not be surprising, as most of Beijing's urban core is severely congested during rush hour.

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, which is 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 as shown in Table 6. For example, under driving restrictions and congestion pricing, people who drive are more likely to switch to the subway and taxis. Similarly, driving, taxi, and bus trips are affected disproportionately more under the subway expansion. According to our preferred estimate, the subway expansion from 2008 to 2014 that doubled the length of the subway network boosted the percentage of people who commute via subway from 9.9% to 15.2%, an increase of more than fifty percent.

To evaluate the importance of incorporating heterogeneous preferences, we re-estimate the housing demand and travel mode choices excluding random coefficients. We include observed heterogeneity from variables such as income, age, gender, and education since models with neither demographic controls nor random coefficients cannot fit the data. Appendix Section E.2 provides an in-depth discussion of models with and without random coefficients in terms of model fit and substitution patterns, speed improvement, welfare effects, spatial sorting, and equilibrium prices. The third row of Panel B in Table 7 summarizes the welfare results.

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. Since one of the main goals of the policies studied here is to induce travel mode changes, predicting these adjustments accurately is of first-order importance for this study. For example, subway expansion increases ridership only by a modest 11% 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 number and generates a wrong sign for the welfare effects of subway expansion. 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 driving restrictions and congestion pricing 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 to Amenities Recent literature points out that improved transportation infrastructure benefits consumers in terms of easier access to amenities. According to (Miyauchi et al., 2021; Rao, 2021), the benefit of consumption access is about a third of the welfare benefits of job access in Tokyo and Beijing, respectively. We multiply the baseline consumer surplus by 1.33 and report the welfare changes 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 6. 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*.

7 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, we find 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.
- Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm, and Nikolaus Wolf**, “The Economics of Density: Evidence from the Berlin Wall,” *Econometrica*, 2015, 83 (6), 2127–2189.
- 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.
- 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**, “Putting “Downs’s Law” under the lens of theory: Roads reduce congestion and therefore increase distance traveled,” *Working Paper*, 2022.
- 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.
- , “Subways, Strikes, and Slowdowns: The Impacts of Public Transit on Traffic Congestion,” *American Economic Review*, 2014, 104 (9), 2763–96.
- Anderson, Michael L. and Lucas W. Davis**, “Does Hypercongestion Exist? New Evidence Suggests Not,” 2018. Working Paper.
- Apte, Joshua S, Emilie Bombrun, Julian D Marshall, and William W Nazaroff**, “Global intraurban intake fractions for primary air pollutants from vehicles and other distributed sources,” *Environmental Science & Technology*, 2012, p. 3415–3423.
- Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare**, “New trade models, same old gains?,” *American Economic Review*, 2012, 102 (1), 94–130.
- Arnott, Richard**, “Time for revisionism on rent control?,” *The Journal of Economic Perspectives*, 1995, 9 (1), 99–120.
- , “A bathtub model of downtown traffic congestion,” *Journal of Urban Economics*, 2013, 76, 110–121.
- 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.
- Barwick, Panle Jia, Shanjun Li, Liguang Lin, and Eric Zou**, “From fog to smog: The value of pollution information,” *American Economic Review*, 2023.
- 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.
- and **Matthew E Kahn**, “The effects of new public projects to expand urban rail transit,” *Journal of Public Economics*, 2000, 77 (2), 241–263.
- 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.
- Beck, Thorsten, Ross Levine, and Alexey Levkov**, “Big bad banks? The winners and losers from bank deregulation in the United States,” *The Journal of Finance*, 2010, 65 (5), 1637–1667.
- Becker, Gary S.**, “A Theory of the Allocation of Time,” *The economic journal*, 1965, pp. 493–517.
- 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**, “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.
- Callaway, Brantly and Pedro HC Sant’Anna**, “Difference-in-differences with multiple time periods,” *Journal of Econometrics*, 2021, 225 (2), 200–230.
- Capps, Cory, David Dranove, and Mark Satterthwaite**, “Competition and market power in option demand markets,” *RAND Journal of Economics*, 2003, pp. 737–763.
- Chen, Wei, Hongzhao Tang, and Haimeng Zhao**, “Diurnal, weekly and monthly spatial variations of air pollutants and air quality of Beijing,” *Atmospheric Environment*, 2015, 119, 21–34.
- 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).
- EPA**, "Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866," Technical Support Document, US Environmental Protection AgencyInteragency Working Group on Social Cost of Carbon 2016.
- Apple, 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.
- Goodman-Bacon, Andrew**, "Difference-in-differences with variation in treatment timing," *Journal of Econometrics*, 2021, 225 (2), 254–277.
- Gorback, Caitlin**, "Your Uber has Arrived: Ridesharing and the Redistribution of Economic Activity," Technical Report, Wharton working paper 2020.
- Gu, Yizhen, Jiang Chang, Junfu Zhang, and Ben Zou**, "Subways and Road Congestion," *American Economic Journal: Applied Economics*, 2020. forthcoming.
- , **Naijia Guo, Jing Wu, and Ben Zou**, "Home Location Choices and the Gender Commute Gap," *Journal of Human Resources*, 2021, pp. 1020–11263R2.
- Guevara, C. Angelo and Moshe E. Ben-Akiva**, "Sampling of alternatives in Logit Mixture models," *Transportation Research Part B: Methodological*, 2013, 58, 185 – 198.
- Gupta, Arpit, Stijn Van Nieuwerburgh, and Constantine Kontokosta**, "Take the Q train: Value capture of public infrastructure

- projects," *Journal of Urban Economics*, 2022, 129, 103422.
- 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.
- Huang, Naqun, Jing Li, and Amanda Ross**, "The impact of the cost of car ownership on the house price gradient in Singapore," *Regional Science and Urban Economics*, 2018, 68, 160–171.
- Humbert, Sebastien, Julian D Marshall, Shanna Shaked, Joseph V Spadaro, Yurika Nishioka, Philipp Preiss, Thomas E McKone, Arpad Horvath, and Olivier Jollet**, "Intake fraction for particulate matter: recommendations for life cycle impact assessment," *Environmental Science & Technology*, 2011, p. 4808–4816.
- Hwang, Yujung**, "An Estimable General-Equilibrium Structural Model of Immigrants' Neighborhood Sorting and Social Integration," 2019. Working Paper.
- Jara-Díaz, Sergio R and Jorge Videla**, "Detection of income effect in mode choice: theory and application," *Transportation Research Part B: Methodological*, 1989, 23 (6), 393–400.
- Jerch, Rhiannon, Panle Jia Barwick, Shanjun Li, and Jing Wu**, "Road Rationing Policies and Housing Markets," 2021. Working Paper.
- Kahneman, Daniel, Alan B Krueger, David A Schkade, Norbert Schwarz, and Arthur A Stone**, "A survey method for characterizing daily life experience: The day reconstruction method," *Science*, 2004, 306 (5702), 1776–1780.
- Khanna, Gaurav, Carlos Medina, Anant Nyshadham, Daniel Ramos, Jorge Tamayo, and Audrey Tiew**, "Spatial Mobility, Economic Opportunity, and Crime," *NEUDC presentation*, 2020.
- 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.
- Kreindler, Gabriel E and Yuhei Miyachi**, "Measuring commuting and economic activity inside cities with cell phone records," *The Review of Economics and Statistics*, 2019, pp. 1–48.
- 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), 1007–62.
- , — , and — , "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,"

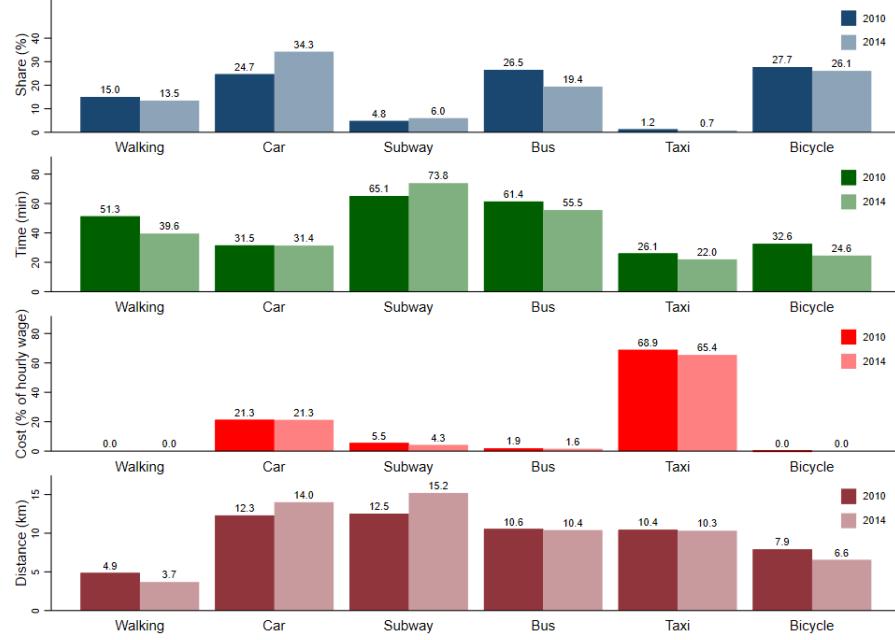
- Brookings-Wharton Papers on Urban Affairs*, 2008, pp. 127–175.
- LeRoy, Stephen F and Jon Sonstelie**, “Paradise lost and regained: Transportation innovation, income, and residential location,” *Journal of Urban Economics*, 1983, 13 (1), 67–89.
- Li, Shanjun**, “Better lucky than rich? Welfare analysis 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.
- Lyu, Xueying**, “Car restriction policies and housing markets,” *Journal of Development Economics*, 2022, 156, 102850.
- McFadden, Daniel**, “The measurement of urban travel demand,” *Journal of public economics*, 1974, 3 (4), 303–328.
- McMillen, Daniel P. and John McDonald**, “Reaction of House Prices to a New Rapid Transit Line: Chicago’s Midway Line, 1983–1999,” *Real Estate Economics*, 2004, 32 (3), 463–486.
- MEP**, “Technical guidelines for compilation of road motor vehicle air pollutant emission inventory.” Technical Report, Science and Technology Standards Department of the Ministry of Environmental Protection, Tsinghua University, & Chinese Academy of Environmental Sciences 2015.
- 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.
- Shertzer, Allison and Randall P Walsh**, “Racial sorting and the emergence of segregation in American cities,” *Review of Economics and Statistics*, 2019, 101 (3), 415–427.
- 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.**, "The scheduling of consumer activities: work trips," *The American Economic Review*, 1982, 72 (3), 467–479.
- 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.
- Suri, Palak**, "Public Transit Infrastructure and Employment Accessibility: The Benefits of the Mumbai Metro," Technical Report, Working Paper 2022.
- Tarduno, Matthew**, "For whom the bridge tolls: Congestion, air pollution, and second-best road pricing," *Unpublished manuscript*, 2022.
- 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.
- Tra, Constant I.**, "A discrete choice equilibrium approach to valuing large environmental changes," *Journal of Public Economics*, 2010, 94 (1), 183 – 196.
- Train, Kenneth and Daniel McFadden**, "The goods/leisure tradeoff and disaggregate work trip mode choice models," *Transportation research*, 1978, 12 (5), 349–353.
- Tsivanidis, Nick**, "The Aggregate and Distributional Effects of Urban Transit Infrastructure: Evidence from Bogotá's TransMilenio," *American Economic Review*, 2018. forthcoming.
- , "Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogota's TransMilenio," 2019. Working Paper.
- Tyndall, Justin**, "The local labour market effects of light rail transit," *Journal of Urban Economics*, 2021, 124, 103350.
- 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**, "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.
- Warner, Pablo Ernesto**, "Transport infrastructure improvements and spatial sorting: Evidence from buenos aires," Technical Report, 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.
- Wu, Wenjie, Wenzhong Zhang, and Guanpeng Dong**, "Determinant of residential location choice in a transitional housing market:

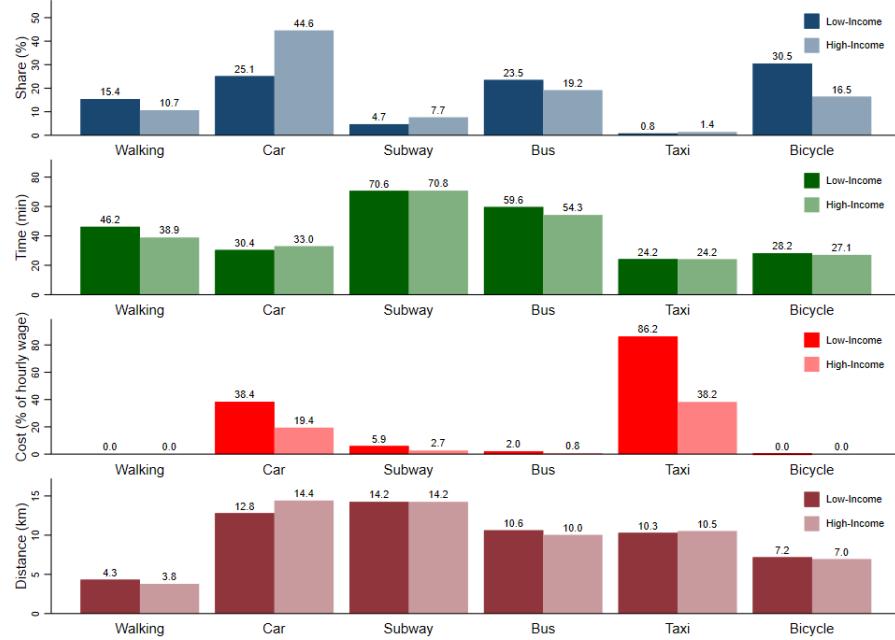
- Evidence based on micro survey from Beijing," *Habitat International*, 2013, 39, 16–24.
- Xiao, Junji, Xiaolan Zhou, and Wei-Min Hu**, "Welfare Analysis of the Vehicle Quota System in China," *International Economic Review*, 2017, 58 (2), 617–650.
- Yang, Jun, Avralt-Od Purevjav, and Shanjun Li**, "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.
- Yinger, John**, "Bumper to bumper: A new approach to congestion in an urban model," *Journal of Urban Economics*, 1993, 34 (2), 249–274.
- 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.
- Zhou, Hui**, "Restricting Used Vehicle Imports: Welfare Effects and Strategic Interactions," *Working Paper*, 2022.
- Zhou, Ying, Jonathan I. Levy, John S. Evans, and James K. Hammitt**, "The Influence of Geographic Location on Population Exposure to Emissions from Power Plants throughout China," *Environment International*, 2006, p. 365–73.

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

(a) Year 2010 vs. Year 2014

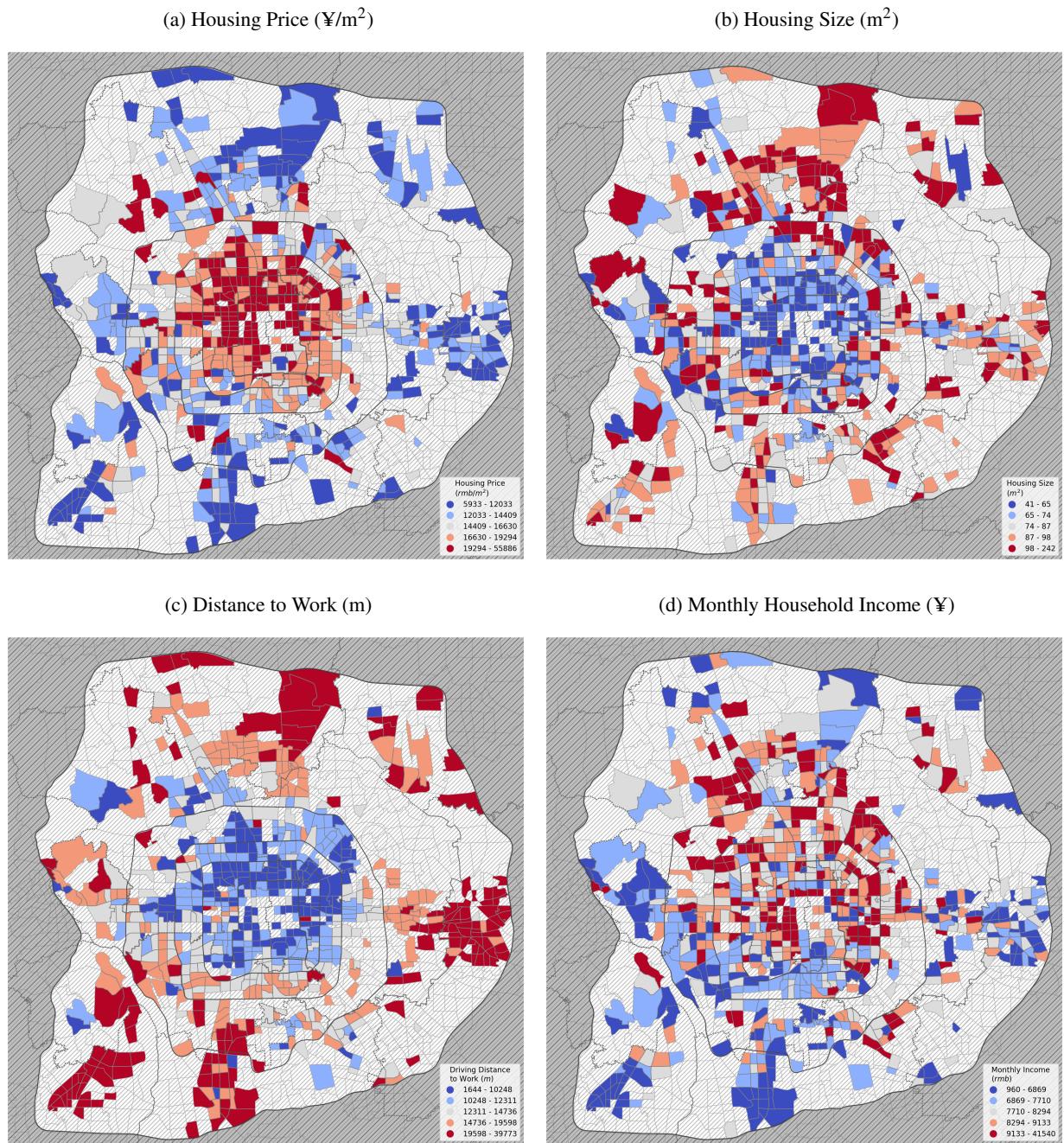


(b) High-income vs. Low-income Households



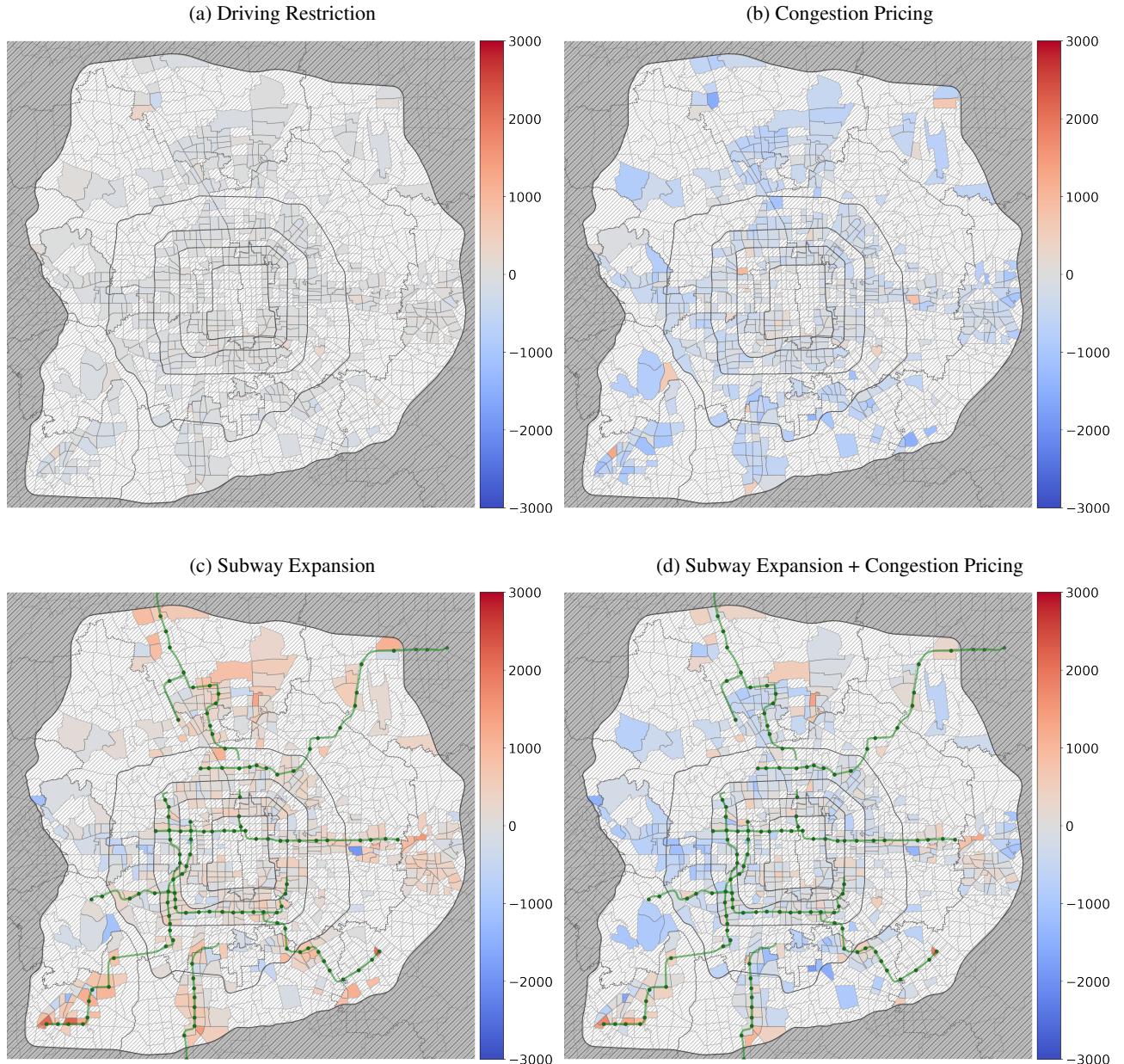
Note: This figure plots the trip share, time, costs, and average distance by travel modes for work commuting trips in the Beijing Household Travel Survey of 2010 and 2014. There are six main trip modes: walk, bike, bus, subway, car, and taxi. Bus and subway trips could include segments with other modes 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.2. 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



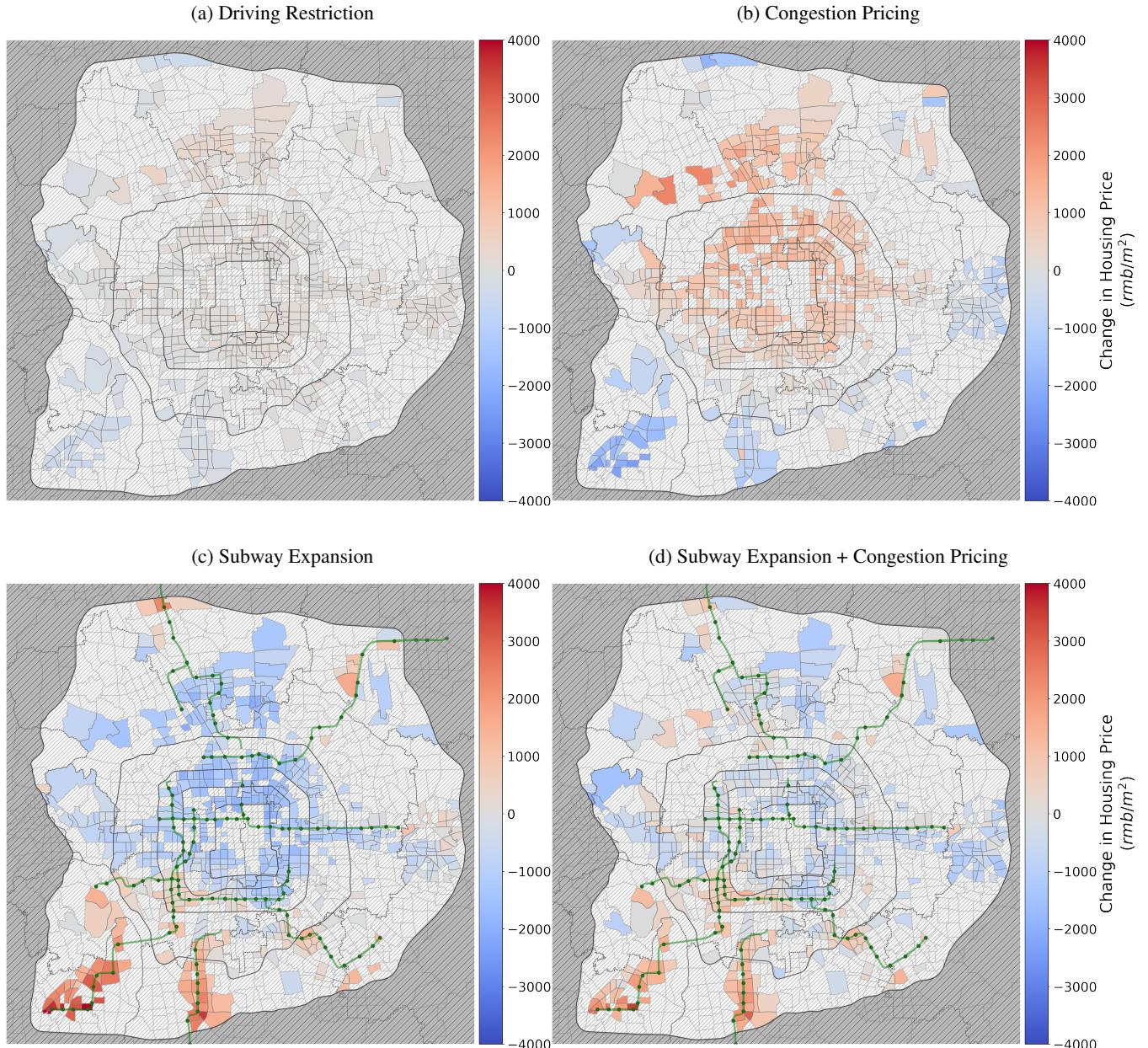
Note: This figure plots the average housing price and size and household commuting distance and monthly income by Traffic Analysis Zones (TAZ) based on the 2006-2014 mortgage data. TAZs are standardized spatial units used by transportation planners. There were 2050 TAZs in Beijing in 2014. Distance to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



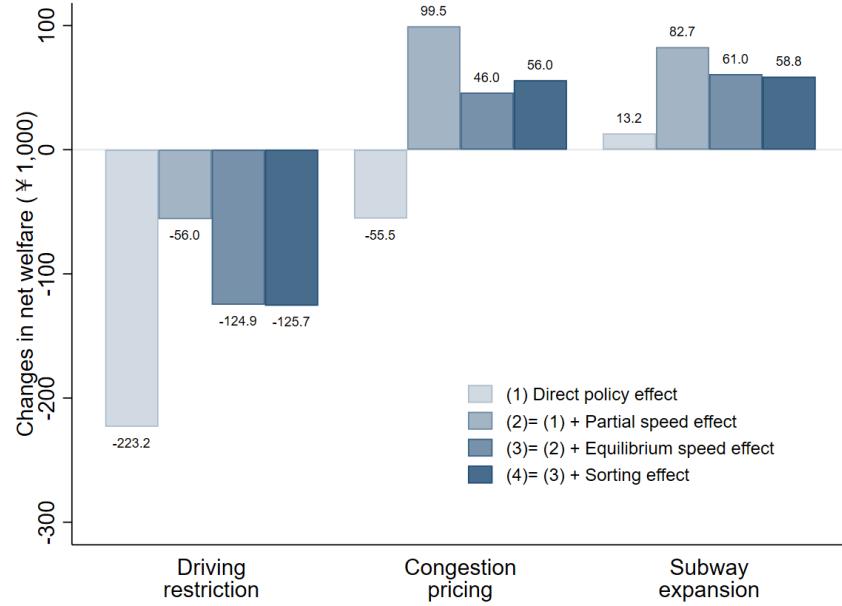
Note: This figure illustrates simulated changes in commuting distances in meters across TAZs under different counterfactual policies (relative to the no policy scenario). The results are based on the simulations in Table 6 that allow for household sorting, fix housing supply, use estimates including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

Figure 4: Changes in Housing Prices from Counterfactual Simulations in $\text{¥}/m^2$



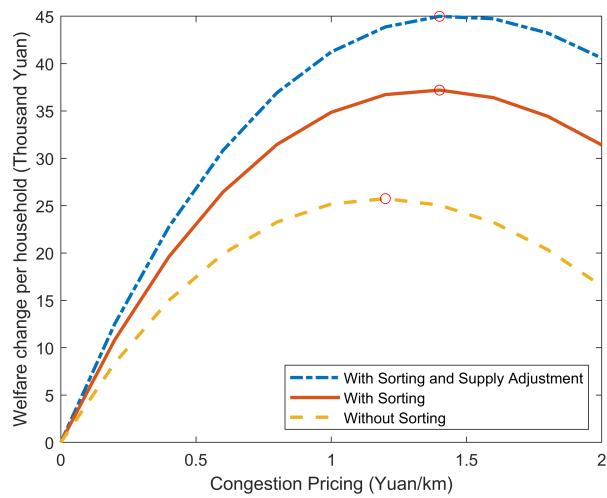
Note: This figure illustrates simulated changes in housing prices in $\text{¥}/m^2$ across TAZs under different counterfactual policies (relative to the no policy scenario). The results are based on the simulations in Table 6 that allow for household sorting, fix housing supply, use estimates including random coefficients, and use a single city-wide congestion index. Warmer colors correspond to increases in commuting distance while colder colors represent decreases. Green lines represent new subway lines built between year 2008 and 2014.

Figure 5: Welfare Decomposition for Different Transportation Policies



Notes: This figure decomposes welfare changes per household along four adjustment margins. Simulations allow for household sorting, fix housing supply, use estimates including random coefficients, and use a single city-wide congestion index. 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 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 6: Optimal Congestion Pricing under the 2014 Subway Network



Note: The figure plots welfare changes over different congestion prices under the 2014 subway network, without household sorting (yellow dotted line), with sorting (orange solid line), and sorting together with housing supply adjustments (blue dashed line). Simulations exclude pollution reduction benefits, allow for household sorting, have fixed housing supply, use estimates including random coefficients, and use a single city-wide congestion index. 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

	N	2010 Mean	SD	N	2014 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 < 2 km	30334	0.25	0.43	42820	0.24	0.43
Distance in [2, 5 km)	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.2. 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 is 73,154. The specifications include an increasingly rich set of fixed effects interacted with travel model dummies. Trip related FEs include trip distance bins (whether the distance is shorter than two km or between 2-5 km) and origin and destination ring road dummies (whether the trip origin is within the fourth ring road, and whether the trip destination is within the fourth ring road). Demographics FEs include a respondent's age, age squared, gender, education, car ownership, and whether the household has more than one commuter. The first three specifications are multinomial logit while the last three add random coefficients. The distribution of preference on travel time is specified as a chi-square distribution (winsorized at the 5th and 95th percentiles) with three degrees of freedom to allow for long tails. The random coefficients on travel mode dummies (driving, subway, bus, bike, and taxi) are assumed to have a normal distribution with a standard deviation of σ_m . The last two rows report the implied mean and median value of time (VOT). Robust standard errors are reported in parentheses.

Table 4: Housing Demand - Nonlinear Parameters from Simulated MLE

	No EV		With EV		EV and random coef.	
	(1) Para	SE	(2) Para	SE	(3) Para	SE
Demographic Interactions						
Price (¥mill.) * ln(income)	0.965	0.007	1.005	0.014	1.030	0.016
Age in 30-45 * ln(distance to key school)	-0.329	0.004	-0.391	0.011	-0.420	0.013
Age > 45 * ln(distance to key school)	-0.074	0.009	-0.111	0.025	-0.123	0.026
Age in 30-45 * ln(home size)	1.343	0.014	1.443	0.026	1.486	0.030
Age > 45 * ln(home size)	2.394	0.028	2.665	0.070	2.746	0.060
EV_{Male}			0.708	0.015	0.755	0.016
EV_{Female}			0.833	0.017	0.893	0.019
Random Coefficients						
$\sigma(EV_{Male})$					0.379	0.019
$\sigma(EV_{Female})$					0.482	0.018
Log-likelihood	-206829		-170057		-168808	

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

Table 5: Housing Demand - Linear Parameters

	OLS (1)	OLS (2)	IV1 (3)	IV2 (4)	IV2+IV3 (5)	ALL (6)
Price (¥mill.)	-2.24 (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)	0.21 (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)	0.95 (0.080)	0.782 (0.137)	0.210 (0.213)	0.323** (0.139)	0.304** (0.121)	0.288** (0.118)
District-month-of-sample FE	Y	Y	Y	Y	Y	Y
Neighborhood FE		Y	Y	Y	Y	Y
First-stage Kleinberg-Paap F			9.88	10.48	14.22	14.22
P value: overidentification test				0.03	0.10	0.19
Avg. Housing Demand Price elasticity			-2.42	-1.40	-1.61	-1.79

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. IV1 is the number of homes that are within 3km from a given home, outside the same complex, and sold in a two-month window. IV2 is the average attributes of these homes (building size, age, log distance to park, and log distance to key school). IV3 is the interaction between IV2 and the winning odds of the license lottery. The winning odds decreased from 9.4% in January 2011 to 0.7% by the end of 2014. Column (6) uses all IVs. Standard errors are bootstrapped to account for the estimation errors in EV terms and clustered at the neighborhood level.

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	-3.4	-3.48	-5.39	-2.14	-1.66	-8.52	-4.62	-5.2	-6.4
Subway	9.02	10.77	1.29	0.7	0.84	0.96	4.62	6.06	5.79	6.44	5.24	6.83
Bus	22.44	30.47	1.78	0.6	0.57	1.24	-1.54	-2.53	0.31	-1.57	-0.76	-1.03
Bike	15.96	24.01	1.6	0.8	0.77	1.78	-0.8	-1.64	0.52	-0.94	-0.13	-0.13
Taxi	2.2	1.32	1.19	0.55	0.63	0.57	-0.16	-0.11	0.89	0.36	0.39	0.36
Walk	8.74	11.99	1.31	0.74	0.67	0.83	0.02	-0.13	1.01	0.32	0.46	0.37
Avg. Speed (<i>km/h</i>)	21.49		3.83		3.83		1.49		5.08		5.29	
Panel B: sorting outcomes												
Distance to work (km)	18.56	15.66	0.01	0.01	-0.17	-0.06	0.36	0.18	0.41	0.17	0.15	0.12
Distance to subway (km)	5.33	4.3	-0.03	0.03	-0.03	0.03	-4.14	-3.44	-4.14	-3.44	-4.14	-3.44
Panel C: welfare changes per household (thousand ¥)												
Consumer surplus (+)	-227.1		-32.7		-98.2		-73.1		220.3		100	
Toll revenue (+)					137.4		137.4				-14	
Subway cost (-)							103		103		103	
Pollution reduction (+)	4.25		4.25		4.25		4.25		1.69		5.79	
Net welfare	-222.8		-28.4		43.5		68.6		119.0		-1.3	
									-111.2		-33.2	
									139.4		59.4	

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. We incorporate household sorting but fix the housing supply and use a single city-wide congestion index. Consumer surplus estimates are recovered from housing transaction prices and should be interpreted as total consumer surplus over a property's life span. Toll revenue is net of the capital and operating costs of revenue collection. Subway cost includes construction and operation costs. Both toll revenue and subway cost are the discounted sum over thirty years (which is approximately the lifespan of a property) and allocated uniformly across households. Welfare benefits from pollution reduction arise from reduced tailpipe emissions. Net welfare is consumer surplus plus toll revenue and environmental benefits minus subway costs. Column (1) reports results when no policy was in place. Columns (2) to (6) present differences from Column (1). Driving restriction prohibits driving in one of five work days. Congestion pricing is set at ¥1.13 per km to generate the same speed improvement as the driving restriction. High-income households are those with income above the median. See Appendix Section D for more details on the simulation procedure.

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

	Driving restriction			Congestion pricing			Subway expansion		
	Δ Speed (km/h)	Δ Welfare (¥1000)		Δ Speed (km/h)	Δ Welfare (¥1000)		Δ Speed (km/h)	Δ Welfare (¥1000)	
Income relative to the median		High	Low		High	Low		High	Low
Panel A: sorting and endogenous congestion									
With sorting (main results)	3.83	-222.8	-28.4	3.83	43.5	68.6	1.49	119.0	-1.3
Without sorting	3.82	-223.1	-26.8	3.61	32.1	59.6	1.76	106.7	16.0
With sorting but without endogenous congestion	5.47	-107.3	-4.4	5.15	118.1	81.0	2.36	158.8	6.4
Panel B: extensions and robustness checks									
With sorting and housing supply response	3.85	-225.5	-27.9	4.02	56.3	72.8	0.95	64.0	-16.0
With region-specific traffic density	3.46	-239.4	-31.8	3.38	24.7	67.5	1.25	104.9	-4.8
Without random coefficients	4.59	-1447.6	-338.0	4.59	-406.9	-3.6	1.61	15.5	-42.5
With migration	4.65	-193.4	-22.6	4.63	77.1	75.7	0.73	82.0	-8.5
With consumption access	3.83	-297.8	-43.5	3.83	56.4	89.8	1.49	191.6	31.6

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

Online Appendix

“Efficiency and Equity Impacts of Urban Transportation Policies with Equilibrium Sorting”

Panle Jia Barwick Shanjun Li Andrew Waxman Jing Wu Tianli Xia

Outline This online appendix is organized into the following sections. In Section A, we compare our dataset to datasets used in recent urban and transportation studies, discuss the methodology used to construct commuting trip cost and travel time and the sample of housing transactions, and examine recall biases. Section B presents empirical evidence on residential sorting in response to transportation policies and that job locations are typically determined before residential locations. In Section C, we provide a detailed description of the estimation process for the travel choice model and housing demand model. We also discuss identification variation and robustness analyses, including the choice set for housing demand and endogenous amenities. Section D explains the algorithm employed for the counterfactual simulation exercise, underlying assumptions, as well as the calculation of costs associated with implementing various transportation policies and environmental benefits resulting from congestion reduction. Lastly, in Section E, we present additional counterfactual results that were not included in the main paper. These include variations in speed reduction with more detailed congestion measures, a comprehensive comparison between models with and without unobserved heterogeneity, and income sorting.

A Data Construction

A.1 Comparison to Datasets in the Literature

We believe the data compiled in this paper is state-of-the-art in that they combine granular data from multiple sources and provide the most comprehensive description of commuting patterns and residential locations in Beijing to date. For commuting patterns, we exploit two waves of a representative household travel survey that report detailed home and work geocodes that were subsequently validated. We combine the travel surveys with historical maps of the subway system in Beijing and GIS software to construct historical commute time and costs from home doorstep to workplace for different travel mode options. This involved millions of Application Programming Interface (API) inquiries (Gaode for buses and Baidu Maps for driving, biking, and walking) to trace commuting routes for all travel modes and all households in the travel survey, and again separately the commuting routes in our mortgage data for all households and properties in their choice sets.

To improve the accuracy of predicted speed responses in counterfactual simulations, we estimate the speed-density relationship using real-time traffic volume and speed data in two-minute intervals from over 1,500 remote traffic microwave sensors covering all major roads for 2014. We follow the best practice to model the effect of congestion that has been adapted to transportation economics from the engineering litera-

ture (Yang et al., 2020; Anderson and Davis, 2018; Anderson, 2014b; Arnott, 1995; Kreindler, 2018; Tarduno, 2022).

Our use of API requests echoes the spate of recent top papers using Google Maps to construct representative commuting data (Akbar et al., 2018; Akbar, 2020; Gorback, 2020). However, these studies draw random latitude and longitude points to match the distribution of home and work locations in a city rather than using real home and work location pairs for actual households. It is difficult for these studies to ground commuting decisions with actual household preferences as we do in our study.

Some studies avail themselves of representative household travel survey data for travel mode choices but are limited to aggregated geographical information for constructing commuting patterns. For example, the US National Household Travel Survey (NHTS) provides researchers with confidential micro-data at the zip code level. The US Census Longitudinal Employer-Household Dynamics (LODES) database is aggregated to Census blocks (Tyndall 2021; Severen 2019). In this regard, the combination of detailed household demographics, commuting mode choices, and exact home and work addresses makes our data unique not just for Beijing but for most transportation studies that model households' commuting decisions.

Other studies use granular cellphone data to recover the origins and destinations of commuting trips but make assumptions on demographics based on trip locations (Miyauchi et al., 2021; Kreindler and Miyauchi, 2019; Gupta et al., 2022). These data have extremely fine spatial granularity and provide a unique opportunity to describe the economic effects of transportation on neighborhoods within cities. On the other hand, they have limited ability to accurately predict individuals' commuting mode choices or decompose welfare effects by different socioeconomic groups (as demographic information is unobserved and instead inferred).

A.2 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 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. 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 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.3 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, they under-represent 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:

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

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

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

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

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

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 $E_{\text{resales}}[(X_j - \mu_j)^r]$ 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 re-weighted mortgage data match the larger representative datasets quite well. In all the empirical analysis, we use the re-weighted the mortgage data to better represent home buyers in Beijing.

A.4 Recall Bias

Beijing's travel survey records self-reported travel time. In Appendix Figure A12, we compare the self-reported travel time with the travel time we constructed using Gaode/Baidu as a function of distance. Notably, the self-reported travel time exhibits a considerably flatter trend compared to the constructed travel time. It shows a modest increase as the trip length grows and tilts downward for trips longer than 35 km. More importantly, the self-reported travel time is significantly lower than the constructed travel time for the majority of

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.

commuting trips. These patterns suggest that people generally hold a positive bias towards the transportation mode they choose and tend to underestimate the actual travel time for longer trips. In light of this, we use the constructed travel times in most of the analyses and use the reported time as a robustness check.

B Reduced-form Evidence

B.1 Evidence on Sorting

To provide direct evidence on sorting in response to transportation policies, we examine whether people are more likely to purchase properties in neighborhoods near new subway stations post subway expansion via an event study. The dependent variable is the fraction of home buyers who purchase a house in neighborhood n conditional on buying a house at time t :

$$Share_{nt} = \frac{\text{#Houses sold in neighborhood } n \text{ at time } t}{\text{#Total houses sold at time } t}.$$

We define two groups of neighborhoods: neighborhoods with new subway stations in the sample period 2008-2014 (treated), and those without (never treated). The key regressor is variable $\mathbf{1}\{subway_{nt}\}$, which takes the value one if a new subway station opens within 1 km of neighborhood n in year t . Results are robust to different cutoffs such as 0.5 km and 2 km. We regress the fraction of home buyers purchasing from neighborhood n on periods before and after nearby subway station openings and control for neighborhood and year fixed effects in a two-way fixed effects (TWFE) model:

$$Share_{nt} = \sum_{k=-4, k \neq -1}^{k=2} \mathbf{1}\{subway_{n,t-k}\} \beta_k + \lambda_n + T_t + \epsilon_{nt},$$

where the time periods that are two years or more after the expansion are grouped as $k = 2$, and time periods that are four years or more before the expansion are grouped as $k = -4$. Given the well-known bias in TWFE models ([Goodman-Bacon, 2021](#)) in the presence of treatment effect heterogeneity, we report the CSDID estimates following [Callaway and Sant'Anna \(2021\)](#) (CSDID) in Figure A13.

The introduction of new subway stations in a neighborhood led to a significant increase in the proportion of total housing purchases attributed to that neighborhood. Prior to the subway expansion, the average market share of the neighborhood was represented by the dashed blue line at 0.7%. However, after the subway expansion, neighborhoods that gained subway stations experienced a treatment effect of 0.5 percentage points, which translates to a 70% increase.

We acknowledge that this estimate has certain limitations. The observed increase is likely influenced by housing supply responses as well as spillover effects across neighborhoods, with houses in areas without subway stations becoming less desirable. Despite these limitations, this figure provides compelling evidence of sorting. The adjustments in housing supply and the spillovers across neighborhoods are consequences of, and responses to, household sorting in relation to the availability of subway access.

B.2 Work Locations and Home Locations

The analysis in the main text assumes that work locations are exogenous. This is driven by data limitations and computational feasibility. Here we provide evidence that property purchases are often consequences of job changes, not vice versa, lending support to our modeling assumption where a household chooses housing location conditional on the job location. For this exercise, we merged our mortgage data with a restricted-access dataset on linked employer-employee history covering the majority of employees in Beijing's formal sectors to examine both job changes and property purchases, building on (Gu et al., 2021). The restricted-access dataset includes 268 million employee-month observations for 6.5 million employees. See (Gu et al., 2021) for more details.

In raw data, 60% of the home buyers started a new job within three years *prior to* purchasing the home. Below we present event studies where an event is either a job change or a home purchase. We denote a job change for employee i in month t as $1\{Job\}_{it} = 1$ and a home purchase (more accurately, the month when i initiates a mortgage application) by $1\{Home\}_{it} = 1$.

The event study on job changes follows Beck et al. (2010), which regresses the job change indicator on lags and leads for home purchase and controls for employee and month fixed effects:

$$1\{Job\}_{it} = \sum_{s=-12}^{12} \beta_s \cdot 1\{Home\}_{i,t-s} + \alpha_i + \delta_t + \varepsilon_{it},$$

where the periods before -12 and after 12 are grouped with $s = -12$ and $s = 12$, respectively. The standard errors are clustered by employee. For comparison, we repeat the regression with $1\{Home\}_{it}$ as the dependent variable and dynamic lag and leads of job change as the control.

Panels (a) and (b) of Figure A14 present the raw data patterns. It is evident that job changes were much more likely to occur prior to home purchases. The probability of a job change decreased sharply upon home purchase and remained low afterwards. Mortgage applications display a nearly opposite trend: the probability of buying a home remains steady before a job change and increases substantially after a job change.

The event studies in panels (c) and (d) confirm these raw data patterns. The probability of a job change exhibits a break and a sharp decline surrounding home purchases, while the probability of a home purchase jumps following a job change. There seems to be an anticipation effect, as some individuals began purchasing properties a few months before the actual job change.

These results provide evidence that job changes tend to precede home purchases, rather than the other way around. While the likelihood of a job change is significantly higher before a home purchase, the opposite does not hold true for home purchases. The probability of a home purchase before a job change is lower, not higher. In other words, our assumption of endogenous home locations while treating work locations as exogenous appears to be supported by the data.

Lastly, it is worth noting that while we hold work locations fixed, we incorporate neighborhood fixed effects in the housing demand analysis. This accounts for the fact that certain houses are more centrally

located and closer to job clusters, making them potentially more valuable to households.

C Estimation Details

C.1 Estimating Travel Mode Choices

To recover preference parameters in travel model choices, we use simulated maximum likelihood. The notation follows that in Section 3.2 of the main text. The log likelihood function is defined as:

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

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}_m^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: $\bar{u}_{ijmh} = \theta_{imh} + \gamma_{1ih} \cdot \text{time}_{ijm}(v_{ij}) + \gamma_2 \cdot \text{cost}_{ijm}/y_i + \mathbf{w}_{ijm}\eta$, where h and $\{\theta_{imh}, \gamma_{1ih}\}$ 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 the ease-of-commuting EV for each i and j pair based on the observed housing and job locations via:

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

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.

C.2 Identification of Travel Demand and Measurement Errors

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

One common issue encountered in the estimation of travel mode choices pertains to the simultaneous relationship between equilibrium mode choices and travel times. If households in a particular neighborhood share similar preferences for driving, it would result in a high driving share and subsequently low driving speed due to congestion. To tackle this, Table 3 in the main text incorporates a comprehensive set of controls. These include mode-by-trip-length bins fixed effects (public transportation may be less reliable for longer trips), mode-by-trip-attribute fixed effects (trips originating or ending in the city center may face more congestion), mode-by-year fixed effects (public transportation may improve over time), and mode-by-demographic fixed effects (older workers and households with multiple commuters may be more likely to drive).

To further investigate the issue of simultaneity, we include the interaction of the driving dummy with fine spatial controls. These regressors absorb correlated preferences in local areas and can alleviate endogeneity concerns. Column (1) of Table A2 replicates Column 6 of Table 3 in the main text, our preferred specification with the most saturated set of regressors. Column (2) adds driving and district fixed effects interactions, and Column (3) further includes driving and neighborhood fixed effects interactions. Estimating these specifications, particularly Column (3) with hundreds of additional fixed effects within a nonlinear framework, takes much longer. However, the resulting estimates closely resemble those of the baseline, and the model's overall fit only improves marginally. This suggests that the extensive set of controls included in our baseline model should adequately address potential endogeneity concerns.

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

To address measurement errors in travel times, we conducted several robustness checks beyond controlling for a rich set of trip-related fixed effects. First, we construct alternative travel time variables based on the average driving speed either at the ring-road-band level or at the ring-road-quadrant level. The average local speed might better reflect households' expectations. Results for both the average ring-road speed and the average ring-road-quadrant speed are similar. We report results based on the average ring-road speed in Tables A3 below. The parameter estimates and the implied value of time (VOT) are both comparable to the baseline results in the main text. For instance, the mean and median VOT are 93% and 82% of the hourly wage, respectively, which is similar to the mean and median of 96% and 85% reported in Table 3 in the main text.

Second, we re-estimated the mode choice model using the self-reported travel time for each trip (Table A4).³⁷ The parameter estimates are comparable to those reported in the main text, though the value of time (at 124% of the hourly wage) is higher than the baseline estimate (at 96% of the hourly wage). This is driven

³⁷Travel time for non-chosen modes uses the measures constructed from GIS software.

by the fact that self-reported values tend to underestimate the actual travel times, the so-called recall biases as shown in Appendix Figure A12, and therefore inflate the implied value of time. Overall, our findings remain robust to measurement errors.

C.3 Estimating Housing Demand

The housing demand model is estimated using a two-step procedure. The notation follows Section 4.2 of the main text. Nonlinear parameters (denoted by θ_1) are estimated via simulated Maximum Likelihood with a nested contraction mapping in the first step and linear parameters (denoted by 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 a value of one if household i chooses housing j and w_i is the weight of household i (obtained from the entropy balancing described in Section A.3 to make the mortgage data 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. Nonlinear parameters θ_1 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):

$$\begin{aligned} &\max_{\theta_1, \{\delta_j\}_j} \ln \mathcal{L}(\theta_1, \delta_j) \\ \text{s.t. } &\sum_{i \in C^{-1}(j)} \frac{1}{Q} \sum_{q=1}^Q w_i \frac{\exp[\mu_{ijq}(\theta_1) + \delta_j]}{\sum_k \exp[\mu_{ikq}(\theta_1) + \delta_k]} = w_j, \forall j \in J \end{aligned} \quad (\text{A2})$$

The left-hand-side of Equation (A2) 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 $\{\delta_j\}_{j=1}^J$ that satisfies

constraint (A2):

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

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 baseline 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 $\{\hat{\delta}_j\}_{j=1}^J$, 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 next.

C.4 Identification of Housing Demand Parameters and Robustness

Our identification strategy for the housing demand closely follows the IO literature, where a common practice is to use the attributes of close substitutes as instruments for prices. In our context, this identification strategy translates to comparing two identical houses situated in close proximity to other houses of varying quality. In addition, we exploit city-wide shocks induced by exogenous policy changes that make certain areas more attractive and affect local housing demand.

Specifically, we have constructed three sets of IVs. The first IV is the number of properties that are located in a separate complex within 3 km of unit j and sold within a two-month window around property j 's sale.³⁸ We exclude properties that lie within the same housing complex as property j since properties in the same complex might share correlated unobserved attributes. This first IV is arguably exogenous and correlated with housing price p_j because the availability of many properties in close proximity exerts downward pressure on p_j through competition.

The second set of instruments consists of the average physical and location attributes of the same set of properties, including the average building size, age, the logarithm of distance to the nearest park, and the logarithm of distance to the nearest key school. The first two sets of IVs are commonly referred to as “BLP instruments” (Berry et al., 1995). In the housing literature, these IVs are sometimes called “donut instruments” (Bayer et al., 2007), because the instruments are constructed from properties that are located between concentric circles around a given house.

³⁸ Alternative cutoffs of 1 km or 5 km deliver similar results.

Our third set of instruments is the interaction between the second set of IVs and the odds of winning the license lottery. In 2011, Beijing implemented a mandatory quota system to limit the number of new vehicle licenses. Only households that win the license lottery are permitted to purchase vehicles. The probability of winning the license lottery experienced a significant decline throughout our sample period, dropping from 9.4% in January 2011 to 0.7% by the end of 2014.

The increased difficulty in acquiring a vehicle has resulted in heightened demand and elevated housing prices for properties located in desirable areas such as those near the subway or city center. The effect of Beijing's license lottery on the housing market has been demonstrated in previous studies (Lyu, 2022). Similar findings have also been reported for Singapore (Huang et al., 2018).

In essence, the time-varying odds of winning the license lottery can be a valid instrument for housing prices in the housing demand analysis. It is exogenous because the policy-induced variation is unlikely to be correlated with the unobserved housing characteristics ξ_j . Moreover, it is excluded from the housing demand equation because its impact on demand for property j is solely through the price p_j (any remaining effects are directly accounted for through the ease-of-commuting variable, EV_{ij}). We interact the winning odds with the second set of IVs to generate more local variation. As we demonstrate below, the third set of IVs are good instruments with strong first-stage F-statistics and pass the Hensen's J-test.

Table A7 below presents the housing demand estimates for all different combinations of IVs, the F-statistics and the Hensen's J-statistics (the overidentification test of all instruments), as well as the average housing demand price elasticities. The last column that uses all three sets of IVs is the preferred specification that we reported in Table 5 in the main text. The results are reassuring: the parameter estimates are robust across all columns, with the same sign, significance, and similar magnitudes for all coefficients. Second, while IV1 and IV2 and IV1+IV2 have borderline first-stage F-statistics, the other columns all pass the weak-IV test as well as the Sargan-Hansen J-test. These results confirm that our choice of instruments is valid and that the parameter estimates are robust to the choice of instruments.

The different choices of instruments in Table A7 deliver roughly two sets of estimates for housing demand elasticity: from -1.61 to -1.79 in Columns (4), (5), and (7) and -2.3 in Columns (3) and (6). The counterfactual analyses in Table 6 use our preferred choice of IVs in Column (7) and the associated housing demand elasticity of -1.79. To examine the sensitivity of the counterfactual analysis with respect to the choice of instrument, we repeat the exercise using the estimates in Column (6) of Table A7 with the associated housing demand elasticity of -2.3 and report the findings in Table A8. Nearly all results are robust to the choice of instruments, including travel-mode substitution patterns, speed improvement, sorting patterns, etc. The only exceptions are changes in consumer surplus, which are around 15% - 30% smaller than Table 6 across different policy simulations. These patterns are expected, as more elastic demand indicates that households can more easily substitute across different options and hence the welfare damage of removing some options is moderated.

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.

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 [A9](#) and [A10](#). The average price elasticity is -1.91 with the 0.5% random sample and -1.79 with the 1% random sample. The parameter estimates and implied willingness to pay for housing attributes are quite similar across the two samples and hence robust to the size of the choice set. Note that we include district-month-of-sample fixed effects in the second-stage of the housing demand estimation to account for temporal market effects.

C.5 Endogenous Amenities

We do not model endogenous amenity responses to transportation policies (more restaurants and shops near subway stations) as we lack suitable measures of amenities. Here we use housing data to examine whether amenities improved in our sample period as a result of transportation policies. We focus on the effect of subway expansion, as it creates more local variation than driving restrictions (and there is a larger literature on the endogenous amenity responses to subway connections) and congestion pricing was under discussion and not yet implemented during our sample period.

To quantify the effect of subway expansion on amenities, we first use the property-specific mean utility $\hat{\delta}_{jt}$ that is estimated in the first step of the housing analysis to recover neighborhood-period quality:

$$\hat{\delta}_{jt} = \alpha_1 p_{jt} + \mathbf{x}_{jt} \bar{\beta} + \psi_{n(j)t} + \xi_{jt},$$

where p_{jt} and \mathbf{x}_{jt} are property j 's price and attributes, $n(j)$ indexes the neighborhood (i.e., *jiedao*) of property j , and $\psi_{n(j)t}$ represents time-varying neighborhood quality (which are essentially coefficients of neighborhood-year fixed effects). Then we regress neighborhood quality on an indicator variable of the subway connection and two-way fixed effects:

$$\psi_{n(j)t} = \mathbf{1}\{\text{subway}\}_{nt} \beta + \lambda_n + T_t + \varepsilon_{nt} \quad (\text{TWFE}),$$

where $\mathbf{1}\{\text{subway}\}_{nt}$ takes the value of one if neighborhood n experienced new subway expansions in time t . Neighborhoods that never experienced subway expansion serve as a control group, the “never treated” group. We control for neighborhood λ_n and year T_t fixed effects and report the estimates in Column (1) of Table [A11](#) below. Column (2) controls for district-year fixed effects to absorb potential city-wide trends that affect

housing demand.

Given the well-known concerns about using two-way fixed effects models to estimate staggered difference-in-differences research designs (Goodman-Bacon, 2021), we also report the CSDID estimator in Columns (3) and (4) of Table A11 following the approach proposed by Callaway and Sant'Anna (2021). The TWFE and CSDID estimates are qualitatively similar. According to Column (4), subway expansion resulted in an 16.9% increase in neighborhood quality (the average quality of never treated neighborhoods is 2.37). While these estimates are consistent with the hypothesis that subway expansion is associated with improved local amenities, the standard errors are large and neither TWFE nor CSDID estimate is statistically significant. This might reflect the fact that our sample contains few repeated transactions (properties that are transacted both before and after subway expansion), which could make it difficult to detect changes in neighborhood qualities.

As we lack direct measures of amenities (e.g., retail shops, restaurants, entertainment facilities), we do not model how amenities respond to transportation policies implemented in Beijing. Our counterfactual analysis maintains the property quality measures δ_{jt} at their estimated values throughout the simulations and does not allow amenities to change in response to transportation policies. If local amenities improve, as Table A11 provides suggestive supporting evidence, our counterfactual simulations would underestimate the welfare effects of transportation investments. To address this limitation, we conducted a calibration exercise in Panel B of Table 7 in the main text that incorporates the welfare benefits from amenity access, in particular consumption access, using estimates from the existing literature (Miyauchi et al., 2021; Rao, 2021).

D The Simulation Algorithm

D.1 Simulation Algorithm

The notation in this section follows that in the main text. Household-trip characteristics $\{\mathbf{w}\}$, housing attributes $\{\mathbf{X}\}$, and driving 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 consider three congestion measures: city-wide congestion, ring-road-band congestion, and ring-road-quadrant congestion. For the last congestion measure, we divide Beijing into 15 ring-road quadrants as shown in Figure A6. For both ring-road-band congestion and ring-road-quadrant congestion, we use GIS software to split the commuting trip route into corresponding regions. Accordingly, we break down the driving distance of each trip to the region-specific driving distance, where a region is either the entire Beijing city, a ring-road band, or a ring-road-quadrant.

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 “ $D_{T,r}^0$ ” for each region). We define $dist_{ijr,drive}$ as the total driving distance of the trip from house j to the household i’s working place that takes place in region r. We use the observed data values as the starting point. Repeat the following steps until convergence:

1. Based on $D_{T,r}^t$ in region r and price vector \mathbf{p}^t for iteration t ($D_{T,r}^0$ and \mathbf{p}^0 for the first iteration):
 - (a) Update the driving speed for every household:

$$\frac{v_{r,ij}^t - v_{r,ij}^0}{v_{r,ij}^0} = e_{T,r} * \frac{D_{T,r}^t - D_{T,r}^0}{D_{T,r}^0},$$

where $v_{r,ij}^t$ is updated driving speed for household i’s work commute from home j in region r and $e_{T,r}$ is the speed-density elasticity for region r (Table A12.)

(b) Use $v_{r,ij}^t$ from step (a) to revise each commuter’s total driving time, which is the sum of the region specific driving time: $time_{ijk,drive}^t = \sum_r \frac{dist_{ijk,r,drive}}{v_{ijk,r}^t}$. Repeat this for the commuting time via taxi: $time_{ijk,taxi}^t$, which is also affected by congestion.

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

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

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:

$$\sum_{i \in C^{-1}(j)} \frac{1}{Q} \sum_{q=1}^Q w_i \frac{\exp \left(\alpha_i p_j^t + \mathbf{X}_j \beta_i + \sum_k \phi_{kq} EV_{ijk}^t + \xi_j \right)}{\sum_{s \in C(i)} \exp \left(\alpha_i p_s^t + \mathbf{X}_s \beta_i + \sum_k \phi_{kq} EV_{isk}^t + \xi_s \right)} = S_j, \forall j \in J \quad (\text{A3})$$

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 baseline 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 neighborhood n 's housing supply to respond to changes in neighborhood average housing prices, we use a constant supply-price elasticity at 0.53 following Wang et al. (2012):

$$\Delta S_n^t \% = 0.53 * \Delta p_n \%$$

Solving equilibrium price vector \mathbf{p}^t requires iterating the market clearing condition (A3) 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 (A3)), 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 $\mathbf{p}^t = \mathbf{p}^{l+1}$, the fixed point from the inner loop.

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

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

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

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

where the summation inside the large brackets $[.]$ 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 $\|\tilde{D}_{T,r} - D_{T,r}^t\| < \varepsilon_{tol}^{\text{outer}}$ where $\varepsilon_{tol}^{\text{outer}}$ is a pre-set tolerance level, stop. Otherwise, set $D_{T,r}^{t+1} = \varphi D_{T,r}^t + (1 - \varphi) \tilde{D}_{T,r}$ for some $\varphi \in (0, 1)$ and return to step 1.

D.2 Assumptions for the Welfare Analysis

We now detail key assumptions underlying the welfare analysis. First, we assume the lifespan of a property is thirty years, consistent with the empirical evidence in Section 4.2. We also assume that transportation policies last over a property's life span, namely thirty years. It is important to note that the thirty-year assumption does not affect estimates of changes in consumer surplus. It only affects the estimation of costs and revenue associated with implementing transportation policies, which are the total congestion revenue collected, the expenses related to toll collections, and the costs of operating the subway. All three terms are discounted over thirty years at a discount rate of 0.98, see more details below.

We acknowledge that the thirty-year assumption, while supported by the empirical evidence in Section 4.2, may be arbitrary. In order for the congestion pricing policy (with revenue recycling) to be welfare-enhancing, the policy needs to remain in effect for a minimum of eighteen years. Furthermore, it takes another seven years for the toll revenue to fully cover the construction cost of the subway expansion (which is ¥34,000 per household, see below). Altogether, it takes roughly twenty-five years for the congestion pricing policy to be welfare-enhancing and for the toll revenue to fully cover the construction cost of the subway expansion. Subsequently, only a quarter of the annual toll revenue is required to cover the annual operating expenses of the subway.

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, but 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 (¥2 per ride * 8% likelihood to take subway * 500 rides/year * 30 years), orders of magnitude smaller than the net welfare per household as reported in the main text.

D.3 Costs of Implementing Congestion Pricing and Expanding the Subway

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.³⁹ 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

³⁹ <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_CPreport_1.4.18_medium.pdf.

operating costs up by the population of Beijing relative to that of Singapore. To facilitate policy comparison, we calculate the total cost per household (assuming 7.2 million households in Beijing) that include both the initial setup cost and the 30-year discounted annual operating cost at a discount rate of 0.98. This amounts to ¥3000, about 2.5% of the total toll revenue per household during the same period.

The subway costs 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. The annual operating cost is estimated to be ¥1246 per household, which translates to a 30-year discounted total operating cost of ¥69,000 per household (at a discount rate of 0.98). Together, the construction and discounted total operating cost amount to ¥103,000 per household.

D.4 Calculation of Pollution Externalities

Air pollution is an important motivating factor in policy decisions and a relevant consideration for the implementation of anti-congestion policies. Here we discuss how to calculate the welfare benefits of reduced air pollution. Specifically, the expected air pollution damage caused by household i in counterfactual simulations can be measured by:

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

$$B_{ij} = \sum_{k=1}^K EF_{ijk} \times VKT_{ij} \times MD_k, \quad (\text{A4})$$

where B_{ij} is the pollution damage if household i resides in property j . It consists of three terms: EF_{ijk} is the emissions factor that converts the kilometers driven by household i into grams of pollutant k , VKT_{ij} denotes the commuting distance, and MD_k indicates the marginal damage per gram of pollutant k . We explain how each of these three terms is constructed below.

The emissions factor EF_{ijk} , which represents the amount of pollutant emitted per kilometer of driving, is calculated by multiplying a baseline emissions factor BEF_k by a speed adjustment factor γ_{ij} :

$$EF_{ijk} = BEF_k \times \gamma_{ij},$$

The speed adjustment factor accounts for the impact of different driving speeds on emissions, with lower speeds generating greater emissions, all else being equal. The baseline emissions factors and speed adjustment factors are derived from the China National Motor Vehicle Pollutant Emission Standards implemented by the Beijing government in 2013, as specified in [MEP \(2015\)](#).

The baseline emissions factors per kilometer of driving distance are as follows:

$$BEF = \begin{pmatrix} CO_2 & NOx & PM2.5 \\ 248g/km & 0.017g/km & 0.003g/km \end{pmatrix}.$$

The speed adjustment factors γ_{ij} depend on the households' driving speeds v_{ij} and are defined as follows:

$$\gamma_{ij} = \begin{cases} 1.69 & \text{if } v_{ij} < 10\text{km}/h \\ 1.69 - \frac{v_{ij}-10}{25-10} \times (1.69 - 1.26) & \text{if } v_{ij} \in [10\text{km}/h, 25\text{km}/h) \\ 1.26 - \frac{v_{ij}-25}{30-25} \times (1.26 - 0.79) & \text{if } v_{ij} \in [25\text{km}/h, 35\text{km}/h) \\ 0.79 - \frac{v_{ij}-30}{60-30} \times (0.79 - 0.36) & \text{if } v_{ij} \in [35\text{km}/h, 60\text{km}/h) \\ 0.36 & \text{if } v_{ij} \geq 60\text{km}/h \end{cases}$$

The second term in Equation (A4), VKT_{ij} , represents the total driving distance in kilometers for household i residing in property j . The expected daily kilometers traveled by households is calculated as the commuting distance weighted by the driving probability. To estimate the expected lifelong driving distance for household i residing in property j , we multiply the daily commuting distance by 2 trips per day, 250 working days per year, and the assumed housing tenure of 30 years:

$$VKT_{ij} = 2 \times 250 \times 30 \times \Pr(Driving_{ij}) \times Distance_{ij},$$

where $\Pr(Driving_{ij})$ represents the probability that household i living in property j drives to work, and $Distance_{ij}$ denotes the commuting distance.

The third term in Equation (A4), MD_k , represents the marginal damage in dollars (\$) per gram of pollutant and is assumed as follows:

$$MD = \begin{pmatrix} CO_2 & NOx & PM2.5 \\ \$41/ton & \$94,000/ton & \$503,724/ton \end{pmatrix}$$

The marginal damage for CO_2 (also known as the social cost of carbon) is of a global nature and is taken from [EPA \(2016\)](#) at \$41 per ton in 2014 dollars. The marginal damages for NOx and $PM2.5$ are specific to China and are obtained from [Zhou \(2022\)](#). These estimates are derived using an intake fractions approach, as reported in ([Humbert et al., 2011](#); [Zhou et al., 2006](#); [Apte et al., 2012](#)). This approach incorporates local population density and estimates the amount of emitted pollution that is inhaled by the local population. Multiplying the estimated intake value with the concentration-mortality response relationship and the value of statistical life (VSL) delivers marginal damage estimates.

In the counterfactual analyses, we begin by calculating changes in the marginal damage resulting from each household's reduced driving and then aggregate these values over households to obtain total welfare estimates. Dividing the aggregate benefits by the number of Beijing households delivers the welfare benefits per household. Note that the social damage caused by CO_2 emissions is global in nature and affects regions and countries beyond Beijing's borders. Our calculation includes all environmental benefits associated with CO_2 reduction, which represents an upper bound of the benefits accruing to Beijing residents (as some of the benefits from CO_2 reduction extend to regions outside Beijing).

The reduction in emissions and its environmental benefit arise from two main factors. First, households

drive less frequently, and in the case of congestion pricing, they experience shorter commutes, leading to a decrease in vehicle kilometers traveled (VKT_{ij}). Second, the presence of fewer vehicles on the road reduces traffic density, resulting in improved driving speeds, increased fuel efficiency, and better speed adjustment factors.

Among the different policy scenarios, the combined implementation of congestion pricing and subway expansion yields the largest reduction in emissions, amounting to 6% of the aggregate welfare effects. These findings indicate that the benefits of emissions reduction are significant but remain orders of magnitude smaller than the baseline welfare estimates. As a result, pollution benefits do not change the qualitative findings of the paper.

E Additional Counterfactual Analysis

E.1 Ring-road-quadrant Congestion and Speed Variability

We examine three congestion measures: city-wide congestion, ring-road-band congestion, and ring-road-quadrant congestion. For city-wide congestion, we use city-wide speed-density elasticity. For ring-road-band congestion and ring-road-quadrant congestion, we use ring-road speed-density elasticities. Results are similar whether we use city-wide or ring-road-specific elasticities, as this relationship exhibits limited variation within Beijing's urban core (Table A12).

Table A16 presents the results with ring-road-quadrant congestion, which show that introducing greater spatial heterogeneity in congestion has a negligible effect on driving shares relative to the baseline results in Table 6 which assumes city-wide congestion. All differences are less than 0.5 percentage points. The average speed improvements and welfare implications are also similar to those obtained with city-wide congestion measure. These findings are perhaps not surprising, given that most areas in Beijing's urban core are severely congested during rush hour.

Next we examine speed variability across regions. Table A17 below shows the percentage change in speed for each ring-road-quadrant under different policy scenarios. Each row corresponds to a ring-road quadrant, and the policies are presented in the same order as in Table A16. Specifically, Column (1) reports the baseline speed without any policy intervention, Columns (2)-(4) denote driving restriction, congestion pricing, and subway expansion, and the last two columns refer to subway expansion + driving restriction and subway expansion + congestion pricing, respectively.

The baseline rush hour speed varies to some extent across regions of Beijing. The inner city (within the 2nd road) has the lowest baseline speed of 15.8 km/h, while most other regions have a baseline speed of around 21-27.4 km/h. The southwest area between the fifth and sixth ring road (a less populated area) has a higher baseline speed of 31 km/h. The area outside the sixth ring road, the rural suburb of Beijing city, has the highest baseline speed of 41.3 km/h.

Consistent with our main findings, subway expansion has the least impact on speeds, while congestion

pricing combined with subway expansion provides the greatest congestion relief in most areas. While driving restrictions result in greater speed increases within the first two ring roads, they generally produce smaller increases beyond those areas compared to congestion pricing, reflecting the fact that congestion fees are distance-based but the driving restriction is not.

In terms of geographic variation, the percentage speed improvement is highest for the 4-5th and 3-4th ring road areas, followed by the within 2-3 and within 2nd ring road areas. The percentage improvement is lowest for areas outside of the 6th ring road. This pattern is expected, as traffic within the 5th ring road is closer to road capacity, making the anti-congestion effects of transportation policies more pronounced. When individuals move closer to work (which is concentrated in the city center), many trip segments in suburban areas are eliminated, accounting for the greater congestion reduction in the 4-5th ring road area. Subway expansion provides significantly greater congestion relief in the southwestern area of Beijing than in other neighborhoods, as the former lacked subway coverage prior to 2014. Once the area was connected to the subway network, many households switched to public transportation, resulting in fewer cars on the road and a larger speed improvement.

In conclusion, our analysis indicates that incorporating spatially varying speed effects at a more granular level yields findings that align with the intuition outlined in Section 6.1 of the paper. Nonetheless, it does not fundamentally alter the main conclusions of the study.

E.2 Comparisons Between Models with and without Random Coefficients

A key contribution of our empirical exercise is to incorporate a rich set of heterogeneity in both travel mode choices and housing demand. Here we discuss the differences between our preferred model and models without heterogeneity in terms of model fit and substitution patterns, speed improvement, welfare effects, spatial sorting, and equilibrium prices.

There are two types of household heterogeneity that we incorporate in our model. The first is observed heterogeneity, such as age, income, gender, education, and vehicle ownership. The second is unobserved heterogeneity, which consists of households and their individual members' idiosyncratic preferences that cannot be captured by observed demographic differences. The unobserved heterogeneity is represented by random coefficients, and we use the two terms "random coefficients" and "unobserved heterogeneity" interchangeably.

Models with neither demographic controls nor random coefficients do a very poor job of fitting the data, hence we will not consider such models in our discussions below to save space. Instead, we focus on comparing our model which has both types of heterogeneity as in Column 6 of Table 3 and Column 3 of Table 4 in the main text with the model that has a rich set of observed heterogeneity but no random coefficients, namely Column 3 of Table 3 and a housing demand specification that is similar to Column 2 of Table 4 except that the ease-of-commuting is calculated using Column 3 of Table 3 and does not contain random coefficients.

Model fit and substitution patterns. Our model with random coefficients delivers a better fit than the model without random coefficients, which is expected since the former has more parameters. However, as

reported in Table 3 and Table 4, the improvement in log-likelihood is substantial. For example, the sample likelihood in the travel mode analysis improved from -91,119 without random coefficients to -77,706 with five random coefficients, despite the fact that both models contain a rich set of interaction terms between travel mode dummies and observed demographics and trip attributes. As we demonstrate below, the random coefficients not only significantly improve model fit but also play important roles in model predictions.

As the IO literature has emphasized, models without random coefficients behave similarly to logit models and produce unrealistic substitution patterns. In particular, conditioning on demographics, models without random coefficients would predict that commuters substitute “proportionally” toward other modes when their preferred travel mode is unavailable as in Table A18. This means that the probabilities of choosing other modes are roughly proportional to other modes’ market shares. For example, under driving restrictions and congestion pricing, commuters who used to drive are more likely to switch to bus and biking, which are the two most popular commuting modes after driving. They are also less likely to use the subway, which is the 4th commonly used commuting mode. In a similar manner, as the subway had a modest market share of close to 10% in 2008, models without random coefficients predict that subway expansion between 2008 and 2014 pulls a modest number of commuters from other competing travel modes.

In contrast, our model with random coefficients predicts more sensible substitution patterns, where people switch between choices with similar attributes to a larger extent, such as travel time and costs. Under driving restrictions and congestion pricing, people who drive are more likely to switch to the subway and taxi. Similarly, subway expansion pulls more drivers off the road than that predicted by models without random coefficients. According to our preferred estimate, subway expansion from 2008 to 2014 that doubled the length of the subway network boosted the percentage of people who commute via subway from 9.9% to 15.2%, an increase of more than fifty percent. In contrast, models without random coefficients predict a mere one percentage point increase in subway shares.

In many structural urban papers, the reported equilibrium outputs are often too aggregate to discern the fact that models without random coefficients predict unrealistic substitution patterns. As we demonstrate below, these unrealistic substitution patterns have important implications for speed improvement and welfare effects.

Speed implications. Similar to differences in substitution patterns, models with or without random coefficients generate very different predictions regarding the effectiveness of transportation policies in congestion relief.

Models without random coefficients overestimate driving restriction’s effects on congestion reduction while underestimating the effectiveness of congestion pricing and subway expansion. In models with random coefficients, some people have a strong preference for driving and are *more* likely to drive during the days when they are not subject to driving restrictions. This rebound effect is much less pronounced in models without random coefficients, which then overstate the aggregate effectiveness of driving restrictions on congestion reduction. Specifically, the speed improvement from driving restrictions is 4.59 km/h in models

without random coefficients vs. 3.83 km/h in our model with random coefficients.

In a similar vein, as models without random coefficients underestimate the subway usage after its expansion, the predicted speed improvement is minimal at 0.16 km/h. Our model predicts that subway expansion is associated with a speed improvement of 1.49 km/h, which is much more reasonable considering the substantial scale of subway investment in Beijing.

Another important difference between these models is the travel demand sensitivity. Models without random coefficients rationalize the popularity of driving (which is more expensive than public transportation) by low sensitivity to driving costs. In contrast, models with random coefficients rationalize high driving shares with idiosyncratic preferences: while some commuters prefer to drive regardless of the costs (hence the high mode shares), the population on the whole is sensitive to driving costs. The estimated driving elasticity confirms this. A 1% increase in driving costs reduces driving share by 0.038 percentage points without random coefficients and 0.07 percentage points with random coefficients. Consequently, models without random coefficients underestimate the effect of congestion pricing on congestion: a toll charge of ¥1.13 per km improves speed by 2.97 without random coefficients compared to 3.83 km/h with random coefficients.

Welfare effects. The IO literature has shown that models without random coefficients tend to exaggerate the welfare implications of policy changes. As shown by simulation exercises in [Petrin \(2002\)](#), logit errors in models without random coefficients play a bigger role in explaining differences in choices by observably similar households. These logit errors enter into the welfare calculation and constitute part of welfare gain/losses. The role of logit errors is much reduced in models with random coefficients, which predict more reasonable welfare effects. Our results confirm this conventional wisdom.

Table [A19](#) below reports changes in consumer surplus resulting from transportation policies. Driving is the most popular travel mode and accounts for 42% of commuting trips by rich households and 21% of commuting trips by the poor. In models without random coefficients, restricting driving either through driving restrictions or congestion pricing (without recycling toll revenue) would lead to huge losses for consumers at about 5-10 times the magnitude predicted by models with random coefficients.

To graphically illustrate the differences in welfare implications, Figure [A15](#) plots the distribution of predicted welfare changes with and without random coefficients. There are two noticeable patterns for the model without random coefficients relative to those with random coefficients. First, the mode and average of welfare losses under driving restrictions and congestion pricing are much higher than those from the model with random coefficients. Second, the dispersion of welfare losses is much wider. Both patterns echo [Petrin \(2002\)](#) and illustrate the outsized role of logit errors.

There are two countervailing forces that drive different predictions on the welfare effects of subway expansion between models with and without random coefficients. On the one hand, logit errors imply that models without random coefficients would overestimate the welfare benefits of subway expansions. On the other hand, subway expansion shortened the commuting time for subways by more than half. As documented above in the discussion on speed adjustment (and in Table 3 on the estimated Value of Time), models with-

out random coefficients underestimate travel demand's sensitivity to commuting costs and underestimate the Value of Time. The attenuation bias on the benefits of shorter commuting time dominates the effect of logit errors. Consequently, the welfare gains of subway expansion from models without random coefficients are only half of those from models with random coefficients.

Sorting patterns. The mechanisms underlying sorting responses are more complex than the mechanisms for substitution patterns or speed changes. In the scenario with a fixed housing supply, households shuffle residential locations and swap properties among themselves. The average distance to the subway system across all households is the same whether or not the model has random coefficients. If rich households move closer to subway stations, by construction poor households would be displaced and move farther away from subway stations. We report the average distance to subway separately for the rich and poor in the paper to illustrate this displacement effects. Since models with and without random coefficients both contain income and a rich set of observed demographics, they predict similar displacement effects. For this reason, our discussion here focuses on the distance to work (Table A20).

In response to transportation policies, households can either adjust travel modes or change their residential locations. Models without random coefficients exaggerate welfare losses under congestion pricing in terms of travel mode adjustments, which leads to stronger incentives for households to move closer to work to mitigate the negative impacts of congestion tolls. As a result, the associated sorting responses are much stronger in models without random coefficients than those in models with random coefficients. Similarly, as models without random coefficients underestimate gains from riding the subway, the sorting response from subway expansion is muted in models without random coefficients relative to models with random coefficients.

Both models, with and without random coefficients, predict modest sorting responses under driving restrictions, with the model without random coefficients showing slightly larger effects. This is because driving restrictions create conflicting forces. On the one hand, they encourage relocation closer to workplaces. On the other hand, the speed improvement disproportionately benefits longer-distance trips. These opposing effects largely offset each other, resulting in minimal sorting responses.

Equilibrium price changes. Here we examine changes in the price-distance gradient as a result of transportation policies. Traditional urban literature uses the distance to the central business district (CBD) to measure a location's centrality. Since we observe job locations, we define CBD as the centroid of all observed job locations, which is a better measure of the centrality of work locations.

We first predict equilibrium housing prices with no policy and then predict housing prices separately for each transportation policy. We repeat this exercise for both our preferred model with random coefficients and the model without random coefficients and report the estimated changes in the price-distance gradient in Table A21 below. All regressions control for property fixed effects and policy fixed effects.

Echoing the discussions on sorting responses, under congestion pricing, the model without random coefficients predicts a very large price appreciation for properties that are centrally located compared to our model

with random coefficients. Under subway expansion, the model without random coefficients predicts a modest urban-sprawl effect and a small increase for properties near new subway stations (which are farther away from the central districts) relative to the model with random coefficients. As the sorting responses are the smallest under driving restrictions, the changes in the price-distance gradient are modest for both models.

E.3 Income Sorting

The reduced-form results in Appendix Section B.1 present evidence that neighborhoods with newly acquired access to the subway system experience a greater fraction of property transactions. Here we examine changes in household income in the simulation exercises.

Specifically, we simulate equilibrium outcomes for the 2014 mortgage cohort before and after the subway expansion and examine changes in household income in areas near new subway stations. We report results in Table A22. All regressions include property fixed effects and policy fixed effects. As expected, the household income for properties that gained access to the expanded subway network increased significantly. These patterns echo the evidence in Appendix Figure A10, where the simulated housing price gradient with respect to the subway distance steepens under the 2014 network. This is because the 2014 network is larger and hence the proximity to this network is more valuable to commuters. These results indicate that people actively sort in our framework, and high-income households outbid low-income households to move closer to the subway stations in response to transportation policies.

Finally, we replicate this exercise for all transportation policies and plot changes in household income for each TAZ in Appendix Figure A9 using Beijing's map. As expected, household income in central regions increases under driving restrictions and congestion pricing and decreases under subway expansion.

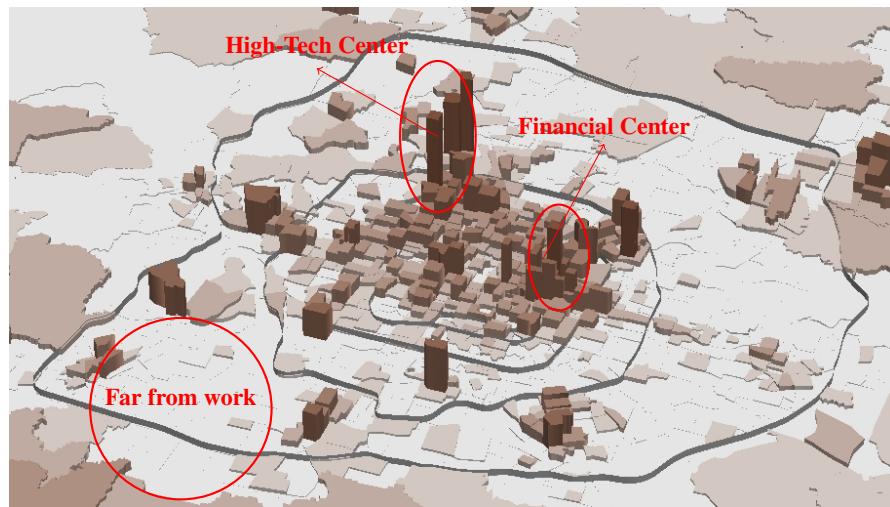
F Figures and Tables

Figure A1: Subway Network Expansion in Beijing



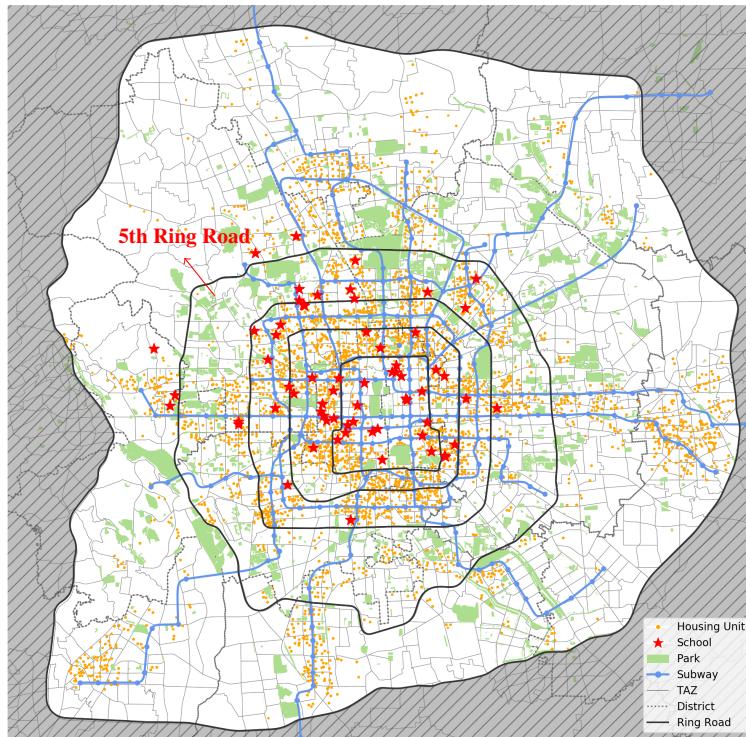
Note: The subway system in Beijing expanded from 2 lines to 22 lines from 1999 to 2019. From 2007 to 2018, 16 new subway lines were built with a combined length of over 500km. By the end of 2019, the Beijing Subway was the world's longest and busiest subway system with a total length of nearly 700km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

Figure A2: Job Density



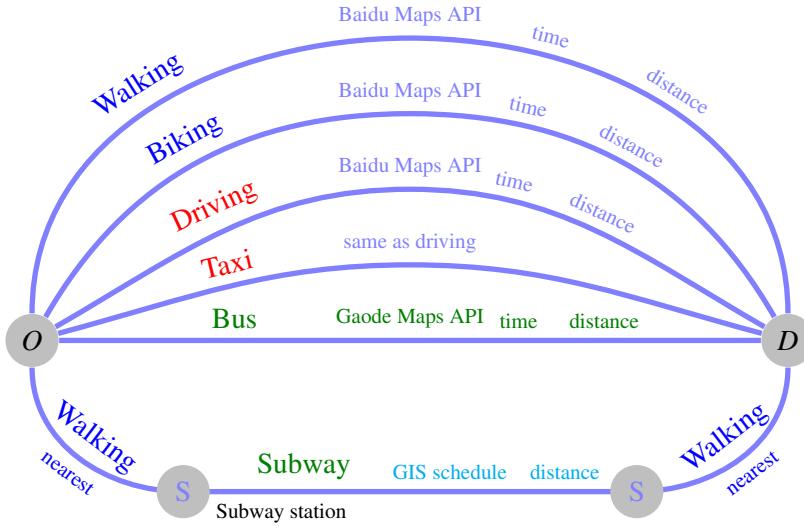
Note: This figure plots work density by TAZ based on work locations from the mortgage data. Darker colors/taller shapes indicate greater work density.

Figure A3: Housing, Amenities, and Transportation Network



Note: The figure shows 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



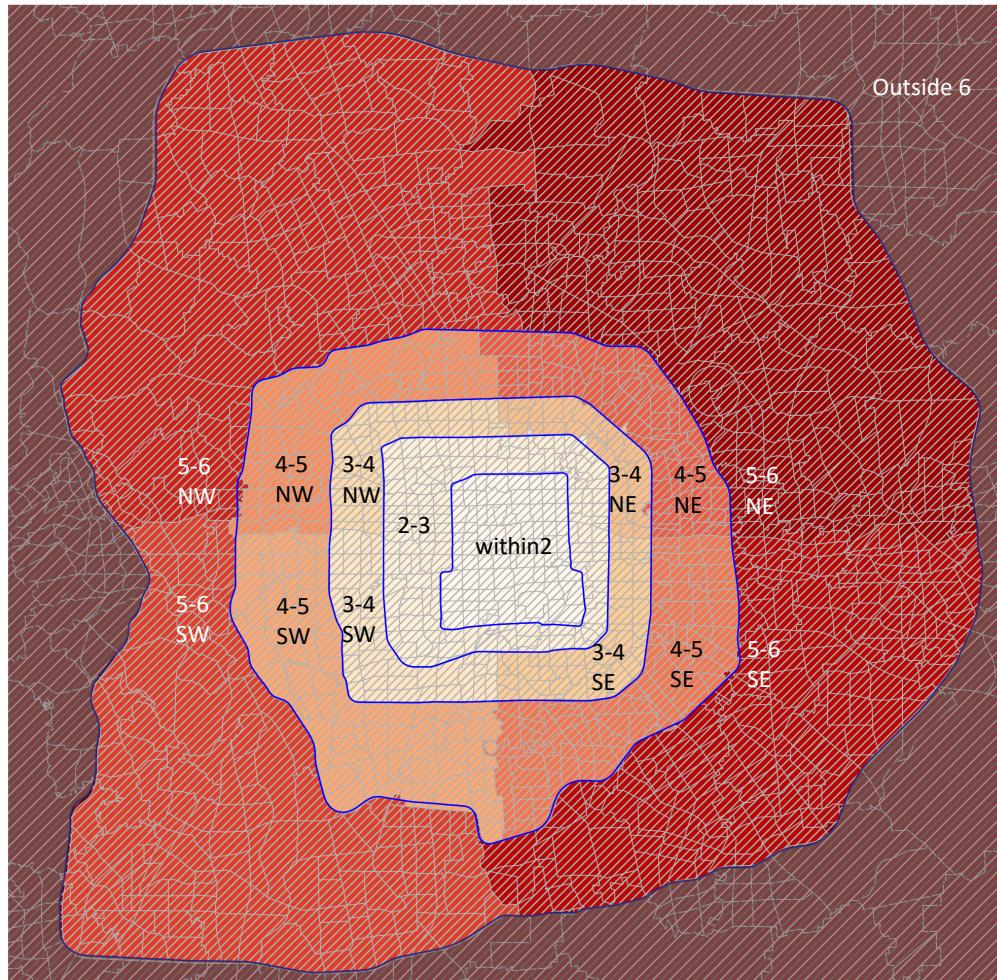
Note: Travel time and distance using walk, bike, car, and taxi are constructed using the Baidu Maps API based on the survey reported departure time and day of the week. Bus travel time and distance are constructed using Gaode Maps API because it provides the number of transfers and walking time between 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 subway trips, 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 timetables in 2010 and 2014.

Figure A5: an Example of Travel Routes



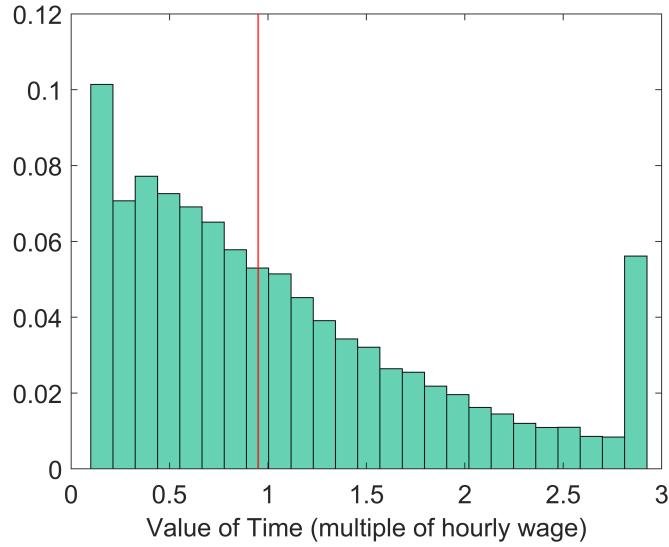
Note: The figure shows travel time and cost for a particular trip that started at 7:09am on 9/12/2010 for each mode. The chosen mode was subway. The left panel shows the straight-line direction of travel, while the right panel shows the time, monetary cost, and distance for each travel mode and the corresponding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

Figure A6: Traffic Analysis Zones and Ring-Road-Quadrant Regions



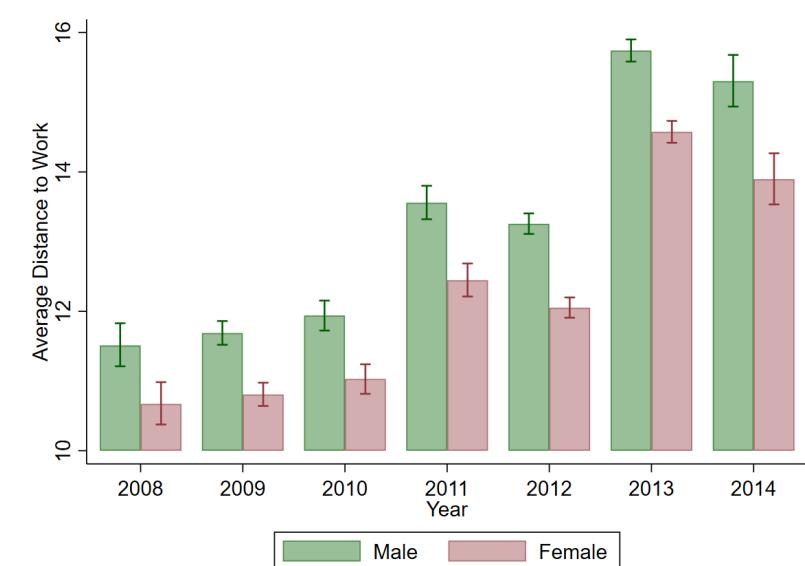
Note: This map illustrates the fifteen localized congestion regions that we create by aggregating TAZs based on their location within each ring road and quadrant (NW, SW, NE, SE). For each of the three areas encircled by the 3rd to the 6th ring road, we divide them into four regions. We also create one region for TAZs between the 2nd and 3rd ring roads, one region for TAZs within the 2nd ring road, and one region for TAZs outside the 6th ring road.

Figure A7: Implied Value of Time Distribution from the Mode Choice Estimation



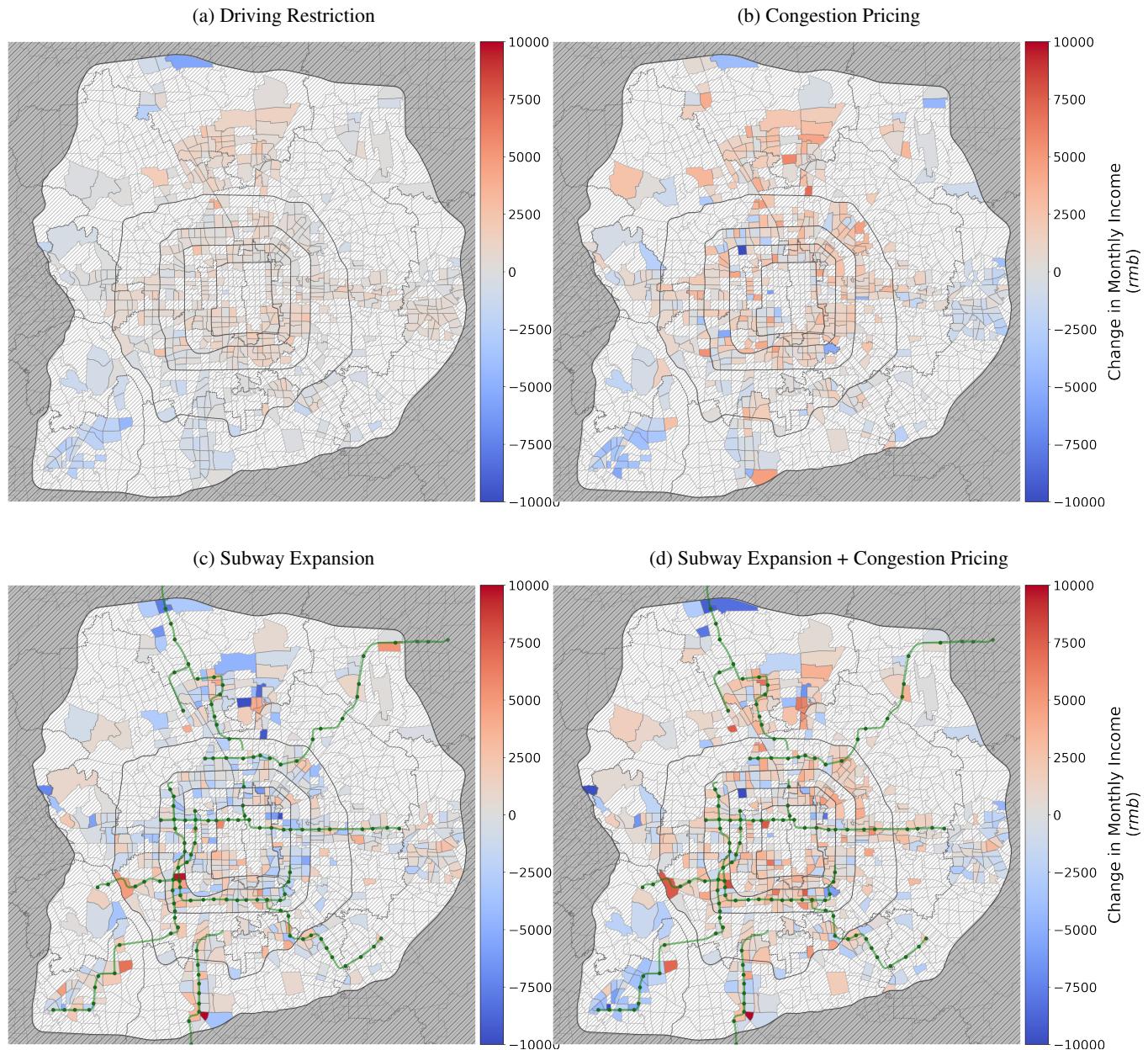
Note: The figure plots the estimated distribution of value of time (VOT) in terms of hourly wage that is based on Column (6) of Table 3. VOT is measured by the ratio of the preference for travel time over the preference for monetary travel cost. The preference for travel time has a winsorized (at the 5th and 95th percentiles) chi-square distribution with three degrees of freedom, while the preference for monetary travel cost 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 A8: Distance to Work by Gender in km



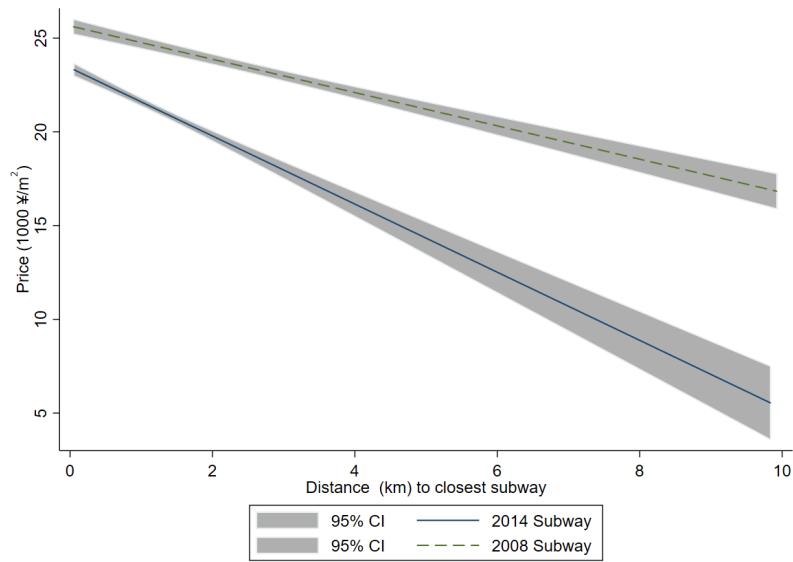
Note: the figure displays the average distance to work by year based on the mortgage data for male (green bars) and female (red bars) household members, separately. The whiskers denote 95% intervals. Males have longer commutes than females. The increasing commuting distance over time reflects the expansion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

Figure A9: Changes in Income from Counterfactual Simulations (in ¥)



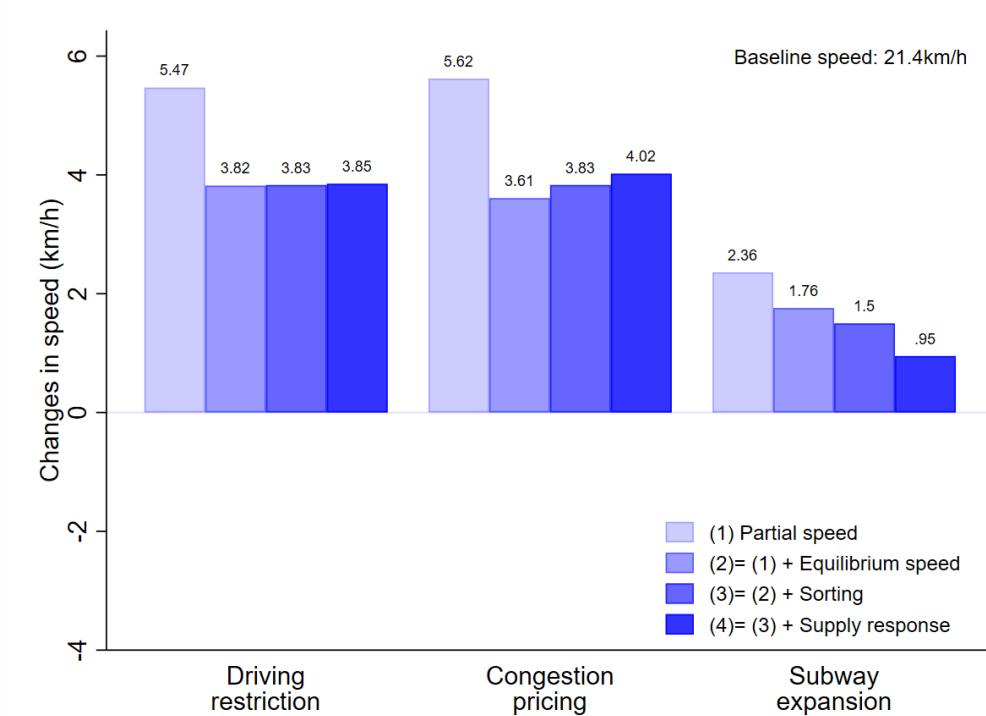
Note: This figure illustrates simulated changes in household annual income (in ¥) across TAZs under different counterfactual policies (relative to the no policy scenario). The results are based on the simulations in Table 6 that allow for household sorting, fix housing supply, use estimates including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

Figure A10: Price Gradient under the 2008 and 2014 Subway Network



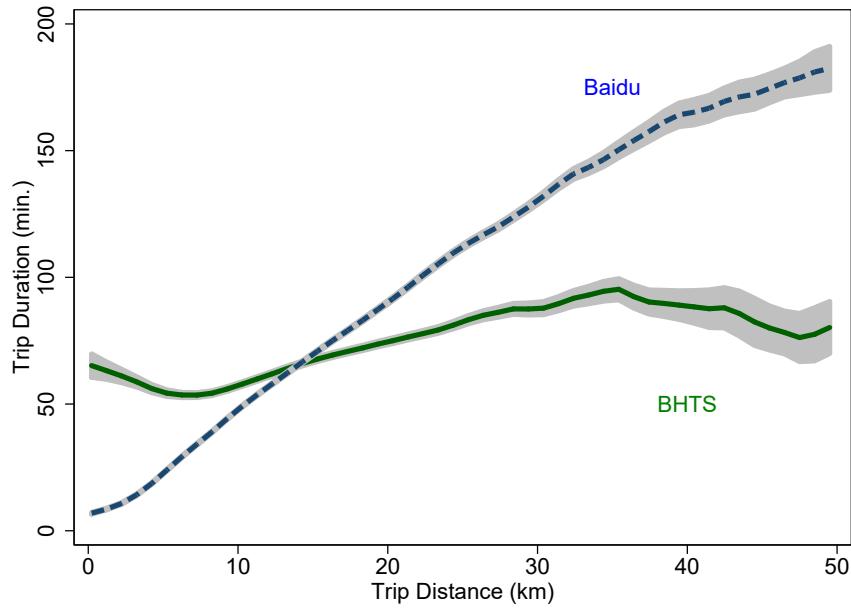
Note: This plot shows the simulated bid-rent curve with respect to subway distance under the 2008 and 2014 network, respectively. The results corresponds to Columns (1) and (4) of Table 6. The gradient of the bid-rent curve under the 2014 subway system (-¥1900/m² per km) is steeper than the 2008 subway 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 A11: Speed Adjustment Decomposition for Different Transportation Policies



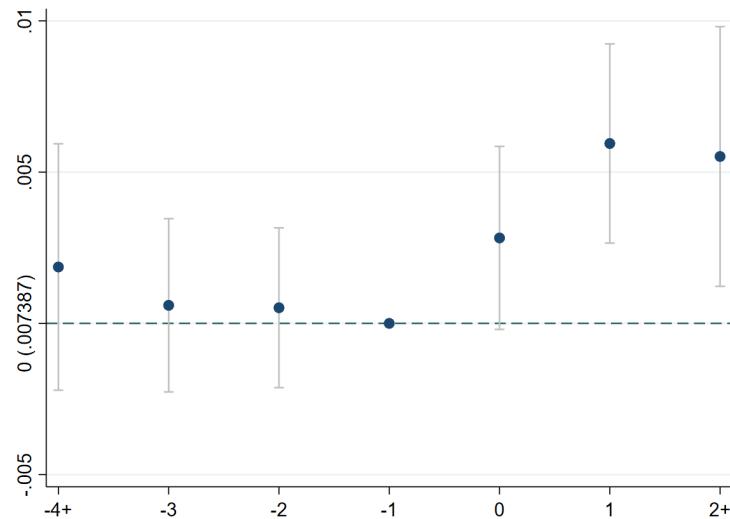
Note: This figure decomposes speed changes along four adjustment margins. For each policy, the bars display the cumulative speed changes incorporating previous margins. The partial speed effect allows the driving speeds to adjust one time via Equation (13), but does not impose the transportation sector's clearing condition. The second bar additionally incorporates the full equilibrium speed effect and additional changes in speed when traffic speeds adjust further to clear the transportation sector. The third bar includes household sorting. The last bar allows housing supply to adjust at the neighborhood level in addition to the three channels above.

Figure A12: Recall Bias



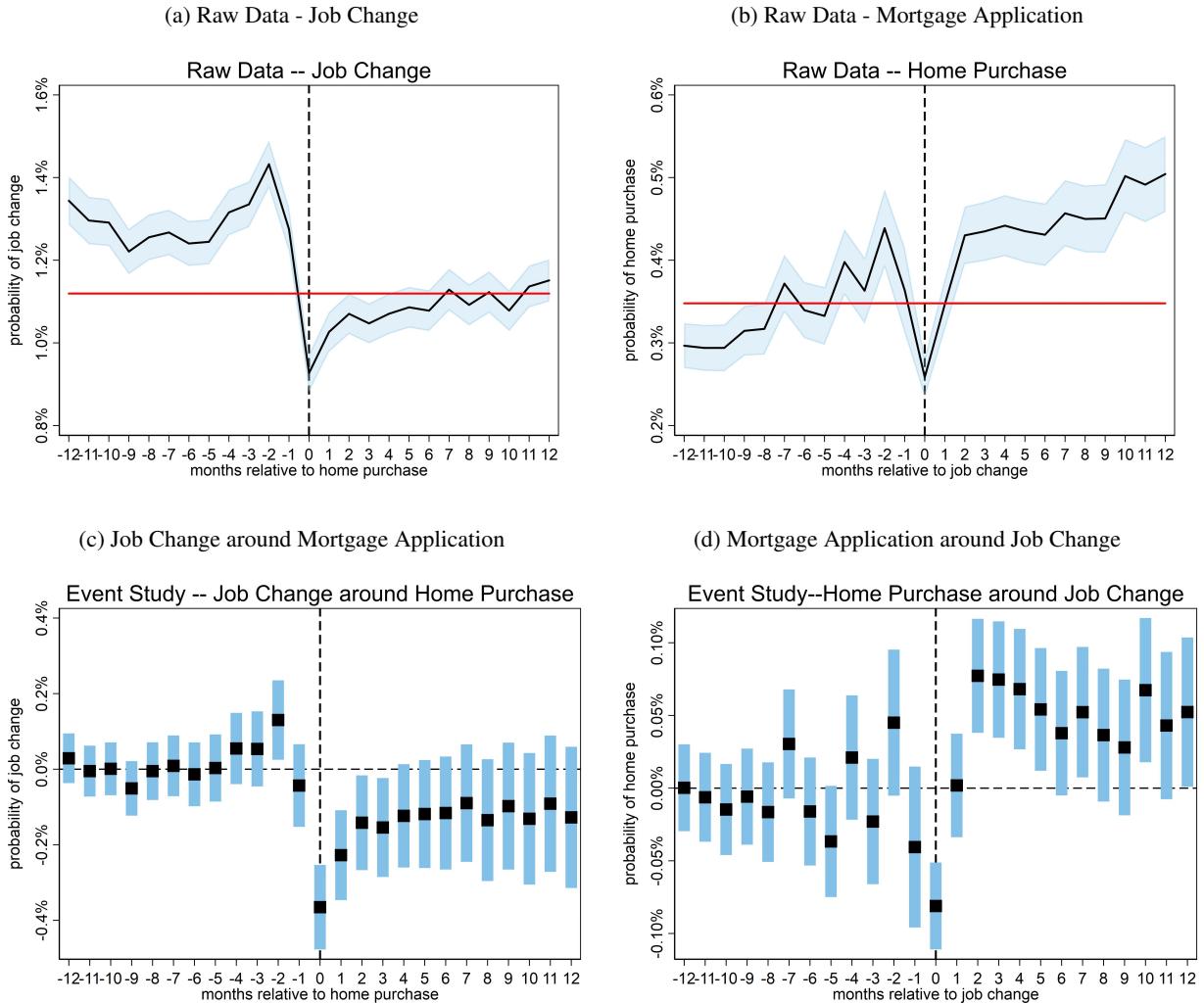
Note: This figure shows the relationship between travel time and travel distance. The solid green line is self-reported subway travel time in the Beijing Household Travel Survey (BHTS). The dashed line is the travel time obtained via Baidu Map API (Baidu). The differences reflect the recall bias. We use calculated travel time from Baidu in most of our analyses and only use self-reported travel time as a robustness analysis.

Figure A13: Propensity to Purchase Houses in Neighborhoods with New Subway Stations



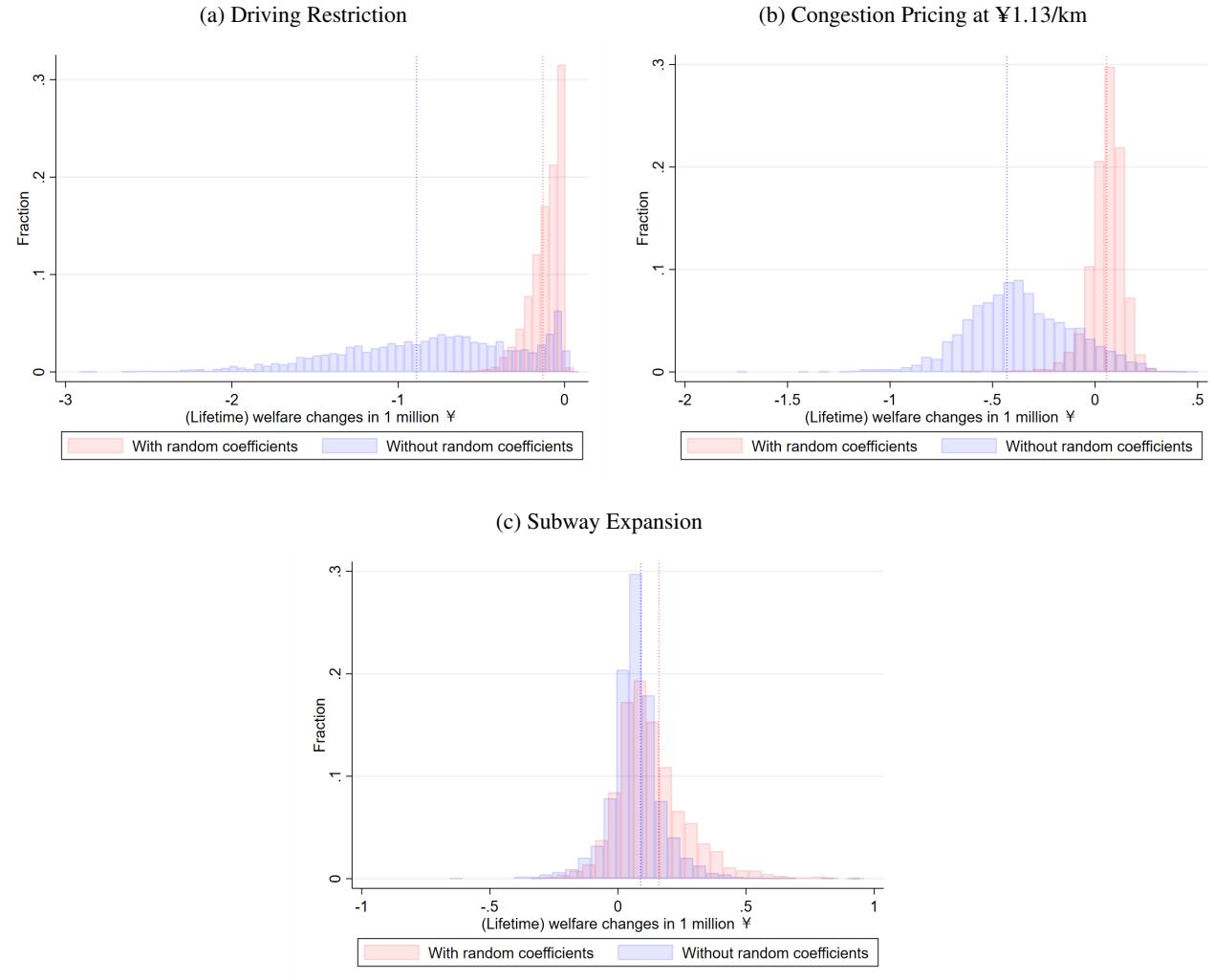
Note: This figure uses mortgage data to illustrate changes in the fraction of property transactions that are accounted for by neighborhoods that gained new subway stations after the subway expansion. The number of observations is 1,395 neighborhood years (9 years with 155 neighborhoods). There were five major waves of subway expansions affecting 90 out of 155 neighborhoods in Beijing from 2008 to 2014. We fail to reject the null hypotheses of no significant pre-trend ($p\text{-value}=0.71$).

Figure A14: Job Change and Housing Purchase



Note: Panel (a) shows the raw trend of the job change frequency before and after home purchase (denoted as month zero). The shadow area represents the 95 percent confidence interval. The red horizontal line refers to the monthly average probability of job change during the sample period. Panel (b) shows the raw trend of home purchases with month zero denoting a job change. Panels (c) and (d) are event studies with year-month and employee fixed effects, where the 13-18 months before the event is taken as the baseline period. Standard errors are clustered by employee. The bars represent the 95 percent confidence intervals.

Figure A15: Distribution of Welfare Changes with and without Random Coefficients



Note: This figure plots the distribution of welfare changes per household under different policies in a model with random coefficients (our preferred specification) and a similar model without random coefficients. Bars denote the fraction of observations in each bin and vertical dashed lines denote average welfare changes per household. Welfare is consumer surplus plus toll revenue and environmental benefits minus subway costs. Consumer surplus estimates are recovered from housing transaction prices and should be interpreted as total consumer surplus over a property's life span. Toll revenues and subway costs are the discounted sum over thirty years (which is approximately the lifespan of a property).

Table A1: Index of Mathematical Notation in the Sequence of Appearance in the Text

Symbol	Description
<i>i</i>	Household
<i>k</i>	Household member (e.g., borrower/co-borrower)
<i>j</i>	House
<i>t</i>	Year
ℓ	Elements of housing attributes
<i>m</i>	Commuting travel mode
<i>n</i>	Neighborhoods (Jiedao) in Beijing
<i>r</i>	Regions of Beijing
<i>s</i>	Road segments
<i>Variables</i>	
<i>U</i>	Utility from housing demand
x	Housing attributes
<i>p</i>	Housing price
<i>EV</i>	Ease-of-commute measure
ξ	Unobserved housing attributes
<i>y</i>	Household income
ζ_i	Idiosyncratic preference for ease-of-commute
<i>P</i>	Housing choice probability
<i>u</i>	Utility from commuting
<i>time</i>	Commuting time
<i>v</i>	Commuting speed
<i>cost</i>	Commuting cost
w	Commuting mode-commuter attributes
<i>R</i>	Commuting mode choice probability
<i>D</i>	Housing demand
<i>S</i>	Housing supply
<i>D_T</i>	Driving demand
<i>S_T</i>	Traffic density that can be sustained at a certain speed under the existing road capacity
<i>dist</i>	Commuting distance
X	Variables affecting speed-density relationship
W	Households' welfare measure in ¥
<i>Parameters on</i>	
α	Housing price
β	Housing attributes
ϕ	Ease-of-commute
θ_{im}	Commuting mode-specific random coefficient
γ	Time & cost of commuting
η	Commuting mode-commuter attributes
ϵ	Price elasticity of housing supply
ϵ_T	The elasticity of traffic density that can be sustained under existing road capacity w.r.t speed
e_T	Speed-density elasticity
β_T	Speed-density estimation covariates
μ	Household-specific utility from housing demand
δ	Housing demand mean utility

Table A2: Travel Mode Choices Estimation with Fine Spatial Controls

	(1) Baseline	(2) (1)+ Drive*District FE	(3) (1)+ Drive*Neighborhood FE
Travel Cost/Hourly Wage (γ_2)	-2.531 (0.065)	-2.552 (0.065)	-2.508 (0.067)
Random coefficients on travel time (μ_γ)			
Travel Time	-0.931 (0.012)	-0.930 (0.012)	-0.904 (0.012)
Random coefficients on mode-specific constants (σ_m)			
Driving	3.391 (0.054)	3.358 (0.054)	3.377 (0.054)
Subway	4.470 (0.142)	4.535 (0.144)	4.308 (0.139)
Bus	3.851 (0.056)	3.853 (0.056)	3.774 (0.056)
Bike	3.887 (0.054)	3.881 (0.054)	3.849 (0.054)
Taxi	4.203 (0.353)	4.212 (0.353)	3.081 (0.279)
Mode*Year FE	Yes	Yes	Yes
Mode*Trip related FE	Yes	Yes	Yes
Mode*Demographic FE	Yes	Yes	Yes
Drive*District FE		Yes	
Drive*Neighborhood FE			Yes
Log-likelihood	-77706	-77609	-76938
Implied mean VOT	0.956	0.947	0.937
Implied median VOT	0.846	0.838	0.829

Note: The number of observations is 73,154. Column (1) replicates Column (6) of Table 3. Column (2) adds interactions of driving and district fixed effects. Column (3) further includes interactions of driving and neighborhood fixed effects. The other controls are the same as in Table 3. Robust standard errors are reported in parentheses.

Table A3: Travel Mode Choices Estimation with Average Driving Speed

	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-specific constants (σ_m)						
Driving					3.424 (0.051)	3.412 (0.054)
Subway						4.381 (0.142)
Bus						3.889 (0.056)
Bike						3.869 (0.055)
Taxi						4.829 (0.401)
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	-116318	-109944	-91131	-87592	-85275	-77832
Implied mean VOT	1.203	0.337	0.333	1.781	1.600	0.932
Implied median VOT	1.203	0.337	0.333	1.576	1.415	0.824

Note: The model estimates use the average driving speed in each ring road to construct the travel time measure. The number of observations is 73,154. The specifications follow those in Table 3 and include an increasingly rich set of fixed effects interacting with travel model dummies. Trip-related FEs include trip distance bins (whether distance is shorter than two km or between 2-5 km) and origin and destination ring road dummies (whether the trip origin is within the fourth ring road, and whether the trip destination is within the fourth ring road). Demographics FEs include a respondent's age, age squared, gender, education, car ownership, and whether the household has more than one commuter. The first three specifications are multinomial logit while the last three add random coefficients. Robust standard errors are displayed below parameter estimates.

Table A4: Travel Mode Choices Estimation with Reported Travel Time

	Logit			Random Coefficient		
	(1)	(2)	(3)	(4)	(5)	(6)
Travel Time (γ_1)	-0.513 (0.005)	-0.262 (0.006)	-0.186 0.007			
Travel Cost/Hourly Wage (γ_2)	-0.426 (0.023)	-0.778 (0.027)	-0.558 0.034	-1.343 (0.040)	-1.383 (0.049)	-2.432 (0.064)
Random coefficients on travel time (μ_γ)						
Travel Time				-0.920 (0.008)	-0.851 (0.008)	-0.872 (0.012)
Random coefficients on mode-specific constants (σ_m)						
Driving					4.182 (0.059)	4.471 (0.059)
Subway						2.899 (0.117)
Bus						3.514 (0.055)
Bike						3.697 (0.056)
Taxi						11.786 (1.158)
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	-109249	-103742	-85147	-84852	-80851	-74327
Implied mean VOT	1.307	0.888	1.017	1.528	1.836	1.239
Implied median VOT	1.307	0.888	1.017	1.352	1.624	1.096

Note: The model estimates use the reported travel time from the BHTS travel survey data for the chosen mode and GIS-constructed travel time for non-chosen modes as the travel time measure. The number of observations is 73,154. The specifications follow those in Table 3 and include an increasingly rich set of fixed effects interacting with travel model dummies. Trip-related FE includes trip distance bins (whether distance is shorter than two km or between 2-5 km) and origin and destination ring road dummies (whether the trip origin is within the fourth ring road, and whether the trip destination is within the fourth ring road). Demographics FE includes a respondent's age, age squared, gender, education, car ownership, and whether the household has more than one commuter. The first three specifications are multinomial logit while the last three add random coefficients. Robust standard errors are displayed below parameter estimates.

Table A5: Housing Demand Without EV Terms - Linear Parameters

Variables	OLS (1)	OLS (2)	IV1 (3)	IV2 (4)	IV2+IV3 (5)	All IVs (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: This table is similar to Table 5 except the dependent variable is the recovered population-average utility $\{\delta_{ji}\}_{jt}$ when EV is excluded from housing attributes. The number of observations is 79,894. 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. IV1 is the number of homes that are within 3km from a given home, outside the same complex, and sold in a two-month window. IV2 is the average attributes of these homes (building size, age, log distance to park, and log distance to key school). IV3 is the interaction between IV2 and the winning odds of the license lottery. The winning odds decreased from 9.4% in January 2011 to 0.7% by the end of 2014. Standard errors are clustered at the neighborhood level.

Table A6: Alternative Specification for First-Stage Housing Demand Model with More Covariates

	(1) Main Table		(2) Demo interactions		(3) (2)+EV interactions		(4) (3)+Edu interactions	
	Para	S.E.	Para	S.E.	Para	S.E.	Para	S.E.
Demographic Interactions								
Price (in 1 million RMB)*ln(income)	1.030	0.016	0.753	0.021	0.759	0.022	0.697	0.022
Ln(distance to key school)*age in 30-45	-0.420	0.010	-0.341	0.012	-0.358	0.014	-0.370	0.014
Ln(distance to key school)*age > 45	-0.123	0.021	0.026	0.024	-0.045	0.028	-0.097	0.028
Ln(home size)*age in 30-45	1.486	0.029	1.119	0.041	1.111	0.042	1.114	0.042
Ln(home size)*age > 45	2.746	0.061	1.927	0.087	1.873	0.087	1.880	0.088
Price (in 1 million RMB)*age in 30-45			0.276	0.023	0.281	0.023	0.317	0.023
Price (in 1 million RMB)*age > 45			0.600	0.045	0.611	0.044	0.758	0.046
Ln(distance to key school)*ln(income)			-0.280	0.011	-0.203	0.013	-0.183	0.013
Ln(home size)*ln(income)			0.573	0.040	0.603	0.040	0.600	0.041
Price (in 1 million RMB)*college							0.526	0.030
Ln(home size)*college							0.006	0.052
Ln(distance to key school)*college							-0.177	0.017
EV and interactions								
EV_{Male}	0.755	0.006	0.755	0.040	0.407	0.046	0.424	0.047
EV_{Female}	0.893	0.006	0.890	0.084	0.506	0.047	0.481	0.049
$EV_{Male} * \ln(\text{income})$					0.072	0.010	0.052	0.010
$EV_{Female} * \ln(\text{income})$					0.086	0.010	0.072	0.010
$EV_{Male} * \text{Age in 30-45}$					0.002	0.010	0.009	0.011
$EV_{Male} * \text{Age} > 45$					0.032	0.011	0.065	0.020
$EV_{Female} * \text{Age in 30-45}$					-0.038	0.020	-0.032	0.011
$EV_{Female} * \text{Age} > 45$					-0.206	0.011	-0.180	0.022
$EV_{Male} * \text{College}$							0.091	0.013
$EV_{Female} * \text{College}$							0.107	0.013
Random Coefficients								
$\sigma(EV_{Male})$	0.379	0.013	0.371	0.033	0.359	0.013	0.346	0.014
$\sigma(EV_{Female})$	0.482	0.012	0.475	0.017	0.465	0.012	0.450	0.012
Log-likelihood	-168807.8		-168375		-168259		-167554	
Avg. Price Elasticity	-1.79		-1.74		-1.75		-1.75	

Note: The table reports MLE estimates of the non-linear parameters in the housing demand model with an increasing number of covariates, using mortgage data from 2006-2014 with 77,696 observations. Column (1) reproduces our preferred specification of Column (3) in Table 4. Column (2) interacts all property attributes with age and income. Column (3) adds interactions between EVs and age and income. Column (4) further includes interactions between the college dummy and all property/EV attributes.

Table A7: Housing Demand Linear Parameters with Different Instruments

	(1) IV1	(2) IV2	(3) IV3	(4) IV1+IV2	(5) IV2+IV3	(6) IV1+IV3	(7) All IVs
Price (¥mill.)	-7.091*** (1.640)	-6.283*** (0.867)	-7.023*** (0.570)	-6.562*** (0.722)	-6.454*** (0.583)	-7.023*** (0.570)	-6.596*** (0.534)
Ln(home size)	4.721 (2.927)	3.331** (1.505)	4.623*** (1.071)	3.818*** (1.289)	3.631*** (1.022)	4.623*** (1.071)	3.879*** (0.969)
Building age	-0.144*** (0.040)	-0.125*** (0.020)	-0.142*** (0.015)	-0.132*** (0.017)	-0.129*** (0.014)	-0.142*** (0.015)	-0.132*** (0.013)
Floor area ratio	-0.019 (0.036)	-0.023 (0.032)	-0.024 (0.036)	-0.023 (0.033)	-0.023 (0.033)	-0.024 (0.036)	-0.023 (0.034)
Ln(dist. to park)	-0.475** (0.222)	-0.389*** (0.117)	-0.471*** (0.117)	-0.420*** (0.115)	-0.408*** (0.101)	-0.471*** (0.117)	-0.424*** (0.103)
Ln(dist. to key school)	0.210 (0.213)	0.323** (0.139)	0.241* (0.124)	0.292** (0.129)	0.304** (0.121)	0.241* (0.124)	0.288** (0.118)
Year-Month-District FE	Y	Y	Y	Y	Y	Y	Y
Neighborhood FE	Y	Y	Y	Y	Y	Y	Y
Kleibergen-Paap rk Wald							
F statistic	9.88	10.48	23.03	11.28	14.22	23.03	14.22
P-value for Hansen's J-test	.	0.03	0.97	0.08	0.10	0.97	0.19
Avg. Price Elasticity	-2.42	-1.40	-2.34	-1.75	-1.61	-2.34	-1.79

Note: The number of observations is 77,696. The dependent variable is the population-average utility δ_{jt} recovered using parameter estimates in Column (3) of Table 4. IV1 is the number of properties that are within 3km from property j , excluding the chosen complex, and sold in a two-month window around property j 's sale. IV2 constitute the average attributes of these homes (building size, age, log distance to park, and log distance to key school). IV1 and IV2 are often called the “BLP” IVs following (Berry et al., 1995). IV3 are the interactions between IV2 and the winning odds of the license lottery. The winning probability decreased from 9.4% in January 2011 to 0.7% by the end of 2014. Columns (4)-(7) use different combinations of IVs. Standard errors are clustered at the neighborhood level.

Table A8: Simulation Results with Alternative Housing Demand Elasticity

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	41.58	21.39	-7.09	-3.36	-3.41	-5.37	-2.10	-1.62	-8.46	-4.57	-5.16	-6.38
Subway	9.03	10.78	1.24	0.67	0.79	0.94	4.64	6.13	5.81	6.51	5.27	6.89
Bus	22.47	30.49	1.78	0.62	0.58	1.26	-1.57	-2.55	0.28	-1.59	-0.79	-1.04
Bike	15.98	24.03	1.58	0.79	0.76	1.77	-0.82	-1.68	0.49	-0.99	-0.16	-0.16
Taxi	2.19	1.32	1.20	0.55	0.64	0.57	-0.16	-0.11	0.90	0.36	0.40	0.36
Walk	8.75	12.00	1.29	0.73	0.65	0.83	0.00	-0.17	0.99	0.27	0.44	0.33
Speed (km/h)	21.40		3.83		3.83		1.54		5.13		5.35	
Panel B: sorting outcomes												
Distance to work (km)	18.59	15.67	0.01	0.01	-0.16	-0.07	0.32	0.21	0.38	0.20	0.13	0.15
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 (+)	-156.4		-21.5	-73.8	-59.0	163.5	82.3	-3.0	53.7	79.9	24.7	
Toll revenue (+)			103.7		103.7						96.0	96.0
Subway cost (-)					103.0		103.0	103.0	103.0	103.0	103.0	103.0
Pollution Reduction (+)	4.25		4.25	4.25	4.25	1.69	1.69	5.79	5.79	6.03	6.03	
Net welfare	-152.1		-17.3	34.1	48.9	62.2	-19.0	-100.2	-43.5	78.9	23.8	

Note: This table is the same as Table 6, except it uses linear housing demand estimates from Column (6) of the previous table (which corresponds to a higher demand elasticity). 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 the previous table. Consumer surplus estimates are recovered from housing transaction prices and should be interpreted as total consumer surplus over a property's life span. Toll revenue is net of the capital and operating costs of revenue collection. Subway cost includes construction and operation costs. Both toll revenue and subway cost are the discounted sum over thirty years (which is approximately the lifespan of a property) and allocated uniformly across households. Net welfare is consumer surplus plus toll revenue and environmental benefits minus subway costs. Column (1) reports results when no policy was in place. Columns (2) to (6) present differences from Column (1). Driving restriction prohibits driving in one of five work days. Congestion pricing is set at ¥1.13 per km as in Table 6. High-income households are those with income above the median.

Table A9: Housing Demand - Nonlinear Parameters with Alternative Sampling

	0.5% Sample		1% Sample	
	(1)	Para	(2)	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	0.015	0.379	0.013
$\sigma(EV_{Female})$	0.408	0.014	0.482	0.012
Log-likelihood	-128,976		-168,808	

Note: The table examines robustness to the sampling choice set and reports MLE estimates of the non-linear parameters in the housing demand model 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 A10: Housing Demand - Linear Parameters with Alternative Sampling

Variables	0.5% Sample (1)	1% Sample (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	-1.91	-1.79

Note: This table examines robustness to the sampling choice set 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 A11: Endogenous Amenities from Subway Expansion

Dependent Var.	(1)	(2)	(3)	(4)
	ψ_{nt}		ψ_{nt}	
	TWFE	TWFE	CSDID	CSDID
$\mathbf{1}\{\text{subway}\}$	0.25 (0.24)	0.15 (0.23)	0.43 (0.29)	0.40 (0.30)
Year FE	Yes		Yes	
District-Year FE		Yes		Yes
Observations	1,144	1,144	1,144	1,144

Note: The dependent variable ψ denotes the neighborhood-year FEs obtained from regressing δ_{jt} on neighborhood and year interactions, partialling out property price and attributes. Independent variable $\mathbf{1}\{\text{subway}\}$ takes the value of 1 if neighborhood n experienced new subway expansions in time t . Each observation is a neighborhood-year. Standard errors clustered by neighborhood are reported in parentheses.

Table A12: 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 $\ln(\text{speed in km/h})$, log of speed in km per hour, 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 segments 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.

Table A13: 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 × 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 2014 cohort in the mortgage data with 7,136 observations. We simulate the equilibrium housing prices under the 2008 network for two scenarios: with and without driving restrictions. We then regress the simulated housing prices in ¥1,000/ m^2 on subway distance in km, which is the observed distance to the nearest subway station based on the 2008 subway network. Standard errors clustered at the neighborhood level. Driving restrictions steepen the price gradient with respect to subway access, consistent with [Jerch et al. \(2021\)](#).

Table A14: 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	103.0
Pollution reduction (+)		4.24	4.24	3.98	3.98		1.98	1.98	6.14	6.14	6.06	6.06
Net welfare		-223.1	-26.8	32.1	59.6		106.7	16.0	-129.0	-14.7	114.1	68.1

Note: This table is similar to Table 6 but shuts down sorting. That is, households are not allowed to change residential locations. Consumer surplus estimates are recovered from housing transaction prices and should be interpreted as total consumer surplus over a property's life span. Toll revenue is net of the capital and operating costs of revenue collection. Subway cost includes construction and operation costs. Both toll revenue and subway cost are the discounted sum over thirty years (which is approximately the lifespan of a property) and allocated uniformly across households. Welfare benefits from pollution reduction arise from reduced tailpipe emissions. Net welfare is consumer surplus plus toll revenue and environmental benefits minus subway costs. Column (1) reports results when no policy was in place. Columns (2) to (6) present differences from Column (1). Driving restriction prohibits driving in one of five work days. Congestion pricing is set at ¥1.13 per km as in Table 6. High-income household are those with income above the median.

Table A15: Simulation Results with Housing Supply Adjustment at the Neighborhood Level

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	-7.33	-3.53	-3.60	-5.43	-2.45	-1.89	-8.75	-4.79	-5.42	-6.61
Subway	9.02	10.77	1.38	0.81	0.92	1.10	5.02	6.58	6.21	6.99	5.50	7.22
Bus	22.44	30.47	1.81	0.59	0.52	1.06	-1.46	-2.43	0.37	-1.48	-0.71	-0.96
Bike	15.96	24.01	1.62	0.79	0.78	1.71	-0.81	-1.74	0.46	-1.08	-0.14	-0.21
Taxi	2.20	1.32	1.18	0.54	0.62	0.57	-0.21	-0.16	0.84	0.30	0.37	0.33
Walk	8.74	11.99	1.35	0.79	0.77	0.99	-0.09	-0.35	0.86	0.06	0.40	0.23
Speed (km/h)	21.49		3.85		4.02		0.95		4.50		4.98	
Panel B: sorting outcomes												
Distance to work (km)	18.52	15.61	0.02	0.02	-0.35	-0.24	0.86	0.72	0.96	0.74	0.47	0.51
Distance to subway (km)	5.28	4.24	-0.07	0.01	-0.14	-0.08	-4.08	-3.38	-4.08	-3.38	-4.09	-3.38
Panel C: welfare changes per household (thousand ¥)												
Consumer surplus (+)	-229.8		-32.2		-83.3		-66.7		165.9		86.0	
Toll revenue (+)					135.0		135.0					
Subway cost (-)							103.0		103.0		103.0	
Pollution reduction (+)	4.29		4.29		4.51		4.51		1.08		1.08	
Net welfare	-225.5		-27.9		56.3		72.8		64.0		-16.0	
									-163.1		-46.0	
									100.7		50.0	

Note: This table is similar to Table 6 but incorporates endogenous housing supply at neighborhood level. 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. Consumer surplus estimates are recovered from housing transaction prices and should be interpreted as total consumer surplus over a property's life span. Toll revenue is net of the capital and operating costs of revenue collection. Subway cost includes construction and operation costs. Both toll revenue and subway cost are the discounted sum over thirty years (which is approximately the lifespan of a property) and allocated uniformly across households. Welfare benefits from pollution reduction arise from reduced tailpipe emissions. Net welfare is consumer surplus plus toll revenue and environmental benefits minus subway costs. Column (1) reports results when no policy was in place. Columns (2) to (6) present differences from Column (1). Driving restriction prohibits driving in one of five work days. Congestion pricing is set at ¥1.13 as in Table 6. High-income households are those with income above the median.

Table A16: Simulation Results with Ring-Road-Quadrant Congestion

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 outcomes												
Drive	41.74	21.51	-7.40	-3.59	-3.76	-5.58	-2.31	-1.76	-8.84	-4.85	-5.70	-6.77
Subway	9.00	10.76	1.33	0.72	0.90	0.99	4.67	6.07	5.88	6.50	5.37	6.90
Bus	22.39	30.43	1.89	0.71	0.70	1.34	-1.47	-2.48	0.45	-1.45	-0.56	-0.88
Bike	15.92	23.98	1.68	0.89	0.87	1.86	-0.74	-1.59	0.63	-0.84	0.03	0.02
Taxi	2.21	1.33	1.16	0.52	0.60	0.54	-0.18	-0.12	0.84	0.32	0.36	0.33
Walk	8.74	11.99	1.33	0.74	0.70	0.84	0.03	-0.12	1.04	0.32	0.50	0.40
Speed	22.46		3.46		3.38		1.25		4.48		4.69	
Panel B: housing market outcomes												
Distance to work (km)	18.57	15.66	-0.01	0.03	-0.19	-0.05	0.36	0.17	0.39	0.18	0.13	0.13
Distance to subway (km)	5.34	4.30	-0.04	0.04	-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 (+)	-243.7		-36.2		-120.0		-77.1		206.2		96.6	
Toll revenue (+)					140.3		140.3				129.0	
Subway cost (-)							103.0		103.0		103.0	
Pollution Reduction (+)	4.35		4.35		4.31		4.31		1.67		5.77	
Net welfare	-239.4		-31.8		24.7		67.5		104.9		-4.8	
									-134.2		-38.0	
											110.4	
											55.2	

Note: This table is similar to Table 6, except that congestion is measured at the ring-road-quadrant level. Simulations are conducted using data from the 2014 cohort, which consists of households who purchased homes in 2014. The parameters used in the simulations are reported in Column (6) of Table 3, Column (3) of Table 4, and Column (6) of Table 5. We incorporate household sorting but keep the housing supply fixed. Consumer surplus estimates are recovered from housing transaction prices and should be interpreted as total consumer surplus over a property's life span. Toll revenue is net of the capital and operating costs of revenue collection. Subway cost includes construction and operation costs. Both toll revenue and subway cost are the discounted sum over thirty years (which is approximately the lifespan of a property) and allocated uniformly across households. The net welfare is calculated as the consumer surplus plus toll revenue and environmental benefits minus subway costs. Column (1) reports the results when no policy was in place. Columns (2) to (6) present differences from Column (1). The driving restriction policy prohibits driving on one of five workdays. Congestion pricing is set at ¥1.13 per km as in Table 6. High-income households are those with income above the median.

Table A17: Region-Specific Speed Changes under Different Policy Scenarios

Zone	Baseline (km/h)	Percentage changes compared to baseline				
		(1) Restriction	(2) Congestion	(3) Subway	(4) Sub+R	(5) Sub+C
Within 2	15.8	16.7%	11.1%	3.2%	19.6%	14.4%
Within 2-3	21.0	16.4%	15.0%	4.4%	20.3%	19.4%
SW 3-4	25.5	16.7%	18.0%	8.1%	23.3%	25.9%
SE 3-4	23.5	16.4%	16.6%	6.1%	21.3%	22.9%
NW 3-4	21.3	17.4%	15.4%	3.6%	20.4%	19.0%
NE 3-4	22.1	15.9%	14.6%	2.5%	18.2%	17.4%
SW 4-5	27.4	17.8%	20.1%	13.4%	28.3%	32.6%
SE 4-5	25.2	16.5%	18.5%	6.6%	22.0%	25.0%
NW 4-5	22.9	16.9%	16.6%	4.7%	20.7%	21.0%
NE 4-5	21.7	16.1%	15.1%	4.3%	19.9%	19.4%
SW 5-6	31.0	9.6%	10.6%	8.7%	16.2%	18.5%
SE 5-6	24.8	8.5%	9.8%	4.3%	12.0%	13.9%
NW 5-6	23.0	8.8%	9.3%	3.5%	11.6%	12.4%
NE 5-6	23.4	8.5%	10.1%	3.0%	11.1%	12.8%
Outside 6	41.3	10.1%	11.8%	5.6%	14.4%	16.9%

Note: This table displays the percentage changes in speed for each ring-road-quadrant relative to the baseline under different transportation policies. Each row denotes a separate ring-road-quadrant. For example, ‘Within 2’ stands for the area within the 2nd ring-road, and ‘SW 3-4’ denotes the Southwestern quadrant between the third and fourth ring roads. Values are colored in different shades according to their magnitudes. Simulations are conducted using data from the 2014 cohort, which consists of households who purchased a house in 2014. The results are based on simulations with ring-road-quadrant congestion.

Table A18: Substitution Patterns with and without Random Coefficients

Income relative to the median	(1)		(2)		(3)		(4)	
	No Policy Baseline levels		Driving restriction Δs from (1)		Congestion pricing Δs from (1)		Subway expansion Δs from (1)	
	High	Low	High	Low	High	Low	High	Low
Panel A: changes in travel model choices, with random coefficients (the preferred specification)								
Drive	41.65	21.44	-7.17	-3.40	-3.48	-5.39	-2.14	-1.66
Bus	22.44	30.47	1.78	0.60	0.57	1.24	-1.54	-2.53
Bike	15.96	24.01	1.60	0.80	0.77	1.78	-0.80	-1.64
Subway	9.02	10.77	1.29	0.70	0.84	0.96	4.62	6.06
Walk	8.74	11.99	1.31	0.74	0.67	0.83	0.02	-0.13
Taxi	2.2	1.32	1.19	0.55	0.63	0.57	-0.16	-0.11
Avg. Speed (km/h)	21.49		3.83		3.83		1.49	
Panel B: changes in travel mode choices, without random coefficients								
Drive			-9.15	-5.41	-2.48	-3.28	-0.43	-0.35
Bus			3.37	1.92	0.82	1.22	-0.24	-0.51
Bike			2.67	1.63	0.67	0.98	-0.14	-0.31
Subway			1.35	0.81	0.42	0.50	0.92	1.32
Walk			1.50	0.92	0.49	0.51	-0.08	-0.13
Taxi			0.26	0.13	0.08	0.07	-0.02	-0.02
Avg. Speed (km/h)			4.59		2.97		0.16	

Note: The model without random coefficients produces counter-intuitive substitution patterns and very different speed improvements from our preferred specification. Column (1) reports results when no policy was in place. Columns (2)- (4) present differences from Column (1). Congestion pricing is fixed at ¥1.13 per km for both panels, as in Table 6 in the manuscript. High-income households are those with income above the median.

Table A19: Welfare Implications with and without Random Coefficients

	(1) No Policy		(2) Driving restriction		(3) Congestion pricing Toll: ¥1.13/km		(4) Subway expansion	
	Baseline levels		Δs from (1)		Δs from (1)		Δs from (1)	
	High	Low	High	Low	High	Low	High	Low
Panel A: welfare changes per household in ¥1,000, with random coefficients (the preferred specification)								
Consumer surplus (+)			-227.1	-32.7	-98.2	-73.1	220.3	100.0
Toll revenue (+)					137.4	137.4		
Subway cost (-)							103.0	103.0
Pollution reduction (+)			4.3	4.3	4.3	4.3	1.7	1.7
Net welfare			-222.8	-28.4	43.5	68.6	119.0	-1.3

Panel B: welfare changes per household in ¥1,000, without random coefficients

Consumer surplus (+)	-1447.6	-338.0	-797.6	-394.2	118.5	60.5
Toll revenue (+)			156.2	156.2		
Subway cost (-)					103.0	103.0
Pollution reduction (+)	6.3	6.3	3.9	3.9	0.2	0.2
Net welfare	-1441.3	-331.7	-637.5	-234.1	15.5	-42.5

Note: The model without random coefficients produces very different welfare implications from our preferred specification. Consumer surplus estimates are recovered from housing transaction prices and should be interpreted as total consumer surplus over a property's life span. Toll revenue is net of the capital and operating costs of revenue collection. Subway cost includes construction and operation costs. Both toll revenue and subway cost are the discounted sum over thirty years (which is approximately the lifespan of a property) and allocated uniformly across households. Welfare benefits from pollution reduction arise from reduced tailpipe emissions. Net welfare is consumer surplus plus toll revenue and environmental benefits minus subway costs. Column (1) refers to the baseline when no policy was in place. Columns (2)-(4) present differences from Column (1). Congestion pricing is fixed at ¥1.13 per km for both panels, as in Table 6.

Table A20: Sorting Patterns with and without Random Coefficients

	(1)	(2)	(3)	(4)				
No Policy	Driving restriction	Congestion pricing	Subway expansion					
Baseline levels	Δs from (1)	Δs from (1)	Δs from (1)					
High	Low	High	Low	High	Low	High	Low	
Panel A: housing market outcomes, with random coefficients (the preferred specification)								
Distance to work (km)	18.987	13.674	0.013	0.014	-0.187	-0.071	0.329	0.146
Panel B: housing market outcomes, without random coefficients								
Distance to work (km)	18.987	13.674	-0.037	0.101	-0.657	-0.258	0.116	0.001

Note: The model without random coefficients produces different sorting patterns from our preferred specification. Column (1) reports results when no policy was in place. Columns (2)-(4) present differences from Column (1). Congestion pricing is set at ¥1.13 per km as in Table 6.

Table A21: Effect of Transportation Policies on Price Gradient w.r.t. Distance to Work

	Dependent variable: housing Price (1000¥)		
	Driving restriction	Congestion pricing	Subway expansion
	(1)	(2)	(3)
Panel A: with random coefficients (the preferred specification)			
Distance to work centroid $\times 1\{\text{Restriction}\}$	0.842*** (0.063)		
Distance to work centroid $\times 1\{\text{Toll}\}$		-6.129*** (0.124)	
Distance to work centroid $\times 1\{\text{Subway}\}$			18.647*** (0.445)
Panel B: without random coefficients			
Distance to work centroid $\times 1\{\text{Restriction}\}$	0.269*** (0.074)		
Distance to work centroid $\times 1\{\text{Toll}\}$		-19.464*** (0.299)	
Distance to work centroid $\times 1\{\text{Subway}\}$			12.781*** (0.334)
House FE and Policy FE	Yes	Yes	Yes

Note: The number of observations is 7,138. The table presents the estimated changes in the price-distance gradient under congestion pricing and subway expansion using simulated equilibrium prices. Each cell is a separate regression. ‘Distance to work centroid’ is a property’s distance in km to the centroid of all job locations. The coefficient of ‘Distance to work centroid $\times 1\{\text{Policy}\}$ ’ measures changes in the price-distance gradient under the corresponding policy. Congestion pricing is set at ¥1.13 per km as in Table 6.

Table A22: Effect of Subway Expansion on Income Sorting

	Dependent variable: household income (¥)			
	(1)	(2)	(3)	(4)
Within .5km of a new subway station	914.022*			
	(554.819)			
Within 1km of a new subway station		1200.453***		
		(343.442)		
Within 1.5km of a new subway station			938.593***	
			(335.135)	
2014 Subway* Distance to the nearest new Subway				-215.796**
				(96.544)
House FE and Policy FE	Yes	Yes	Yes	Yes

Note: The analysis is based on the 2014 cohort of the mortgage data with 7,136 observations. We simulate the equilibrium outcomes under the 2008 subway network and separately under the 2014 network. We then regress the simulated household income on a property's distance to the nearest subway station built between 2008 and 2014. The coefficient 'within x km of a new subway station' measures changes in annual household income post the subway expansion for properties that are located within x km of a new subway station that was built between 2008 and 2014. Standard errors clustered at the neighborhood level.