Efficiency and Equity Impacts of Urban Transportation Policies with Equilibrium Sorting

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

June 2021

Abstract

We estimate an equilibrium model of residential sorting with endogenous traffic congestion to evaluate the efficiency and equity impacts of urban transportation policies. Leveraging fine-scale data on household travel diaries and housing transactions with home and work locations in Beijing, we jointly estimate travel mode and residential location decisions. The estimation highlights the importance of incorporating work commute in housing decisions and features preference heterogeneity for the ease of work commute by gender. Counterfactual simulations show that while different policies can attain the same level of congestion reduction, their impacts on residential sorting and social welfare are drastically different. First, a driving restriction intensifies income-stratified urban structure where high-income households live closer to subway and work. Distance-based congestion pricing reduces the spatial separation between residence and workplace across income levels, while subway expansion does the opposite. Second, residential sorting strengthens the effectiveness of congestion pricing in improving traffic conditions but undermines that of the driving restriction and subway expansion. Third, the driving restriction is welfare reducing as it leads to large distortions on travel choices. Congestion pricing improves welfare but is regressive, highlighting the need to recycle revenue to address the associated equity concern. Finally, congestion pricing and subway expansion when combined deliver the largest congestion relief and efficiency gain and at the same time achieve self-financing, with revenue from congestion pricing fully covering the cost of subway expansion.

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

JEL Classification Codes: H41, R21, R41

*Barwick: Department of Economics, Cornell University and NBER, panle.barwick@cornell.edu; Li: Dyson School of Applied Economics and Management, Cornell University, NBER and RFF, SL2448@cornell.edu; Waxman: LBJ School of Public Affairs, University of Texas at Austin, awaxman@utexas.edu; Wu: Hang Lung Center for Real Estate, Tsinghua University, email: ireswujing@tsinghua.edu.cn, Xia: Cornell University, tx58@cornell.edu. We thank seminar participants at Boston University, Brazilian School of Economics and Finance, Cornell University, Duke University, Jinan University, MIT, Peking University, Shanghai University of Finance and Economics, University of Illinois, University of Maryland, University of Texas-Austin, the World Bank, and the 2019 Urban Economics Association meeting for helpful comments. We acknowledge the excellent research assistance from Avralt-Od Purevjav and Ziye Zhang and the financial support from the International Initiative for Impact Evaluation (3ie) under project DPW1.1106, and the Center for Transportation, Environment, and Community Health at Cornell University.

1 Introduction

Transportation plays a crucial role in urban spatial structure and the organization of economic activity (Allen and Arkolakis, 2019; Tsivanidis, 2019; Heblich et al., 2020; Gorback, 2020). In most fast-growing developing countries, rapid urbanization and motorization, together with poor infrastructure, have created unprecedented traffic congestion with severe consequences for economic outcomes (Akbar et al., 2018; Harari, 2020).[1] To address this challenge, local governments around the world have implemented a suite of policies, including driving restrictions, gasoline taxes, public transit investment, and congestion pricing. In the short term, the effectiveness of these policies on alleviating congestion crucially hinges on the substitutability among travel modes and the sensitivity of travel demand to changes in the cost of commuting. In the medium to long run, these policies are likely to have broader impacts on the urban spatial structure through residential location adjustment, which, in turn, could mediate the effectiveness of these policies and have important distributional consequences. This paper aims to understand the efficiency and equity impacts of urban transportation policies while accounting for the interaction between these policies and residential location decisions. 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 been routinely ranked as one of the most congested cities in the world. Severe congestion exacerbates air pollution, reduces time allocation efficiency, and negatively affects the quality of urban life (Kahneman and Krueger, 2006; Anderson et al., 2016). Beijing's municipal government has adopted several policy interventions to aggressively combat traffic congestion and air pollution. It has implemented a driving restriction policy since 2008 that restricts vehicles from driving one day per week during weekdays based on the last digit of the license plate. It also invested a staggering \$100 billion in transportation infrastructure between 2007 and 2018. The 16 newly-added subway lines with a total length of 523 kilometers, together with more than 200 additional bus lines, constitute a major upgrade of Beijing's public transit network. Despite these efforts, Beijing has only seen a modest reduction of peak-hour traffic congestion since 2015. The city's experience, common among urban centers around the world, echoes long-standing concerns from economists that without appropriate road pricing, effective congestion management is unlikely (Vickrey, 1959, 1963).

Beijing's policies to combat congestion – driving restrictions and subway expansion – together with its proposed policy on congestion pricing embody three general approaches to regulating unpriced externalities: command-and-control, supply-side, and market-based approaches, respectively. To understand the primary channels by which these policies affect travel mode and residential location choices, we proceed in several steps. We first develop a stylized theoretical model based on LeRoy and Sonstelie (1983) and Brueckner

[1] Based on real-time GPS traffic data in 403 cities from 56 countries in 2018, the TomTom Traffic Index shows that the ten most congested cities were all from developing and emerging economies. The top five cities were Mumbai, Bogota, Lima, New Delhi, and Moscow. Drivers in Mumbai spent 65% more commuting time on average than they would have under the free flow condition, while drivers in Beijing (the number 30 on the list) spent 40% extra time on the road. Four cities in China were among the top 30 on the list. Los Angeles, the most congested city in the US, was ranked 24^{th} with a congestion index of 41%. The full ranking based on the TomTom Traffic index 2018 is available at `https://www.tomtom.com/en_gb/traffic-index/ranking.`

(2007) while accounting for endogenous congestion and heterogeneity in income and commuting technologies. The model illustrates the differential impacts of urban transportation policies on the spatial pattern of residential locations and highlights countervailing forces at play between travel mode choices and housing locations. It also suggests that the general equilibrium effects of transportation policies on the housing market can have efficiency and distributional consequences. Lastly, the model highlights the ambiguity in qualitative comparative statics even for simple models with two income types and two commuting technologies, demonstrating the need for subsequent empirical analyses to understand how these factors play out in an actual urban setting.

Our empirical analysis leverages fine spatial resolutions from two unique data sets that allow us to jointly model the residential locations and commuting choices. The first is the Beijing Household Travel Survey (BHTS) from 2010 and 2014, a large representative survey that records households' home and work locations, trips made in a 24-hour window and other demographic and transportation-related information. We complement this data by constructing the counterfactual commuting choices from home to work using historical Geographical Information System (GIS) maps and the Application Programming Interface (API) from online mapping services. This exercise allows us to compile the commuting route, travel time, distance and pecuniary travel cost for each trip-mode combination (walking, biking, taking a bus, subway, car, or taxi). The second data set contains mortgage transactions from a major government-run mortgage program and provides a large representative sample of Beijing home buyers. Critical to our analysis, the housing data contain not only home location information but also work locations of both the primary and secondary borrowers. Using this information, we construct over 12 million hypothetical trip-mode combinations of the home-work commute for both the primary and second borrowers using the same GIS and API procedure as was done for the travel survey data.

We then build an equilibrium model of residential sorting with endogenous congestion that we estimate using the two datasets discussed above. The model incorporates preference heterogeneity and allows for the general equilibrium feedback between housing and commuting decisions. In the model, households choose both a residential property and travel modes for their commuting trips. A key consideration in a household's choice of a residential property is the location's ease-of-commute for both working members of the household. This ease-of-commute attribute is an equilibrium object that crucially depends on congestion and is determined by the location choices and travel decisions of all households. Once estimated, the model allows us to conduct counterfactual simulations to predict new equilibrium outcomes for both marginal and non-marginal policy changes in terms of travel mode choices, household locations, the congestion level and housing prices, as well as the welfare implications.

We use a two-step strategy to estimate the equilibrium sorting model. The first step is to recover heterogeneous preferences for travel time and cost (therefore the value of time) using the household travel surveys and the time and pecuniary cost information of all hypothetical commuting trips that we construct. We utilize the estimated parameters from this step to construct the ease-of-commute attribute that is discussed above, which corresponds to the expected maximum utility of commuting from a property's location to a buyer's workplace via available travel modes. We construct this measure for all houses in a buyer's consideration set. It accounts for preference heterogeneity and takes into consideration the commuting distance, proximity to public transit, level of traffic congestion in different parts of the city, as well as upgrades in the transportation system at the time of home purchase. For married couples with two jobs, we include a separate ease-of-commute attribute for each spouse, allowing for gender-specific preference on a location's attractiveness in terms of work commute.

The second step recovers preferences for housing attributes using observed purchases. The ease-of-commute index is included as an observed (buyer-specific) house attribute. The key challenge in estimating housing demand is the potential correlation between unobserved housing attributes and the housing price as well as the ease-of-commute index. The latter two variables are equilibrium outcomes determined by observed and unobserved housing attributes. To address this challenge, we construct three sets of instrumental variables in the spirit of Berry et al. (1995) and Bayer et al. (2007). These instruments include the average housing and neighborhood attributes within a reasonable distance from a given property, the number of houses sold in a three-month window around the sales date, and the time-varying odds of winning a license lottery to purchase a vehicle. The lottery's time-varying winning odds is a powerful IV and shifts demand for houses in premium locations, such as places that are close to subways or in the city center. We allow both observed and unobserved preference heterogeneity, control for property fixed effect, and estimate parameters through maximum-likelihood estimation with a nested contraction mapping that is combined with IVs (Train and Winston, 2007).

Utilizing these estimates, we then simulate equilibrium residential sorting and transportation outcomes based on the three policies of interest in our study: the license plate-based driving restriction, subway expansion from 2008 to 2014, and congestion pricing. Since the first two policies were enacted during the sample period, we begin with a no-policy counterfactual and then compare this no-policy baseline with different policy combinations. We also compare partial equilibrium outcomes that do not allow residential sorting with general equilibrium outcomes that allow households to relocate and housing supply to adjust in response to changes in equilibrium housing prices and congestion.

Our policy simulations yield four key findings. First, although different transportation policies can attain the same level of congestion reduction, they exhibit different and sometimes opposite impacts on the equilibrium housing price and the spatial patterns of residential locations. Both the driving restriction and congestion pricing increase the price premium of houses near the city center and subway stations, as high-income households outbid low-income households for these desirable locations. Driving restriction reduces the distance to work for high-income households but increases that for low-income households. In contrast, distance-based congestion pricing reduces the distance to workplace for both income groups, while subway expansion does the opposite and entails longer commutes for both groups.

Second, transportation policies generate equity implications and could either exacerbate or alleviate economic inequality (Waxman, 2017; Akbar, 2020). Congestion pricing provides a larger welfare gain for low-income households than high-income households if its revenue can be uniformly recycled. Without recycling, however, congestion pricing is regressive and leads to a larger welfare loss for low-income households. This distributional concern is an important impediment to congestion pricing adoption in practice. The driving restriction policy is progressive and leads to less distortion on low-income households' travel decisions, potentially explaining the wider adoption of this policy than congestion pricing around the world.

Third, residential sorting can either strengthen or undermine the congestion-reduction potential of transportation policies as well as the welfare impacts. Sorting strengthens the effectiveness of congestion pricing in both congestion reduction and welfare gains as households are incentivized to live closer to work locations and drive less. On the other hand, sorting in response to subway expansion would lead to a farther separation between residential and work locations, damping the congestion reduction effect and welfare gains from infrastructure investment. The implication of sorting on the effectiveness of driving restrictions is more nuanced and depends on the extent of reduction in average commuting distance and the increased propensity of driving for more distant trips. Our analysis finds that sorting only slightly weakens the effectiveness of the driving restriction.

Finally, the actual and hypothetical combinations of transportation policies have different implications on aggregate welfare. Beijing's rapid subway expansion from 2008 to 2014 leads to an increase in consumer surplus and aggregate welfare despite modest congestion reduction. In contrast, driving restriction is welfare reducing in spite of a larger congestion reduction. Congestion pricing and subway expansion in tandem deliver the largest improvement in traffic speed and welfare gain. In addition, the revenue from congestion pricing could fully finance the capital and operating costs of subway expansion, eliminating the need to fund the expansion from distortionary taxes. These results showcase the sorting model's strength in capturing various adjustment margins and its ability to compare different policy scenarios in a unified framework that accounts for general equilibrium welfare effects with preference heterogeneity.

Our study makes three main contributions to the literature. First, while quantitative spatial economics has made considerable advances to explore the role of transportation in urban systems (see Redding and Rossi-Hansberg (2017) for a review), there has been limited attempt in the empirical urban literature to explore the role of preference heterogeneity and congestion externalities in mediating the welfare effects of different transportation policies. Accounting for preference heterogeneity could be crucial for understanding the equilibrium and distributional impacts of urban policies, in light of the growing concerns about the role of location in economic opportunity (Chetty et al., 2014). The rich preference heterogeneity incorporated in our framework allows us to identify winners and losers from urban transportation policies, in the spirit of the theoretical and reduced form work by LeRoy and Sonstelie (1983), Glaeser et al. (2008) and Brueckner et al. (1999). As some more recent quantitative spatial models (Allen and Arkolakis, 2019; Fajgelbaum and Gaubert, 2020), we explicitly model endogenous congestion to capture its increasing marginal external cost, as highlighted by Anderson (2014). Critically, our dual-market approach provides a micro-foundation for the hedonic models that study the capitalization of transportation investment in the housing market (Baum-Snow and Kahn, 2000; Gibbons and Machin, 2005; Zheng and Kahn, 2013) and demonstrates considerable distributional impacts.

Second, our analysis contributes to a large equilibrium residential sorting literature by incorporating endogenous work commuting decisions in residential location choices. 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; Kuminoff et al., 2013).[2] Most existing papers treat both the distance to work and congestion as exogenous attributes. An exception is Kuminoff (2012), which models household decisions in both the work and housing markets and endogenize the commuting distance while keeping congestion as exogenous. Our paper is to our knowledge the first in the empirical sorting literature that explicitly models both congestion and distance to work as equilibrium outcomes that are simultaneously determined by household locations and travel mode choices.

Third, our paper relates to the literature on transportation policies that address the negative congestion externality (Parry et al., 2007).[3] Studies in this literature commonly focus on short-run or partial equilibrium effects of transportation policies on travel choices, traffic congestion, and air pollution. By characterizing the underlying travel and housing choices, our equilibrium sorting framework provides a micro-foundation for the reduced-form impact evaluation studies. More importantly, the unified framework offers a common yardstick to evaluate actual and counterfactual policies over a wide range of outcomes including congestion reduction, urban spatial structure, social welfare, and distributional consequences. Another approach in the literature allows for the general equilibrium feedback effects between the transportation and housing sectors in a calibrated computable general equilibrium framework without estimating the underlying consumer preferences (Anas and Kim, 1996; Langer and Winston, 2008; Parry and Small, 2009; Basso and Silva, 2014). Compared with these studies, our framework is internally consistent in that the estimation of structural parameters and the policy simulations are based on the same model.

Section 2 uses a stylized model to explain the key forces underlying the interaction between housing and transportation. Section 3 describes the data and provides reduced-form evidence on the effect of Beijing's driving restriction on the housing market to motivate and ground subsequent analysis. Section 4 lays out the equilibrium sorting model and the estimation strategy. Estimation results are presented in Section 5. Section 6 conducts simulations to examine the impacts of transportation policies and compare their welfare consequences. Section 7 concludes.

[2] See for example Bayer et al. (2007); Ferreyra (2007); Epple and Ferreyra (2008); Epple et al. (2012) on school quality, Sieg et al. (2004); Bayer et al. (2009); Kuminoff (2009); Tra (2010); Bayer et al. (2016) on air quality, Timmins and Murdock (2007); Walsh et al. (2007); Klaiber and Phaneuf (2010) on open space and recreation, Bajari and Kahn (2005); Bayer et al. (2007); Bayer and McMillan (2012); Hwang (2019) on racial and ethnic composition, Calder-Wang (2020) on the distributional impacts of the sharing economy in the housing market. Several recent studies seek to incorporate dynamic models of housing demand into sorting, allowing for households to make choices based on expectations about the evolution of location amenities, prices and wages over time on endogenous amenities as a bundle (Almagro and Domínguez-Iino, 2020; Murphy, 2015; Wang, 2020). We do not incorporate dynamics into our model but consider their influence in our interpretation of the results.

[3] Various policies have been evaluated. See Parry and Small (2005); Bento et al. (2009); Knittel and Sandler (2013); Li et al. (2014) on gasoline taxes, Bento et al. (2005); Parry and Small (2009); Duranton and Turner (2011); Anderson (2014); Basso and Silva (2014); Li et al. (2019); Severen (2019); Gu et al. (2020) on public transit subsidies and expansion, Davis (2008); Viard and Fu (2015); Carrillo et al. (2016); Zhang et al. (2017) on driving restrictions, and Langer and Winston (2008); Anas and Lindsey (2011); Hall (2018); Yang et al. (2019); Kreindler (2018) on congestion pricing.

2 Theoretical Framework

We motivate our setup with a graphical presentation of the welfare effects of two transportation policies that are examined in our empirical analysis: congestion pricing and driving restrictions. There are two dimensions of this effect captured: that in the primary market for commuting reflecting the direct effect of an unpriced externality on the marginal social cost of driving, and a secondary effect on a related market, housing, where changes in commuting costs are capitalized into housing prices and can induce changes in commuting mode choice and housing location.[4]

Figure 1 illustrates the welfare effects of these two policies in the primary "market" for vehicle road traffic. Here, the economic cost induced by congestion can be demonstrated in terms of the level of traffic volume, V as the difference between marginal social cost (MSC) and Average Social Cost (ASC) and the unpriced externality created by the marginal external cost of congestion (MEC). Both congestion pricing and a driving restriction result in a reduction of traffic volumes from the unregulated level, V^0, to the socially optimal level, V^*. However, congestion pricing reduces the trips with the lowest marginal benefit while the driving restriction, due to the fact that its design does result in sorting based on differences in the value of time, could reduce trips with various levels of marginal benefit.[5] If the length of commuting trips are reduced in a random manner, the driving restriction will lead to welfare loss equivalent to the blue triangle in Figure 1. The size of the triangle is positively related to the degree of heterogeneity in the marginal benefit of trips. The figure illustrates that while congestion pricing leads to welfare gain, the welfare impact of a driving restriction is ambiguous.

However, Figure 1 is only a partial equilibrium analysis that does not take into account the potential impact of transportation policies on proximate markets. It also tells us little about differences in these incomes across individuals. To understand these additional effects, consider a monocentric city model where households with different incomes sort into different locations in response to transportation policies based principally on their deterministic preferences for housing, other goods and time.[6] This model includes three key components of recent applied work in urban economics: endogenous congestion, mode choice, and residential sorting. We fully develop this model in Appendix A including presenting key comparative statics building on the approach from Brueckner (2007), but here we summarize some key properties, standard in this class of models, and elaborate on a set of stylized outcomes that illustrate heterogeneous welfare effects via capitalization.

Model Primitives Here are the key model primitives:

- The monocentric city is linear with a fixed population (N) of rich, N_R, and poor, N_P, residents.

[4]Following Roback (1982), there is also a capitalization in the labor market but this lies outside the scope of our study.

[5]In this sense, driving restrictions correspond to classic command-and-control or mandate-based quantity restriction form of environmental policy. On the other hand, congestion pricing corresponds to a traditional market-based approach. We use these designations interchangeably in this paper.

[6]In our empirical analysis, we will relax assumptions of monocentricity and allow for random utility.

- All residents work at the urban center (CBD) at location 0, where wage income for the rich is larger: $y_R > y_P$.[7]
- The rest of urban space is occupied by homes with lot sizes normalized to 1 and where land rents are remitted to absentee landlords.
- Households maximize utility via housing and non-housing consumption subject to a budget constraint that includes commuting costs and varies between rich and poor based on their value of time (higher for the rich).
- Housing consumption (in square meters) is provided by perfectly competitive developers.
- Beyond the residential area is agricultural land, which returns rental value p_a.

Commuting Technology Several key features characterize the nature of commuting technology:

- Two commuting modes exist in the city: personal vehicles with higher fixed costs and lower variable costs relative to the alternative commuting mode, subway.
- Variable costs, denoted $w_{d,m}(x)$ for mode $m = car, subway$ and group $d = R, P$, include time and pecuniary costs.
- Travel time is monetized by the value of time (VOT): $v_R > v_P$.[8]
- We begin by assuming that the subway network covers the entire urban area and then relax this assumption when considering the role of public transportation infrastructure.
- Car commuting suffers from endogenous congestion determined by the commuting choices of all other households in the city. We ignore the role of congestion in public transportation and focus solely on its effect on car travel.
- The model is a closed-city model with intracity, but not intercity migration.[9]

A feature of the model that usefully simplifies the analysis is that changes in commuting cost will not affect the overall size of the city as reflected by the location of the urban boundary, $\bar{x}$, since the population is fixed and land use per household is also fixed.

[7] Given the linear structure of the city, we assume roads take up no space and all land goes towards housing.

[8] We also assume that fixed commuting costs are larger but variable costs (without congestion) are lower for car relative to subway.

[9] Public transit congestion and closed-city assumptions could be relaxed without affecting the key predictions of the model. Brueckner (1987) provides an analysis of a monocentric city model with a perfectly competitive supply side for both cases of a closed and open city.

Equilibrium Properties Given a mass of rich and poor households residing and working in the city, a spatial equilibrium is determined by a bid rent function $p^*(x)$ that is the envelope of individual willingness-to-pay for housing based on mode and housing type at each point, x, in the city,[10] keeping the utility for each income type fixed at $\bar{u}_d$, $d = R,P$:

$$p^*(x) = \max_{d,m} \left\{ p\left(y_d - \theta_m - w_{d,m}(x), \bar{u}_d\right) \right\} \qquad d = R,P; m = car, subway. \tag{1}$$

For subway commuters, who are assumed to experience no congestion, the slope of the bid rent function does not change. In residential regions with car commuting, moving from right to left across the region means adding additional car commuters, further increasing per kilometer commuting time costs and steepening the bid rent function.

Model Calibration and Policy Analysis In the Appendix A, we explain and show the result of calibration yielding the urban configuration presented in Figure 2: rich subway commuters live closest to the CBD, then poor subway commuters, then rich car commuters, then poor subway commuters. This outcome is not unique and is purely illustrative to yield a pattern of sorting that roughly approximates the qualitative pattern for Beijing described in section 3.[11]

In Figure 2, we show the effect of two transportation policies, congestion pricing and a driving restriction, on the equilibrium bid-rent envelope corresponding to (1). Colored lines in Figure 2 reflect the gradient of housing prices in equilibrium after the indicated transportation policy, where the colors correspond to the bid-rent for the group of commuters indicated below the horizontal axis. Gray lines in both panels reflect the same no-policy baseline price gradient.

The additional welfare effects of transportation policies in the housing market is reflected by the areas between these envelopes. A key principle in urban economics underlying the Rosen (1974) and Roback (1982) approach as well as the Henry George Theorem is that investments in public goods or reductions in negative externalities should be capitalized into housing values and so comparing differences in the sum of housing values can be used to approximate welfare changes under housing and land markets characterized by perfect competition.

Key Model Takeaways In Figure 2, we can see two principal effects of both congestion pricing and driving restrictions on rents and therefore the capitalization of transportation policies: one, the value of proximity to the CBD (i.e., workplace) rises for those nearby (specifically subway commuters), while for those farthest away, it falls. Capitalization gains are larger than losses for congestion pricing (where revenues are recycled lump-sum) because wealthy drivers gain from time savings net of tolls, while the poor who move to the

[10] The full set of market clearing conditions are presented in the Appendix A.

[11] In reality, Beijing does not have a single CBD and there are varying patterns of proximity of relatively wealthy to relatively less wealthy across Beijing.

subway commuting area benefit from revenue recycling and shorter commutes net of higher housing costs.[12] Two, rent losses for longer commutes are larger under the driving restriction. This is because those who would ordinarily drive are forced to take the subway for very long commutes on restricted days of the week.

In summary, while the primary effect of transportation policies on commuting costs in the market for driving is straightforward, the secondary effect on housing via price capitalization can be large and depends on relative differences in the marginal cost of commuting, income heterogeneity and preferences for housing and time. While illustrative, this simple model ignores a host of important features important for understanding the economic effects of transportation policies in Beijing such as polycentricity of the city, travel modes beyond driving and subway, and variation in the availability of housing across the city. For these reasons, we now turn to our equilibrium sorting model to empirically evaluate these policies.

3 Policy Background, Data Description and Reduced-form Evidence

3.1 Policy Background

During the last four decades, China has witnessed the largest rural to urban migration in human history: urban population increased from 171 million (about 18% of total population) in 1978 to 823 million (nearly 60% of total population) by the end of 2018. After the turn of the century, vehicle ownership increased dramatically: total sales of new passenger vehicles grew from 0.7 million units in 2001 to nearly 24 million units in 2018. The rapid growth in urbanization and vehicle ownership has overwhelmed the road infrastructure and public transit, leading to serious traffic congestion and exacerbating severe air pollution in all major urban areas across the country. Similar challenges are observed in large cities in other developing and especially fast-growing countries as well. Beijing has been ahead of most other urban centers in China in terms of growth in population, household income, and vehicle ownership. Between 2001 and 2018, Beijing experienced a 56 percent increase in population while household disposable income grew from about $1,500 to nearly $9,000 per year, and the vehicle stock increased from one million to over six million.[13]

The central and municipal governments in China have pursued a series of policies to address traffic congestion. In Beijing, these policies include a driving restriction, a vehicle purchase restriction, and a massive subway and rail transportation infrastructure investment boom. The driving restriction started as part of Beijing's effort to prepare for the 2008 Summer Olympics. It initially restricted half of the vehicles from driving on a given weekday based on their license plate. After the Olympics concluded, the restriction was relaxed to one day a week during weekdays depending on the last digit of the license plate number.[14] In an attempt to

[12] Panel (a) shows a small, but negligible loss of rents for poor car commuters reflecting the fact that given low values of time, congestion reduction may not fully compensate for the fee.

[13] The consumer price index increased by 50% from 2001 to 2018. Among the six million vehicles in Beijing, about five million are owned by households. The household vehicle ownership rate is about 0.6 cars per household in Beijing, compared to 0.46 in New York city and 1.16 in the U.S. based on the 2010 U.S. Census.

[14] Police vehicles, buses, municipal street cleaning vehicles, and taxi are exempt from the policy. Athens, Greece implemented the first of driving restrictions in 1982 and since then, at least a dozen other large cities in the world have adopted similar policies

curb the growth in vehicle ownership, the Beijing municipal government adopted a purchase quota system for new vehicles in 2011 by capping the number of new vehicle sales. About 20,000 new licences were distributed each month through non-transferable lotteries during 2011 and 2013 and the monthly quota was reduced to about 12,000 after 2013. Winning the lottery became increasingly difficult: the winning odds decreased from 1:10 in early 2012 to nearly 1:2000 in 2018 as the pool of lottery participants increased dramatically while the number of licenses fell over time (Xiao et al., 2017; Li, 2018; Liu et al., 2020). Along with demand-side strategies, 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 500km (See Appendix Figure A1 for subway maps over time). By the end of 2019, the Beijing Subway is the world's longest and busiest subway system with a total length of nearly 700km, and daily ridership over 10 million.[15]

Despite these policy efforts, traffic congestion continues to be a pressing issue in Beijing: the average traffic speed was 24.6km/h during peak hours (7-9am and 5-7pm) in 2019 according to the 2020 Beijing Transportation Report. From a neoclassical microeconomic perspective, the aforementioned policies fail to directly address the root cause of traffic congestion: the mispricing of road capacity. By recognizing traffic congestion as a classic externality, (Vickrey, 1959, 1963) proposed congestion pricing as the first-best policy to address traffic congestion. Despite being continuously advocated by economists since then, congestion pricing has only limited adoption in practice with many failed attempts around the world largely due to technical feasibility and especially political acceptability.[16] The Beijing municipal government recently announced a plan to introduce road pricing in the near future while soliciting feedback from experts and the general public (Yang et al., 2019).

3.2 Data Description

To compare the impacts of different transportation policies on commuting and housing location decisions, we construct the most comprehensive data on work-commute travels and housing transactions ever used in the context of equilibrium sorting models. We rely on two main data sets for our analysis: Beijing Household Travel Survey in 2010 and 2014 and housing mortgage data over 2006-2014 with detailed information on household demographics as well as the work address of home buyers.

---

including Bogotá, Mexico City and New Delhi. The impacts of these policies on reducing congestion and air pollution have been mixed (Davis, 2008; Viard and Fu, 2015; Carrillo et al., 2016; Zhang et al., 2017).

[15]Many other cities in China are also rapidly building and expanding their subway systems. 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 kilometers of by the end of 2019. One intended effect of these expansions is to slow the growth of personal vehicle use by making public transportation more accessible. See Anderson (2014); Yang et al. (2018); Gu et al. (2020) for recent analysis on the impact of subway expansion on traffic congestion.

[16]Vickrey (1963) asserted: "... in no other major area are pricing practices so irrational, so out of date, and so conducive to waste as in urban transportation." This statement remains largely true today. Singapore first adopted congestion pricing in 1975 and is now transitioning to the 4th generation GPS-based system in 2020. During the last 15 years, several European cities (London, Milan, Stockholm, and Gothenburg) have successfully implemented various congestion pricing schemes. After several proposals over the years, New York State legislature has approved a congestion pricing plan for New York City. Pending approval by the Federal Highway Administration, New York City will become the first city in the US to enact congestion pricing, potentially as soon as late 2021.

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

Our analysis focuses on 73,154 work commuting trips (home to work and work to home). Work trips are likely to be the most important trips in housing purchase decisions. They account for 53% and 59% of weekday trips among the working-age respondents and 62% and 75% of total travel distances in 2010 and 2014, respectively.

Table 1 provides summary statistics for variables used in the analysis by the survey year. Household income increased dramatically from 2010 to 2014, with the share of the lowest income group ($<$ 50k annually) decreasing from 48 to 18 percent while shares for higher income groups grew. Vehicle ownership increased from 44 to 62 percent. Both the share of respondents living within the 4th ring road (which proxies for the city center) and that working within the 4th ring road decreased by about 10 percentage points from 2010 to 2014, reflecting the increased spatial dispersion of housing and work locations. Other individual attributes (gender, age, and education) are similar across the two years.

To understand a commuter's travel mode choices, we need attributes for all travel modes in his choice set. We focus on six travel modes: *Walk*, *Bike*, *Bus*, *Subway*, *Car*, and *Taxi*, as other modes (motorcycles, company shuttles, and unlicensed taxis) collectively account for less than 4% of all trips. The travel survey only reports the actual mode that is taken. We complement the travel survey and construct commuting time, distance, and cost for each of the six modes and validate our calculations using information from reported trips.

Appendix Figures A3 and A4 provide two illustrations. We use Baidu Map API to calculate the travel time and distance for walking, biking, car and taxi. Baidu maps incorporate predicted congestion level based on the time of day and day of week. We query Baidu API at the same departure time that is recorded in the travel survey (e.g., 7am). To account for changes in the average congestion level between the survey year and the year we queried Baidu API, we adjust the predicted travel time based on the annual traffic congestion index across different regions in Beijing. For bus travel time, we use Gaode Map API as Baidu does not provide information on the number of transfers and walking time between bus stops, which can be substantial for longer trips. To take into account the subway expansion in our sample period, we use historical subway maps and an GIS software to reconstruct the historical subway network. The subway travel time is calculated based on the published time schedules of subway lines. Our calculation assumes that commuters use the subway stations that are closest to their trip origin and destination and incorporates the walking distance to the subway stations as well as the corresponding time in the total trip distance and duration. Appendix **??** provides more details on the full procedure.

Figure 3 plots each travel mode's observed share of commuting trips, as well as the *constructed* travel time, cost, and distance by each mode. Panel (a) contrasts travel patterns in 2010 with those in 2014 and presents several notable patterns. First, walking account 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 kilometers. Second, from 2010 to 2014, the shares of walk, bike, and especially bus see a reduction while the share of car (i.e., driving) and subway have increased, reflecting the increase in vehicle ownership and the expansion of the subway network. Third, walking and subway trips are the longest in duration, while the subway and car trips are longest in distance. While car trips have slightly longer duration and distance than taxi trips, they are cheaper. Overall, the trade-off between time and cost is clear: trips by walking are slowest but also the cheapest. Car and tax trips are faster but more expensive than other modes.

Panel (b) of Figure 3 shows the data by high- and low-income groups. High-income households are more likely to drive, use subway, take taxis and less likely to use other modes, compared with low-income households. As a percentage of the hourly wage, car and taxi trips are much more expensive for low-income households than for high-income households. In terms of travel distance, there is very little difference across the two income groups except among car trips. This is consistent with the housing data below that display no evidence on strong income-delineated residential sorting patterns.

Housing Transactions Data Data on housing transactions come from a major government-sponsored mortgage program in Beijing from July 2006 to July 2014. As is reflective of the housing supply in urban China, almost all of the housing units are within housing complexes analogous to condominiums in the United States. The interest rate for this mortgage program is subsidized and more than 30% lower than the commercial mortgage rates for eligible borrowers. Virtually all home buyers apply for mortgages through this program first before going to other loans. There are no refinancing activities and each mortgage contract represents a housing transaction.

The final data set includes 79,884 mortgage transactions. Table 2 provides summary statistics of the data.[17] The mortgage data include information on household demographics including income, age, marital status, residency status (*hukou*), and critically for our analysis, the work address for the primary borrower (and that of the co-borrower if present). The data also contain information on housing attributes such as the size, home age, street address, transaction price, and date when the mortgage was signed. We geocode the home and work locations. The mortgage data represent a subset of housing transactions and may be subject to selection issues. Hence, we re-weight the mortgage data to match the population distribution of housing price, size, age, and distance to city center using entropy balancing (Hainmueller, 2012). Appendix B.2 discusses the re-weight procedure in more detail and additional data patterns (such as differences in commuting distance by gender). We use the weighted sample in our benchmark analysis and the unweighted sample in robustness

[17] We remove transactions with missing or zero reported price, price per square meter less than ¥5,000, buyers with no reported income, and an address outside of the 6th ring road. The data set for the structural analysis is slightly smaller than this as we lose some data in constructing the choice set.

checks.

Appendix Figures A5 maps the location of all home transactions in the mortgage data overlaid with black ring roads and blue subway lines (as of 2015). Beijing's spatial structure largely represents a monocentric city possessing multiple ring roads, with some notable exceptions. The expanding set of concentric ring roads layouts the city center. For example, the 2nd ring road largely traces out the contour of the old Beijing prior to the 1980s. On the other hand, there are several large work clusters across the city, such as the financial cluster between the 2nd and 4th ring roads on the east side of the city and a high-tech cluster towards the northwest between the 3rd and 5th ring roads. Government-designated signature schools are denoted by red stars and the government designated parks by green areas, which are important amenities that affect housing purchase decisions.[18] Signature schools are concentrated within the 4th ring road while the parks are more dispersed across the city. Beijing has a total of 18 districts, each containing on average 8 Jiedaos (or neighborhoods).

Figure 4 shows the spatial pattern of housing and household attributes by Traffic Analysis Zone (TAZs) based on the mortgage transactions from 2006 and 2014, with a warmer color representing a higher value. TAZs are spatial units defined by the government used for Beijing transportation planning purposes. They are one to two square kilometers on average and smaller when they are closer to the center of Beijing. There are 2050 TAZs in 2014. Housing prices tend to be higher in the city center while housing size smaller, as predicted by the classical monocentric city models. Distance to work is shorter for those living close to the city center, reflecting a higher concentration of work closer to the city center. There are exceptions: some TAZs in the northwest outside the 5th ring road exhibit short work distance, due to a high-tech center in that area as shown in Figure A6. Relative to housing price, housing size, and commuting distance, the pattern of household income is more mixed. Some high-income households opt to live in larger homes with a lower unit price to the north of the city center, reflecting the classic distance-housing size tradeoff illustrated in the monocentric city model in Section 2. In addition, the northern parts of the city attract high-income households with better amenities and more work opportunities (Appendix Figures A5 and A6).

To incorporate commuting into housing decisions, we need to construct work commute attributes (i.e., time and out-of-pocket costs). In theory, one could construct the attributes of different travel modes for each potential housing choice for a given home buyer/work location as illustrated in Appendix Figure A3. However, this is technically infeasible given the vast number of home-work-mode pairs and the query restrictions by the Baidu and Gaode APIs. The large choice set is a common empirical challenge in the housing demand literature, and in our case, it is further compounded by the fact that there are multiple travel modes associated with each potential housing choice. To reduce the computational burden, we use a choice-based sampling strategy to limit the size of the housing choice set following McFadden (1978), Wasi and Keane (2012), and Guevara and Ben-Akiva (2013). The choice set for a given transaction in the mortgage data is composed of the purchased home and a one percent sample of homes randomly chosen from those sold during a two-month

[18]There are 65 signature schools in Beijing that are designated by the Beijing municipal government as the 'key' elementary schools. These schools have better resources and better student performance. The enrollment in these schools in most cases guarantees a seat in the the top middle schools and subsequently top high schools.

window (30 days before and after) around the purchase date. Beijing's real estate market was fluid during our data period: the median days-on-market for a home seller is only 8 and 13 in 2013 and 2014, respectively, with the average days on market 22 and 38 days. For each home in a household's choice set, we construct the travel mode attributes for both the male borrower's and the female borrower's work commute, based on their respective work locations. The construction of the mode attributes involves over 13 million route-mode combinations.

3.3 Reduced-form Evidence

Before proceeding to the structural sorting model, we examine whether changes in the transportation system are capitalized into housing prices and the residential sorting in response to these changes is meaningful. Specifically, we examine the housing market response to the car driving restriction policy (CDR). The CDR was implemented in July 2008 and prohibited car-owners from driving one day a week based on the last digit of their license plates. The theoretical model in Section 2 predicts both an adjustment in travel choices (substitution away from driving towards the public transportation) and relocation of residential locations. Driving restrictions induce greater demand for homes closer to the public transportation, increasing the price of these homes. In addition, wealthier households with potentially higher values of time are more likely to sort into these units, the so-called transit-induced gentrification.

Figure 5 shows scatter plots of home prices in ¥1,000/m^2 against the distance to the nearest subway station before and after the CDR.[19] The top panel uses raw data, while the bottom panel shows residualized plots after controlling for year-by-month and neighborhood fixed effects.[20] The price gradient becomes steeper post CDR, suggesting that homes close to subways command a higher price premium after the policy. Consistent with theoretical predictions, the driving restriction increases the price premium of the homes near subway stations. Appendix Section C provides an event study analysis, a falsification test and additional evidence.

We next examine residential sorting by regressing the distance from home to subway (and distance between home and work) on the interaction between the CDR dummy and household income (Appendix Section C. Driving restriction policy reduced the distance to the nearest subway station and distance to work much more for high income households. This provides suggestive evidence that high-income households sorted closer to premium locations and low-income households sorted away, potentially because they were priced out.

The reduced-form analyses above confirm the importance of the housing market capitalization and sorting in response to transportation policies. We now turn to an equilibrium sorting model of housing demand and commuting choice that features preference and spatial heterogeneity to quantify the channels by which

[19] We focus on two years before and two years after the starting date of the program in balancing the tradeoff between the sample size and the potential for confounding changes in housing and transportation.

[20] A neighborhood is defined as a Jiedao, an administrative unit that is similar to a census tract and is 2.8 square kilometers in size on average. Each district of Beijing contains an average of 8 Jiedaos. We use Jiedao to denote neighborhoods where homes share similar observed and unobserved amenities.

transportation policies affect these choices. This allows us to analyze the efficiency and equity impacts of different policies in a unified framework.

4 Empirical Sorting Model

We now lay out an empirical equilibrium sorting model that incorporates commuting choices into housing decisions. We first describe the model and then discuss identification and estimation of model parameters.

4.1 Model Overview

Our sorting model characterizes the determinants of individual commuting choices and residential location decisions. It also specifies the joint equilibrium conditions for the traffic congestion (a key amenity in our analysis) and the housing market. On the one hand, residential locations determine households' commute distances and affect the driving demand and hence traffic congestion. On the other hand, traffic congestion in turn affects the attractiveness of a residential location and consequently the housing demand. For example, high congestion levels increase demand for premium locations (places close to subways and the city center). The equilibrium nature of our sorting model allows for counterfactual simulations and provides direct comparative statics of housing prices, residential locations, and congestion levels from marginal or non-marginal policy changes.

In practice, the choice of housing location could be part of a joint decision of work and home locations that may be simultaneous or sequential. We assume they are sequential and take work locations as given in our analysis for three reasons. First, for many households, the choice of work location is likely to be the outcome of a longer-term process of labor supply and migration decisions. Second, employment opportunities in the same industry tend to be clustered in Beijing and switching jobs may not entail meaningful changes of work locations. Third, while the mortgage data provide rich information on locations of housing and current employment, they do not have information on job opportunities (e.g., available job openings at the time of searching). Therefore, incorporating work location decisions would necessitate additional data. Similarly, we do not model firm locations which could be affected by transportation policies in the long run. Our analysis therefore does not model the potential positive spillovers from the agglomeration of firms as modeled in some recent studies using the quantitative spatial models.[21]

Our approach contrasts with the emerging literature using quantitative spatial models that have made considerable advances in modeling the joint processes of work and residential locations. Improvements in the transportation system could translate into higher productivities (through better allocation of time and labor market matching), a margin of adjustment not incorporated in our model. This literature uses observed worker flows and wages to recover iceberg commuting costs via a gravity equation framework. However, with

[21] Diamond (2016) is a micro-founded study that bridges this gap by incorporating housing and labor markets into the evaluation of the heterogeneous welfare consequences of the movement of workers between US cities, although it does not endogenize congestion from the transportation sector.

few exceptions (Fajgelbaum and Gaubert, 2020), this literature tends to recover the cost of commuting through an origin-destination-specific disamenity rather than incorporating individual commuting decisions and congestion externality. In contrast, While taking work locations as given, our equilibrium sorting framework can predict the endogenous congestion level and welfare impacts from unpriced externality across commuters under different policy scenarios.[22]

4.2 Housing Demand

We specify a characteristics-based housing demand model, in which preferences over housing are parameterized as a function of both observed and unobserved household attributes (Lancaster, 1971; McFadden, 1978; Berry et al., 1995). Our data are longitudinal, but we suppress time t to ease exposition. Conditioning on the work locations, utility for household i choosing home unit j can be written as:

$$\max_{\{j \in J\}} U_{ij} = \alpha_i p_j + \mathbf{x}_j \beta_i + \sum_k \phi_{ik} EV_{ijk} + \xi_j + \varepsilon_{ij}, \tag{2}$$

where J is the choice set for household i and the construction of the choice set is discussed below. The household-specific price coefficient α_i is related to the log of household income y_i:

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

We use p_j to denote the price of home j, $\mathbf{x}_j$ to denote a vector of housing attributes such as size and the number of bedroom, and β_i to denote household preferences over housing attributes. The marginal utility for each housing attribute is decomposed into an individual-specific component and a population average, i.e., for each element s in β_i:

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

where $\mathbf{z}_i$ are household demographics. ξ_j captures unobserved housing attributes and ε_{ij} is an i.i.d. error term that reflects unobserved preferences over each housing choice.

Household members with commuting needs are denoted by $k \in \{Male, Female\}$. EV_{ijk} is the expected utility for member k in household i that is derived from the best commuting alternative. It characterizes the attractiveness of home j in terms of member k's work commute. This ease-of-commute term is constructed from a discrete choice model of commuting mode that we describe below. It is affected by the congestion level, which is determined in equilibrium by all households' travel mode choices and residential locations. Preference for ease-of-commute ϕ_{ik} differs across gender and households and is characterized by random

[22] One limitation of the quantitative spatial model is that welfare improvements only result from changes in real income due to gains from trade via an increase in market access. This benefit is mediated directly through the elasticity of imports with respect to variable trade costs (Arkolakis et al., 2012). In the context of urban transportation, this seems potentially limiting because spatial mismatch and wasteful commuting due to pre-existing distortions, like congestion, may leave open opportunities for Pareto improvements without a change in the level of market access.

coefficients:

$$\phi_{ik} = \bar{\phi}_k + \phi_k \nu_{ik}, k \in \{Male, Female\}, \text{ and}$$
$$\begin{pmatrix} \nu_{i,male} \\ \nu_{i,female} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right].$$

In subsequent formulations, we suppress subscript k for the ease of exposition and use EV_{ij} to denote the commuting utility for both household members $\sum_k \phi_{ik} EV_{ijk}$. The commuting utility is the key innovation relative to residential sorting models that incorporate commuting based on fixed distances. It allows for a joint consideration of the heterogeneous impact of transportation policies on commuting costs across individuals and homes.

4.3 Travel Mode

Utility-maximizing individuals within a household choose a commuting mode based on the time and financial cost associated with each of the six modes: Walk, Bike, Bus, Subway, Car, and Taxi. In this subsection, we abuse the notation and use i to denote an individual within a household rather than a household. This is consistent with the level of aggregation in the travel survey that reports the travel mode choices for each commuting member within a household.[23] Preferences vary across individuals, such as the enjoyment of driving a car, perceived "greenness" of using public transportation, or health benefits of biking and walking. We include mode-specific random coefficients to account for these considerations. Individual i's utility of commuting from home j to work using mode choice m is specified as:

$$\max_{m \in M_{ij}} u_{ijm} = \theta_m + \gamma_{1i} time_{ijm} + \gamma_2 cost_{ijm}/y_i + w_{ijm}\eta + \varepsilon_{ijm}, \tag{3}$$

where M_{ij} is the set of transportation modes available to individual i. We allow for a mode-specific random coefficient, θ_m, that has a normal distribution with mean μ_m and variance σ_m. The mode-specific random coefficient for walking is normalized to zero. These random-coefficients capture heterogeneous mode-specific (dis)amenities, scheduling or inconvenience costs that do not scale with the time or distance traveled. Variable $time_{ijm}$ denotes the commute duration between i's work location and home j via mode m. Note that driving time $time_{ij,car}$ is affected by the congestion level, an endogenous outcome as discussed above.[24] Time preference γ_{1i} follows a chi-square distribution with three degrees of freedom and mean μ_γ. The chi-square

[23] For the purpose of our analysis we consider the mode choices of different individuals within a household as independent because of limited information on the joint decision processs. When it comes to housing location, we are also modeling hypothetical commutes, for which the exact process of trip-chaining is not likely to be ex-ante clear to most households. Rather they may have a general sense of the relative cost of commuting for one home relative to another as specified in the EV term.

[24] Road congestion affects travel times for buses in addition to driving and taxi, however this effect is more complicated as it depends on the local characteristics of the roadway, the design of bus schedules, and location of bus stops. For the purpose of our analysis we treat buses as if they were in dedicated lanes unaffected by congestion, which may result in a an over-prediction of bus mode shares from our estimates.

distribution allows all individuals to have a positive value of time. The monetary cost of the trip is denoted as $cost_{ijm}$. Individual's sensitivity to the monetary costs of commuting is assumed to decrease in income: γ_2/y_i. Finally, variable w_{ijm} captures mode-commuter specific controls (such as the driving dummy interacted with commuter's gender) and ε_{ijm} is the i.i.d. error term.

The utility function makes it straightforward to calculate the value of time (VOT), which is $\frac{\gamma_{1i}}{\gamma_2} \cdot y_i$. In this formulation, the financial burden of travel scales with income and VOT can be conveniently expressed as a share of hourly income. VOT is a fundamental concept in transportation analysis and its empirical measurement is crucial for travel demand analysis and the evaluation of public policies. Under Becker (1965)'s framework of time allocation, the travel time should be valued relative to the after-tax wage rate assuming that time can be freely transferred between work and non-work activities (e.g., leisure or travel).

Expected utility from commuting, EV_{ij} in equation (2), is defined as the following:

$$EV_{ij} = \mathbb{E}_{\varepsilon_{ij}} \left(\max_{m \in M_{ij}} u_{ijm} \right), \tag{4}$$

where the expectation is over the set of i.i.d. draws ε_{ijm} across travel modes. Once we obtain the expected commuting utility for all commuting members within a household, we aggregate it to the household level and use it to measure house location j's ease-of-commute given individual/household i's work locations.

4.4 Market Clearing Conditions and the Sorting Equilibrium

This section defines the sorting equilibrium and the market clearing conditions for two interrelated markets in our model: the housing market and the market for driving. In the housing market, choices of individual households aggregate to the total housing demand. Housing prices adjust to equate demand and supply (housing supply is specified in Section 6.1). In the market for driving, the equilibrium driving speed and hence congestion level is jointly determined by driving demand through all individuals' travel mode choices and road capacity (the supply of the driving market). The two markets interact in two dimensions: the spatial location of households affects the distance of work commute and the travel mode, and hence the market for driving. At the same time, the level of traffic congestion that is determined in the driving market affects the attractiveness of residential locations through the 'ease of commuting' index as discussed above, and in turn the spatial distribution of households.

Commuting Mode Choice Conditional on home location j, the probability that individual/household i chooses mode m for his work commute is defined as:

$$R_{ijm}|i,j = r(\text{cost}_{ij}/y_i, \text{time}_{ij}(v_{ij}), \text{w}_{ijm}; \theta) \tag{5}$$

where cost_{ij}/y_i and $\text{time}_{ij}(v_{ij})$ denote the vector of travel cost as a share of individual i's hourly wage and the vector of travel time for each travel mode, respectively. The travel time by cars is a function of the driving speed between individual i's work location and home location j: v_{ij}, which is endegeneously determined in equilibrium.[25] Lastly, w_{ijm} captures all other individual-trip-mode specific characteristics and θ denotes all relevant parameters in travel mode choices: $\theta = \{\{\theta_m\}_m, \gamma_{1i}, \gamma_2, \eta\}$. We use $\mathbf{R} = \{R_{ijm}|i, j\}$ to denote all households' commuting mode choices.

Housing Choice The probability that household i chooses home j is determined by the distribution of random utility as specified in the housing demand model and is denoted as:

$$P_{ij} = h(\mathbf{EV}_i(\mathbf{v}), \mathbf{p}, \mathbf{X}, \boldsymbol{\xi}, \mathbf{z}_i), \tag{6}$$

where $\mathbf{EV}_i(\mathbf{v})$ is a vector of the 'ease-of-commute' index for all potential home locations giving household i's work locations. It links the housing market with the commuting mode choices, whose element is defined above in equation (4). We have made it explicit that the ease-of-commute index depends on driving speed $\mathbf{v}$. The following triplet, $\mathbf{p}$, $\mathbf{X}$, and $\boldsymbol{\xi}$, denotes prices, observed housing attributes, and unobserved housing quality for all homes in household i's choice set. The last term, $\mathbf{z}_i$, represents household i's demographic attributes. We use $\mathbf{P} = \{P_{ij}\}$ to denote all households' residential choice probabilities.

The aggregation of households' choice probabilities P_{ij} gives rise to the aggregate housing demand: $D_j = \sum_i P_{ij}(\mathbf{p}, \mathbf{v}), \forall j$. Note that the aggregate housing demand depends on both housing prices $\mathbf{p}$ and driving speed $\mathbf{v}$ through the east-of-commute index. Housing market clears when the aggregate demand is equal to aggregate supply (that also depends on housing price):

$$D_j = \sum_i P_{ij}(\mathbf{p}, \mathbf{v}) = S_j, \forall j \tag{7}$$

Driving Speed Demand for driving is determined by both housing locations and travel mode choices. Intuitively, mode choices determines the extensive margin of the driving demand (to drive or not), while housing choices determines the intensive margin of the driving demand (the commuting distance). Total driving demand and hence traffic density is the aggregation of all households' location and commuting decisions, which ultimately depends on the housing price $\mathbf{p}$ and driving speed $\mathbf{v}$:

$$D^v = \sum_i \sum_j R_{ij,car} \cdot \text{dist}_{ij,car} \cdot P_{ij} = g(\mathbf{p}, \mathbf{v}). \tag{8}$$

where $R_{ij,car}$ is the driving probability for household i living in location j, $\text{dist}_{ij,car}$ is the commuting distance, and P_{ij} is the probability that household i chooses location j.

[25] Specifically, $time_{ij,drive} = \frac{dist_{ij,drive}}{v_{ij}}$. Note that $time_{ijdrive}$ in w_{ij} is also determined by v_{ij}.

For a fixed road capacity, the driving speed $\mathbf{v}$ decreases in traffic density:[26]

$$\mathbf{v} = f\left(S^{v}\right)$$

Intuitively, to sustain a higher travel speed, a transportation system has to limit the traffic density to a greater extent. Traffic market clears when the aggregate traffic demand equals to traffic supply:

$$D^{v} = g\left(\mathbf{p}, \mathbf{v}\right) = S^{v}(\mathbf{v}) \tag{9}$$

Sorting Equilibrium A sorting equilibrium is defined as a set of housing choice probabilities $\mathbf{P}^{*}$, the vector of housing prices $\mathbf{p}^{*}$, a set of travel choice probabilities $\mathbf{R}^{*}$, and speed, $\mathbf{v}^{*}$, such that:

1. The housing market clears according to equations (6) and (7), and
2. The travel sector clears according to equations (8) and (9).

Our model follows the class of equilibrium horizontal sorting models with local spillovers studied in Bayer and Timmins (2005) and more closely in Bayer et al. (2004). If the error terms in both the housing equation (2) and the commuting mode choice equation (3) are from continuous distributions (such as the type I extreme value distribution), then the equation system (6), and (7), (8), (9) is continuous. The existence of such a sorting equilibrium follows from 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 (6) and (7), conditional on a set of observed and unobserved housing attributes ($\mathbf{X}$ and $\boldsymbol{\xi}$) as well as EVs. At the same time, (8), (9) implicitly define traffic speed $\mathbf{v}$ as a continuous mapping of a compact and convex set. Any fixed point of this mapping determines EVs and is associated with a unique vector of housing prices $\mathbf{p}^{*}$. The equilibrium housing choice probabilities $\mathbf{P}^{*}$ and travel choice probabilities $\mathbf{R}^{*}$ directly follow from the sorting equilibrium.

In this class of sorting models, the presence of a negative spillover due to traffic congestion leads to a unique equilibrium. In the simplest example when the driving speed is uniform across all locations (though the speed level negatively depends on the density of cars), this is a traditional demand system with a unique allocation (choice probabilities), as shown in Berry (1992). With heterogeneous congestion effects that differ across space, one can show that the equation system (6), and (7), (8), (9) remains a contraction mapping, and hence accommodates a unique fixed points (i.e., a unique equilibrium). It is worth noting that if there are positive spillovers (e.g., agglomeration effects), uniqueness is not guaranteed. Sufficiently strong positive spillovers could alter the rank-order of the location choices and give rise to multiple equilibria. A proof of existence and uniqueness is provided in Appendix D.[27]

[26] This gives rise to a congestion externality, because the driving decisions of others reduce the driving speed for household i. Figure 1 depicts the congestion externality and how the equilibrium congestion level is determined. We formulate the parametric relationship between driving speed and traffic density in Section 6.1.

[27] The proof follows Bayer and Timmins (2005). In the presence of positive spillovers, the unique equilibrium is more likely to arise with strong consumer heterogeneity, weak spillover effects, or a larger number of choices.

4.5 Identification and Estimation

Choice Set We first expand on the construction of the housing choice set discussed in Section 3.2. Computational and data limitations often, and in our case, require a restriction on the number of alternatives included for empirical estimation. Nevertheless, overly restrictive culling of the choice set can be problematic as documented by Banzhaf and Smith (2007). While it may be logical to restrict the choice set to a set of affordable or nearby homes, this literature suggests that this approach may unnecessarily bias estimation due to unobserved heterogeneity in the choice set definition, and so we eschew any restriction of the choice set based on attributes. In our implementation, we rely on choice-based sampling by taking one percent random sample from homes on the market 30 days before and after the sale date of the chosen home.[28] The consistency of choice-based sampling methods in multinomial logit and mixed logit models is formalized in McFadden (1978), Wasi and Keane (2012), and Guevara and Ben-Akiva (2013).

Identification & Estimation in Travel Mode Model To estimation of the parameters specified in the housing demand and travel choice models follows a two-stage process. In the first stage, we estimate the mode choice model via simulated maximum likelihood estimation (MLE) based on household travel survey data. The key parameters of interest are preferences for time and monetary costs. We include mode-specific random coefficients to control for mode-specific (dis)amenities or qualities that do not scale with the time or distance traveled. To further control for unobservables that could be correlated with travel time and monetary costs, the model also includes mode-specific fixed effects interacting with year fixed effects, district fixed effects, and household demographics (income categories, age). These interactions control for a rich set of time-varying and location specific unobservables by travel mode.

The key identification assumption in estimating mode choices is that the error term is not correlated with time and monetary costs. This assumption would be violated if, for example, the route-specific quality of public transit service (e.g., in terms of delay, comfort, or safety) is correlated with the route-specific monetary cost or travel time. Cost is likely to be exogeneous because the public transit is run by Beijing's public transit authority which sets bus and subway fares according to a uniform rule across all routes. Although some bus or subway routes are more congested than others, the fares do not vary by the level of congestion on-board. In addition, time is unlikely to correlated with unobserved shocks because during our sample period (2010 and 2014) real-time traffic apps are not widely used so people may be more likely to make travel decisions based on ex-ante estimate of travel time. This ex-ante estimate is likely orthogonal to the realization of traffic shocks on a particular day, hence requires no instruments.

Identification & Estimation in Housing Location Choice Model In the second stage, we first construct the ease-of-commuting index EV_{ij} for every house j in household i's choice set using the logsum expression in Equation (4) and parameter estimates from travel mode choices. This step consists of a large set of EV_{ij},

[28] To reduce the computational burden, we treat each year as a distinct market and this allows us to conduct the contraction mapping year by year. The average size of the choice set is 27 with a range of 3 to 56. We drop households with a choice set less than 5.

one for each home-work pair, including pairs between i's work location and homes that they did not choose but considered. The calculation of EV_{ij} is computationally intensive, requiring us to construct travel time and cost for all available modes for each home location, following the discussion in the household travel survey and Appendix Figure A3. While the application of this two-stage approach to residential sorting is new to our knowledge, similar approaches of nesting the logsum values from random utility models have been used by Capps et al. (2003) and Phaneuf et al. (2008) in healthcare and recreational demand respectively. The estimated $\widehat{EV}_{ij}$ enters the housing choice model as an observed housing attribute.

The parameters in the housing demand model themselves are estimated using a two-step procedure, with the first step being a simulated MLE with a nested contraction mapping and the second step being a linear IV/GMM. The two-step strategy follows Berry et al. (1995) and Bayer et al. (2007) in order to address unobserved attributes that could be correlated with housing prices. Unobserved attributes (e.g., quality) ξ_j in Equation (2) could render the price variable endogeneous and bias the price coefficient toward zero. The nested contraction mapping algorithm isolates price endogeneity into a linear framework which permits the usage of the IV strategy. Following the structural demand literature we re-organize Equation (2) into a sum of household-specific utility μ_{ij} and mean utility δ_j (or alternative-specific constants) which absorbs variation from unobserved housing attributes ξ_j:

$$U_{ij} = \mu_{ij}(\theta_2) + \delta_j(\theta_1) + \varepsilon_{ij} \tag{10}$$

$$\mu_{ij}(\theta_2) = \alpha_2 ln(y_i) p_j + \sum_k X_{jk} z_{ik} \beta_k + \phi_m EV_{ijm} + \phi_f EV_{ijf} \tag{11}$$

$$\delta_j(\theta_1) = \alpha_1 p_j + \mathbf{x}_j \bar{\beta} + \xi_j. \tag{12}$$

Further details about the estimation can be found in Appendix E.

Once θ_2 and δ_j are estimated, we then use linear regressions to estimate preference parameters (θ_1) in mean utilities δ_j as specified in Equation (12) using three different sets of variables. First, the average housing and neighborhood attributes (excluding price and the EV term) within 3 kilometers outside the same complex sold within two-month time window from a given house. The identification assumption is that the average attributes of housing choices outside the same complex from housing choice j are not correlated with the unobserved housing attributes ξ_j and have no direct effect on the utility from housing choice j. But due to the housing market competition, they are correlated with the price of housing choice j. Consider, for example, two identical housing choices in the same neighborhood being sold at two different points in time (or from two different neighborhoods with the same amenities but sold at the same time), prices may be different due to varying intensity of competition (e.g., the availability of other housing choices) on the market faced by the two housing choices at the time of their sales.

The second set of instruments is the interaction between the first set of IVs and the winning odds of the vehicle licence lottery policy. The winning odds have decreased dramatically from 9.4% in Jan. 2011 to 0.7% by the end of 2014. The interaction terms capture the likely impact of license lottery policy on the nature of housing market competition and price setting. Decreasing winning odds push up demand (and hence prices) for houses in desirable locations, such as places close to the subways or the city center. The third set of IVs is the number of homes sold in the real-estate listing dataset in the two-month time window of a given home, which is also a proxy for market competition.

5 Estimation Results

We begin by presenting estimation results for the commuting mode choice model. Then we construct the ease-of-commuting index, which captures the value of commuting options for work trips based on home and work locations. Taking the ease-of-commuting index as an observed housing attribute, we then estimate the housing demand model.[29]

5.1 Commuting Mode Choice

We assume that the error terms in both the housing equation (2) and the commuting mode choice equation (3) have the type I extreme value distribution. Table 3 presents parameter estimates for six specifications of the multinomial logit model based on work commutes from household travel survey in 2010 and 2014. The first three specifications do not have random coefficients and the heterogeneity comes only from the interaction between the travel cost and income. The last three specifications include random coefficients on travel time to capture unobserved consumer heterogeneity. The value of time is represented as the percentage of the hourly wage, and is defined by the ratio of the parameter on travel time and that on travel cost.

Column (1) begins with interactions between year dummies (2010 or 2014) and mode-specific constants (car, taxi, bus, subway, walking, and biking). The implied VOT from these estimates is 0.757 times the hourly wage. Column (2) adds the interactions between mode-specific constants and trip characteristics including trip distance and ring road dummies of the trip origin and destination. Trip distance could affect mode choices because important trip characteristics such as uncertainty in travel time will likely scale with the length of a trip. This uncertainty, typically called travel time reliability in the transportation literature, has been shown to be an important factor in travel decisions (Brownstone and Small, 2005; Small et al., 2005; Tseng et al., 2009). Ring road dummies for trip origin and destination may also capture differences in the frequency or quality of the public transit, which could affect travel model choice. Including these sets of interactions dramatically affects the coefficient estimate on the travel cost variable, resulting an implied VOT being 0.342 of the hourly wage.

Column (3) further adds the interactions of model-specific constants with household demographics including age, gender, education, vehicle ownership, the number of workers, and household size. Adding these

[29] A key assumption underlying our approach is that, after accounting for location and demographic differences, preferences for commuting mode choice from the travel survey are representative of those for home buyers in the mortgage data. A large literature in environmental economics considers conditions under which the approach of transferring preferences for non-market amenities is valid (Boyle and Bergstrom, 1992; Rosenberger and Loomis, 2003).

variables greatly improves the fit of the model and better captures the heterogeneity in mode choices across demographic groups. The VOT estimate is 33.9% of hourly wage. Columns (4) to (6) use a chi-square distribution with three degrees of freedom to approximate heterogeneous preference on travel time following Petrin (2002).[30] In addition to the random coefficient on travel time, Column (5) also allows a random coefficient on the mode of driving. Column (6) further incorporates random coefficients for each of the five travel modes (with walking as the reference group).

Our preferred specification is Column (6). The preference heterogeneity for different travel modes is assumed to be i.i.d. normal and captures the impact of unobserved demographics on mode choices. For example, some commuters choose driving or taxi not because of their high VOT but because of scheduling constraints. Some commuters choose walking or biking for their exercise benefit. The dispersion on the preference parameters for all transit modes is quite large, suggesting significant heterogeneity for different modes. Adding these random coefficients leads to much stronger consumer sensitivity to travel cost. The average (median) VOT estimate is 95.6% (84.6%) of hourly wage and these estimates are within the range typically found in the recent literature.[31]

5.2 Housing Location Choice

We now turn to the estimation results of the housing demand model described in Section 4.2. We first construct the ease of commuting (EV_{ij}) for each member within household i based on parameter estimates from the travel choice model:

$$EV_{ijk} = \mathbb{E}_{\varepsilon_{ijk}}\left(\max_{m\in M_{ijk}} u_{ijkm}\right) = \log\left(\sum_{m\in M_{ijk}} \exp\left[\theta_m + \gamma_{1ik} time_{ijkm} + \gamma_2 cost_{ijkm}/y_i + w_{ijkm}\eta\right]\right), k \in \{Male, Female\}.$$

For each home in household i's choice set, we generate this measure separately for male (61% main borrower) and female (39% main borrower) members based on their work locations.[32] These two variables enter the housing demand as additional household-specific attributes. We first present the MLE estimates of household-specific preference parameters, and then discuss the IV estimates for coefficients in the mean utility.

[30] We winsorize the top and bottom 5% of the distribution to bound the distribution and to minimize the impact of outliers. The three degrees of freedom provide the best model fit.

[31] Appendix Figure A11 depicts the histogram of the VOT estimate in our sample. The empirical estimates of VOT in the literature vary as they come from different contexts and methodology. In the context of travel demand, the estimates typically range between 30% and 100% of hourly income (Small et al., 2007; Small, 2012). The US Department of Transportation recommends 50% of the hourly income as VOT for local personal trips (e.g.,work commute and leisure but not business trips) to estimate the value of travel time savings (VTTS) for transportation projects (USDOT, 2015). Using a discrete choice framework similar to ours, Small et al. (2005) estimate the median VOT at 93% of hourly wage for commuters in Los Angeles based on data from both travel surveys and choice experiments. Leveraging the tradeoff between vehicle driving speed and gasoline usage, Wolff (2014) estimates the average VOT of 50% of hourly wage based on traffic speed data in eight rural locations in Washington State. Buchholz et al. (2020) use the tradeoff between wait time and price among users on a large ride-hail platform in Prague and find the average VOT to be equal to users' wage during work hours. Goldszmidt et al. (2020) find an average (median) VOT of 75% (100%) of hourly (after-tax) wage based on a large-scale field experiment by the ridesharing company Lyft in 13 US cities by leveraging the random variation in customer wait time and fare.

[32] If a family member is unemployed, we set $EV_{ij} = 0$ for that member, effectively ignoring this term in the decision process.

Table 4 reports the estimates of heterogeneous preference parameters for three specifications: without the EV terms, with the EV terms, and with random coefficients on the EV terms. The coefficient estimates from these three specifications are by and large similar, except for the coefficient on the interaction between age great than 45 and distance to key schools. Housing price is interacted with household income, which is used as an proxy for household wealth.[33] As expected, high-income households tend to be less price sensitive.

Both EV terms in the second and third specifications have positive and significant coefficients estimates. The log-likelihood value increases substantially from Column (1) to Columns (2) and (3), indicating strong explanatory power of the EV terms. The estimates imply that households prefer homes with better ease-of-commute measures, i.e., more convenient for work trips, for both family members. The coefficient estimates in our preferred specification (3) suggest that an average household is willing to pay ¥18,000 (21,000) more on a home to save ¥1 in the male (female) member's work commute, or ¥185,000 (219,000) more to shorten the male (female) member's work commute by 10 minutes. The coefficient estimate on the EV term for the female member is 18% larger than that for the male member, suggesting that households prioritize the female member's ease-of-commute in housing choices. This is consistent with the descriptive evidence that females tend to live closer to their work locations (Appendix Figure A7). In addition, there is significant preference heterogeneity across households on the EV terms, e.g., due to unobserved household demographics.

We interact the age group dummies with the distance to the nearest signature elementary school. Enrollment to these top schools is restricted to the residents in the corresponding school district, and homes in these districts command a high premium. The baseline group is those with the primary borrower younger than age 30. The interaction coefficients in all specifications are negative and highly significant, though borrowers between age 30 and 45 exhibit the strongest preference for proximity to key schools, as they are most likely to have school age children.

We do not observe the household size. To capture preference heterogeneity on home size due to variation in household size, we use the age of the primary borrower as a proxy and interact age group dummies with home size. Older households have stronger preference for home size. The group with age over 45 has the strongest preference, likely due to the presence of both children or elderly grandparents living in the household, a common household structure in China.

Estimates for coefficients in the mean utility are reported in Table 5. Columns (1) and (2) use OLS, while columns (3)-(6) are from IV regressions. All regressions month-of-sample interacted with district fixed effects to capture time-varying changes in market conditions and amenities that could vary across the 18 districts in Beijing. Columns (2)-(6) also include neighborhood (158 different jiedaos) fixed effects to capture unobserved time-invariant neighborhood amenities. We use the three sets of IVs for housing prices that are discussed in Section 4.5: the average attributes of the homes within 3km outside the sample complex sold in a two-month time window of a given home; the interaction between the first set of IVs and the winning odds of the vehicle licence lottery; the number of homes sold in a two-month window. Our preferred specification

[33] The interaction itself only captures part of the consumer price sensitivity since housing price also enters the mean utilities (household-invariant utilities).

is Column (6) with all instruments. The first stage F- statistics is 14.22, and the over-identification test cannot be rejected at 10% level.

Across all columns, the price coefficient estimate is negative and statistically significant. The IV estimates are larger (in magnitude) than the OLS estimates, consistent with the findings in the demand literature that unobserved product attributes bias OLS estimate toward zero. The average price elasticities vary from -1.34 to -1.94 in Columns (4)-(6), suggesting elastic housing demand.[34] The coefficient estimates from IV regressions in columns (3) to (6) are all intuitively signed. Households prefer larger homes and homes closer to the signature schools, but dislike older buildings and homes that are far away from parks.

Based on parameter estimates from the last columns in 4 and 5, the sample average of the implied income elasticity of housing demand and the income elasticity of marginal driving cost is 0.10 and 0.78, respectively. To our knowledge, these are the first estimates of the two elasticities based on data from China. Our estimate of the income elasticity of demand for housing size is somewhat smaller than those based on the US data while the elasticity of marginal driving cost is largely consistent with the literature. Using 2003 American Housing Survey, Glaeser et al. (2008) find the elasticity of lot size to be from 0.25 to 0.5 and they argue that these estimates likely provide an upper bound on the true income elasticity of land demand with respect to housing prices. In addition, our elasticity of housing demand is with respect to the (condo) interior size rather than the lot size.

6 Counterfactual Simulations

We now utilize the estimates from the housing and travel mode choices to conduct counterfactual simulations. We examine five policy scenarios: the driving restriction, congestion pricing, subway expansion, driving restriction and subway expansion, and congestion pricing and subway expansion. The first two are demand-side policies while the third is a supply-side policy. The last two counterfactual analyses examine combinations of different policy mix. The driving restriction scenario follows the actual policy employed in Beijing during our sample: a vehicle is prohibited from driving in one of the five working days. Under congestion pricing, which is hypothetical, we choose a distance-based congestion charge to achieve the same level of congestion reduction as the driving restriction to facilitate comparison. The key difference between these two policies is that driving restriction is a command-and-control approach while congestion pricing is a market-based policy that affects the price that drivers pay. The subway expansion simulation compares the subway network in 2008 and 2014. During this period, the length of the subway network increased from 100km to 486km, with 8 new lines opened for operation. The details of the simulation approach are provided in Appendix F.

Before we present simulation results, we first validate the structural model by comparing the model's

[34]To evaluate the impact of ignoring the work commuting attribute (the EV terms) on price elasticities, Appendix Table A4 reports the second-stage regression results based on the first specification without the EV terms in Table 4. The price coefficient estimates and the price elasticities are smaller in magnitude, consistent with the downward bias due to unobserved attributes. Timmins and Murdock (2007) find a 50% downward bias in the estimation of consumer welfare from recreation sites when congestion on site is ignored in demand estimation.

predictions to the mortgage data. We simulate the market equilibrium under the 2008 subway network with and without the driving restriction. Then we examine the effect of the driving restriction on the housing price gradient with respect to the subway access using the model's predicted equilibrium housing price. The results are reported in Appendix Table A5. Consistent with the reduced-form evidence in Table A1 based on observed data, the driving restriction steepens the price gradient with respect to subway access. The coefficient estimate on the interaction between subway distance and the driving restriction policy dummy is -0.01 compared with -0.023 (with a standard error of 0.013) in the regression with observed data.

6.1 Travel and Housing Choices and Equilibrium Prices

We begin by considering the simulation results with sorting in Table 6. To incorporate sorting, we allow households to re-optimize their housing locations and solve for the new market equilibrium under each scenario. The first three columns are under the 2008 subway network, while the next three are under the 2014 subway network, reflecting the effect of subway expansion. Column (1) shows the baseline results while columns (2) to (6) present the differences relative to the no-policy baseline in column (1). Panel A reports changes in the share of mode choices and equilibrium traffic speed under each scenario, Panel B displays key housing market outcomes, and Panel C presents the welfare results. The results are shown separately for two income groups: households with income above the median (rich) and those with income below the median (poor) to reflect distributional considerations following Section 2.

Several features of endogenous congestion and household sorting have important implications on the policy effectiveness in terms of congestion relief. Sorting introduces two countervailing forces under a driving restriction. On one hand, a driving restriction incentivizes households to live closer to work. This reduction in the driving distance further magnifies the alleviation of congestion as a result of travel mode changes when households substitute away from driving. On the other hand, congestion reduction from the driving restriction improves the driving speed, makes driving less costly, and hence disproportionally increases driving among those with a long commute. As the equilibrium congestion is affected by both intensive and extensive margins, the increase in the extensive margin dampens the effectiveness of a driving restriction policy.

In contrast, as a distance-based policy, congestion pricing affects driving in the same direction for both the intensive and extensive margins and delivers a larger congestion reduction with sorting.[35] Congestion pricing induces mode shifting in the same directions as a driving restriction, but there are two key differences. First, while congestion pricing generates a modest driving reduction among high-income households, it leads to a much bigger reduction among low-income households. The large response from the low-income group is driven by the fact that low-income households are more sensitive to the travel cost and hence to congestion charges. Second, although the level of congestion reduction is the same under the two policies, the share of commuting trips via driving remains higher under congestion pricing. This is because distance-based congestion pricing reduces congestion through an intensive margin by disproportionally reducing longer driving

[35] We set the congestion price to be ¥0.92/km to achieve the same congestion reduction as that under a driving restriction.

trips.

Despite the large increase in subway usage, subway expansion leads to the smallest congestion reduction. Column (4) presents the impact of subway expansion from the 2008 network to the 2014 network. Traffic speed increases by about 7 percent with sorting, only about 40% of what is achieved under the driving restriction and congestion pricing, the two demand-side policies. Our estimate is lower than what is implied by recent empirical studies that focus on the short-run impact of the subway system on traffic congestion (Anderson, 2014; Yang et al., 2018; Gu et al., 2020).[36] Our analysis shows that the effectiveness of subway expansion is attenuated by sorting as illustrated in Figure 6. Both income groups move farther away from work and commute longer distances with a more extensive subway network. This additional induced travel demand from transportation infrastructure investment undermines the objective of congestion reduction, a result consistent with the previous literature (Downs, 1962; Vickrey, 1969; Duranton and Turner, 2011). The results on the further separation of workplace and residence from subway expansion corroborate with the evidence in Gonzalez-Navarro and Turner (2018) and Heblich et al. (2020).

Nonetheless, the expansion dramatically increased subway access for both income groups: the distance to the nearest subway station from home is reduced by about 80% for both groups. As a result, the expansion increases subway ridership by 51% and 56% among high- and low-income groups, respectively. Subway expansions reduce the share of all other travel modes, though the reduction in taxi and bus trips is more pronounced. Low-income groups are much more likely to substitute from other travel modes toward subway, due to their larger price sensitivity. Overall, the reduction in the driving share of commuting trips as a result of the subway expansions is about 43% of that observed under driving restriction and congestion price, leading to smaller congestion relief. While the substitution away from bus, bike, and walk trips toward subway trips does not alleviate traffic congestion, it improves welfare by offering quicker and hence better commuting choices for some trips.

We calibrate the congestion charge so that both congestion pricing and driving restrictions achieve the same level of congestion reduction under the 2008 subway network. At the same level of congestion charges, congestion pricing is more effective than driving restrictions under the 2014 subway network. That is, the market-based demand policy and the supply side policy exhibit complementarity by producing a stronger aggregate impact.[37] The results reflect two underlying countervailing forces. On the one hand, subway expansion increases the attractiveness of using subways and hence reduces the share of driving. This leaves a smaller room for and reduces the impact of demand-side policies among an average driver. On the other hand, the demand-side policy could be more effective in affecting the infra-marginal drivers who now have a better subway network to switch to. The first force appears to dominate under the driving restriction but the

[36] Using a regression discontinuity (in time) approach, Anderson (2014) finds a 47% increase in highway traffic delays during the peak hours from the shutdown of the Los Angeles bus and rail lines for 35-days. Yang et al. (2018) shows that the subway expansion in Beijing from 2009 to 2015 reduces traffic congestion by 15% on average using a 120-day window surrounding subway opening. 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 45 subway lines opened across 42 Chinese cities during 2016 and 2017.

[37] As demonstrated in Akbar et al. (2018), the supply-side constraint (poor transport infrastructure) is a key determinant in traffic speed across cities India, highlighting the importance of transport infrastructure provision.

second force is stronger under congestion pricing. Congestion pricing affects both the extensive and intensive margins, both of which could be reinforced by the subway expansion.

Column (5) presents the results from the combination of subway expansion and driving restriction while column (6) shows the combination of subway expansion and congestion pricing. The impacts on driving under each of these two columns are similar to the sum of the impacts from the two individual policies. There are two countervailing forces at play under the combination of supply-side and a demand-side policies. First, the policies could have redundant impacts in reducing driving trips: some of the driving trips would be reduced under either the supply-side or the demand-side policy, leading to a smaller aggregate impact than the sum of the impacts from individual policies. Second, the supply-side policy could enhance the demand-side policy in that the larger subway network makes substitution away from driving easier under driving restriction or congestion pricing. Indeed, as the subway becomes a more attractive option, both driving restriction and especially congestion pricing lead to a larger substitution from driving to subway under the 2014 network than the 2008 network.

It is instructive to compare these results to those without sorting that are presented in Appendix Table A6.[38] The comparison illustrates that sorting reinforces the impact of congestion pricing on congestion reduction but weakens the impact of subway expansion. Finally, the optimal congestion price (with revenue recycling) that maximizes consumer surplus is ¥1.2/km (the speed increase is 3.81km/h) without sorting compared to ¥1.4/km (the speed increase is 4.70km/h) with sorting under the 2014 subway network (as shown in Panel (b) of Figure 8 below).

Spatial Distribution of Changes While driving restrictions and congestion pricing both reduce congestion, the impacts on the spatial distribution of households differ under sorting. Consistent with the reduced-form evidence, a driving restriction induces high-income households to move closer to subway, while pushing low-income households to move farther away from work and subway. In contrast, congestion pricing induces both high- and low-income groups to move closer to work, hence reducing "wasteful commuting" for both groups. As congestion pricing is distance-based, it induces a stronger sorting than driving restriction. As illustrated in Figure 6, driving restrictions lead to small changes in commuting distance and often in opposite directions across neighborhoods, but congestion pricing leads to a larger reduction in commuting distance in nearly all neighborhoods relative to the no-policy scenario. In terms of subway access, both driving restrictions and congestion pricing make high-income households move closer to the subway while low-income households move further away from the subway compared to the baseline scenario. The exact opposite effect on distance between home and subway for the two income groups is driven by the fact that the subway network is the same for all households and therefore households have to compete for the closeness to subways in a zero-sum game, but work locations differ across households so a Pareto improvement in commuting distance is possible.

In Figure 7, while both driving restriction and congestion pricing increase the prices of homes that are closer to work centers, the impact is stronger under congestion pricing driven by the distance-based nature of

[38] The congestion price is kept at ¥0.92/km as in Table 6 with sorting to facilitate comparison.

congestion pricing. Under congestion pricing, housing prices in northwest parts of the city (near work centers) would increase by about 2,000 ¥/m^2 while those in some parts of southeast that are far from work centers and public transit would decrease by 2000 ¥/m^2 (from a baseline average price at 24022 ¥/m^2). Subway expansion has opposite spatial impacts on housing price: the price increase is mainly observed among homes farther away from the city center (where the public transportation is poor prior to the expansion) but also along the new subway lines as shown by the green lines in Panel (c). Home prices increase by as much as 4000 ¥/m^2 in some southwest parts of the city, where the subway expansion is greatest and the prices have been the lowest historically. With both the subway expansion and congestion pricing, the price impacts of subway expansion dominate those from congestion pricing.

To understand the differential impact on home prices with respect to access to subway, Appendix Figure A12 plots the the housing price gradient with respect to the subway distance for 2008 subway network, and 2014 subway network, respectively. The bid-rent curve is steeper under the 2014 network (-¥1900/m^2 per km) than 2008 network (-¥700/m^2 per km) because the 2014 network is larger and hence the proximity to this network is more valuable to commuters. The bid-rent curve under the 2014 network shifts down, reflecting the composition change of the homes whereby the subway expansion reaches to cheaper homes farther away from the city center.[39]

With Sorting and Supply Adjustment Finally, we present the results with sorting in Table 7 and additionally allowing the housing supply to adjust. Housing supply is modeled as a constant elasticity function of local home prices.[40] To fix ideas of how incorporating the supply-side changes our previous findings, think about a home whose price increases after a policy. Now housing supply of that home will increase. Hence in the new equilibrium, we expect part of housing price changes to be absorbed by the increasing housing supply. The additional increases in the housing supply of attractive homes will allow more efficient sorting, hence magnifying the sorting effects compared to the case when the housing supply is fixed.

This intuition is supported by our simulation results. In terms of the driving speed, housing supply adjustments make the congestion pricing policy more effective (the driving speed improvement increases from 3.13km/h to 3.26 km/h), while attenuating the gains from subway constructions (from 1.49km/h to 1.13 km/h). These changes could be driven by both the intensive and the extensive margin. With a varying housing supply, distance to work under congestion pricing further decreases relative to the case where the supply is fixed (from -0.15km to -0.26 km for high-income male member and from -0.07 km to -0.19 km for low-income male member), while the distance to work under subway constructions increases relative to the case where the

[39] Similarly, from Figure A13, we find *change* in the subway price gradient from congestion pricing is larger (-¥80/m^2 per km) than that from driving restriction (-¥10/m^2 per km). While the driving restriction makes homes close to the subway more attractive for everyone, congestion pricing, being a distance-based policy, makes homes close to subway more attractive for those who live far from work, and for those who are sensitive to a cost increase. The differential impact across households from congestion pricing therefore steepens the bid-rent curve more. We also find that the increase in the price premium from subway proximity due to congestion pricing is smaller under the 2014 network than that under the 2008 network under either policy. As the subway network becomes more attractive, fewer commuters use driving as the travel mode under 2014 network, implying less competition for the homes close to the subway.

[40] We set the elasticity to be 0.52 following Wang et al. (2012)'s estimates on housing supply elasticity in Beijing.

supply is fixed (from +0.33 km to +0.76 km for high-income male member and from +0.15 km to +0.61 km for low-income male member). Driving probability, on the other hand, does not change significantly when supply adjustment is allowed. Hence we conclude that the changes in speed should mostly be contributed to intensive margin. In other words, the supply-side adjustment under congestion pricing increases the supply of homes in the city center and allows people to live even closer to their work, which shortens their commuting trips and magnifies the anti-sprawling effects. The supply-side adjustment under subway construction increases the supply of homes in suburban areas, making people live farther away from their work and exacerbating the sprawling effects. We elaborate on the welfare consequences of the supply adjustment in the next section.

6.2 Welfare Analysis

Panel C in Table 6 and 7 as well as Appendix Tables A6 presents the welfare results under the five scenarios, relative to the baseline scenario of no policy and the 2008 subway network. To construct net welfare, we keep a balanced government budget. Subway construction and operation are funded by a head tax while the revenue from congestion pricing is recycled back to households via a lump sum. The discussion in this section focuses on our benchmark results that are presented in Table 6.

Since transportation infrastructure such as subways are durable, we assume a 30-year time-span during which the capital cost should be recouped. The choice of the time span matters for the magnitude of the welfare but does not qualitatively affect the comparison across the policy scenarios. To be conservative, we assume that households only benefit from commuting trips and ignore utilities they derive from non-commuting trips.[41] A related issue is the utility function specification where the total housing price rather than rental price is used. So the numbers on consumer surplus reported below should be considered as a discounted lifetime utility over a period of 30 years.[42]

There are several key findings. First, driving restrictions reduce consumer welfare especially for the high-income group despite the reduction in traffic congestion in both Table 6 and Table A6.[43] There are two opposing effects as illustrated in Figure 1. On the one hand, the policy should reduce the deadweight loss from congestion relief. On the other hand, the policy leads to losses in consumer surplus as it removes the choice of driving from households' choice set one day a week (equivalent to shifting the driving demand curve downward). The second effect dominates: driving restrictions are associated with a ¥92 thousand loss per household. The welfare loss is larger for high-income households because they are more likely to have cars and commute via driving (a ¥165 thousand loss per high-income household and a ¥18 thousand loss per low-income household). Sorting exacerbates the welfare loss by 1% and 2% for high- and low-income groups, respectively.

[41] Work trips account for about 60% of all trips and 75% of the total travel distance in the 2014 travel survey.

[42] Our parameter estimates suggest a much larger marginal utility of housing price than the marginal utility of (per-trip) travel costs. On average, a 30-year housing tenure would imply 467 trips per year for male borrower ad 577 trips for female borrower. The implied trip numbers are plausible.

[43] The average household income in Beijing is about ¥167k in 2014. The average is ¥204k and ¥110k for high- and low-income groups, respectively.

In contrast to the driving restriction policy, congestion pricing disproportionally affects low-income households more in mode choices and consumer surplus given that these households are more price-sensitive. This distributional concern could hinder the political acceptability of congestion pricing and explain the limited adoption of congestion pricing despite it being continuously advocated by economists and urban planners. With the recycling of the congestion revenue that is uniform across income groups, congestion pricing leads to a welfare gain overall. This highlights the role that revenue from congestion pricing can play in addressing equity concerns. Sorting strengthens congestion reduction (an additional 0.13 km/h increase with sorting, a 4% increase) and enhances welfare gain from congestion pricing (an additional ¥5.8 thousand per household from sorting, a 10% increase), consistent with the finding in Langer and Winston (2008) based on a cross-section analysis for 98 US cities.

Figure 8 shows the welfare gain from different levels of congestion pricing under different assumptions. The optimal congestion price is 1.6 ¥/km and 1.4 ¥/km under the 2008 and 2014 subway networks, respectively. At the optimal levels of congestion pricing, sorting would increase consumer welfare by 20%-30%, and supply adjustment contributes to another 10%-20%. The additional gain of sorting and supply-side adjustment both stems from the reduced deadweight loss from the congestion externality due to the further increase in traffic speed. This result highlights another reason for incorporating sorting to understand the impacts and cost-effectiveness of different transportation policies. Sorting also shifts the optimal congestion pricing level to the right, achieving higher level of equilibrium traffic speed (26.2 km/h with sorting versus 25.3 km/h without sorting). The figure also shows that under a wide range of levels of congestion pricing (<¥2.5/km) and different sorting assumptions, consumer welfares are always positive. This indicates that congestion pricing is likely to be an effective tool even when governments cannot gauge the exact optimal pricing level a priori.

In our simulation table, we assume away the implementation cost of the congestion pricing system in Beijing because the congestion pricing has yet to be implemented. Singaporean government's implementation of a satellite based road pricing system in 2021 provides a back of envelope calculation on the cost of a congestion toll system.[44] The system costs Singaporean government around $ 400 million, which is roughly what would cost Beijing to adopt a similar system. This cost translates into around 1,000 RMB per household in Beijing, likely to be negligible in our welfare analysis (under optimal pricing, one-year operation of congestion pricing will create 3,000 RMB toll revenue per household, already enough to cover the cost).

Third, although subway expansion from 2008 to 2014 does not achieve the same level of congestion reduction as driving restriction and congestion pricing, it leads to a larger increase in consumer surplus, especially for the high-income group with and without sorting. The large increase in consumer surplus is consistent with the fact that the share of subway trips increases more than half after the expansion. Sorting slightly reduces the aggregate welfare gain due to the reduction in traffic speed. To gauge the magnitude of net consumer surplus, we calculate the construction cost and the operating cost during a 30-year period. Assuming that the cost are financed through a uniform lump-sum tax across households, consumer surplus for

[44] For more information on Singapore's road pricing syste, refer to https://www.zdnet.com/article/singapore-readies-satellite-road-toll-system-for-2021-rollout/

high-income households exceeds their tax burden in both with-sorting and without-sorting scenarios, while consumer surplus of low-income households exceeds their tax burden without sorting, and marginally get hurt with sorting.

Fourth, the combination of congestion pricing and subway expansion achieves the largest congestion reduction and has the potential to achieve the largest welfare gain across five policy scenarios. The results in Column (6) of Table 6 also show that the revenue from congestion pricing (107.8 thousand ¥per household) could fully cover the costs of subway expansions (103.0 thousand ¥per household). Earlier studies have shown that the revenue from optimal road pricing could be used to fully finance the capital and operating costs of transportation infrastructure under the condition that capacity and (pavement) durability costs are jointly characterized by constant returns to scale (Mohring and Harwitz, 1962; Winston, 1991; Verhoef and Mohring, 2009). In our simulation, the cost of subway construction can indeed be covered by congestion pricing. The comparison highlights the advantage of congestion pricing from congestion reduction, welfare, and fiscal perspectives.

7 Conclusion

Transportation plays a critical role in determining residential locations, while at the same time, the pattern of residential locations affects the efficiency of the transportation system and policies. This study provides, to our knowledge, the first unified equilibrium sorting framework with endogenous congestion to empirically evaluate the efficiency and equity impacts of various transportation policies taking into account the interaction between the transportation system and the housing market. Our empirical analysis leverages spatially disaggregate data on travel behavior and housing transactions with information on residential and work locations in Beijing from 2006 to 2014. We first estimate a flexible travel mode choice model and then construct measures of ease-of-commuting for different homes based on the job locations of each working members. This home-work pair specific measure is determined by traffic congestion and transportation infrastructure, and enters the housing demand model as an observed housing attribute. Based on the estimates of model parameters, we conduct counterfactual simulations to examine the impacts and welfare consequences of various transportation policies from both the demand- and supply- sides: a driving restriction, congestion pricing, subway expansion, and combinations of demand-side and supply-side policies.

The parameter estimates from the flexible travel mode choice model imply the median value of time being two thirds of the hourly wage. The estimates from housing demand illustrate the importance of incorporating work commute in the model: doing so improves the model fit dramatically and affects preference estimates on other housing attributes. An average households is willing to pay 20% more in exchange for an easier work commute for the female member than for an equivalent improvement for the male member. The optimal congestion pricing taking into account sorting and supply-side responses is estimated to be ¥1.6/km and 1.4/km under 2008 and 2014 subway network, respectively. Allowing for equilibrium sorting could have significant implications on welfare estimates of urban transportation policies: sorting accounts for over 20-

30% of the welfare gain from optimal congestion pricing.

While different policies can be designed to attain the same level of congestion reduction, they lead to different spatial patterns of residential location. A driving restriction leads to an income-stratified structure that favors high-income households with respect to access to subways and work, which could disadvantage low-income households in the long run. Congestion pricing incentivizes residents to live closer to their work locations and the equilibrium sorting leads to a more compact city with shorter commutes to work for both income groups. Subway expansion does the opposite by increasing the separation of residence and workplace.

In addition to residential locations, different policies generate drastically different efficiency and equity consequences. While the driving restriction reduces social welfare due to the large distortion in travel choices, congestion pricing is welfare improving for both income groups with a uniform recycling of congestion revenue. A driving restriction generates a larger welfare loss among high-income households, while congestion pricing hurts low-income households more in the absence of revenue recycling, pointing to the underlying difference in political acceptability between the two policies. These results underscore both the distributional concern and the efficiency gain from congestion pricing relative to the driving restriction. The combination of congestion pricing and subway expansion stands out as the best policy among all policy scenarios from congestion reduction, social welfare, and fiscal perspectives. With the congestion pricing of ¥0.92/km and the observed subway expansion from 2008 to 2014, the policy mix generates the largest improvement in both traffic speed (about 25%) and welfare (¥43,000 per household). In addition, it is self-financing in that the revenue from congestion pricing could fully cover the cost of the subway expansion.

Our analysis does not consider the potential impacts of policies on intercity migration and the labor market. Both could be additional margins of adjustments that affect traffic congestion and urban spatial structure. Future research could relax these assumptions to capture even broader general equilibrium effects. Incorporating these channels in our current framework with rich heterogeneity and endogenous congestion would necessitate additional data and computational resources. Nevertheless, such a framework would allow the existence of both congestion and agglomeration forces, which could affect the nature of the interaction between transportation policies and urban spatial structure.

References

Akbar, Prottoy, "Who Benefits from Faster Public Transit?," Technical Report, working paper 2020.

Akbar, Prottoy A, Victor Couture, Gilles Duranton, and Adam Storeygard, "Mobility and Congestion in Urban India," Technical Report, National Bureau of Economic Research 2018.

Albouy, David and Aarsh Farahani, "Valuing Public Goods More Generally: The Case of Infrastructure," 2017. Upjohn Institute working paper.

Allen, Treb and Costas Arkolakis, "The welfare effects of transportation infrastructure improvements," Technical Report, National Bureau of Economic Research 2019.

Almagro, Milena and Tomás Domínguez-Iino, "Location Sorting and Endogenous Amenities: Evidence from Amsterdam," 2020. Working Paper.

Anas, Alex and Ikki Kim, "General Equilibrium Models of Polycentric Urban Land Use with Endogenous Congestion and Job Agglomeration," *Journal of Urban Economics*, September 1996, *40* (2), 232–256.

_ **and Robin Lindsey**, "Reducing urban road transportation externalities: Road pricing in theory and in practice," *Review of Environmental Economics and Policy*, 2011, p. req019.

Anderson, Michael L, "Subways, strikes, and slowdowns: The impacts of public transit on traffic congestion," *The American Economic Review*, 2014, *104* (9), 2763–2796.

_ **, Fangwen Lu, Yiran Zhang, Jun Yang, and Ping Qin**, "Superstitions, street traffic, and subjective well-being," *Journal of Public Economics*, 2016, *142*, 1–10.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare, "New trade models, same old gains?," *American Economic Review*, 2012, *102* (1), 94–130.

Arnott, Richard J and Joseph E Stiglitz, "Aggregate land rents, expenditure on public goods, and optimal city size," *The Quarterly Journal of Economics*, 1979, *93* (4), 471–500.

Bajari, Patrick and Matthew E Kahn, "Estimating Housing Demand With an Application to Explaining Racial Segregation in Cities," *Journal of Business & Economic Statistics*, 2005, *23* (1), 20–33.

Banzhaf, H. Spencer and V. Kerry Smith, "Meta-analysis in model implementation: choice sets and the valuation of air quality improvements," *Journal of Applied Econometrics*, September 2007, *22* (6), 1013–1031.

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.

Baum-Snow, Nathaniel 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.

Becker, Gary S, "A Theory of the Allocation of Time," *The economic journal*, 1965, pp. 493–517.

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, Amit Gandhi, and Philip Haile, "Connected Substitutes and Invertibility of Demand," *Econometrica*, September 2013, *81* (5), 2087–2111.

— **, James Levinsohn, and Ariel Pakes**, "Automobile prices in market equilibrium," *Econometrica*, 1995, pp. 841–890.

Boyle, Kevin J and John C Bergstrom, "Benefit transfer studies: myths, pragmatism, and idealism," *Water Resources Research*, 1992, *28* (3), 657–663.

Brownstone, David and Kenneth A. Small, "Valuing time and reliability: assessing the evidence from road pricing demonstrations," *Transportation Research Part A: Policy and Practice*, 2005, *39* (4), 279 – 293.

Brueckner, Jan K, "The structure of urban equilibria: A unified treatment of the Muth-Mills model," *Handbook of regional and urban economics*, 1987, *2*, 821–845.

— , "Urban growth boundaries: An effective second-best remedy for unpriced traffic congestion?," *Journal of Housing Economics*, 2007, *16* (3-4), 263–273.

— **, Jacques-Francois Thisse, and Yves Zenou**, "Why is central Paris rich and downtown Detroit poor?: An amenity-based theory," *European economic review*, 1999, *43* (1), 91–107.

Buchholz, Nicholas, Laura Doval, Jakub Kastl, Filip Matejka, and Tobias Salz, "The Value of Time: Evidence From AuctionedCab Rides," 2020. Working Paper.

Calder-Wang, Sophie, "The Distributional Impact of the Sharing Economy on the Housing Market," 2020. Working Paper.

Capps, Cory, David Dranove, and Mark Satterthwaite, "Competition and market power in option demand markets," *RAND Journal of Economics*, 2003, pp. 737–763.

Carrillo, Paul E, Arun S Malik, and Yiseon Yoo, "Driving restrictions that work? Quito's Pico y Placa Program," *Canadian Journal of Economics*, 2016, *49* (4), 1536–1568.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez, "Where is the land of opportunity? The geography of intergenerational mobility in the United States," *The Quarterly Journal of Economics*, 2014, *129* (4), 1553–1623.

Davis, Lucas W, "The effect of driving restrictions on air quality in Mexico City," *Journal of Political Economy*, 2008, *116* (1), 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, "The fundamental law of road congestion: Evidence from US cities," *American Economic Review*, 2011, *101* (6), 2616–52.

Epple, Dennis and Holger Sieg, "Estimating Equilibrium Models of Local Jurisdictions," *Journal of Political Economy*, 1999, *107* (4), 645–681.

_ **and Maria Marta Ferreyra**, "School finance reform: Assessing general equilibrium effects," *Journal of Public Economics*, 2008, *92* (5), 1326 – 1351.

_ **, 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 Cecile Gaubert, "Optimal spatial policies, geography, and sorting," *The Quarterly Journal of Economics*, 2020, *135* (2), 959–1036.

Ferreyra, Maria Marta, "Estimating the effects of private school vouchers in multidistrict economies," *American Economic Review*, 2007, *97* (3), 789–817.

Gibbons, Stephen and Stephen Machin, "Valuing rail access using transport innovations," *Journal of urban Economics*, 2005, *57* (1), 148–169.

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.

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.

Guevara, C. Angelo and Moshe E. Ben-Akiva, "Sampling of alternatives in Logit Mixture models," *Transportation Research Part B: Methodological*, 2013, *58*, 185 – 198.

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

Hall, Jonathan D., "Pareto improvements from Lexus Lanes: The effects of pricing a portion of the lanes on congested highways," *Journal of Public Economics*, 2018, *158*, 113 – 125.

Harari, Mariaflavia, "Cities in bad shape: Urban geometry in India," *American Economic Review*, 2020, *110* (8), 2377–2421.

Heblich, Stephan, Stephen J Redding, and Daniel M Sturm, "The Making of the Modern Metropolis: Evidence from London," *The Quarterly Journal of Economics*, 05 2020, *135* (4), 2059–2133.

Hwang, Yujung, "An Estimable General-Equilibrium Structural Model of Immigrants' Neighborhood Sorting and Social Integration," 2019. Working Paper.

Kahneman, Daniel and Alan B. Krueger, "Developments in the Measurement of Subjective Well-Being," *Journal of Economic Perspectives*, March 2006, *20* (1), 3–24.

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.

Knittel, Christopher R and Ryan Sandler, "The welfare impact of indirect pigouvian taxation: Evidence from transportation," Technical Report, National Bureau of Economic Research 2013.

Kreindler, Gabriel E., "The Welfare Effect of Road Congestion Pricing: Experimental Evidence and Equilibrium Implications," 2018. Working Paper.

Kuminoff, Nicolai V., "Decomposing the structural identification of non-market values," *Journal of Environmental Economics and Management*, March 2009, *57* (2), 123–139.

Kuminoff, Nicolai V, "Partial identification of preferences in a dual-market sorting equilibrium," 2012. Working Paper.

Kuminoff, Nicolai V., V. Kerry Smith, and Christopher Timmins, "The New Economics of Equilibrium Sorting and Policy Evaluation Using Housing Markets," *Journal of Economic Literature*, December 2013, *51* (4), 1007–1062.

Lancaster, Kelvin, *Consumer Demand: A New Approach*, Columbia University Press, 1971.

Langer, Ashley and Clifford Winston, "Toward a Comprehensive Assessment of Road Pricing Accounting for Land Use," *Brookings-Wharton Papers on Urban Affairs*, 2008, pp. 127–175.

LeRoy, Stephen F and Jon Sonstelie, "Paradise lost and regained: Transportation innovation, income, and residential location," *Journal of Urban Economics*, 1983, *13* (1), 67–89.

Li, Shanjun, "Better lucky than rich? Welfare analysis of automobile license allocations in Beijing and Shanghai," *Review of Economic Studies*, 2018, *85* (4), 2389–2428.

_ , Joshua Linn, and Erich Muehlegger, "Gasoline taxes and consumer behavior," *American Economic Journal: Economic Policy*, 2014, *6* (4), 302–42.

_ , Yanyan Liu, Avralt-Od Purevjav, and Lin Yang, "Does subway expansion improve air quality?," *Journal of Environmental Economics and Management*, 2019, *96*, 213 – 235.

Liu, Youming, Shanjun Li, and Caixia Shen, "The Dynamic Efficiency in Resource Allocation: Evidence from Vehicle License Lotteries in Beijing," Working Paper 26904, National Bureau of Economic Research March 2020.

McFadden, Daniel, "Modeling the choice of residential location," *Transportation Research Record*, 1978, (673).

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.

_ , Margaret Walls, and Winston Harrington, "Automobile externalities and policies," *Journal of economic literature*, 2007, *45* (2), 373–399.

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.

Redding, Stephen J. and Esteban Rossi-Hansberg, "Quantitative Spatial Economics," *Annual Review of Economics*, 2017, *9* (1), 21–58.

Roback, Jennifer, "Wages, rents, and the quality of life," *Journal of political Economy*, 1982, *90* (6), 1257–1278.

Rosen, Sherwin, "Hedonic prices and implicit markets: product differentiation in pure competition," *Journal* *of political economy*, 1974, *82* (1), 34–55.

Rosenberger, Randall S and John B Loomis, "Benefit transfer," in "A primer on nonmarket valuation," Springer, 2003, pp. 445–482.

Saiz, Albert, "The geographic determinants of housing supply," *The Quarterly Journal of Economics*, 2010, *125* (3), 1253–1296.

Severen, Christopher, "Commuting, Labor, and Housing Market Effects of Mass Transportation: Welfare and Identification," 2019. Working Paper.

Sieg, Holger, V Kerry Smith, H Spencer Banzhaf, and Randy Walsh, "Estimating the general equilibrium benefits of large changes in spatially delineated public goods," *International Economic Review*, 2004, *45* (4), 1047–1077.

Small, Kenneth A., "Valuation of travel time," *Economics of Transportation*, December 2012, *1* (1-2), 2–14.

Small, Kenneth A, Clifford Winston, and Jia Yan, "Uncovering the distribution of motorists' preferences for travel time and reliability," *Econometrica*, 2005, *73* (4), 1367–1382.

_ , Erik T Verhoef, and Robin Lindsey, *The economics of urban transportation*, Routledge, 2007.

Stiglitz, Joseph E, "The theory of local public goods," in "The economics of public services," Springer, 1977, pp. 274–333.

Timmins, Christopher and Jennifer Murdock, "A revealed preference approach to the measurement of congestion in travel cost models," *Journal of Environmental Economics and management*, 2007, *53* (2), 230–249.

Tra, Constant I., "A discrete choice equilibrium approach to valuing large environmental changes," *Journal of Public Economics*, 2010, *94* (1), 183 – 196.

Train, Kenneth E. and Clifford Winston, "Vehicle Choice Behavior and the Declining Market Share of U.S. Automakers," *International Economic Review*, 2007, *48*, 1469–1496.

Tseng, Yin-Yen, Erik Verhoef, Gerard de Jong, Marco Kouwenhoven, and Toon van der Hoorn, "A pilot study into the perception of unreliability of travel times using in-depth interviews," *Journal of Choice Modelling*, 2009, *2* (1), 8 – 28.

Tsivanidis, Nick, "Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogota's Trans-Milenio," 2019. Working Paper.

USDOT, "evised 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, "Statement on the pricing of urban street use," *Hearing of US Joint Committee on Metropolitan Washington Problems*, 1959, *11*, 42–65.

—, "Pricing in Urban and Suburban Transport," *American Economic Review*, 1963, *53* (2), 452–465.

Vickrey, William S, "Congestion theory and transport investment," *The American Economic Review*, 1969, *59* (2), 251–260.

Walsh, Randy et al., "Endogenous open space amenities in a locational equilibrium," *Journal of urban Economics*, 2007, *61* (2), 319–344.

Wang, Songtao, Su Han Chan, and Bohua Xu, "The Estimation and Determinants of the Price Elasticity of Housing Supply: Evidence from China," *Journal of Real Estate Research*, 2012, *33* (3), 311–344.

Wang, Wen, "Environmental Gentrification," *Working Paper*, 2020.

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

Waxman, Andrew, "The Long Road to Work: The Divergent Effects of Transportation Policies by Worker Skill in a Locational Sorting Model," *working paper*, 2017.

Winston, Clifford, "Efficient Transportation Infrastructure Policy," *The Journal of Economic Perspectives*, 1991, *5* (1), 113–127.

Wolff, Hendrik, "Value of time: Speeding behavior and gasoline prices," *Journal of Environmental Economics and Management*, 2014, *67* (1), 71–88.

Xiao, Junji, Xiaolan Zhou, and Wei-Min Hu, "Welfare Analysis of the Vehicle Quota System in China," *International Economic Review*, 2017, *58* (2), 617–650.

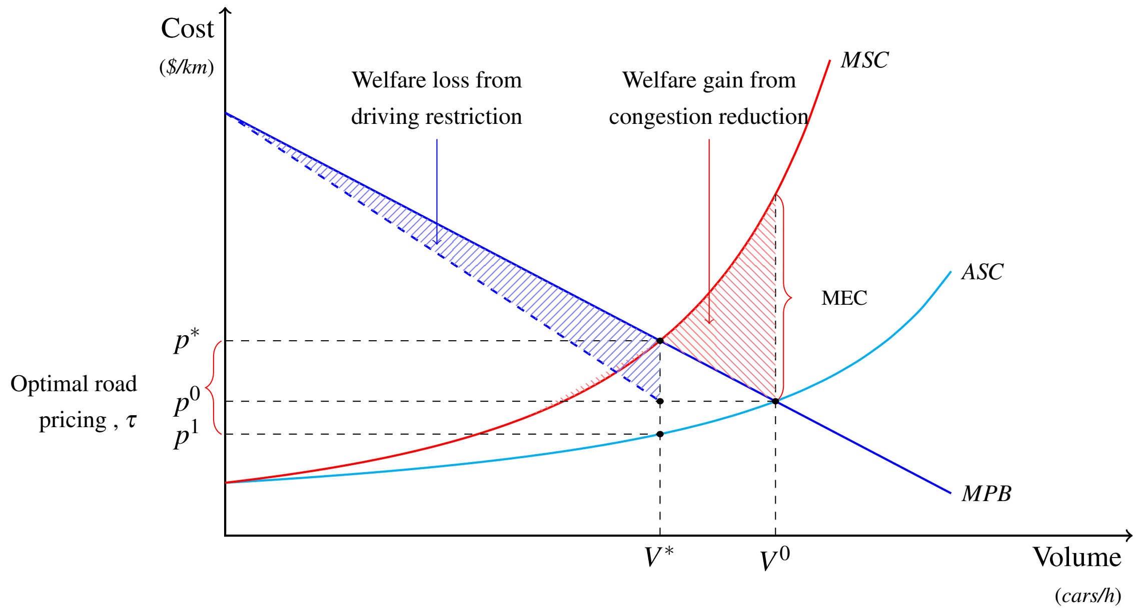
Yang, Jun, Avralt-Od Purevjav, and Shanjun Li, "The Marginal Cost of Traffic Congestion and Road Pricing: Evidence from a Natural Experiment in Beijing," *American Economic Journal: Economic Policy*, 2019. forthcoming.

—, Shuai Chen, Ping Qin, Fangwen Lu, and Antung A Liu, "The effect of subway expansions on vehicle congestion: Evidence from Beijing," *Journal of Environmental Economics and Management*, 2018, *88*, 114–133.

Zhang, Wei, C-Y Cynthia Lin Lawell, and Victoria I Umanskaya, "The effects of license plate-based driving restrictions on air quality: Theory and empirical evidence," *Journal of Environmental Economics and Management*, 2017, *82*, 181–220.

Zheng, Siqi and Matthew E Kahn, “China’s bullet trains facilitate market integration and mitigate the cost of megacity growth,” *Proceedings of the National Academy of Sciences*, 2013, *110* (14), 1248–1253.

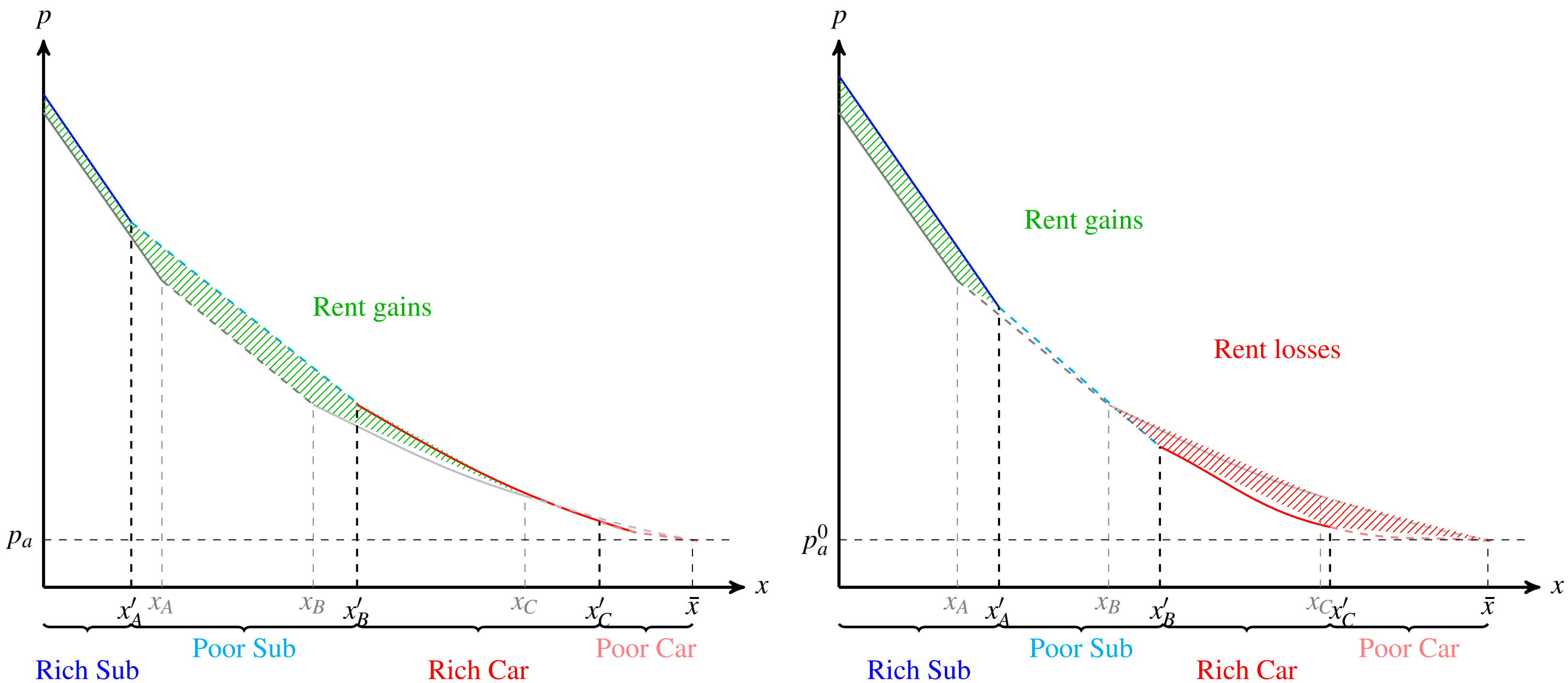
Figure 1: Traffic under Congestion Pricing and Driving Restriction



Note: The figure illustrates the welfare impacts of optimal congestion pricing and driving restriction. The x-axis denotes traffic volume (or throughput measured in the number of cars per hour passing the point). The marginal private benefit *MPB* curve represents the demand curve for driving (willingness to pay for driving). The average social cost *ASC* curve reflects the private cost of driving, which is experienced based upon the average time cost across the vehicles on the road. The difference between *ASC* and the marginal social cost, *MSC*, is the congestion externality (or the marginal external cost of congestion, *MEC*). In the absence of any intervention, equilibrium occurs at V^0 compared to the social optimal level of traffic volume is V^*. The shaded area on the right (red area) shows the deadweight loss due to excess congestion. A Pigouvian tax, τ, can be imposed to achieve the optimal level V^*. Alternatively, a driving restriction can be adopted to achieve the same level of congestion reduction but it would incur a welfare loss denoted by the shaded area on the left (blue area) assuming that the reduction of trips is random among all privately beneficially trips. Therefore, the welfare impact of the driving restriction is ambiguous *a priori*.

Figure 2: Welfare Effects of Transportation Policies with Sorting

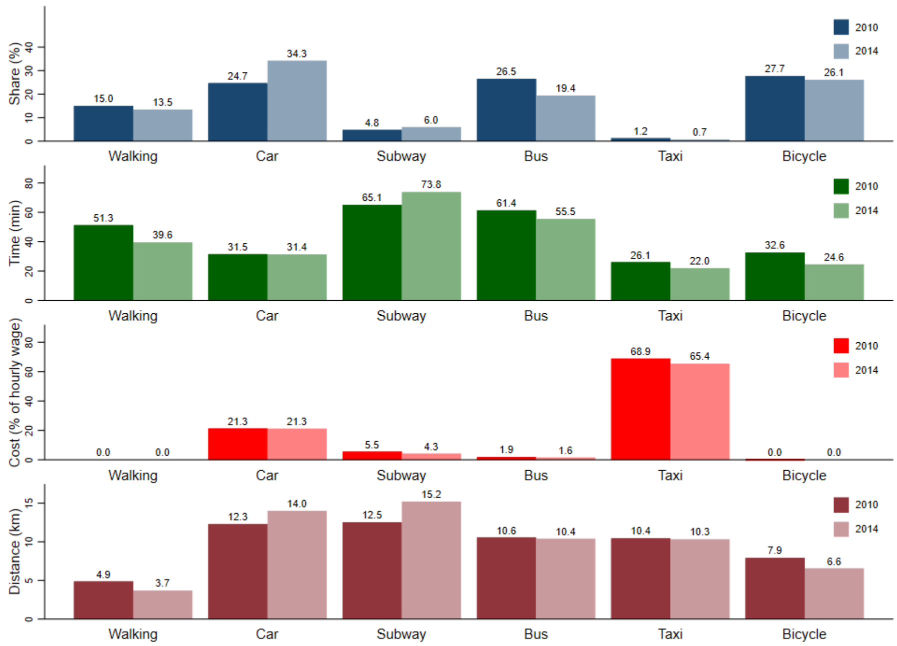
(a) Congestion Pricing (b) Driving Restriction



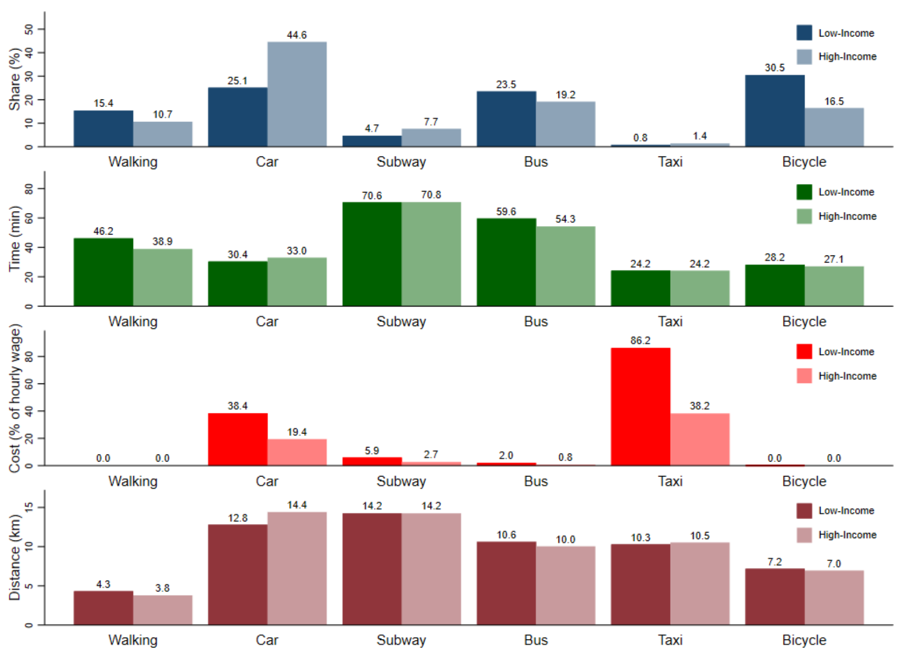
Note: This figure illustrates the capitalization of commuting cost changes into the housing market with sorting. A full exposition of the individual bid-rent curves in the diagram is provided in the Appendix as well as the underlying assumptions and equilibrium conditions. The left panel shows the effect of a distance-based congestion charge. Colored lines refer to the equilibrium bid rent functions lying along the envelope corresponding to rich subway, poor subway, rich driving, poor driving moving from the CBD to the urban boundary. Grey lines correspond to the no-policy baseline bid-rent envelope. The boundaries marked with a prime indicate the change in spatial structure induced by sorting, namely that fewer rich and more poor take the subway. The area between the policy and no-policy bid-rent envelopes reflects the capitalization effect of transportation policies, which corresponds to net welfare improvement. Congestion pricing induces a welfare increase for subway commuters, based largely on the effect of recycled revenues. It also induces a improvement in welfare for rich drivers reflecting the impact of lower congestion net of the cost of the congestion fee. It also induces an almost negligible decrease in welfare for poor drivers reflecting their smaller values of time. In contrast, the driving restriction in the right panel shows how the rich are induced to increase subway commuting and the poor do, albeit by a trivial amount. The driving restriction induces a gain for subway commuters, but a larger loss for those driving because of the time costs associated with using subway on long commuters during restricted days.

Figure 3: Travel Patterns for Commuting Trips from Beijing Household Travel Survey

(a) 2010 vs. 2014



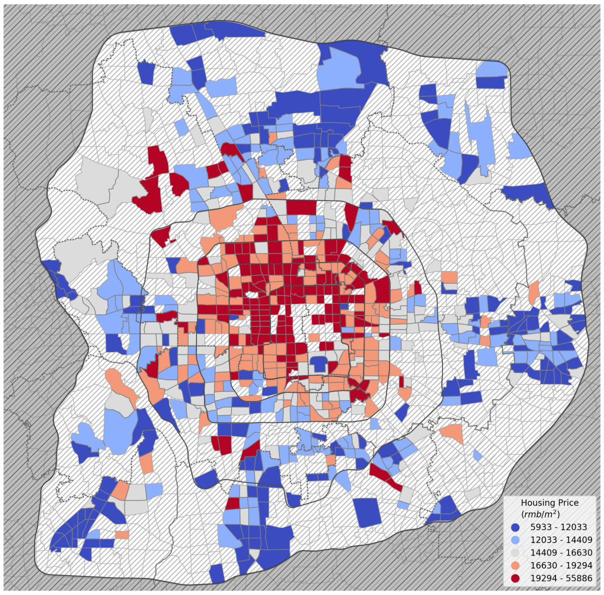
(b) High income vs. Low Income



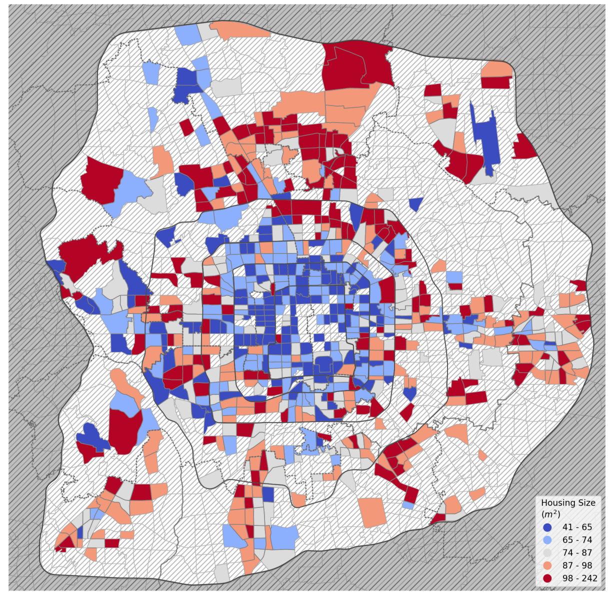
Note: This figure plots trip share, time, and costs by different modes for work commuting trips in the Beijing Household Travel Survey of 2010 and 2014. There are six main trip modes: walk, bike, bus, subway, car, and taxi. For bus and subway trips, they could include segments with other modes but we characterize them as the bus and subway trips. Trips using both bus and subway are rare (less than 3% in the data and we drop them in the analysis.) The mode shares are based on chosen modes in the data. Travel time, cost (defined as % of hourly wage), and distance are constructed as shown in Appendix **??**. The numbers in the figures are for the chosen modes. High-income households are defined as households whose income level is greater than the median in the survey year.

Figure 4: Housing and Household Attributes from Housing Mortgage Data

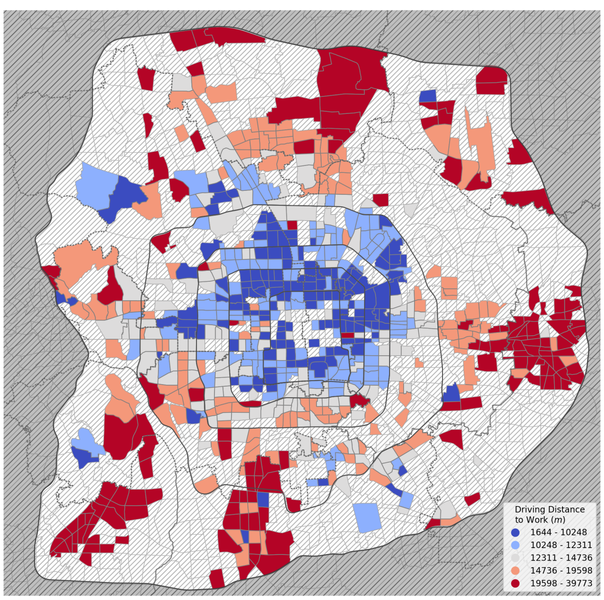
(a) Housing Price (¥/m^2)



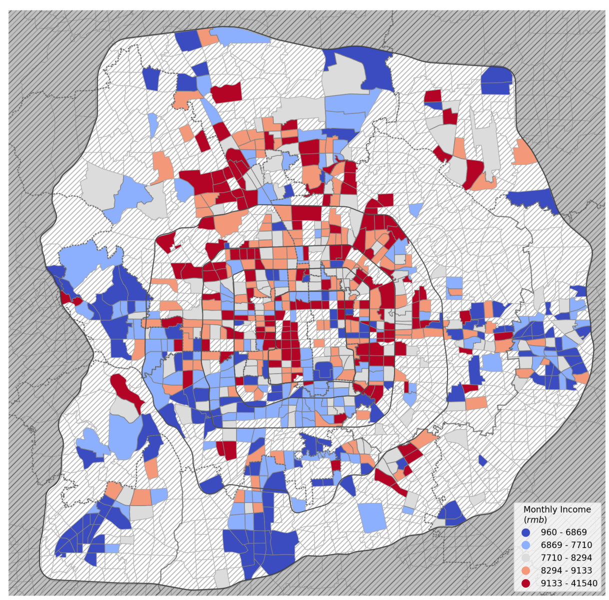
(b) Housing Size (m^2)



(c) Distance to Work (m)

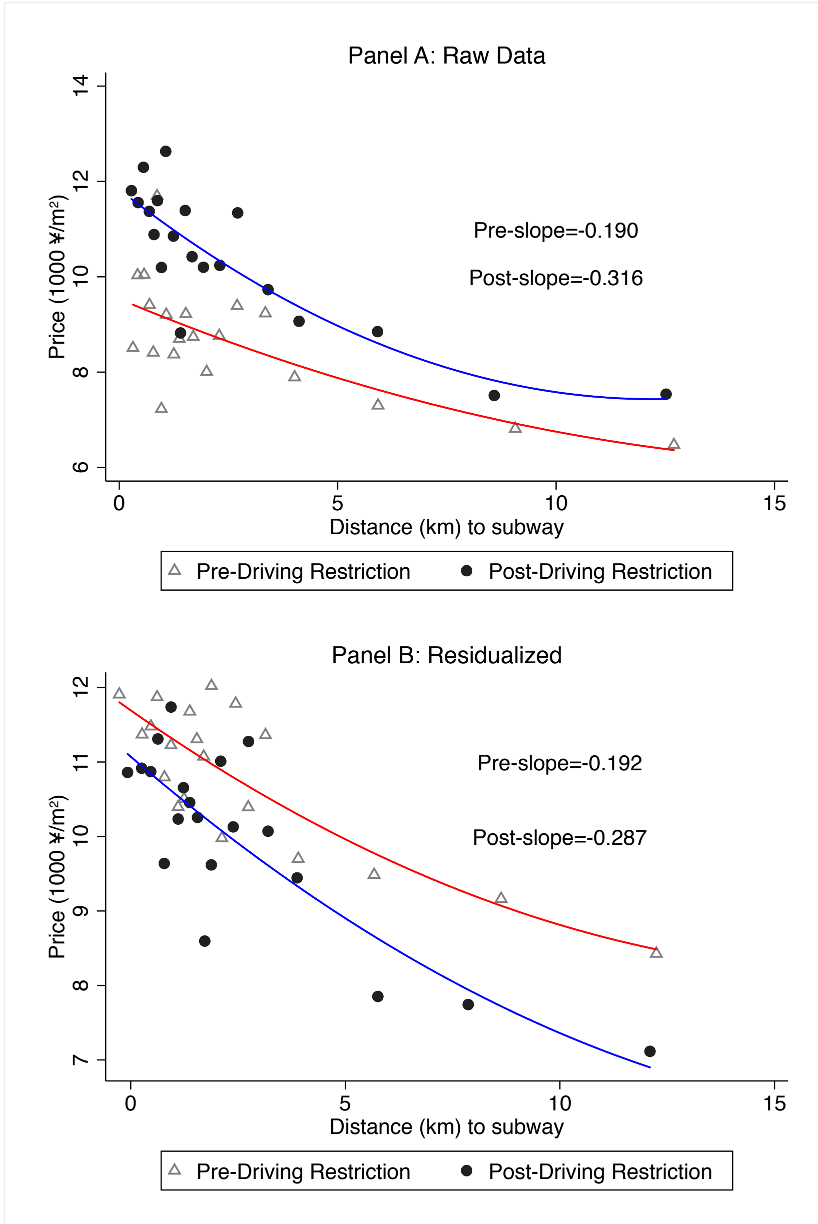


(d) Monthly Household Income (¥)



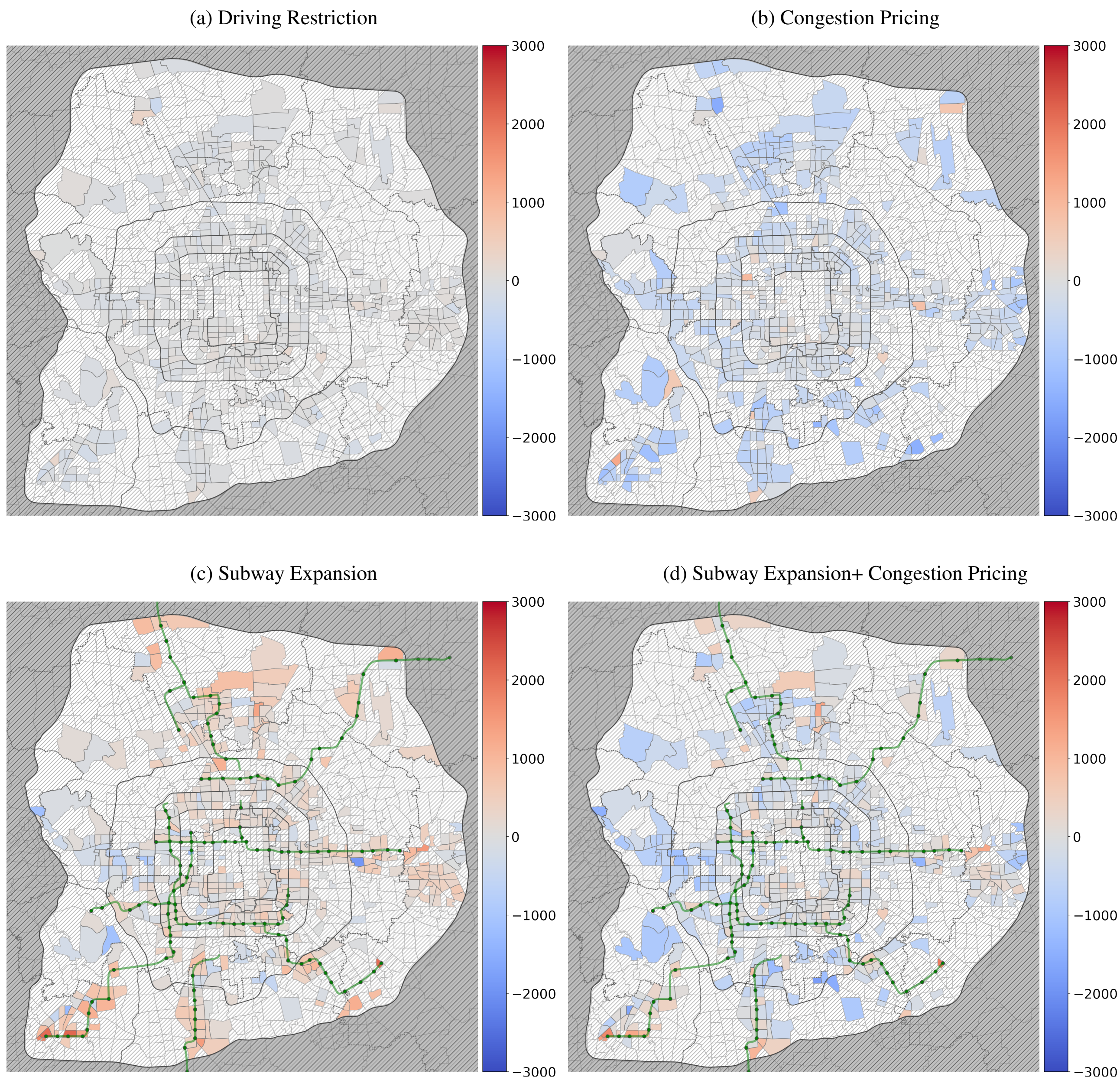
Note: This figure plots the averages of key housing and household attributes by Traffic Analysis Zone (TAZ) based on mortgage data from 2006 to 2014. The values are the averages across homes in the TAZ from the mortgage data during the data period. Distance to work is the driving distance to work for all borrowers in the data (including primary and secondary borrowers when both are present). Monthly household income is based on the income of the households at the time of purchase in the given TAZ. Values are classified into five quintiles: the red color corresponds to larger values while the blue color for low values. The white color represents no observations in the TAZ.

Figure 5: Reduced Form Evidence: Housing Price Gradient before and after Driving Restriction



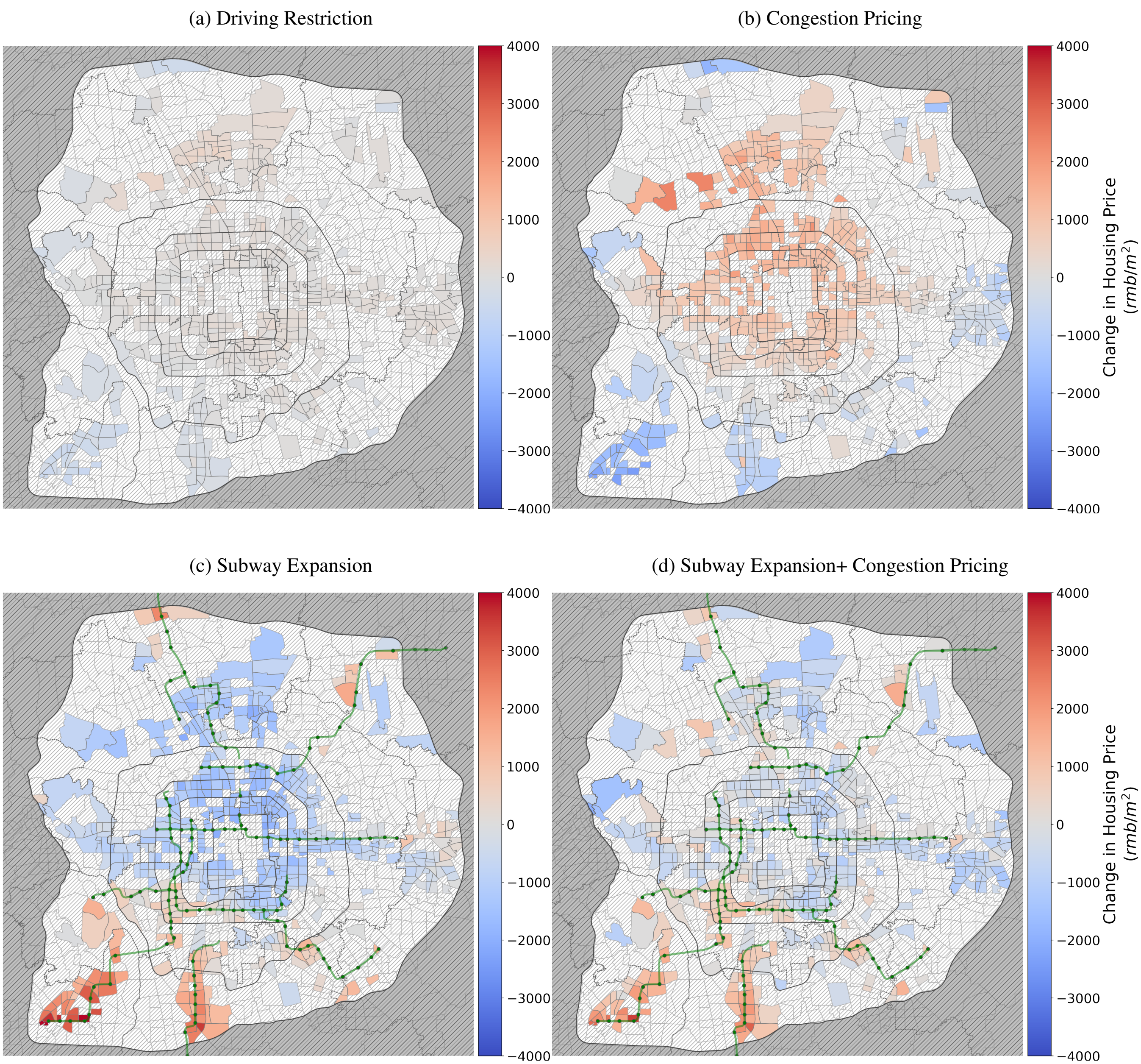
Notes: These binned scatterplots show housing price per square meter against distance to subway before and after the driving restriction goes into effect in Beijing. The sample spans 24 months before and after the policy starting point (July 2008). The top panel is the binned scatter plot based on the raw data of price per m^2 and the distance to the nearest subway station. The bottom panel controls for neighborhood fixed effects, and year by month fixed effects. The slopes denoted on the figure are based on quadratic fits.

Figure 6: Changes in Commuting Distance from Simulated Policies (in meters)



Note: This figure illustrates simulated changes in commuting distance under different counterfactual policy scenarios (relative to the baseline scenario of no policy). The results are based on the simulations in Table 6. Each cell represents a TAZ. A warm color corresponds to an increase in distance while a cold color represents a decrease. Cells in white have no observations in the simulation sample period in the TAZ. Green lines represent new subway lines from year 2008 to 2014.

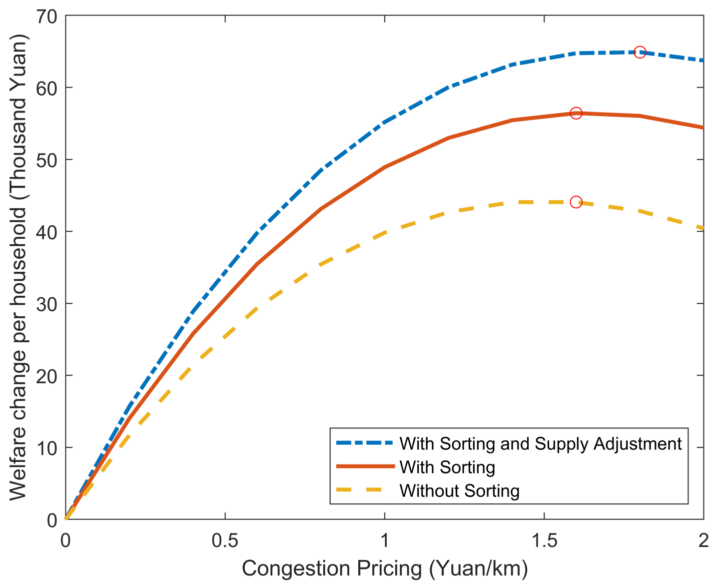
Figure 7: Changes in Housing Prices from Simulated Policies (¥$/m^2$)



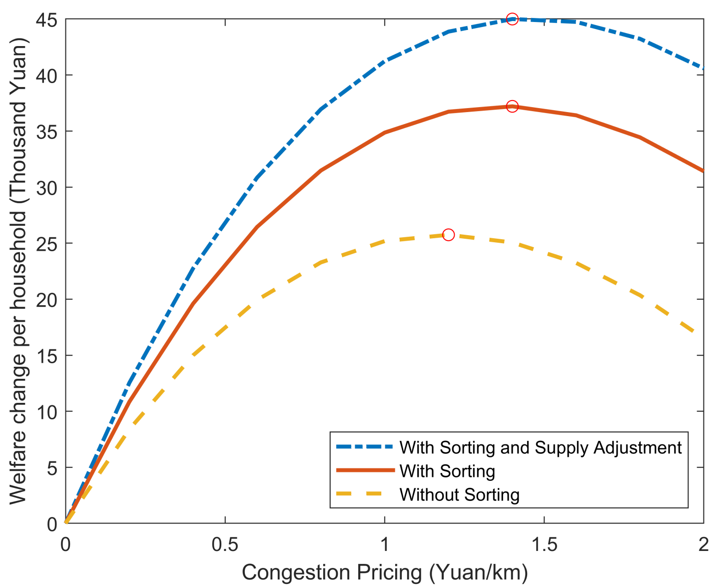
Note: This figure illustrates simulated changes in home prices under different counterfactual policy scenarios (relative to the baseline scenario of no policy). The results are based on the simulations shown in Table 6. Each cell represents a TAZ. Warmer colors corresponds to an increase in price while a cold color represents a decrease. Cells in white have no observations in the simulation sample period in the TAZ. Green lines represent new subway lines from year 2008 to 2014.

Figure 8: Optimal Congestion Pricing under 2008/2014 Subway Network

(a) 2008 Subway Network



(b) 2014 Subway Network



Note: The plot shows the welfare change with respect to congestion pricing under 2008/2014 subway network, without sorting (PE scenario, yellow dotted line), with sorting (GE scenario, orange solid line), or with sorting and supply adjustment (blue dashed line). Without sorting, the optimal congestion pricing is ¥1.4/km for 2008 subway system and ¥1.2/km for 2014. Sorting shifts the optimal toll to the right. The optimal congestion pricing is ¥1.6/km for 2008 subway system and ¥1.4/km for 2014 under sorting. The difference of "with sorting" and "without sorting" welfare shows the welfare gain from household sorting. The difference of "with sorting" and "without sorting and supply adjustment" welfare shows the welfare gain from supply adjustment.

Table 1: Summary Statistics of Household Travel Survey

	2010			2014		
	N	Mean	SD	N	Mean	SD
Respondent characteristics						
Income: <50k	14780	0.48	0.50	20573	0.18	0.38
Income: [50k, 100k)	14780	0.39	0.49	20573	0.44	0.50
Income: >=100k	14780	0.13	0.34	20573	0.38	0.49
Having a car (=1)	14780	0.44	0.50	20573	0.62	0.49
Female (=1)	14780	0.44	0.50	20573	0.43	0.50
Age (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	30334	0.87	1.06	42820	0.74	0.98
Travel cost	30334	2.47	5.55	42820	3.83	6.96
Distance<2km	30334	0.25	0.43	42820	0.24	0.43
Distance within 2-5km	30334	0.27	0.45	42820	0.26	0.44

Note: The table reports respondent and trip characteristics of all work commuting trips within the 6th ring road from 2010 and 2014 Beijing Household Travel Survey.Travel time and travel cost variables are those associated with the chosen modes, and are constructed as shown in Appendix **??**. Distance<2km and Distance within 2-5km denotes straight-line distance and captures short to medium-distance commuting trips. The travel mode shares are shown in Figure 3.

Table 2: Summary Statistics of Housing Data

	Mean	SD	Min	Max
Housing attributes				
Transaction year	2011	1.89	2006	2014
Price/m^2 (¥'000s)	19.83	9.56	5.00	68.18
Unit size (m^2)	92.68	40.13	16.71	400.04
Household annual income (¥'000s)	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 to land area ratio	2.56	1.12	0.14	16.00
No. 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 mortage dataset over 2006-2014. The number of housing transactions is 79,884, all of which are within the 6th ring road. The dataset is weighted to match the statistics of real-estate listings. Housing complex is defined as a group of building in the same development. Distance to key school is the distance of home to the nearest elementary school with a key school designation. Distance to park is the distance to the nearest park or green space. Home-work travel variables are constructed following the same method as outlined in the household travel survey. Home to subway distance is the distance from home 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 of Travel Mode Choices

	Logit			Random Coefficient		
	(1)	(2)	(3)	(4)	(5)	(6)
Travel Time (γ_1)	-1.194	-0.270	-0.191			
	(0.082)	(0.006)	(0.006)			
Travel Cost/Hourly Wage (γ_2)	-1.578	-0.788	-0.565	-1.411	-1.424	-2.531
	(0.324)	(0.028)	(0.034)	(0.041)	(0.052)	(0.065)
Random coefficients on travel time (μ_γ)						
Travel Time				-0.955	-0.885	-0.931
				(0.008)	(0.008)	(0.012)
Random coefficients on mode-specific constants (σ_m)						
Driving					3.394	3.391
					(0.049)	(0.054)
Subway						4.470
						(0.142)
Bus						3.851
						(0.056)
Bike						3.887
						(0.054)
Taxi						4.203
						(0.353)
Mode*Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Mode*Trip related FE		Yes	Yes	Yes	Yes	Yes
Mode*Demographic FE			Yes	Yes	Yes	Yes
Log-likelihood	-116287	-109929	-91119	-87353	-85099	-77706
Implied mean VOT	0.757	0.342	0.339	1.760	1.615	0.956
Implied median VOT	0.757	0.342	0.339	1.557	1.429	0.846

Note: The number of observations are 73,154. All six specifications include a rich set of fixed effects interacting with mode-specific constants (travel model dummies). Trip related FE includes trip distance bins and the origin and destination dummies (e.g., if the origin is within 2nd ring row). Demographics FE includes respondent's age, gender, education, and car ownership. The first three specifications are multinomial logit while the last three add random coefficients to the model. 200 randomized Halton draws are used to estimate the random coefficients in the last three specifications. The distribution of the preference on time in the last three specifications is specified as a chi-square distribution (winsorized at 5th and 95th percentile) with degrees of freedom equals three: $\mu_\gamma \chi^2(3)$ so as to capture the long tail of VOT distribution. The estimates of μ_γ are provided in the table. The random coefficients on travel mode dummies (driving, subway, bus, bike, and taxi) are assumed to have normal distribution (walking is taken as the baseline group). The estimates of σ_m of those normal distributions for each travel mode are provided in the table. The last row provides the implied median value of time for each specification. Standard errors clustered at the respondent level are below estimates: $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$.

Table 4: Estimation Results of Housing Choices - Nonlinear Parameters

	No EV		With EV		EV and Random Coef.	
	Para	SE	Para	SE	Para	SE
Demographic Interactions						
Price (in 1 million RMB)*ln(income)	0.965	0.007	1.005	0.008	1.030	0.016
Age in 30-45*ln(distance to key school)	-0.329	0.004	-0.391	0.005	-0.420	0.010
Age > 45*ln(distance to key school)	-0.074	0.009	-0.111	0.010	-0.123	0.021
Age in 30-45*ln(home size)	1.343	0.014	1.443	0.015	1.486	0.029
Age > 45*ln(home size)	2.394	0.028	2.665	0.031	2.746	0.061
EV_{Male}			0.709	0.026	0.755	0.006
EV_{Female}			0.833	0.026	0.893	0.006
Random Coefficients						
$\sigma(EV_{Male})$					0.379	0.013
$\sigma(EV_{Female})$					0.482	0.012
Log-likelihood	-206829		-170057		-168808	

Note: The estimation uses weighted mortgage plan data from Year 2008-2014. The number of observations is 79,884. The results are from MLE. The first specification does not include EV (ease of commuting); the second specification does; the third specification further controls random coefficients on EV terms. EV is constructed by taking observed household demographics into travel model estimates (Column 6 of Table 3). The big decrease in log-likelihood from the first to the second specification indicates strong explanatory power of EV term.

Table 5: Estimation Results of Housing Choices - Linear Parameters

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

Note: The number of observations is 79,884. The dependent variable is the mean utilities recovered from the first stage. The first two columns are from OLS and the last four are from IVs. The floor-area ratio of the complex, a measure of complex density, is the size of the total floor area over the size of the parcel that the complex is located on. Distance to key school is the distance to the nearest key elementary school. Column (3) and (4) use IV1 as price instruments, i.e. the average attributes of homes (building size, age, log distance to park, and log distance to key school) that are within 3km outside the same complex sold in a two-month time window from a given home. Column (4) and (6) additionally use IV2, i.e. the interaction between the distance related instruments defined in Column (3) and the winning odds of the vehicle licence lottery as instruments. The winning odds decreased from 9.4% in Jan. 2011 to 0.7% by the end of 2014. Column (5) uses number of homes transacted in the three-month time window in the real estate listings dataset. Standard errors clustered at the neighborhood-year level are in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Simulation Results: with Sorting

	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)	
Household Income Relative to Median	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low
Panel A: Travel outcomes												
Drive	41.65	21.44	-6.07	-2.59	-2.90	-4.54	-2.14	-1.66	-7.44	-3.84	-4.59	-5.60
Subway	9.02	10.77	1.09	0.66	0.75	0.89	4.62	6.06	5.59	6.42	5.12	6.70
Bus	22.44	30.47	1.51	0.26	0.42	0.96	-1.54	-2.53	-0.06	-1.95	-0.92	-1.28
Bike	15.96	24.01	1.54	0.63	0.63	1.48	-0.80	-1.64	0.52	-1.08	-0.27	-0.38
Taxi	2.20	1.32	0.70	0.34	0.51	0.46	-0.16	-0.11	0.42	0.17	0.29	0.27
Walk	8.74	11.99	1.23	0.70	0.59	0.74	0.02	-0.13	0.98	0.28	0.37	0.29
Speed	21.49		3.13		3.13		1.49		4.44		4.61	
Panel B: Housing market outcomes												
Male member's distance to work (km)	19.45	18.88	0.01	0.02	-0.15	-0.07	0.33	0.15	0.37	0.15	0.15	0.08
Female member's distance to work (km)	17.54	11.95	0.01	0.01	-0.12	-0.05	0.39	0.21	0.43	0.21	0.24	0.17
Distance to subway (km)	5.33	4.30	-0.03	0.03	-0.02	0.02	-4.14	-3.44	-4.14	-3.44	-4.14	-3.44
Panel C: Welfare analysis per household (thousand ¥)												
Consumer surplus (+)			-165.3	-19.6	-76.8	-61.3	220.3	100.0	47.8	77.6	131.9	40.2
Toll revenue (+)					115.9	115.9					107.8	107.8
Subway cost (–)							103.0	103.0	103.0	103.0	103.0	103.0
Net welfare			-165.3	-19.6	39.1	54.5	117.3	-3.0	-55.2	-25.4	136.7	44.9

Note:Simulated results based on estimated model parameters in Column (6) of Table 3 and Column (6) of Table 5. The simulation shows counterfactual results for 2014 sample households and homes. The detailed simulation procedure can be found in F. This table shows results with sorting and supply adjustment. In particular, we allow housing supply to adjust with a price elasticity of two (implying that a ¥1,000 price increase would induces a 0.12% increase in housing supply on average). Column (1) shows the baseline results while columns (2) to (6) shows the differences from column (1). Driving restriction prohibits driving in one of five work days. Congestion pricing is ¥0.92 per km to generate same reduction as driving restriction. High-income household are those with income above the median household income. Subway cost per household includes the construction cost and the 30-year operating cost equally shared among 7.2 million households. We apportion 100% of it to work commute in our welfare analysis. Toll revenue is the revenue per household from congestion pricing during a 30-year period (to keep a balanced government budget, the toll revenue is recycled uniformly to each household). Net welfare is consumer welfare per household after revenue recycling or tax-funded subway construction and operation.

Table 7: Simulation Results: with Sorting and Housing-Supply Response

	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)	
Household Income Relative to Median	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low
Panel A: Travel outcomes												
Drive	41.65	21.44	-6.08	-2.60	-2.90	-4.52	-2.32	-1.80	-7.55	-3.93	-4.71	-5.72
Subway	9.02	10.77	1.10	0.68	0.78	0.94	5.00	6.57	5.98	6.95	5.42	7.13
Bus	22.44	30.47	1.52	0.27	0.38	0.87	-1.52	-2.48	-0.06	-1.91	-0.92	-1.25
Bike	15.96	24.01	1.54	0.62	0.62	1.45	-0.86	-1.78	0.42	-1.26	-0.33	-0.51
Taxi	2.20	1.32	0.70	0.34	0.51	0.47	-0.20	-0.15	0.38	0.13	0.27	0.24
Walk	8.74	11.99	1.23	0.69	0.62	0.80	-0.10	-0.36	0.83	0.02	0.28	0.10
Speed	21.49		3.13		3.26		1.13		4.04		4.39	
Panel B: Housing market outcomes												
Male member's distance to work (km)	19.45	18.88	0.02	0.03	-0.26	-0.19	0.76	0.61	0.84	0.63	0.47	0.44
Female member's distance to work (km)	17.54	11.95	0.02	0.02	-0.21	-0.14	0.76	0.60	0.82	0.62	0.52	0.48
Distance to subway (km)	5.33	4.30	-0.04	0.01	-0.09	-0.06	-4.13	-3.45	-4.14	-3.44	-4.14	-3.45
Panel C: Welfare analysis per household (thousand ¥)												
Consumer surplus (+)			-165.4	-18.7	-65.7	-59.1	187.4	105.3	18.3	84.8	108.4	45.9
Toll revenue (+)					115.2	115.2					108.9	108.9
Subway cost (–)							103.0	103.0	103.0	103.0	103.0	103.0
Net welfare			-165.4	-18.7	49.5	56.0	84.4	2.3	-84.7	-18.2	114.3	51.9

Note: Simulated results based on estimated model parameters using 2014 housing data. We allow housing supply to adjust with a price elasticity of two (implying that a ¥1,000 price increase would induces a 0.12% increase in housing supply on average). Column (1) shows the baseline results while columns (2) to (6) show the differences from column (1). The driving restriction prohibits driving in one of five workdays. Congestion pricing is ¥0.92 per km to generate the same reduction as the driving restriction. High-income household are those with income above the median household income. Subway cost per household includes the construction cost and the 30-year operating cost equally shared among 7.2 million households. We apportion 100% of it to work commute in our welfare analysis. Toll revenue is the revenue per household from congestion pricing during a 30-year period (to keep a balanced government budget, the toll revenue is recycled uniformly to each household). Net welfare is consumer welfare per household after revenue recycling or tax-funded subway construction and operation.

Online Appendices

A Theoretical Model Details

A.1 Spatial Structure

This appendix provides the details for the monocentric city model used to provide illustrative comparative statics for housing market capitalization of transportation policies in Section 2. We consider a monocentric, linear city with a fixed population (N) of rich, N_R, and poor, N_P, residents. All residents work at the urban center (CBD) at location 0, where wage income for the rich is larger: $y_R > y_P$.[45] The rest of urban space is occupied by homes with lot size normalized to 1 and where land rents are remitted to absentee landlords. Housing consumption (in square meters) is provided by perfectly competitive developers facing constant returns to scale. Beyond the residential area is agricultural land, which returns rental value p_a. The model is a closed-city model with intracity, but not intercity migration. Both of these assumptions could be relaxed without affecting the key predictions of the model.[46] A key feature of the model then is that changes in commuting cost will not affect the overall size of the city as reflected by the location of the urban boundary, $\bar{x}$, since the population is fixed and land use per household is also fixed.

A.2 Household Utility Maximization and Housing Demand

Households consume two goods: a numeraire good, c, with unitary price, and housing, $q(x)$, which varies in quantity depending upon the distance from the CBD x. Households seek to maximize utility given income y_d, fixed (θ_m) and variable commuting cost ($w_{d,m}$), which vary by mode $m \in M$ and by household income via differences in the value of time:

$$u_d = \max_{c_{d,m}, q_{d,m}} \quad u\left(c_{d,m}, q_{d,m}\right) \quad \text{s.t.} \quad c_d + p(x) \cdot q_d = y_d - \theta_m - w_{d,m}(x), \quad d = R, P. \tag{A1}$$

The solution to (A1), $\{c^*_{d,m}, q^*_{d,m}\}_{d=R,P}$, determines the pattern of residential locations as well as household-level mode choices, which both helps to determine and, in turn, is affected by the pattern of endogenous congestion that enters the budget constraint via $w_{d,m}$ based upon where car commuters choose to locate. We demonstrate the nature of this endogenous congestion in the next subsection.

A.3 Travel Mode Choice

Two commuting modes exist in the city: personal vehicles with fixed cost θ_C and variable (per-kilometer) cost $w_{d,C}$, and subway with fixed cost θ_S and variable cost $w_{d,S}$. Variable costs include time and pecuniary costs and travel time is monetized by the value of time (VOT): $v_R > v_P$.[47] We begin by assuming that the

[45] Given the linear structure of the city, we assume roads take up no space and all land goes towards housing.

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

[47] We also assume that fixed costs are larger for car commuting but that variable costs (without congestion) are lower.

subway network covers the entire urban area and then relax this assumption when considering the role of public transportation infrastructure. We ignore the role of congestion in public transportation and focus solely on its effect on car travel.

Car commuting is subject to congestion. Congestion at a given location x depends on the flow of vehicles $n_C(x)$ of rich and poor from the urban boundary to that location:

$$n_C(x) = \int_x^{\bar{x}} \mathbf{1}\{m = C\}_R(x)ds + \int_x^{\bar{x}} \mathbf{1}\{m = C\}_P(x)ds. \tag{A2}$$

What this equation makes clear is that congestion at x depends on the total number of car commuters who live between x and the urban boundary. The flow of car commuters, $n_C(x)$, determines the congestion at point x measured in travel time per unit of travel distance and unit commute cost:

$$t_{d,C}(x) = \nu_d \mathscr{C}\left(n_C(x)\right), \quad d = R, P \tag{A3}$$

where ν_d is the value of time for income group, d, and $\mathscr{C}(\cdot)$ is a congestion function with positive first and second derivative.

Another integral is required to calculate total commuting costs for a commuter living at point x as it includes the level of $t_{d,C}(x)$ at everyone point between x and the CBD:

$$w_{d,C}(x) = \int_0^x t_{d,C}(s)ds. \tag{A4}$$

A.4 Market Clearing Conditions and Spatial Equilibrium

Given a mass of households $N = N_P + N_R$ residing and working in the city, a spatial equilibrium is determined by a bid rent function that is the envelope of individual willingness-to-pay for housing based on mode and housing type, keeping the utility for each income type fixed at $\bar{u}_d$, $d = R, P$

$$p^*(x) = \max_{d,m} \left\{ p\left(y_d - \theta_m - w_{d,m}(x), \bar{u}_d\right) \right\}. \tag{A5}$$

Equations (A1) - (A4) make clear the simultaneous determination of housing location and traffic congestion across the city. Solving for congestion at any location x in the city requires knowing the distribution of car users at all points in the city. To understand the shape of the bid rent functions, it is helpful to express the slope, which is the derivative of $\frac{w_{d,C}}{q_{d,m}(x)}$ with respect to x:

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

For subway commuters, who experience no congestion, the slope of the bid rent function does not change. In residential regions with car commuting, moving from right to left across the region means adding additional car commuters, further increasing per kilometer commuting time costs and steepening the bid rent function.

We now define the spatial behavior of housing prices that a household is willing to pay while maintaining a utility level $\bar{u}_d$ for $d = R, P$. The conditions for a spatial equilibrium are: 1) city population of rich and poor is equal to the sum of the distance within the city that they occupy (given fixed lot size for housing), 2) the equilibrium bid rent and the agricultural rent equate at the urban boundary, 3) budget constraints for each household type and travel mode hold, 4) the equilibrium traffic congestion experienced at each location x in the city is determined by the number of car commuters from x to $\bar{x}$ via $n_c(x)$, and 5) the boundary of residential patterns of rich and poor households using car or subway is determined by the intersections of respective bid rent functions, and the location of these bid rents along the envelope equilibrium bid rent function. The full analytical exposition of these conditions is:

The market clearing conditions for a spatial equilibrium in our model to exist are:

1. All N households are housed within the city, which given fixed unitary lot size per household means that the sum of the subway and car commuting areas yields the total population for rich and poor as: $N_R = x_A + x_C - x_B$ and $N_P = x_B - x_A + \bar{x} - x_C$.

2. Following from the previous condition, since there is fixed lot size per household, the total size of the city, $\bar{x}$ is equal to the total population N, and so the equilibrium bid-rent at the urban boundary $p^*(\bar{x})$ must adjust to equate to the agricultural rent, p_a.

3. Households choose to live in location x so that no alternative location would return higher utility, and given identical preferences, all households in the same income group attain the same level of utility, $\bar{u}_d$:
$$u(c_d, q_d^*(x)) = \bar{u}_d \quad \text{for all} \, x \in [0, \bar{x}], \quad d = R, P.$$

4. The market clearing bid-rent function is the envelope of bid-rent functions across all income and commuting groups:
$$p^*(x) = \max\{p_{d,m}(x)\}, \quad d = R, P, \quad m = C, S.$$

5. Bid rents equate between income and commuting types at some $x \in (0, \bar{x})$:
$$\begin{aligned}
p_{R,S}(x) &= p_{P,S}(x) \\
p_{R,S}(x) &= p_{R,C}(x) \\
p_{R,S}(x) &= p_{P,C}(x) \\
p_{P,S}(x) &= p_{R,C}(x) \\
p_{P,S}(x) &= p_{P,C}(x) \\
p_{R,C}(x) &= p_{P,C}(x),
\end{aligned}$$
where commuting boundaries between each group are defined by intersections that lie along the envelope equilibrium bid-rent function, $p^*(x)$. There must be at least 1 and at most 3 intersections for a spatial equilibrium to exist.

6. Congestion from car commuting at each location $x \in [0,\bar{x}]$ is determined by the density of car commuting between x and $\bar{x}$:

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

The uniqueness of an equilibrium is not guaranteed, and, in particular, the configuration of rich, poor, subway- and car-commuting households will depend upon the relative size of fixed costs of commuting relative to the variable costs, differences in the value of time, and household preferences for housing reflecting the income elasticity of commuting cost and housing demand.

In equilibrium, the bid rent function $p^*(x)$, the housing demand function $q(x)$, the utility level $\bar{u}$, and the urban boundary $\bar{x}$ are determined endogenously based on the level of commuting cost t, incomes y_d, population size N, and agricultural rent p_a.[48] Many urban configurations are possible, though we now focus on a specific baseline one. Given sufficiently high fixed costs for driving relative to subway, high variable costs for subway relative driving, and large enough differences in the value of time between rich and poor, a spatial configuration as in Figure A15 may emerge where a mass of rich households live closest to the CBD commuting by subway. Beyond this group, a mass of poor households also commute by subway, followed by a mass of rich households commuting by car that consume more housing than their subway commuting counterparts ($q_{R,S} < q_{R,C}$) to compensate for longer commuting. Finally a mass of car commuting poor households live at the urban boundary given their lower value time, but also consume more housing than their subway commuting counterparts ($q_{P,S} < q_{P,C}$). Bid rents are steeper for rich households than their poorer counterparts for each respective commuting mode because the rich have higher value of time.

Bid-Rent Functions under Heterogeneous Commuting Technology

The functional forms for bid rent functions illustrated in Figure A15 are described below. The bid rent functions for each transportation technology evaluated at the CBD ($x=0$) are:

$$p^0_{d,m} = \frac{1}{q_{d,m}}\left[y_d - \theta_m - c_d\right], \quad d = R,P; m = C,D,$$

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

The bid-rent function for subway riders is:

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

where ξ is the average subway speed.

[48] Endogenous determination of the urban boundary, $\bar{x}$ comes mechanically through the fixed population size and lot size: lot size multiplied by population yields the boundary where the envelope bid rent function is equal to agricultural rents.

The bid rent function for drivers is:

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

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

Appendix Figure A14 shows the effect of an exogenous increase in vehicle traffic congestion (e.g., due to road conditions that affect the relationship between traffic density and speed) to illustrate the equilibrium pattern of bid rent functions, housing density and mode choices. The bid rent functions are drawn taking into account congestion, reflecting the fact that the slope and curvature of the car commuting bid-rent functions adjusts based on the extent of car commuting across the city. To make this clear, consider the derivative of the rich, car commuting bid-rent function at a small, arbitrary distance ε to the right of x_B, the boundary between poor subway commuting and rich car commuting:

$$\begin{aligned} p'_{R,C}(x_B+\varepsilon) &= \frac{t_{R,C}(x_B+\varepsilon)}{q_{R,C}} = \frac{\nu_R}{q_{R,C}}\left[\mathscr{C}\left(n_C\left(x_B\right)-\varepsilon\right)\right] \\ &\approx \frac{\nu_R}{q_{R,C}}\left[\mathscr{C}\left(n_C\left(x_B\right)\right) - \varepsilon\mathscr{C}'\left(n_C\left(x_B\right)\right)\right] > p'_{R,C}(x_B), \end{aligned}$$

where ε also corresponds to the mass of commuters living along that distance given the assumption of fixed lot size at 1, and where $q_{R,C}$ is fixed as mentioned above. The second line is then an approximation to the change in the bid-rent, which is comprised of the congestion from the mass of all car commuters (rich and poor) at the modal boundary, x_B where commuting switches from car to subway (moving from right to left) less the congestion from the mass of rich car commuters living between x_B and $x_B+\varepsilon$. This demonstrates that a change in endogenous congestion in the model has two effects: it steepens the bid rent curve overall, and the curve itself gets steeper after passing through residential areas with car commuters as the flow of vehicles onto the roadway builds up.

A.5 Policy Approaches

We consider the effect of a set of transportation policies on the urban structure described above to motivate our empirical work (see Appendix 3.1 for the specific policies adopted in Beijing). While there are many potential equilibrium spatial configurations for the city laid out depending on model parameters, we calibrate the parameters of the model to a case in which the rich taking the subway to work live closest to the CBD, the poor taking the subway live in the next area, then the rich taking cars, then the poor taking cars and finally agricultural land beyond the urban boundary as can be seen in panel (a) of Figure A15.[49] LeRoy and Sonstelie (1983) demonstrate that the urban configuration can be explained, in part, by the relation between the income elasticity of commuting costs relative to the income elasticity of housing demand. The figure is consistent

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

with the case where the income elasticity of commuting costs is larger than the income elasticity of housing demand. therefore the rich outbid the poor to live close to the city center to save commuting costs.

Congestion Pricing. First we consider a typical first-best approach: a per-kilometer congestion charge. The optimal level would be equal to the marginal external cost of congestion and can be derived from differentiating (A4) with respect to $n_C(x)$ giving the increased cost from one additional car commuter. Multiplied by the number of car commuters, this yields:

$$\tau_C(x) = \left(v_R \frac{n_{R,C}(x)}{n_C(x)} + v_P \frac{n_{P,C}(x)}{n_C(x)} \right) \mathscr{C}'(n_C(x)) \,. \tag{A6}$$

The revenue from congestion pricing can then be recycled lump sum to each resident of the city. Panel (b) of Figure A15 shows the effect of the congestion pricing on spatial equilibrium in this hypothetical city. Due to the recycling of the revenue, the bid rent curves shift up for subway users. Under the uniform recycling of the revenue, the shift is larger for the poor than for the rich due to the smaller home size among the poor (the denominator of the intercept) as shown in Appendix A. The intercept of the bid rent curves for car users moves up as well. For poor drivers, the slope of the bid rent curve steepens as the congestion toll defined above would be larger than the savings from improved speed (due to their low VOT), hence leading to a higher travel cost per unit of distance (the numerator of the slope). For richer drivers, the bid recent curve would be flatter as the congestion toll would be smaller than the savings from improve speed (due to high VOT). However, the curve to the right of x'_B becomes steeper as it moves to x'_B from the right. This is due to the fact that congestion worsens as it is closer to x'_B from the right, leading to an increasing unit travel cost.

There are two competing forces at work that affect the spatial pattern of residential locations. First, congestion pricing increases the unit travel cost and incentivizes residents to move closer, hence bidding up home prices near the city center. Second, the reduction in congestion leads to time savings and reduces the travel cost. However, due to the differences in VOT, the time saving is more valuable to the rich than to the poor. That is, the second force is relatively stronger compared to the first force for the rich than for the poor. Given the initial spatial configuration, congestion pricing results in some poor residents shifting away from driving to subway while moving their residence from the outer ring to the inner ring. At the same time, some rich residents move out of the inner ring and switch to driving while living in larger homes. If congestion pricing or the cost of driving increases enough, the poor may occupy the city center as in the case when driving is prohibitively expensive for the poor (LeRoy and Sonstelie, 1983).

Subway Expansion. Panel (a) of Figure A16 considers subway expansion. This can be incorporated into this theoretical framework by making subway costs approach infinity beyond a desired distance x. While the choice of reference point is arbitrary, consider the effect of constraining transit to $\bar{x}_B$ in panel (a) of Figure A16. Given the fixed supply of housing, the constraint of public transit from x_B in panel (a) of Figure A15 to $\bar{x}_B$ shifts the mass of poor subway commuters $\bar{x}_B - x_B$ to the poor car commuting region. This increases congestion to the left of x_B for all car commuters, steepening the slope of the bid rent curves to the left of x_B

more for the rich than for the poor due to the high VOT among the rich. In the extreme case of removing all subway from the city, the bid rent curve for the rich would become even steeper as it gets closer to the city as the congestion worsens. Only the rich will occupy the city center. The results are consistent with Glaeser et al. (2008) which calibrates this class of models to corroborate the narrative that in the United States better public transportation leads more poor to live in the city center.

Driving Restriction. Finally, we consider the effect of raising the cost of driving via a driving restriction. In practice, the driving restriction only bans a portion of the cars from driving each day. We assume that during those days, residents need to take the subway to work. This would imply that the car commuters would need to pay the fixed cost of two modes, while the variable cost would be a weighted sum of the two modes. Panel (b) of Figure A16 shows the effect of the driving restriction. Due to the increase in the fixed cost for car commuters, the intercept of the bid rent curves shifts down, more so for the poor than for the rich (the denominator being larger for the rich). The change in the slope for the car commuters is subjects to two countervailing forces. On the one hand, the added variable cost (due to the higher variable cost when using the subway) will increase the slope. On the other hand, congestion reduction to the left of x'_B will reduce the slope, more so for the rich than for the poor. The first force likely dominants.

The impact of the policy is that the rich reduce car commuting ($x'_A - x_A$) by more than the poor ($x'_C - x_C$).

Welfare. Given the stylized nature of this model the welfare implications from this stylized example are probably less informative to real world policy applications than the empirical exercises performed in Section 6.1. That said, Figures A15 and A16 yield an important observation: key welfare effects follow the movement up and down of equilibrium bid rent functions from capitalization of changes in the transportation system. Put simply, an approximation to total welfare is the sum of equilibrium rents paid:[50] $\int_0^{\bar{x}} p^*(s)ds$.

This can be visualized as the area under the equilibrium bid rent envelope in the figures: when the envelope is lower than the baseline under a given policy scenario, then aggregate rents (and thus welfare) is lower, and visa versa. Considering Figure A16 panel (b), the effect of the driving restriction seems to lower the sum of rents by more than it is increased as the envelope is lower beyond x'_B. This accounts for a larger area than the increase in the level of the equilibrium bid-rent to the left of x'_B. In contrast, the overall effect of congestion pricing in Figure A15 panel (b) is to increase the equilibrium bid-rents envelope at all points in the city because congestion reduction flattens bid rents and the redistribution of the toll revenues increases their value for subway commuters. This points to an important general equilibrium effect of transportation policies in cities, where their benefits are capitalized into the housing market and provide additional welfare

[50] To be precise, this provides the sum of willingness-to-pay. Changes in the transportation system may induce changes in consumption with cost implications. If this is provided by perfectly competitive markets, we may assume that equilibrium prices reflect these costs. This logic, the Henry George Theorem, underlies the approach of using changes in land values to uncover the benefits of public good investments (Stiglitz, 1977; Arnott and Stiglitz, 1979). If transportation policies create or remove welfare reducing distortions, this will be reflected in relative changes in bid rent curves. Albouy and Farahani (2017) show that the value of infrastructure could be underestimated using this approach and argue for a more broad method to incorporate imperfect mobility, federal taxes, and non-traded production.

gain relative to a driving restriction beyond a partial equilibrium framework on the transportation sector alone as shown in Appendix Figure 1. Our simulations suggest the welfare impacts of these capitalization effects (and subsequent sorting) may be larger than the direct effects on transportation choice itself.

B Data Construction

B.1 Travel Survey Data

Commuting Mode Choice Set Construction Here we describe the details of choice set construction for commuting mode introduced in Section 5.1. We define the choice set to include six modes for a commuting trip: *Walk*, *Bike*, *Bus*, *Subway*, *Car*, and *Taxi*. In principle, a traveler could take one of the six modes alone or any combination of them, which would expand the choice set significantly. For subway and bus, we allow commuters to walk to and from subway stations and bus stops. For the other four modes, we eschew multi-mode commuting to make the number of alternatives tractable and because trips single-mode trips account for over 95% of all trips in our data. The construction of the travel time and travel distance via API or GIS for each mode is illustrated in Appendix Figure A3.

The same procedure was used to construct the travel choice set for home-work commutes in the housing data. Given the rapid expansion of the subway network during our data period, we construct the subway time and hence EV_{ij} term using the subway network two-year ahead in our baseline analysis. The subway construction needs to go through a long process including the approval from the central government, impacts evaluation, construction, and testing Yang et al. (2018); Gu et al. (2020). It takes 2-5 years from the start of the construction to the operation. The public announcement of subway station locations can be a few months ahead of the construction. We also conduct a robustness check using a one-year projection window in constructing subway time. Appendix Figure A4 shows travel time and cost of six routes for a particular trip based on the procedure.

The size of the choice set varies across commuters and trips based on household demographics and trip characteristics. Car mode is available for a given household member only if the household owns a vehicle and the member has a driver's license.[51] Walk, bike, and taxi modes are available for all trips. Bus availability is determined by the home and work locations and the mode is removed from the choice set if Gaode Maps API fails to provide any bus route, indicating the lack of the public bus service in the vicinity. We allow the subway to be available for each traveler based on the nearest station and assuming that they walk to/from the nearest subway stations to the origin and the destination.

We construct the monetary travel cost for each mode as follows. The monetary cost for walking is zero. For biking, the cost is zero for households who own a bike. For non-bike owners, the cost of biking is the rental price (free for the first hour and then ¥1 per hour with ¥10 as the maximum payment for 24 hours). The bus fare is set based on municipal bus rates at 0 for senior citizens, ¥0.2 for students, ¥0.4 for people

[51] Car rental is not common in Beijing and the mode share of using rental car is nearly zero.

with public transportation cards, and ¥1 for people without public transportation cards. The baseline 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. The taxi fee is based on the following rules: in 2010, ¥10 for the first 3 km, and then additional ¥2 per km and ¥1 as the gasoline fee; and in 2014, ¥13 for the first 3 km, and then additional ¥2.3 per km plus ¥1 as the gasoline fee.

Travel Survey Address Geocoding & Commuting Routes Here we describe the process used to clean the BHTS data as well as what was done to geocode home and work addresses from the BHTS as well as the approach to constructing commuting routes for counterfactual trips in the BHTS and hypothetical trips in the mortgage data.

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.

BHTS is made 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 kilometers on average and their size is inversely proportional to the density of trip origins and destinations: smaller TAZs are closer to the center of Beijing. 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. The sample locations are shown in Appendix Figure A2.

Geocoding addresses was performed using Baidu's API because the quality of Chinese character string matches from addresses was found to be better relative to alterantive 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. In order to validate geocoding results, we compared the traffic analysis zone (TAZ) for the geocoded home or work location to the TAZ reported in the BHTS. In 2010, 36% of respondents had a home or work address that did not generate a valid geocoding and were dropped from our sample. For 2014, 44% of respondents were dropped for this reason. In addition, we drop observations with the same origin and destination, multimode trips, walking trips over 10km, biking over 25km, driving over 50km, and trips between 11pm and 5am.

For route construction for the BHTS sample, we used the departure time for work commutes reported in the survey to determine the timestamp for Baidu and Gaode API driving time predictions. As discussed in the text, driving and taxi travel times were adjusted from the present to the relevant year using historic levels of Beijing's Travel Congestion Index (TCI) to reflect changes in the average level of congestion across the city over time. The TCI values vary between ring roads of the city.

For subway commuting, we identified the nearest subway stations to home at work using ArcGIS maps of the historical subway station. For the BHTS travel survey data, maps correspond to the date the trip was taken, but for the hypothetical trips being considered by home buyers in the mortgage data, we assume buyers are forward-looking and use the subway network two years after the household purchases a home. We used Baidu's API to calculate walking distances from home and work to the nearest subway station, and then used historical subway time tables to calculate travel times between origin and destination stations. We account for additional time when transferring lines.

The constructed travel distances and reported travel distances of chosen modes in the final dataset are highly correlated (correlation= 0.81). [52]

B.2 Mortgage Data Details

Appendix Figure A7 shows the distance to work by gender and how the boxplots of the distance to work by year for males borrowers and female borrowers, separately. On average, female borrowers have a slightly shorter commute than for male borrowers The average distance to work has increased from about 10km to 12km over time, reflecting the expansion of the city and transportation infrastructure.

As part of the social safety net, the mortgage program aims to encourage home ownership by offering prospective homeowners mortgage 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 during the employment period with the employer. As discussed above, only for those with formal employment were eligible for the government-backed mortgage program upon which our data are based. Although the mortgage data may have a good representation of the middle-class in Beijing it under-represents the two ends of the income distribution: low-income households without employment and rich households who do not take loans to finance purchases. To increase the representativeness of the mortgage data, we re-weight them based on two larger data sets we believe to be more representative. These data include resales from real estate listings and new sales from home registrations.

This larger data set includes about 40% of all transactions in Beijing's second-hand market during our data period from the largest real estate brokerage company, Lianjia, present throughout the city and across housing segments. The larger data set also includes all new residential home sales based on real home registration from Beijing Municipal Commission of Housing and Urban-Rural Development.[53] Different from

[52] 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).

[53] The new sales data include "commercial residential properties", accounting for about 90% of all new home sales. The data do not include transactions for employer-provided or subsidized housing.

the mortgage data, the housing transactions in this larger data set do not include information on the work location of the owners, therefore preventing us from using it for the main empirical analysis. To improve the representativeness of the mortgage data, we match the distributions of housing price, size, age, and distance to city center of the mortgage data to those in the larger housing data using entropy balancing (Hainmueller, 2012).

Resales data come from the largest real estate brokerage company that has a 40% market share in the resale market during our sample period. New sales data, accounting for about 90% new home sales, are from real home registration records from Beijing Municipal Commission of Housing and Urban-Rural Development, a government agency that records housing transactions.[54] New sales and resales account for 43% and 57% of the Beijing housing market, respectively during our sample period. Within the 6th ring road, the two shares are 37% and 63%. In the mortgage data, the sales of new homes account for about 29% of all sales throughout the city and 25% within the 6th ring road. To calculate weights based on these dataets, we apply an entropy method following Hainmueller (2012), which solves the following constrained optimization problem to match sample moments between the mortgage data and these other two, more representative datasets:[55]

$$\min_{w_i} H(w) = \sum_i h(w_i) = \sum_i w_i \log(w_i) \tag{A7}$$

subject to balance and normalizing constraints:

$$\begin{gathered}
\textstyle\sum_{i \in \text{new homes}} w_i c_{ri}(X_i) = m_{r,\text{registration data}} \\
\textstyle\sum_{i \in \text{resales}} w_i c_{ri}(X_i) = m_{r,\text{listing data}} \\
\textstyle\sum_i w_i = \text{total number of new homes + resales} \\
w_i \geq 0 \text{ for all } i. \\
\frac{\sum_{i \in \text{new homes}} w_i}{\sum_{i \in \text{resales}} w_i} = \text{true new home-resale ratio.}
\end{gathered}$$

$m_{r,\text{registration data}}$ and $m_{r,\text{listing data}}$ is our matching objective, the true rth covariate moments from registration data and real-estate listing datasets. $c_{ri}(X_{ij}) = (X_{ij} - \mu_j)^r$ is the rth order moment function of each matching covariate j across each i home. Matching covariates include housing prices, sizes, building ages, and distances to city center. We match first two moments (mean and variances). The third constraint normalizes the sum of weights equal to the total number of homes in the mortgage dataset. The fourth constraint requires positive weights. Fianlly, we impose a condition such that the ratio of new homes to resales is the same as what we have for the Beijing housing market in the same period provided by Beijing Municipal Commission of Housing and Urban-Rural Development. We solve for optimal weights, w_i^* using the `entropy` package in STATA.

To further gauge the representativeness of the re-weighted mortgage data, Appendix Figures A8 compares

[54] The new home sales data do not include transactions for employer-provided/subsidized housing.

[55] The 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.

re-weighted the mortgage data with registration and real-estate listings dataset. For both resales and new sales, homes in the mortgage data have a good match with the data in registration data and real-estate listings. Hence in all the analysis, we reweight the mortgage data based on the two larger data sets to improve the representativeness of our sample.

C Reduced-form Evidence

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

Price Gradient w.r.t. Subway Proximity This section provide further details on the reduced form results on the effects of Beijing's driving restriction policy on housing price gradients depicted in Figure 5. A number of confounding factors could undermine identification of the causal relationship between housing price gradients and the driving restriction policy. For example, if amenities improve over time in locations near subway stations more than in locations that are farther away from subway stations, this would result in a larger price increase for homes close to subway stations, leading to an overestimation of the true impact of the policy. On the other hand, if the changes in dis-amenities such as congestion or noise follow the aforementioned pattern, we would underestimate the impact of the policy. Causal identification requires an assumption that the housing price gradient with respect to the distance to the nearest subway station would be unchanged in the absence of the driving restriction policy. We test the plausibility of this assumption by examining trends in price gradients in the periods leading up to the policy based on an event study framework:

$$\ln(\text{Price})_{jt} = \sum_{k=-24}^{24} \beta_k \times \ln \text{D}_{jt} \times \mathbb{1}(t=k) + \text{x}'_{jt}\gamma + \varepsilon_{jt} \tag{A8}$$

where j denotes a home and t denotes a month. The outcome variable is the logarithm of theunit price (¥per m^2). We allow the slope of the price gradient β's to vary over time. The regression includes a flexible set of controls (x_{jt}) that include neighborhood fixed effects, year by month fixed effects, and complex-level attributes.[56] Standard errors are clustered at the neighborhood level to allow for correlations among the homes in the same neighborhood (e.g., due to unobservables).

Appendix Figure A9 shows the coefficient estimates of β_k with the coefficients varying by quarter. There does not appear to have a pre-existing trend before the policy, alleviating the concern of the time-varying and location-specific unobservables discussed above. While there is not a clear relationship between subway proximity and housing price before the policy, there is a clear downward shift in the slope of price gradient. Moreover, the negative relationship between subway distance and housing price becomes stronger over time.

[56] Complex-level attributes include complex age, floor to area ratio, green space ratio, land area of the complex, home management fee (HOA fee), the number of units in the complex, and the number of buildings in the complex.

The increasingly larger impact over time after the policy could be driven by the fact that the uncertainty in terms of the policy is reduced over time and that the enforcement has been tightened over time.

One additional concern for identification is that subway expansion may result in network externalities so that the benefit of a new station occurs not just to those living or working nearby but to all who may use that station. Subway construction usually starts 2-3 years before the opening and the announcements are even earlier. If our results are driven by the subway expansion, there should have been a steepening of the slope before the driving restriction. To further address this issue, we include in the regressions a measure of subway density and it is constructed as the inverse distance weighted number of subway stations from a given location following Li et al. (2019). This measure can be considered as the number of subway stations per unit area centered around a given housing unit and it increases as the subway network expands.

Appendix Table A1 provides the regression results for six specifications. The parameter of interest is the interaction between subway distance and the policy dummy. The third column corresponds to the specification with the event study graph in Appendix Figure A9. Adding neighborhood fixed effects to control for neighborhood amenities significantly changes both coefficient estimates from column (1) to column (2). However, adding the rich set of complex-level variables barely changes the results. The fourth column is a weighted regression in order to make the result more representative of the universe housing transactions. We re-weight the sample based on the large data on real estate listings and new home registrations to match the price distribution the homes using entropy balancing (Hainmueller, 2012) as described in section 3.2. The last two columns include the subway density measure as an additional control. The coefficient estimates on them are positive but not precise. The results from the last five columns are all qualitatively the same: the driving restriction policy increases the price premium for homes that are closer to subway.

Appendix Figure A10 provides a falsification test by randomizing treatment status (before or after the driving restriction policy) of each transaction while keeping the share of post-policy transactions fixed.[57] The figure shows the histogram of the coefficient estimates of the interaction term between distance and treatment dummy from 500 iterations. The estimate from the true sample (-0.019) lies outside of 99 percentile of the distribution. This further alleviates the concern that the estimated impact might be driven by unobservables.

Evidence for Household Sorting The change in the housing price gradient with respect to subway proximity highlights the impact of the policy on household location choices. To understand the underlying sorting process and how the policy impact differs by income groups, we examine household location choices relative to subway and work locations. Appendix Table A2 provides two sets of regressions. The dependent variable in the first three columns is the distance of the home to the nearest subway station while that in the last three columns is the distance of the home to the work location of the borrower(s). The key explanatory variables are household income and its interaction with the policy dummy. The coefficient estimates on the interaction term

[57] We take random draws from a uniform distribution and replace the draws with an indicator variable so that the indicator variable has the same mean as the treatment variable. This indicator variable is the new treatment variable in the estimations. This effectively randomizes the transaction date for the homes.

in the first two columns suggest that after the policy, high-income households tend to purchase homes that are closer to subway than before the policy. The third column is a weighted regression which produces imprecise estimates but with the same directions as the first two columns. The last three columns examine the impact on housing choices with respect to work location. The coefficient estimates suggest that high-income households live further away from where they work compared to low-income households. After the policy, high-income households tend to choose homes closer to work locations relative to before the policy. The results are robust to using the distance to the work location of the primary borrower only.

D Proof of Sorting Equilibrium

Here we abstract from uniqueness associated with endogenous housing prices. Propositions 1 and 2 of Bayer et al. (2004) demonstrate that given the assumptions laid out below regarding the functional form of utility and the error term, a unique vector of prices clears the market and results in a unique sorting equilibrium. The form of that proof applies Brouwer's Fixed Point Theorem as below, but on the basis of an element-by-element inverse of the demand function for prices. The sufficient condition for uniqueness is that the sum of the partial derivatives of inverse demand function with respect to prices are bounded by $(-1,1)$. The unique set of prices then implies a unique set of choice probabilities that imply a unique sorting equilibrium. We proceed to demonstrate existence and uniqueness of the equilibrium in the presence of endogenous congestion.

We assume utility from housing as:

$$U_{ij} = \mathbf{x}_j \beta_i + \phi_i EV_{ij} + \varepsilon_{ij}, \tag{A9}$$

where $x_j =$ are exogenous attributes of housing location j, and EV_{ij} is the ease of commuting for housing location j for household i.[58] EV_{ij} is constructed as the logarithm as the sum of indirect utility from each commuting mode:

$$EV_{ij} = \log\left(\sum_{m\in M_{ij}, m\neq car} \exp\left\{\mathbf{Y}_m \lambda_i\right\} + \exp\left\{\mathbf{Y}_{car}\lambda_i + \gamma_1 \frac{dist_{ijcar}}{\bar{v}_{ijcar}\cdot v\left(R_{car}\right)}\right\}\right) \tag{A10}$$

$Y_m =$ are exogenous attributes of individuals and driving modes m in the mode choice set (cost, alternative specfic constants, and demographic characteristics), M_{ij} of household i at housing location j. The second exponentiated term is the indirect utility from driving, which depends on exogenous attributes and endogenous travel time, which is the ratio of driving distance to baseline travel speeds from Baidu, v_{ijcar}.[59]

Here the term $v\left(R_{car}\right)$ is a single-valued speed adjustment factor that depends on the share of households in Beijing choosing driving as their commuting mode, R_{car}. Travel times therefore adjust in response to

[58] As noted in the preceding paragraph, we omit the housing price from this equation although the results hold with endogenous housing prices by the logic laid out there.

[59] The coefficient vector λ_i here nests the parameters other than those account for travel time in equation (3) of the sorting model.

changes in the share of drivers.

Each exogenous attribute, k varies by individuals

$$\beta_{ik} = \bar{\beta}_k + \mathbf{z}_i \beta_k, \quad \lambda_{ik} = \bar{\lambda}_k + \mathbf{z}_i \lambda k,$$

where $\mathbf{z}_i$ is a vector of household i's attributes.

Assumptions

- Coefficient on driving travel time γ_1 terms and on travel time (which is determined by speed) is the same across individuals
- Unobserved vector of preferences ε_i observed by all other homebuyers through a static, simultaneous-move game following Nash Equilibrium concept
- Continuum of individuals with observed characteristics z_i to allow for integration out
- Speed changes from changes in the number of drivers affect congestion in every part of the city by the same factor $v(R_{car})$
- ε is drawn from a continuous, well-defined distribution function

Probability of choosing car,

$$R_{ijcar} = r_{ij}(\mathbf{z}_i, \mathbf{Y}_{ijm}, \bar{v}_{ij}, v(R_{car}); \lambda, \gamma_1) \quad \forall i, j \tag{A11}$$

is a function of household demographics, mode attributes including baseline speed, and the endogenous speed adjustment factor.

Aggregating these probabilities across individual car commuters yields:

$$R_{car} = \sum_j \sum_i r_{ij}(\mathbf{z}_i, \mathbf{Y}_{ijm}, \bar{v}_{ij}, v(R_{car}); \lambda, \gamma_1) \tag{A12}$$

Definition 1. *A sorting equilibrium is a set of location and commuting decisions that are optimal given the location and commuting decisions of all other individuals in Beijing.*

(A12) defines a single-valued function R that maps $[0,1]$ into itself and therefore a fixed point allowing the following proposition.

Proposition 1. *If U_{ij} is defined as in (A9), a sorting equilibrium exists for all i, j.*

Proof. The equations above show how (A12) is a continuous mapping of a closed and bounded interval onto itself. The existence of a fixed point follows from Brower's Fixed Point theorem. Any fixed point R^* is associated with a unique set of choice probabilities of driving in (A11) that satisfy the conditions for a sorting equilibrium. The existence of this fixed point implies the existence of a sorting equilibrium. ■

To define uniqueness of the equilibrium, it is helpful to define an implicit function of A12

$$\Phi(\mathbf{z}_i, \mathbf{Y}_{ijm}, \bar{v}_{ij}, v(R_{car}); \lambda, \gamma_1) = R_{car} - r(\mathbf{z}_i, \mathbf{Y}_{ijm}, \bar{v}_{ij}, v(R_{car}); \lambda, \gamma_1) \tag{A13}$$

Proposition 2. *If U_{ij} is defined as in (A9), and $\gamma_1 < 0$ a sorting equilibrium is unique.*

Proof. Proof of uniqueness is simplified by the fact that (A12) is a single-valued function of R onto itself. So long as a solution exists, it must be unique. ■

E Estimation Details

Here we elaborate on the simulated maximum-likelihood approach to estimating housing demand discussed in section 4.5. The utility function is re-written as a sum of household-specific utility μ_{ij} and mean utility (or alternative-specific constants) δ_j which absorb variation from unobserved housing attributes ξ_j. The simulated MLE with a nested contraction mapping estimates household-specific parameters (θ_2) and mean utilities (δ_j). The log-likelihood function is defined as

$$\begin{aligned} ln\mathscr{L}(\theta_2, \delta_j) &= \sum_i \sum_j \mathbb{I}^i_j w_i ln P_{ij}(\theta_2, \delta_j), \text{ where} \\ P_{ij}(\theta_2, \delta_j) &= \frac{\exp[\mu_{ij}(\theta_2) + \delta_j]}{\sum_h \exp[\mu_{ih}(\theta_2) + \delta_h]}. \end{aligned} \tag{A14}$$

$\mathbb{I}^i_j$ is an indicator function being one when household i chooses housing j, w_i is the weight of household i, and P_{ij} is the choice probability. Berry et al. (2013) show that under reasonable assumptions, for a given vector of θ_2, there exists a unique vector of δ_j that can perfectly match the predicted market shares and observed market shares:[60]

Therefore, δ_j can be estimated through the nested contraction mapping by matching observed and predicted market shares by inverting shares in each iteration d:

$$\delta_j^{d+1} = \delta_j^d + \ln \sigma_j - \ln \tilde{\sigma}_j(\theta_2, \delta_j^d),$$

$$\sigma_j = \frac{1}{N} \sum_i P_{ij}(\theta_2, \delta_j),$$

where σ_j on the left hand side is observed market share for housing choice j and right hand side is the predicted share. N is the number of potential buyers on the market, which varies over time.[61] By controlling for unobserved housing attributes ξ_j using housing fixed effects δ_j, the estimation can produce consistent estimates on household specific parameters θ_2 including the coefficients on price and the EV term, both of which could be correlated with unobserved housing and neighborhood attributes.

[60] In our case, since we do not have an outside good option, we need to normalize the mean utility of a random home to be zero.

[61] In practice, the market share of each housing choice is $1/N$ times the housing weight and the contraction mapping matches the aggregate choice probabilities to a vector of housing weights, which is essentially the first order condition of the log-likelihood function with respect to δ_j as shown in Bayer et al. (2007).

F Simulation Approach

In the benchmark simulations, we assume Beijing is a "closed city" with no change in population and a fixed housing supply consisting of the units in our sample. While this assumption is clearly runs counter to the facts, it helps us to isolate the direct effects of transportation policies on current Beijing residents as distinct from effects that are mediated by in-migrantion of new residents. We also assume that the transportation network is fixed apart from the expansion of subway network. In the robustness check in Table 7, we allow the housing market supply to adjust to housing prices. This is done by assuming a constant price elasticity of supply of two following Saiz (2010); Wang et al. (2012). We use observations from 2014 to conduct simulations and set the baseline policy scenario the counterfactual outcome without a driving restriction, congestion pricing, and the expansion in Beijing's subway between 2008 and 2014 (i.e., we fix the network to the 2008 system).

To allow driving congestion levels to respond to changes in the pattern on commuting as households change mode and residential location, we first estimate the relationship between traffic speed and density following Yang et al. (2019). This approach leverages the plausibly exogenous variation in traffic density induced by the driving restriction policy in Appendix Table A3.[62] Based on hourly traffic speed and density data from all major roads (freeways and expressways, but not the secondary roads), the elasticity of traffic speed with respect to density is estimated to be -0.62 within the 6th ring road during peak hours. Because the data does not cover secondary roads where traffic congestion tends to be worse, this estimate is likely a lower bound for the elasticity in the whole city. We restrict the sample to observations with traffic density larger than 35 (cars/lane-km), the elasticity estimate becomes -1.1. The average speed among these observations is about 30km/h, close to the city-wide average speed during peak hours. Hence we use -1.1 as the speed-density elasticity in our simulations.[63]

We now describe the simulation algorithm used to simulate outcomes under the baseline and other policy scenarios. The simulation algorithm starts with an initial observed housing price vector and road congestion factor, which are endogenously determined by the algorithm in each iteration. Each iteration has an inner loop to solve for housing location choices. When solving for housing location, the algorithm takes the ease-of-commuting measure for each of the two household workers (borrower and co-borrower), EV_{ij} defined in equation (4) as given and solves for housing market clearing prices. In our setting of a closed city (without outside options), housing prices are only identified up to a constant. As result, we fix the average price of all homes to be the same so as to partial out price changes as asset shock for all consumers to simplify welfare calculations. Next, the outer loop takes residential location choices as given and solves for the level

[62] The policy restricts some vehicles from driving one day per week during weekdays depending on the last digit of the license plate number. The policy follows a preset rotation schedule in terms of which pair of numbers (1&6, 2&7, 3&8, 4&9, 5&0) is restricted on a given day, and it is not adjusted based on traffic conditions. The policy generates exogenous variation on traffic density due to the fact that the distribution of vehicles is not uniform with respect to the last digit of plate numbers. Vehicles with the license plate ending with number 4 only account for about 2% of all vehicles because the number 4 is considered an unlucky number in Chinese culture. Therefore, when numbers four and nine are restricted, more vehicles are on the road, and congestion tends to be worse than other days.

[63] In practice, there could be route-specific congestion responses from commuters, which we abstract away in our simulations for tractability.

of congestion given optimal commuting choices.

Under the driving restriction, we calculate EV_{ij} by assuming driving is not an option one in five days. This would make the driving mode to be less attractive and homes that are closer to subway or work centers more attractive. Congestion pricing enters EV_{ij} by increasing the monetary cost of driving.

A step-by-step simulation algorithm is illustrated below.

We observe the baseline traffic density d_0 and driving speed $\mathbf{v}_0$. The elasticity of traffic speed with respect to density whose estimation procedure described above is denoted by e. Within the algorithm, we recalculate the density on each iteration of the outer loop based on the pattern of demand for car commuting across the city. The demand for driving, in turn, comes from the relative probability of car commuting based on equation 3.

Demand parameters whose estimation is described in Section 4 are: $\{\gamma_1, \gamma_2, \eta, \beta, \alpha, \phi, \theta, \xi\}$. $\mathbf{p}$ is the observed baseline price vector. Household-trip characteristics $\{\mathbf{w}\}$ and housing attributes $\{\mathbf{X}\}$ are observed. We also know the travel distance, time, and cost information for different modes of each house-working place pair. Housing choice sets $C(i)$ are fixed in simulations (constant housing supply). We also fix sets of random draw $v_r, r = 1, ...R$, and $w_q, q = 1, ...Q$. (We use the same random draw as in the estimation. R=200, Q=100.)

For each counterfactual scenario:

1. Guess an initial traffic density level d^0 (e.g., baseline density d_0).

2. Based on density level d^t for iteration t ($= d^0$ for first iteration):

(a) Compute driving speed $\mathbf{v}^t = \mathbf{v}_0 \left(1 + e\left(\frac{d^t}{d_0} - 1\right)\right)$;

(b) Update each member k (borrower and co-borrower) in household i's new commuting time to each house j in household i's choice set based on new driving speed $time^t_{ijk,\text{drive}} = \frac{dist_{ijk,\text{drive}}}{v^t_{ijk}}$(as well as taxi speed $time^t_{ijk,\text{taxi}}$);

(c) Given $time^t_{ijk}$ calculate expected ease-of-commuting measure value of home-work commute for member k conditional on home j :

$$EV^t_{ijk}(time) = \frac{1}{R}\sum_{r=1}^{R}\log\left(\sum_m \exp\left(\frac{\text{cost}_{ijk,m}}{\text{hourly wage}_{ik}}\gamma_1 + time^t_{ijk,m}\gamma_{2r} + \mathbf{w}_{ijk,m}\eta + \theta_{mr}\right)\right)$$

and member k's probability for driving:

$$R(drive|i,j,k) = \frac{1}{R}\sum_{r=1}^{R}\frac{\exp\left(\frac{\text{cost}_{ijk,drive}}{\text{hourly wage}_{ik}}\gamma_1 + time^t_{ij,drive}\gamma_{2r} + \mathbf{w}_{ijk,drive}\eta + \theta_{mr}\right)}{\sum_m \exp\left(\frac{\text{cost}_{ijk,m}}{\text{hourly wage}_{ik}}\gamma_1 + time^t_{ijk,m}\gamma_{2r} + \mathbf{w}_{ijk,m}\eta + \theta_{mr}\right)}$$

If we allow sorting, continue with step (d)-(e). Otherwise skip them and move to step (f);

(d) Given the new expected commuting values, calculate a new home price vector $\mathbf{p}^t$ such that (1) housing demand= housing supply, (2) weighted mean housing prices are kept the same as the baseline level:

$$\begin{aligned}\frac{1}{Q}\sum_{q=1}^{Q}\sum_{i\in C^{-1}(j)} w_i \frac{\exp\left(\sum_k \phi_{kq} EV_{ijk}^t + \alpha_i p_j^t + \mathbf{X}_j\beta_i + \xi_j\right)}{\sum_{s\in C(i)} \exp\left(\sum_k \phi_{kq} EV_{isk}^t + \alpha_i p_s^t + \mathbf{X}_s\beta_i + \xi_s\right)} &= f\left(w_j, p_j\right), \ \forall j \in J \\ \text{and mean}\left(\mathbf{p}^t\right) &= \text{mean}\left(\mathbf{p}\right)\end{aligned}$$

if we keep the supply of housing constant, then

$$w_j^t = f\left(w_j, p_j\right) = w_j$$

if we allow housing supply to adjust, we assume that it adjusts based on a constant elasticity rule:

$$w_j^t = f\left(w_j, p_j\right) = w_j\left(1 + e_{w,p}\left(\frac{p_j^t}{\bar{p}_j} - 1\right)\right)$$

(e) Given new $\mathbf{p}^t, EV^t$, calculate the new housing choice probability

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

(f) Update the new traffic density:

$$\tilde{d} = \sum_i \sum_j \Pr\left(j|i\right)\left[\sum_k R\left(drive|i,k,j\right) \times dist_{ijk,\text{drive}}\right]$$

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

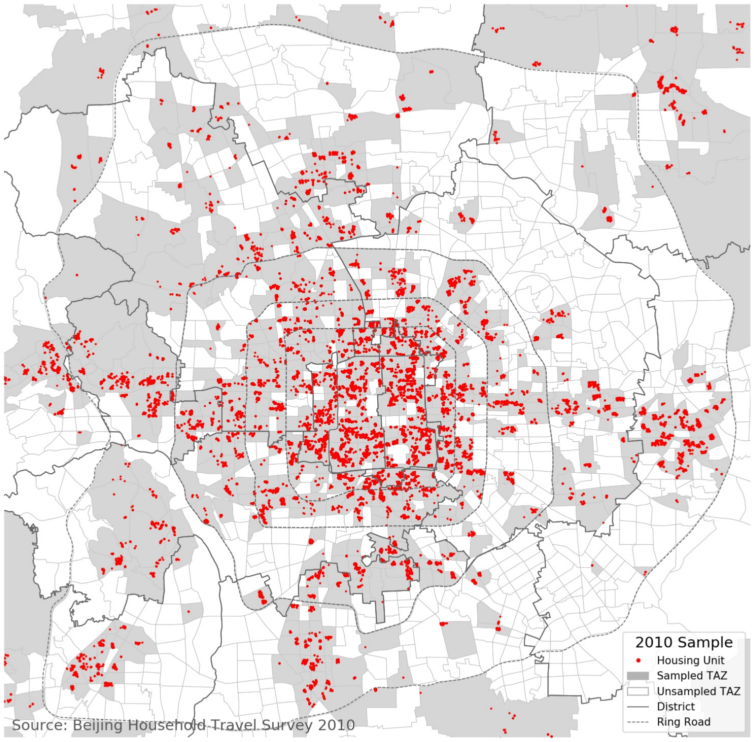
G Figures & Tables

Figure A1: Subway Network Expansion from 2000 to 2020 in Beijing

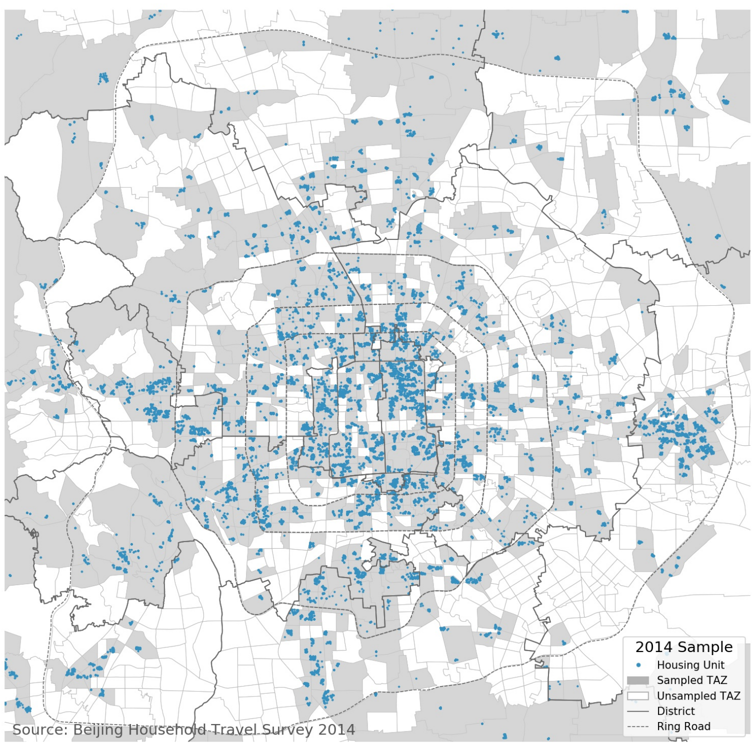
Note: Subway expansion from 1999 to the end of 2019 in Beijing expanded from 2 lines to 22 lines. From 2007 to 2018, 16 new subway lines were built with a combined length of over 500km. By the end of 2019, the Beijing Subway is the world's longest and busiest subway system with a total length of nearly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

Figure A2: Beijing Household Travel Survey Samples

(a) 2010 BHTS

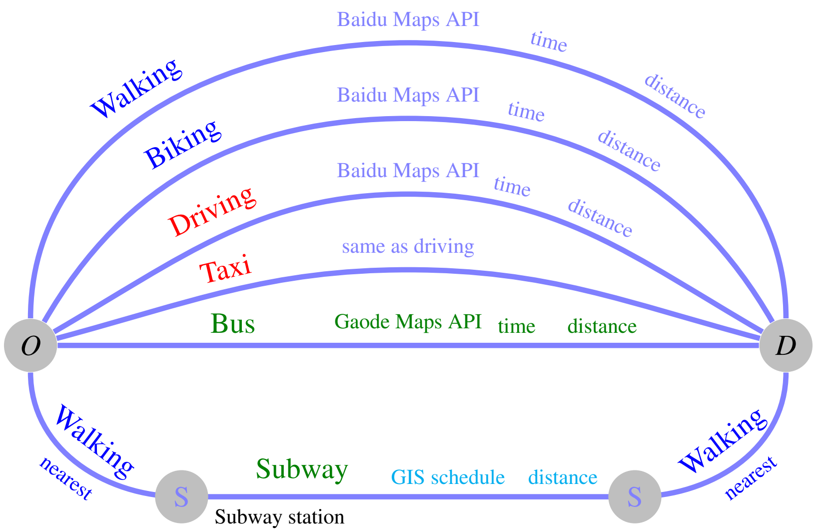


(b) 2014 BHTS



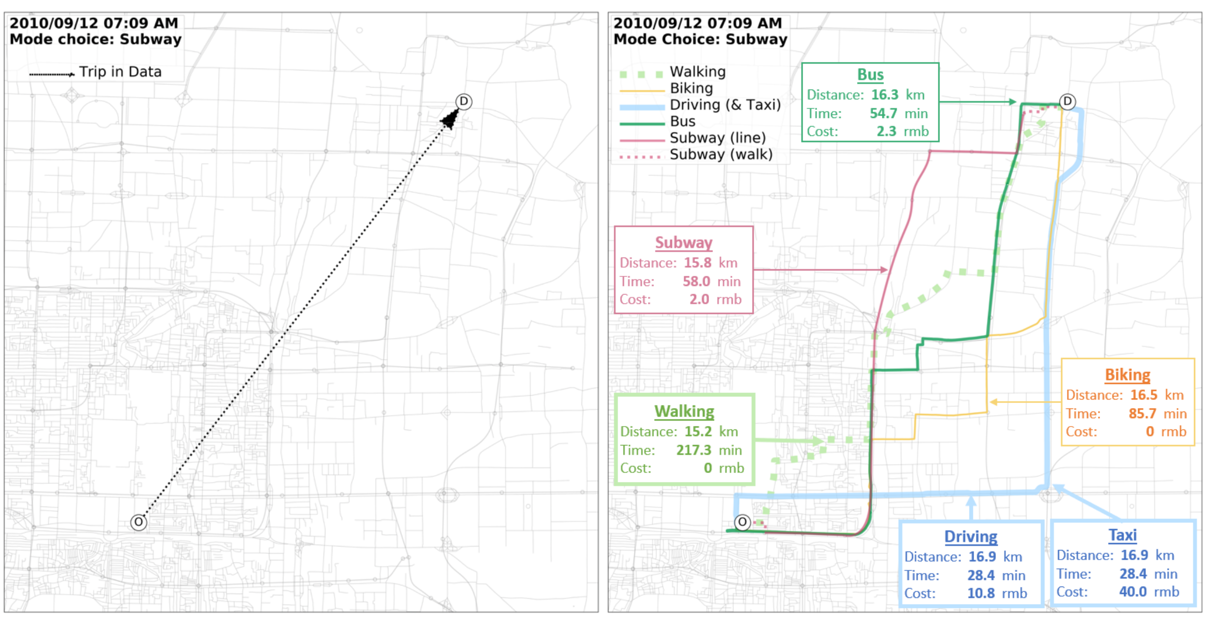
Note: The figures show the home locations of 2010 and 2014 Beijing Household Travel Survey with the 6th ring road, the focus of our analysis. The shaded areas are sampled TAZs. The success rates of geocoding for the home addresses are 97% and 98% for the two years.

Figure A3: Construction of the Travel Choice Set



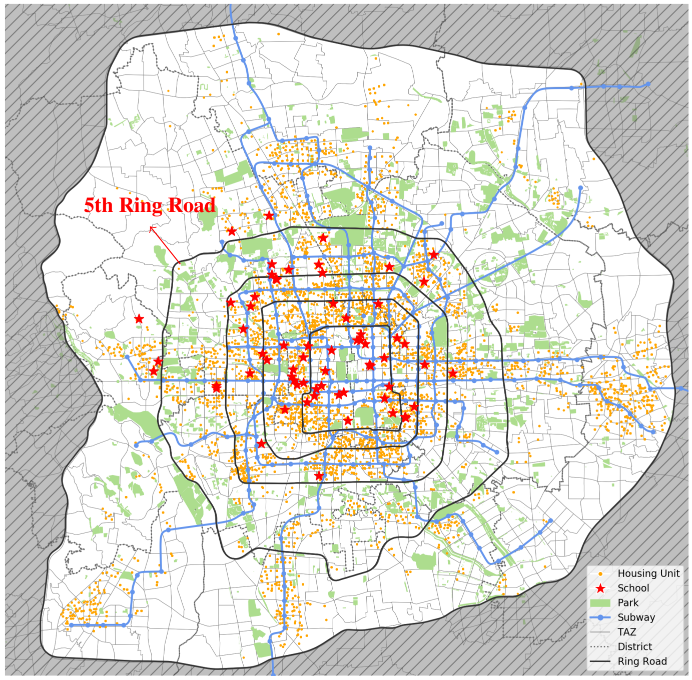
Note: Travel time and distance using walk, bike, car, and taxi are constructed using Baidu Maps API based on the departure time and the day of the week. Travel time and distance are constructed using Gaode Maps API because it provides the number of transfers and walking time between bus stops. Walking time and distance to and from the subway station are provided by Baidu Maps API while the distance from the first station to the last station are calculated using GIS based on the subway network in 2010 and 2014. The corresponding subway time table is used to construct station-to-station travel time. The data queries for car and bus trips are based on the same time and day-of-the-week as the survey. The travel time for bus, car and taxi are adjusted based on the traffic congestion condition on the day of the travel, relative to that on the day of API query.

Figure A4: Sample Routes



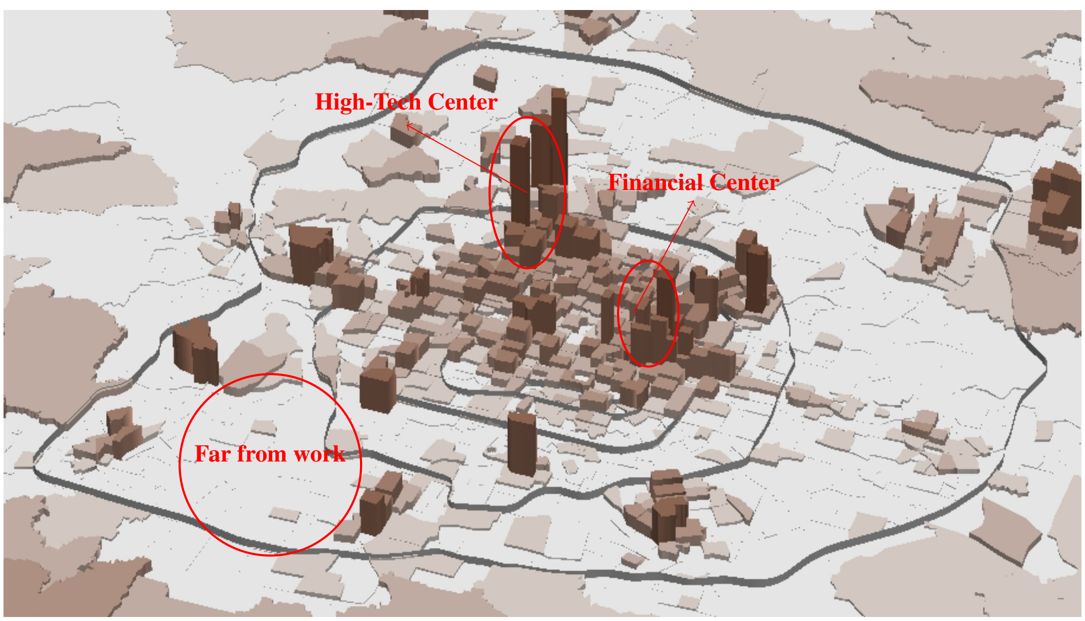
Note: For illustrative purposes, the figure shows travel time and cost of six routes for a particular trip that started from 7:09am on 9/12/2010. The chosen mode was subway. The left panel shows the the straight-line direction of travel, while the right panel shows the sets of routes, time and monetary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

Figure A5: Housing, Amenities and Transportation Network



Note: The figure shows the home locations in the mortgage data overlaid with ring roads (black lines), subway lines in blue (as of 2015), government-designated key schools denoted by red stars(65), and government-designated parks denoted in green.

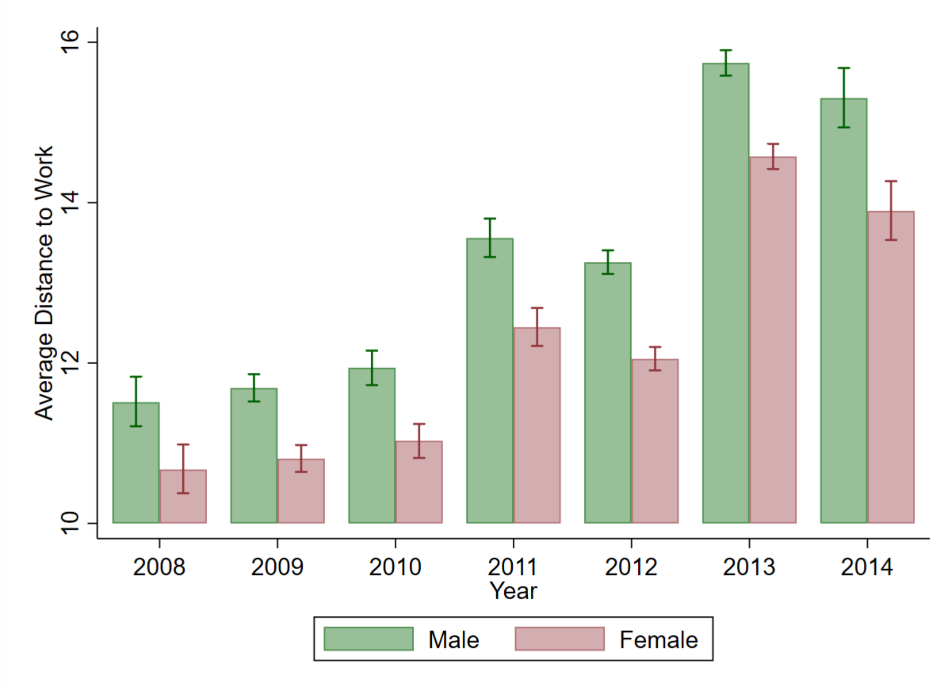
Figure A6: Job Density



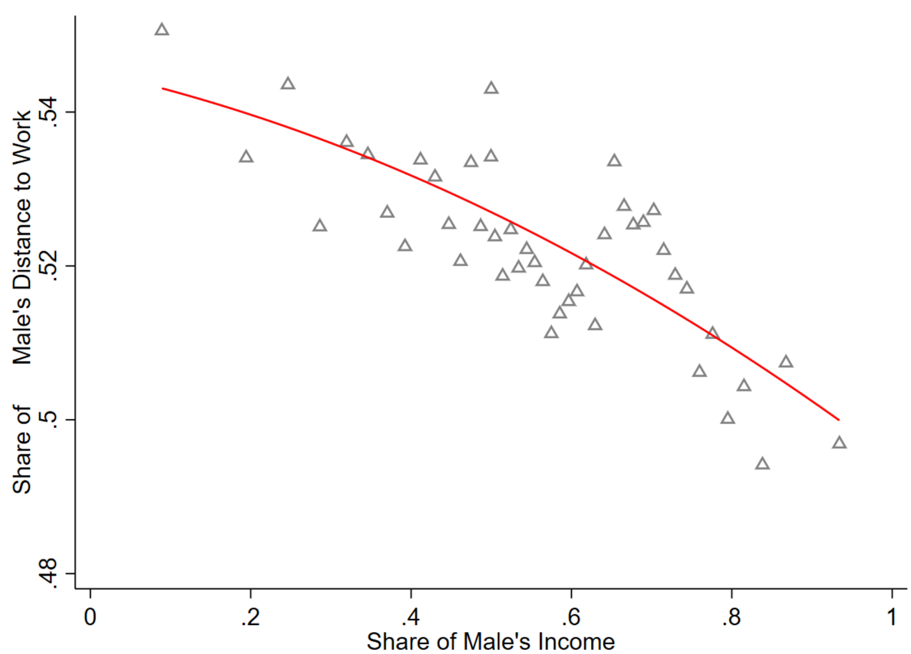
Note: This figure plots work density by TAZ based on counts of work locations from the mortgage data in our sample. Darker colors/taller shapes indicate greater work density.

Figure A7: Distance to Work

(a) Distance to Work for By Gender



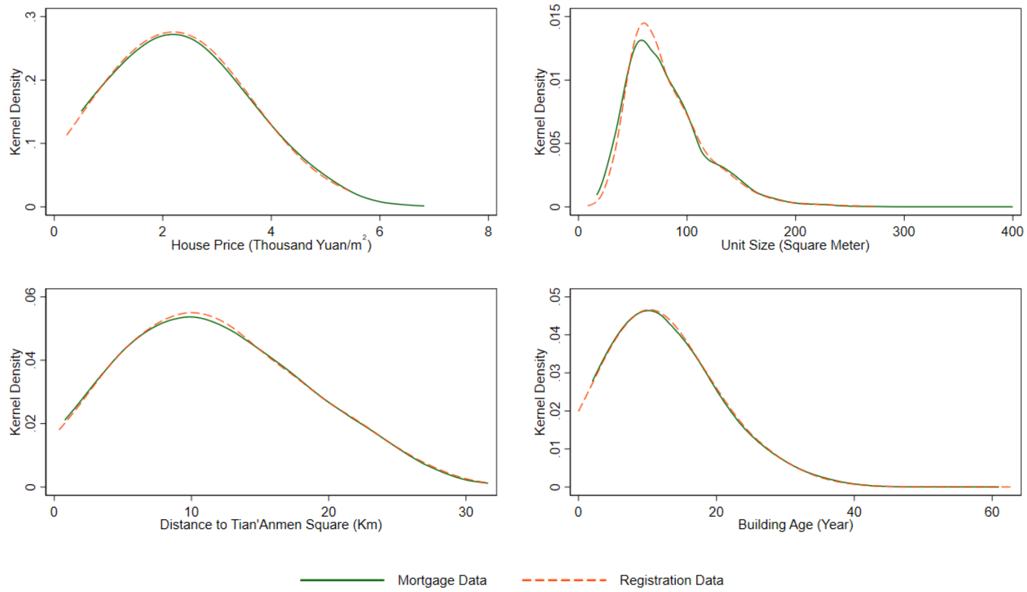
(b) Borrower's Share of Distance to Work with respect to Borrower's Share of Income



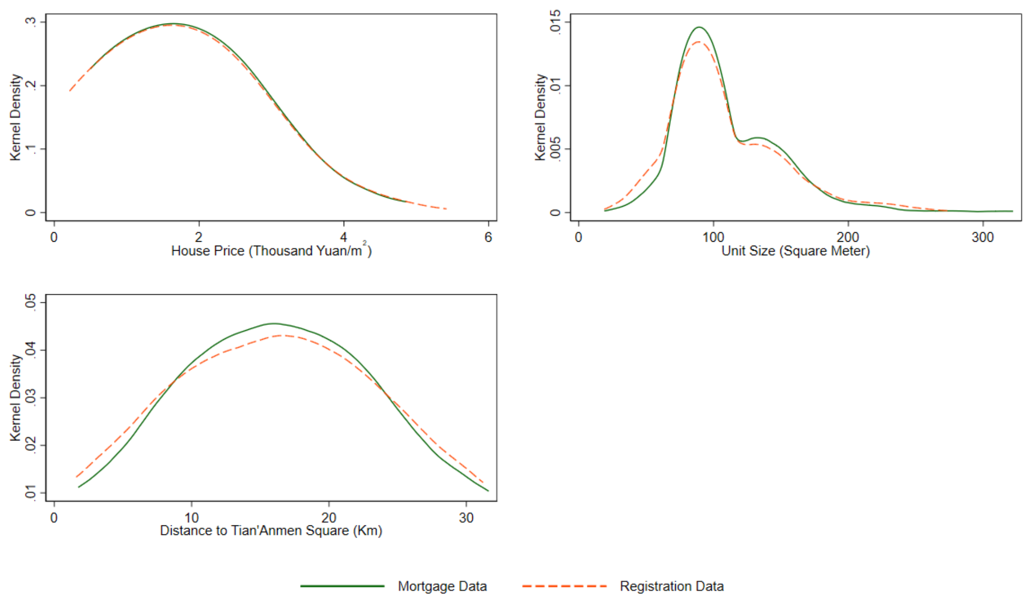
Note: Panel (a) shows how the average distance to work by year for male and female members, separately. The green bars are for males while the red bars are for females, the whiskers denote 95% intervals. Male borrowers have a longer commute than female borrowers. The distance has increased over time reflecting the expansion of the city and transportation infrastructure. Panel (b) is the binned scatter plot showing that the male member's relative income share has a weak negative relationship with his share of distance to work (the male member's distance to work over household's total distance to work).

Figure A8: Comparing Mortgage and Real Estate Listings after Weighting

(a) Resales

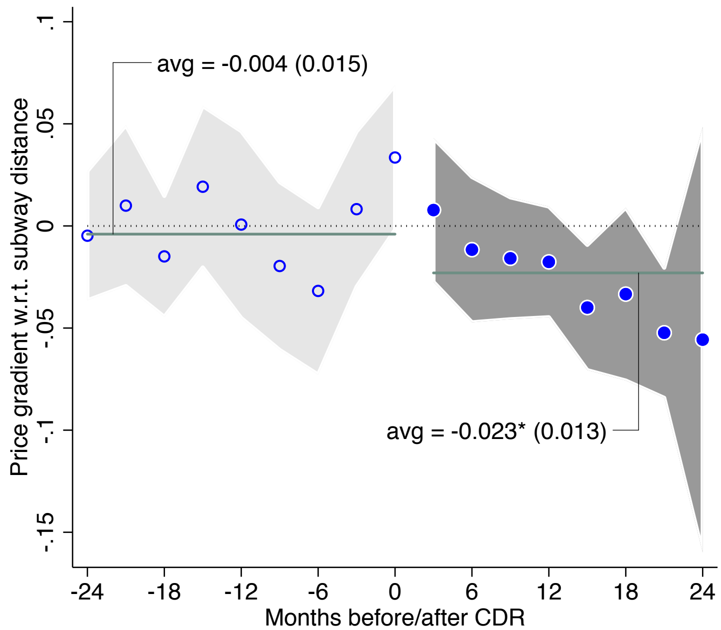


(b) New Homes



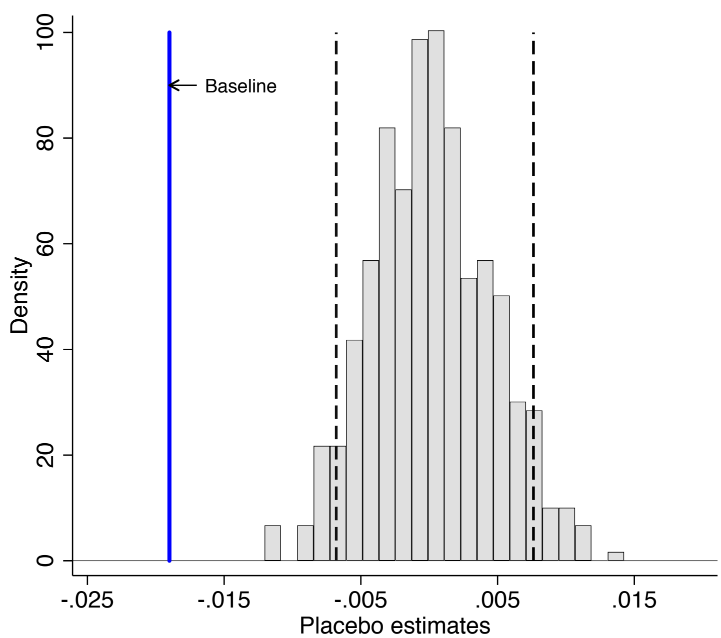
Note: The graphs gauge the representativeness of the mortgage data by comparing homes in the re-weighted mortgage data set with homes in a much larger real estate listing database in terms of prices per square meter, home size, home age, and distance to Tian'anmen Square. The upper panel shows comparisons of resales (58,718 obs in mortgage dataset and 160,836 obs in real estate listings) and the bottom panel shows comparisons of new sales (18,529 obs in mortgage dataset and 309,256 in real estate listings). Solid lines represent the mortgage dats and dashed lines represent real estate listings. As shown in the figure, after the weighting, homes in these two data sets are well balanced.

Figure A9: Event Study of the Price Gradient Estimates



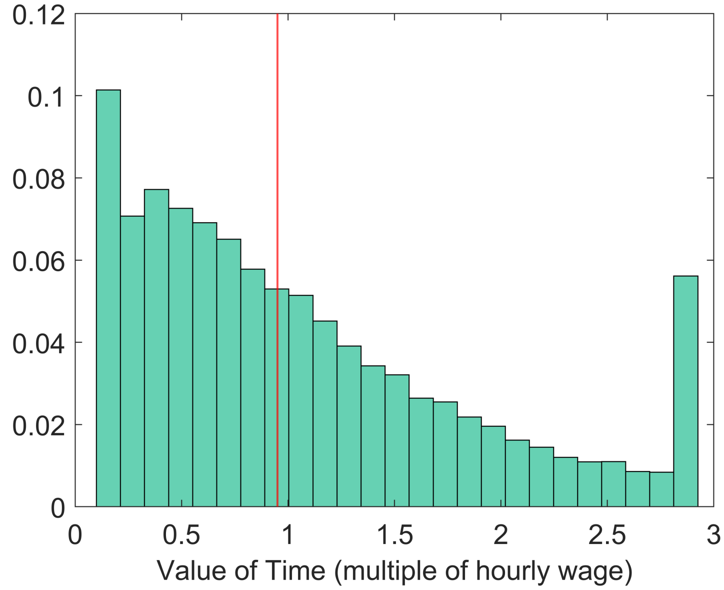
Note: This figure shows estimates from the event study of Beijing's driving restriction on the slope of the housing price gradient shown in Figure 5. Blue dots report the estimates by month relative to the start of the driving restriction in 2008, shaded areas report 95% confidence intervals, green lines indicate the average value of the coefficient before and after the policy, and the dashed line shows the location of zero.

Figure A10: Falsification Test on Price Gradient



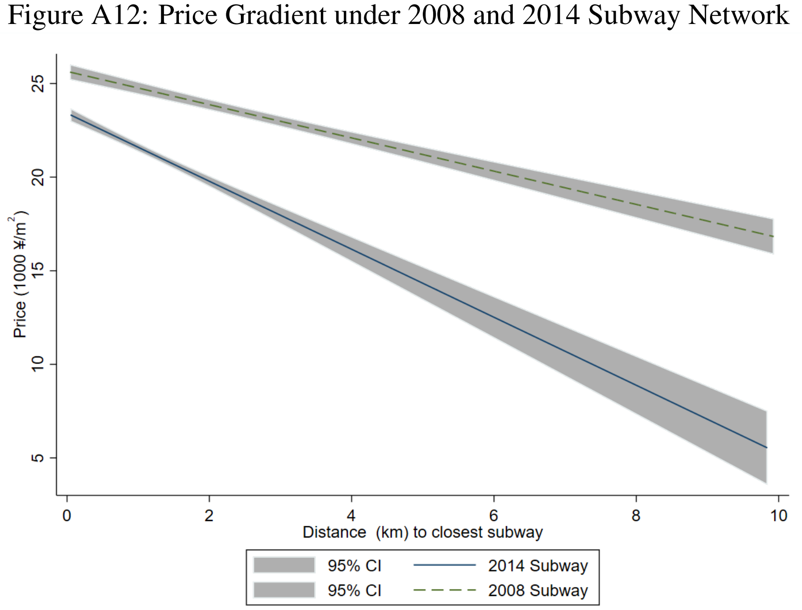
Note: This figure reports the distribution of coefficient estimates from placebo tests with randomized event time. Dashed vertical lines indicate the 95% confidence interval, while the solid vertical line indicates the estimate from our event study.

Figure A11: Implied Value of Time Distribution from Mode Choice Estimation



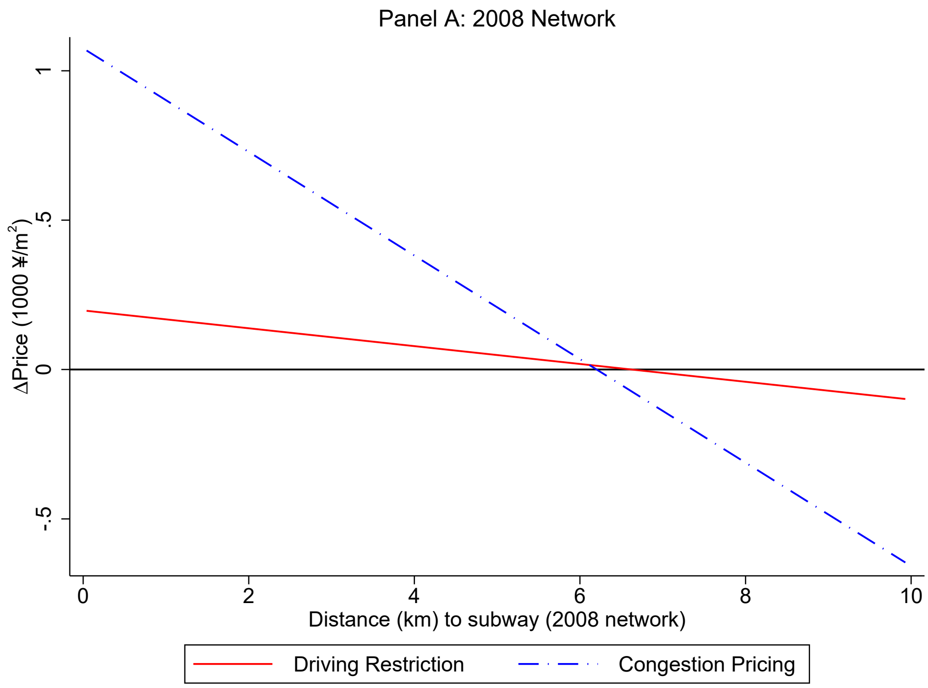
Note: The figure plots the distribution of the implied value of time (VOT) is based on the last specification of our mode choice model in Table 3. The preference on travel time has a winsorized (at 95th and 5th percentile) chi-square distribution with degrees of freedom equal to three while the preference on travel cost is inversely proportional to income. The value of time is in terms of hourly wage. The red line shows the mean VOT (51.1% of hourly wage). The median VOT is 57.1% of hourly wage.

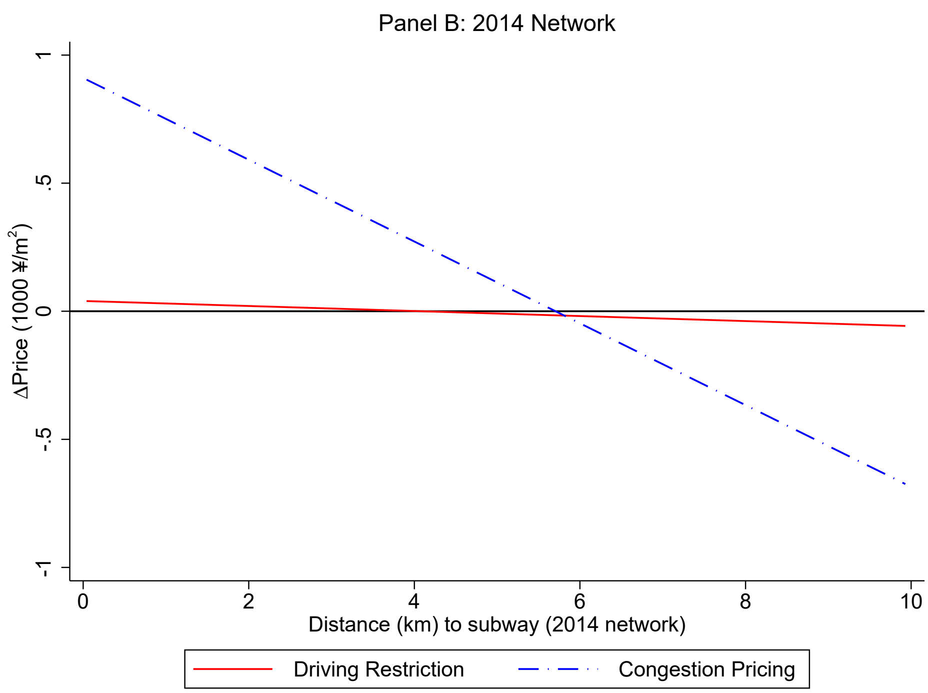
Figure A12: Price Gradient under 2008 and 2014 Subway Network



Note: This plot shows the simulated bid rent curve with respect to subway distance under the 2008 and 2014 network, respectively. The results are based on specifications as shown in 6. The gradient of the bid rent curve 2014 subway system (-¥1900/m^2 per km) is steeper than 2008 subway system (-¥700/m^2 per km). This reflect people's higher WTP for proximity to subway stations. The level shifts down under 2014 system as well, which reflects that subway expansion reaches to cheaper homes farther away from the city center.

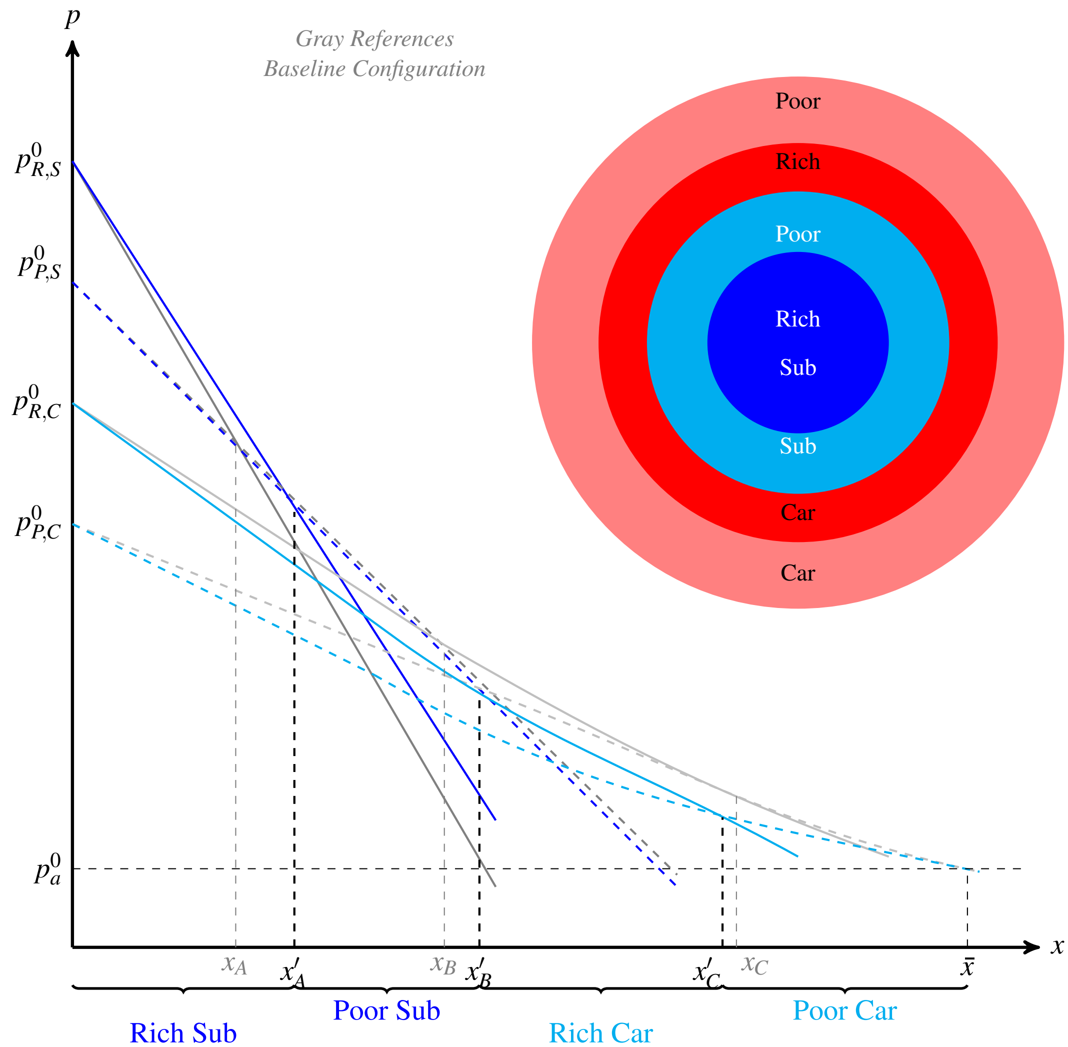
Figure A13: Change in Price Relative to the Baseline





Note: The plot shows the changes in the price gradient (with respect to subway distance) as a result of driving restriction and congestion pricing relative to the price gradient under the no policy scenario. The top graph is under the 2008 subway network and the bottom graph is under 2014 network.

Figure A14: Exogenous Increase of Congestion in Monocentric City Model

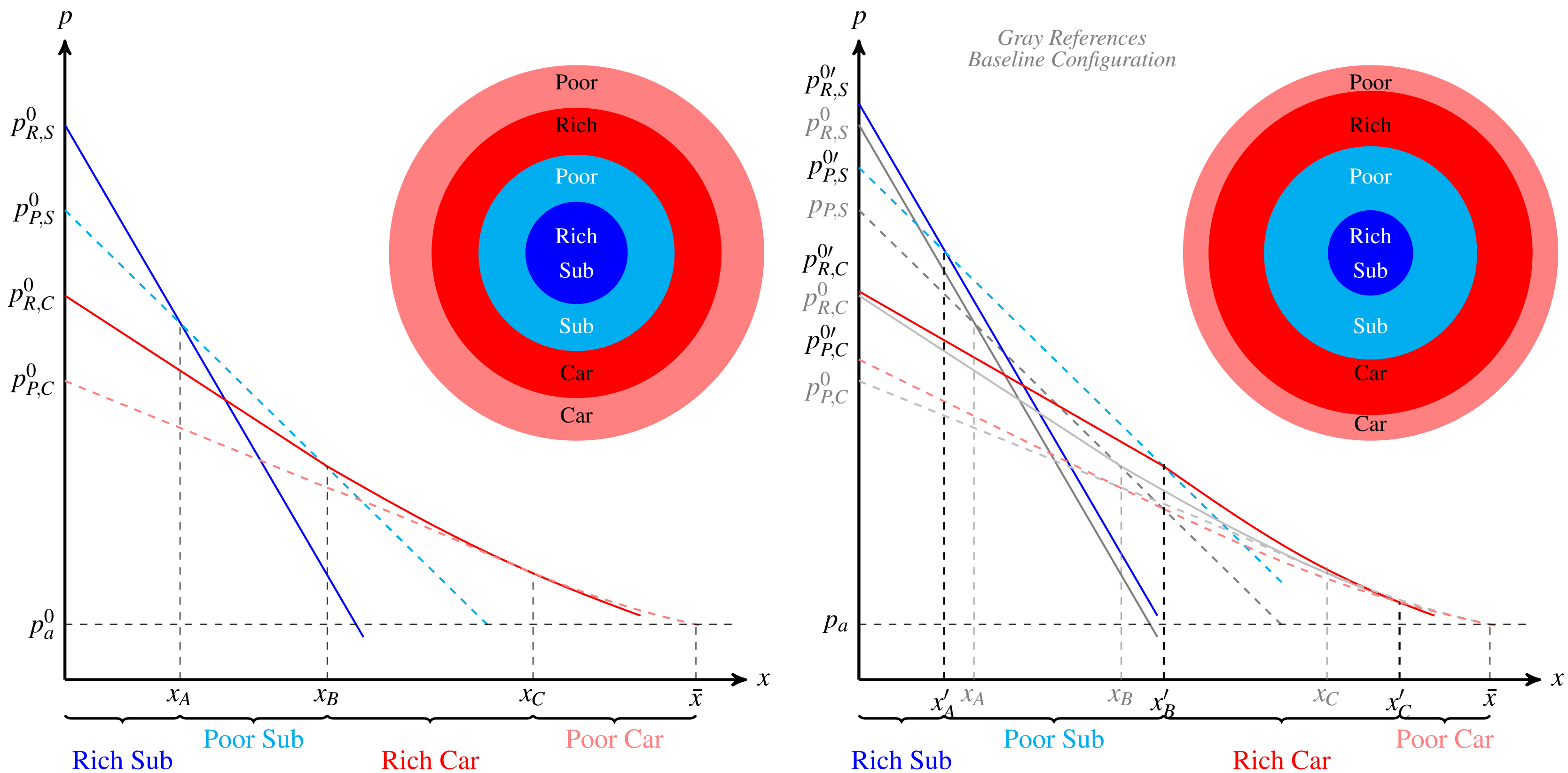


Note: The x-axis denotes the distance to city center and the y-axis is the rental price. Concentric circles show the equilibrium pattern of housing location where the linear city is drawn as radially symmetric. Commuters are in two income groups (Rich, R, indicated with solid lines, and Poor P, indicated with dashed lines) and choose from two modes: car (C, in red) and subway (S, in blue). The model is calibrated to generate an urban configuration as indicated. Bid-rent curves are steepest for the rich because of their higher value of time. Bid-rent curves are flatter for car commuting because the variable time cost is assumed to be smaller. Values indicated by $p^0_{d,m}$ are the value of bid-rent curves at the origin, while those with a $\prime$. Gray curves reference the baseline configure in panel (a) of Figure A15. Here we consider the effect of an exogenous increase in congestion through a decrease in capacity throughout the city. This steeps the bid rent curves for driving, for the rich more than the poor. It also changes the slope of the subway bid-rent curves by an increase in the fixed demand for housing for the rich subway commuting and a small decrease in the fixed demand of housing for poor subway commuters. The result is to shift more of the rich to subway commuting and small share of the poor are induced to car commuting given changes in housing prices.

Figure A15: Spatial Equilibrium with Two Modes & Income Heterogeneity: Baseline & Congestion Pricing

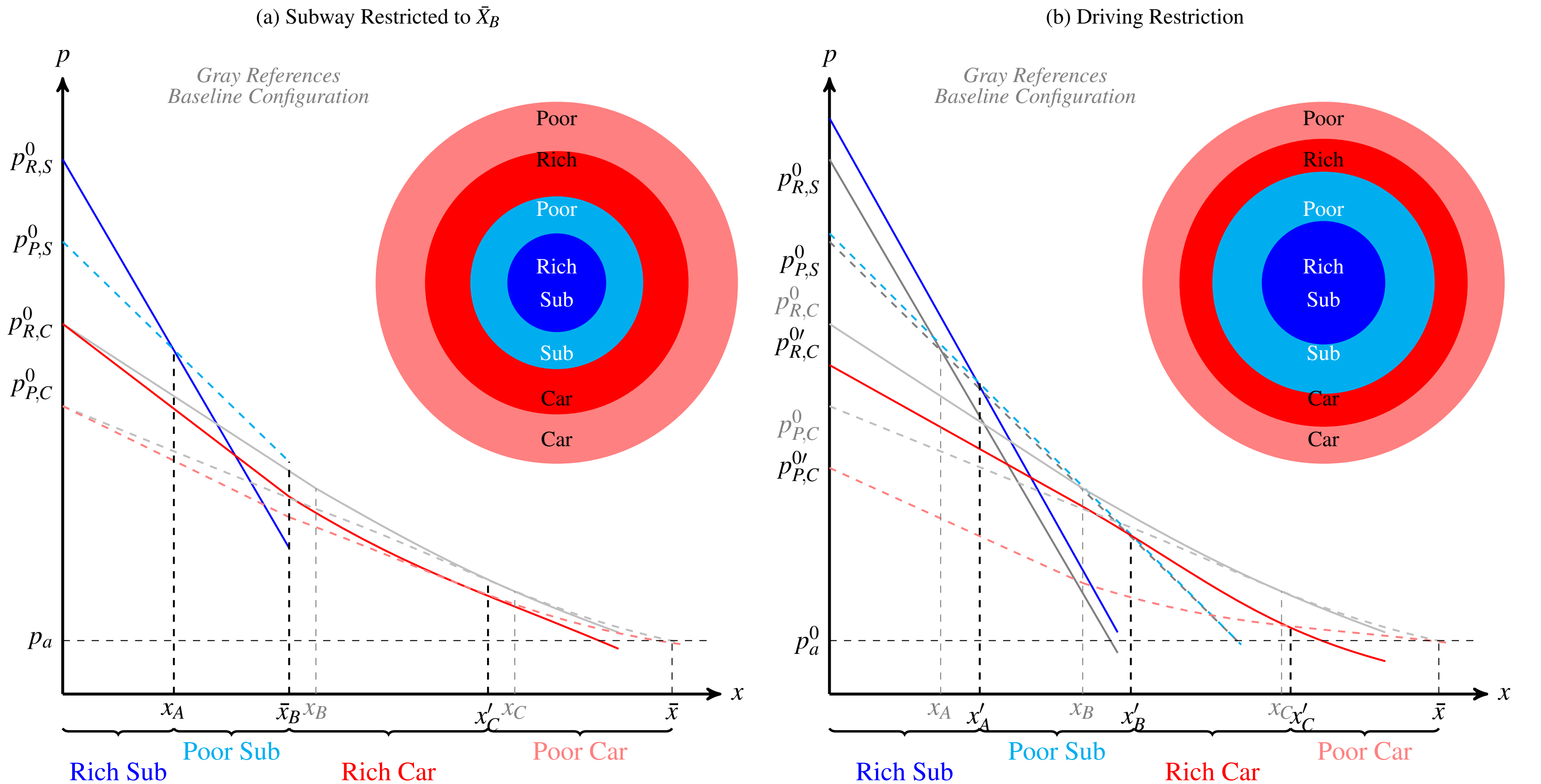
(a) Baseline Configuration

(b) Congestion Pricing



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

Figure A16: Spatial Equilibrium with Two Modes & Income Heterogeneity: Subway & Driving Restriction



Note: The x-axis denotes the distance to city center and the y-axis is the rental price. Concentric circles show the equilibrium pattern of housing location where the linear city is drawn as radially symmetric. Commuters are in two income groups (Rich, R, indicated with solid lines, and Poor P, indicated with dashed lines) and choose from two modes: car (C, in red) and subway (S, in blue). The model is calibrated to generate an urban configuration as indicated. Bid-rent curves are steepest for the rich because of their higher value of time. Bid-rent curves are flatter for car commuting because the variable time cost is assumed to be smaller. Values indicated by $p^0_{d,m}$ are the value of bid-rent curves at the origin, while $p^{0\prime}_{d,m}$, are the new values at the origin after the policy change. Gray curves reference the baseline configure in panel (a) of Figure A15. Panel (a) shows the effect of restricting the subway network to $\bar{x}$. Comparing this panel to panel (a) of Figure A15 shows the effect of expanding the subway network from $\bar{x}_B$ to x_B. There is no effect of expansion on the bid rents for subway except that they do not extend beyond $\bar{x}_B$ in panel (a) of Figure A15 because the subway has not been built that far. Because expansion allows a larger share of commuters to use the subway, here only from the poor, it induces lower congestion, flattening out the bid-rent curves for driving, for the rich more than the poor because of value of time differences. Panel (b) considers a driving restriction policy, which induces increases in both the fixed and variable costs of driving not remitted to households (unlike congestion pricing). Car commuters will need to incur the fixed cost of both driving and and using subway, and when they cannot drive, they will have to use subway which has a higher variable cost due to time. The bid rent curves for car commuters move down and become steeper because of the change in fixed cost and variable cost of commuting, respectively. The increase in commuting costs is larger for the rich due to their high VOT than for the poor, leading to a larger movement of the rich away from driving to subway.

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

	(1)	(2)	(3)	(4)	(5)	(6)
ln(Subway Distance)	-0.111***	-0.005	-0.004	-0.000	-0.001	0.003
	(0.015)	(0.016)	(0.015)	(0.016)	(0.015)	(0.017)
ln(Subway Distance) × CDR	-0.002	-0.018*	-0.019**	-0.023*	-0.019**	-0.023*
	(0.013)	(0.009)	(0.009)	(0.013)	(0.009)	(0.013)
ln(Subway Density)					0.004	0.007
					(0.007)	(0.008)
YearxMonth FE	Y	Y	Y	Y	Y	Y
neighborhood FE	N	Y	Y	Y	Y	Y
Complex-level Controls	N	N	Y	Y	Y	Y
Weighted	N	N	N	Y	N	Y
Observations	9640	9634	9634	9634	9634	9634
Adjusted R^2	0.236	0.522	0.534	0.688	0.534	0.688

Note: Sample spans 24 months before and after the car driving restriction policy (CDR). The dependent variable is log(price per m^2). Subway distance is the distance (in km) from the housing unit to the nearest subway station. Subway density is constructed at the TAZ level as the inverse distance weighted number of subway stations from the centroid of an TAZ. Standard errors clustered at the neighborhood level are in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Weights are constructed based on a data set of housing translations which accounted for about 17% of total housing transactions in Beijing during 2006 to 2014.

Table A2: Effect of CDR on Household Sorting

	ln(Distance to Subway)			ln(Distance to work)		
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Household Income)	0.041*	0.026	0.012	0.203***	0.227***	0.214***
	(0.021)	(0.020)	(0.027)	(0.039)	(0.048)	(0.044)
Ln(Household Income)× CDR	-0.039*	-0.041*	-0.017	-0.122***	-0.109**	-0.125***
	(0.024)	(0.023)	(0.027)	(0.042)	(0.042)	(0.046)
YearxMonth FE	Y	Y	Y	Y	Y	Y
neighborhood FE	Y	Y	Y	Y	Y	Y
Complex-level Controls	N	Y	Y	N	Y	Y
Household Demographics	N	Y	Y	N	Y	Y
Weighted	N	N	Y	N	N	Y
Observations	9634	9634	9634	9634	9634	9634
Adjusted R^2	0.833	0.837	0.790	0.229	0.254	0.213

Note: Sample spans 24 months before and after CDR. Standard errors clustered at the neighborhood level are in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Weights are constructed based on a data set of housing translations which accounted for about 17% of total housing transactions in Beijing during 2006 to 2014.

Table A3: The Elasticity of Speed w.r.t. Traffic Density

	(1)	(2)	(3)	(4)	(5)	(6)
ln(density)	0.032	-0.173***	-0.683***	-1.018***	-1.099***	-0.620***
	(0.026)	(0.037)	(0.076)	(0.066)	(0.089)	(0.030)
Observations	393,634	386,717	412,556	243,302	156,670	1,592,879
Density (cars/lane-km)	< 8	≥ 8 & < 14	≥ 14 & < 23	≥ 23 & < 35	≥ 35	All
Average speed (km/h)	65.3	70.9	63.8	50.0	30.3	60.5

Note: Each column reports results from a 2SLS regression where the dependent variable is ln(speed in km/h) and the key explanatory variable is log(traffic density in the number of cars/lane-km). The unit of observation is road segment by hour during peak hours within 6th ring roads in 2014. The first five columns are based on the observations in each of the quintiles: column (1) is for observations in the first quintile of the density distribution, and column (5) is for observations in the fifth quintile. The last column is for all observations. The IVs are constructed based on the driving restriction policy which has a preset rotation schedule for restricting certain vehicles from driving one day per week based on the last digit of the license plate number. We construct a policy indicator being 1 for the days when vehicles with a license number ending 4 or 9 are restricted from driving. We interact this variable with ring road and hour-of-day dummies. The control variables include temperature (C^o), wind speed (km/h), visibility (km), dummies for (16) wind directions, and dummies for (5) sky coverage at the hourly level. See Yang et al. (2019) for details on data construction and the variables. The time fixed effects include day-of-week, month-of-year, hour-of-day, holiday fixed effects. The spatial fixed effects include road segments (or monitoring stations) fixed effects, and the interactions between ring road dummies (e.g., inside 2nd ring roads, between 2nd and 3rd ring roads) with hour-of-day (Ring roads $\times$Hour). Parentheses contain standard errors clustered by road segments. Significance: *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

Table A4: Sorting Model Estimates Without EV Terms - Second Stage

Variables	OLS (1)	OLS (2)	IV1 (3)	IV1+IV2 (4)	IV3 (5)	All (6)
Price (in 1 million RMB)	-2.073***	-2.062***	-4.224***	-5.031***	-7.101***	-5.356***
	(0.180)	(0.176)	(0.535)	(0.432)	(1.646)	(0.418)
Ln(home size)	-3.590***	-3.657***	0.104	1.512**	5.102*	2.079***
	(0.248)	(0.251)	(0.932)	(0.761)	(2.937)	(0.765)
Building age	-0.032***	-0.026***	-0.076***	-0.095***	-0.144***	-0.103***
	(0.006)	(0.006)	(0.012)	(0.010)	(0.040)	(0.010)
Floor to Area Ratio	0.017	0.001	-0.007	-0.009	-0.009	-0.009
	(0.031)	(0.023)	(0.021)	(0.025)	(0.036)	(0.027)
Ln(dist. to park)	0.167***	0.052	-0.196***	-0.285***	-0.513**	-0.321***
	(0.059)	(0.054)	(0.073)	(0.075)	(0.225)	(0.081)
Ln(dist. to key school)	0.631***	0.555***	0.312***	0.223**	-0.034	0.187**
	(0.060)	(0.091)	(0.086)	(0.089)	(0.213)	(0.091)
Year-Month-District FE	Y	Y	Y	Y	Y	Y
Neighborhood FE		Y	Y	Y	Y	Y
First-stage F			10.48	14.22	9.88	14.22
J Statistics			7.62	8.30		15.75
P value: overidentification test						0.01
Avg. Price elasticity	3.09	3.10	0.94	0.13	-1.94	-0.19

Note: The number of observations is 79,894. The dependent variable is the mean utilities recovered from the first stage (without including EV in the first stage). The first two columns are from OLS and the last four are from IVs. The floor-area ratio, a measure of complex density, is the size of the total floor area over the size of the parcel that the complex is located on. Distance to key school is the distance to the nearest key elementary school. Column (3) and (4) use IV1 as price instruments, i.e. the average attributes of homes (building size, age, log distance to park, and log distance to key school) that are within 3km outside the same complex sold in a two-month time window from a given home. Column (4) and (6) additionally use IV2, i.e. the interaction between the distance related instruments defined in Column (3) and the winning odds of the vehicle licence lottery as instruments. The winning odds decreased from 9.4% in Jan. 2011 to 0.7% by the end of 2014. Column (5) uses the number of homes transacted in the three-month window in the real estate listings dataset. Standard errors clustered at the neighborhood-year level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Model Fit: Effect of CDR on Price Gradient

	(1)	(2)	(3)
ln(Subway Distance)	-0.175***	0.011	
	(0.015)	(0.036)	
ln(Subway Distance) × CDR	-0.010***	-0.010***	-0.010***
	(0.002)	(0.002)	(0.002)
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 homes sold in 2014 in the mortgage data with 7,136 observations. We simulate the equilibrium prices under the 2008 network for two scenarios: with and without the car driving restriction (CDR). We then estimate regressions as in Table A1. The dependent variable is log(price per m^2). Subway distance is the observed distance (in km) from the housing unit to the nearest subway station (based on 2008 network). Standard errors clustered at the neighborhood level are in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The driving restriction steepens the price gradient with respect to subway access consistent with the result in Table A1 based on the observed data.

Table A6: Simulation Results: without Sorting

	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)	
Household Income Relative to Median	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low
Panel A: Travel outcomes												
Drive	41.65	21.44	-6.02	-2.60	-2.91	-4.56	-1.89	-1.49	-7.22	-3.71	-4.40	-5.48
Subway	9.02	10.77	1.05	0.69	0.77	0.93	4.02	5.55	4.95	5.91	4.67	6.25
Bus	22.44	30.47	1.49	0.26	0.48	0.98	-1.48	-2.43	0.02	-1.82	-0.84	-1.16
Bike	15.96	24.01	1.54	0.62	0.63	1.49	-0.68	-1.50	0.66	-0.93	-0.18	-0.23
Taxi	2.20	1.32	0.71	0.34	0.50	0.45	-0.11	-0.07	0.47	0.21	0.32	0.30
Walk	8.74	11.99	1.23	0.69	0.54	0.71	0.14	-0.07	1.11	0.33	0.43	0.32
Speed	21.49		3.12		2.95		1.76		4.73		4.68	
Panel B: Housing market outcomes												
Borrower's distance to work (km)	19.45	18.88										
Co-borrower's distance to work (km)	17.54	11.95										
Distance to subway (km)	5.33	4.30					-4.13	-3.45	-4.13	-3.45	-4.13	-3.45
Panel C: Welfare analysis per household (thousand ¥)												
Consumer surplus (+)			-165.2	-18.0	-87.5	-69.7	207.7	117.0	30.9	95.5	108.2	50.1
Toll revenue (+)					116.9	116.9					107.8	107.8
Subway cost (–)							103.0	103.0	103.0	103.0	103.0	103.0
Net welfare			-165.2	-18.0	29.4	47.2	104.7	14.0	-72.1	-7.5	113.0	54.9

Note: Simulated results based on estimated model parameters in Column (6) of Table 3 and Column (6) of Table 5. The simulation shows counterfactual results for 2014 sample households and homes. The detailed simulation procedure can be found in F. This table shows results without sorting, hence households' housing choice decisions are fixed but we allow travel mode decisions to adjust to clear traffic market. Column (1) shows the baseline results while columns (2) to (6) show the differences from column (1). The driving restriction prohibits driving in one of five work days. Congestion pricing is ¥0.92 per km to generate the same reduction as the driving restriction without sorting in Table 6. High-income household are those with income above the median household income. Subway cost per household includes the construction cost and the 30-year operating cost equally shared among 7.2 million households. We apportion 100% it to work commute in our welfare analysis. Toll revenue is the revenue per household from congestion pricing during a 30-year period (to keep a balanced government budget, the toll revenue is recycled uniformly to each household). Net welfare is consumer welfare per household after revenue recycling or tax-funded subway construction and operation.