SERC DISCUSSION PAPER 116

Does Public Investment Spur the Land Market?: Evidence from Transport Improvement in Beijing

Wenjie Wu (Department of Geography & Environment, London School of Economics and Political Science)

July 2012

Revised April 2014

This work is part of the research programme of the independent UK Spatial Economics Research Centre funded by a grant from the Economic and Social Research Council (ESRC), Department for Business, Innovation & Skills (BIS) and the Welsh Assembly Government. The support of the funders is acknowledged. The views expressed are those of the authors and do not represent the views of the funders.

Does Public Investment Spur the Land Market?: Evidence from Transport Improvement in Beijing

Wenjie Wu*

July 2012

Revised April 2014

* Department of Geography and Environment, London School of Economics and Political Science

Acknowledgements
I thank seminar participants at SERC and the European Real Estate Society Conference (2012), and I thank Steve Gibbons, Paul Cheshire, Olmo Silva, Henry Overman, Simona Iammarino, Guanpeng Dong, and Yaping Wang for helpful comments.

Abstract
Over 140 billion CNY (1GBP=10CNY) has been spent between 2000 and 2012 in Beijing on the construction of new rail transit lines. This massive public investment allows me to examine the consequences of transport improvements for land prices near rail stations. Using unique vacant parcel-specific data, I estimate the significant heterogeneity in the capitalization effects of rail transit development for multiple land uses in Beijing urbanised area. The results show that these transport improvements, identified by the parcel-station distance reductions, give rise to sizeable price premiums in the local residential and commercial land markets. Strikingly, the difference between the increase in the value of residential and commercial land parcels are not distributed evenly. These findings lend to support the evidence that public investment has an essential role to play in spurring the spatially targeted land market and provide implications for further land and transport policy making in China.

JEL Classifications: H41, Q51, R41
Keywords: Land prices; transport improvement; Geographical Information System; China

1 Introduction

Improvement of transport infrastructure is of crucial importance to shaping urban development (World Bank 2008; Adams and Tiesdell 2013). Proponents argue that this transit mode forms the backbone of commuting networks, and promotes urban sustainability and economic viability (Baum-Snow and Kahn 2000; Kahn 2007; Wu et al. 2009). This is a particular concern in countries like China, where there were over 800 billion RMB (1USD≅6.5 RMB) investment in new rail transit constructions over the past decade. Clearly the supply of new rail infrastructure would increase station proximity and reduce commuting time to do work and leisure activities, leading a number of economic benefits: urban growth (Duranton and Turner 2012), sub-urbanisation and commuting patterns (Baum-Snow 2010), and firm productivity (Gibbons et al. 2012).

Despite the importance of public investments, little is known about the overall impact of transport infrastructure investment on economic output in China, and even less is known about the *ex post* effects of multiple transport improvement programs within cities. Two central identification problems are difficult to overcome. First, public resources might be endogenous to local economic outcomes. Variations in transport investments could be confounded with other socioeconomic factors that determine the outcomes. Second, even the standard cost-benefit transport appraisals based on fixed travel time savings may miss benefits that are not visible in measured economic output (Gunn 2000).

Land markets can be used to mostly overcome challenges of valuing rail access (Cheshire and Sheppard 1995; Gibbons and Machin 2005). If real estate developers overvalue a land parcel due to increased rail access they will pay to bid it, spending increases will then lead to increase land prices. This means that if land markets are efficient, land prices will reflect *all* the benefits that a place offers in terms of station proximity. Indeed, there is a substantial literature that used cross-sectional hedonic models to document the popular narratives about the straightforward positive or negative proximity effects on property values. But this cross-sectional approach does not avoid the challenge of obtaining causal effects, especially when new stations are opened or planning to open (Gibbons and Machin 2008).

In this paper we implement a novel research design that uses transport improvement programs as a plausibly exogenous change. We define the treatment as land parcels that have experienced the station-distance reductions due to the building of new rail stations. Although treated parcels that near new stations are likely to differ in both observable and unobservable ways from those that do not, these differences can be minimized by focusing on very close station areas. Thus, a difference-in-difference (DiD) framework is the preferred approach to identify the impacts of transport improvements on land prices near station areas.

Only a handful of papers have used transport improvement programs as sources of identification in hedonic models. Most existing studies, employing before-and-after comparisons, have focused on examining the property price effects of new rail transit lines in the US and UK cities (Davis 1970; Bajic 1983; McMillen and McDonald

2004; Baum- Snow & Kahn, 2005). Recent studies, though less common, have exploited changes in the distances between properties and stations as a result of new stations opened and estimated such impacts on property prices. For example, Gibbons and Machin (2005) developed a precise framework for capturing the changes in distances between houses and tube stations in London when the Jubilee line and Docklands Light Railway (DLR) opened in the late 1990s. They highlighted the fact that difference-in-difference estimates can better avoid the biases inherent in pure cross-sectional empirical studies. In a similar vein, Billings (2011) examined the commercial property price effects of transport network extensions using the difference-in--difference approach.

Our analysis is complicated by the presence of “dynamics” in the treatment effects due to the multiple transport improvement events. We propose the “spatial multiple intervention DiD” estimation strategy that exploits explicit land price dynamics in treated places, and bringing the identification power of a DiD design into economies with fast developing transport networks. As a further extension, we focus on not just how residential land prices causally respond to the station distance changes, but also the effects on commercial land uses. We also allow the proximity effect of rail stations to vary on local demographics that believed to influence the amenity value of rail access (Bowes and Ihlanfeldt 2001; Cheshire and Sheppard 2004; Gibbons 2004).

We apply our estimators to a unique micro-geographic data set combining information on vacant land parcel transaction and demographic characteristics from 1999-2009 in the entire urbanized area of Beijing, rather than pre-designed sample areas. We focus on Beijing because it provides a unique institutional setting, where there is an emerging land market system and dramatic changes in transport infrastructure. It is important to emphasize that unlike in the US, public infrastructure construction are highly centralized and controlled by the Beijing central government. The local communities (zone, *jiedao*) are only responsible for street cleaning and do not have control over public infrastructure construction and service provision. We therefore interpret the impact of increased rail access on land prices as reflecting the effects of transport infrastructure investment.

Our results suggest that new transport systems in Beijing have a big average effect on local land prices, but these proximity effects decay in a possible non-linear trend over space. For example, in 2008 a new rail station added 2.02-4.20% on residential land prices, for areas 1-4km from the station. Our preferred estimates indicate that commercial land parcels would accrue greater benefit than residential land parcels at a closer distance range from a station---by gathering large population flows and high demand for commercial activities. Importantly, shifts in local land values also vary widely according to local demographic characteristics. For example, the value of proximity to new stations rises substantially with employment accessibility and educational attainment level. Our results also suggest complementary effects between public investment and private sector investment, as higher levels of economic activity should translate into higher future tax receipts. More speculatively, it suggests that improving local people's happiness (Wu 2014), alongside physical development may gentrify the neighbourhood dynamics.

This paper proceeds as follows: Section 2 describes the institutional settings and data. Section 3 presents the theoretical framework and econometric models. Section 4 reports the results, and Section 5 concludes.

2 Institutional Settings and Data

2.1 Land development

China's urban transition and land development have been extensively studied. Since the 1978 Reform-and-Opening-up policy in China, tremendous changes had happened in this booming economy, from a central-planned economy towards a market-oriented economy. Within this context, urban real estate markets were reborn in the recent two decades (Wang and Murie 2000; Zheng and Kahn, 2013). In 1988, the Chinese Constitution---which had prohibited land transfers before, was amended to permit land leasing rights while retaining land ownership. In 1990, the State Council formally affirmed such dramatic transformation of the land use system from free allocation toward a leasehold system. By 1992, local government in Beijing had begun to practice the land leasing policy, and it quickly spread to other cities in China.

In Beijing, the Municipal Land Resource Authority is responsible for the land allocations and sales of leasehold right, first through negotiation between developers and governments (during 1992 and 1998), then through partly negotiation and partly competitively open auction (during 1999 and 2003), and recently through the full competitively open auction way (since 2004). This research obtained vacant land transaction data from a database maintained by the Beijing Municipal Land Resource Authority. For each land parcel, the data include the land uses, prices, locations, lot sizes of transacted land parcels from 1999 to 2009.

To generate trustable estimates, we excluded uncompleted land transaction data and parcels that were obtained through negotiation because the strong institutional forces could reduce the market price effectiveness (Cai et al. 2009). The final sample size is 2343 and 1341 parcels[1] for residential and commercial land uses respectively. We used the Geographical Information System (GIS) software to assign parcels to geographic locations, and calculated the straight-line distance from each parcel to the nearest station.

To implement the transport improvement analysis, we group the parcel-level land data into three time periods: before 2003 ($\text{year} < 2003$); during 2003 and 2008 ($2003 \leq \text{year} < 2008$); after 2008 ($\text{year} \geq 2008$). We then define the treatment group by using two selection principles. A land parcel will be assigned to a treatment group if: **Criteria 1:** It experienced the station-distance reductions with the stations opening after 2003 $(station \geq 2003)$; **Criteria 2:** And if the outcome distance to the closest station opening after 2003 is now less than the 2km distance band.

We impose the second criteria because we want to avoid the estimating noise from the parcels that became closer to a station, but still remain a long distance away from the new station. Notably, the choice of a 2 km threshold is based on most existing empirical literature as well as a reasonable walking distance to a station (about 20 minutes).

[1] To mitigate the inflation effect, we have adjusted the land prices by using the CPI index. All monetary figures are constant in 2009 RMB. Also, we have trimmed the land price distribution by only keeping parcels in each year whose price is between the 5th and 95th percentiles of the whole sample price distribution.

Instead of using one fixed distance band, we extend to the literature by using flexible distance bands (1km, 2km, 4km) to select treated parcels. This research design would then allow us to compare the treatment effects using alternative parcel-station distance bands. Such comparison would hold everything the same in the model specification and any changes in land prices would be attributable to the difference in the selection of distance bands. As such, we are able to test the marginal effects of each distance band relative to the larger one[2].

Following the same principles, we further create the treatment groups of $(1km\_station \geq 2008/2009)$, $(2km\_station \geq 2008/2009)$, $(4km\_station \geq 2008/2009)$ when a parcel has been experienced the station-distance reductions with the stations opening after 2008/2009; and the outcome distance to the closest station opening after 2008/2009 is now less than 1km, 2km, 4km respectively. Of necessity, the treatment groups of $(station \geq 2009)$ are nested within the corresponding treatment groups of $(station \geq 2008)$, and the treatment groups of $(station \geq 2008)$ are nested within the corresponding treatment group of $(station \geq 2003)$.

Figures 1-2 show the spatial distributions of treated residential and commercial land parcels respectively. From the GIS map it can be seen a clear spatial differentiation pattern among parcels in the treatment groups of $(station \geq 2009)$, $(station \geq 2008)$ and $(station \geq 2003)$, which gives some confidences that our results

2 Recall that this study simplified the analysis through the use of the flexible distance bands. But it is important to note that there is a long tradition of the land value gradient, suggesting that land prices decline smoothly with distance from a new station and therefore the actual treatment effects are not discrete across either time or space.

are not sensitive to the potential spillover effects within-treatment groups. Below, we will examine the spillover effects both within and across treatment groups in the robustness section.

Geographical information on other localized characteristics is taken from a variety of sources for the use of controllable variables in the regression models (see Appendix Table 1). The local public goods were built long ago in the central-planning economy and seldom change their locations after they are built. Thus, one advantage of using these local public goods as a set of controllable variables is that the location of public goods (such as schools, parks) is exogenously determined in Beijing. School location and quality comes from the Beijing Municipal Committee of Education. The location of bus stops and expressways are used as proxies for the competing commuting modes, and is obtained by a web-based search from the Beijing Municipal Committee of Transport. Parks and green spaces are important environmental amenity that may influence land values (Wu and Zhang 2009; Wu and Dong 2014). Location data of parks in Beijing is taken from the Beijing Municipal Garden Bureau. Air quality is another key indicator of the environmental amenity (Zheng et al. 2013), which can be measured by the recorded air pollution index (API) from the Beijing Municipal Environmental Protection Bureau. Crime rates for the number of violent crimes taking place in each zone are obtained from the Beijing Public Security and Safety Bureau. The 2000 City Population Census reports the basic socio-demographic characteristics such as the population density, resident median education attainment levels, public housing rent ratio, and the percentage of old housings built before 1949.

The 2001 City Employment Census provides the necessary information for calculating the employment accessibility[3].

2.2 Transport infrastructure improvement

Decades of fast economic growth and urbanisation have significantly changed the transport infrastructure in China. Like other large cities in the BRICS countries[4], Beijing is investing heavily in new rail transit constructions, a largely place-based investment process that is of great importance for planners, land developers and policymakers. This section provides the political economy backgrounds of the new rail infrastructure development in Beijing.

Over 140 billion RMB have been spent between 2000 and 2012 in Beijing on the building of new rail transit lines. Table 1 highlights that the constructions of rail transit lines differ with respect to their starting time and completion date[5]. This table also provides differential figures of each line with respect to the construction cost, track length, and station numbers. Figure 3 shows the spatial patterns of the Beijing rail transit network before and after the completion of these new rail transit lines.

---

[3] The employment accessibility is measured by using the non-parametric gravity model (see Ding et al. 2010).

[4] BRICS is the title of an association of leading emerging economies (Brazil, Russia, India, China and South Africa). See http://en.wikipedia.org/wiki/BRICS for details.

[5] The Beijing Municipal Committee of Transport's official website http://www.bjjtw.gov.cn/ contains informative details of subway lines in Beijing. Anticipated announcement effects would give us more insights on the ways that land markets capitalize changes in amenities. However, there were many versions of subway development plans in Beijing. The first version of the plan had been announced in the late 1980s (under this plan, most of the current new subway lines should have been built in the 1990s), but nothing happened until the early 2000s. Meanwhile, the land and housing markets have not been fully transformed from the socialist welfare system into the market-pricing system until the early 2000s. In the robustness result not reported, we do test the anticipation effects for subway stations (i.e. Line 14 and 16) that had been announced but the exact completion time would be no early than 2015. The insignificant coefficients of the results suggest that it is difficult-to-measure the announcement effects of new subway lines given data limits and the large uncertainties associated with the proposed timetable.

Indeed, these new subway lines were considered as the most significant improvement in the Beijing subway network since the 1960s. It is expected that these transport improvements have altered the parcel-station distance for some places, whilst left others unaffected. This provides us with a plausibly exogenous change, from which we can examine the impact of rail access changes on land prices nearby new stations.

To investigate the motivations behind government's infrastructure investment decisions, we searched historical planning documents and interviewed local officials. There are three strategic reasons. The first is to mitigate road traffic congestion and meet the rapid growth of the commuting demand in the central city areas where there are few established stations. For instance, a recent internal report by Beijing Municipal Commission of Urban Planning has clarified that subway line 6 and line 7 are constructed to handle the ridership growth of subway line 1 and the road congestion around the CBD areas. Second, improvements of rail infrastructure are often mentioned as an effective policy lever to gentrify the deprived areas. As such most of the new subway lines are aimed to link the central city with the depressed areas and suburbs with large "bedroom" communities[6] (*Tiantongyuan, Yizhuang, Daxing*, and *Tongzhou*). Third, the Beijing municipal government has decided to built one short subway line (Line 8A, with only four stations) to connect the Olympic park with the main rail transit network. Below, we will control the interactions of time trends with distance to CBD, distance to Olympic Park, and distance to "bedroom"

---

[6] Note that the term of "bedroom communities" represents places where commuters perform most professional and personal activities in another location, maintaining their residence solely as a place to sleep. See http://en.wikipedia.org/wiki/Commuter_town for details.

communities (*Distance to New Residential Area*$_i$) that can help to affirm the robustness of the proximity effects on prices of nearby land parcels.

It is noteworthy that we use the opening of two lines in 2003, four lines in 2008, and eight planned lines opening after 2009 (to be completed before 2012) as the transport improvement programs. Ideally, we could single out the effects of each of these new lines and even go further by measuring each new station's effect individually. Yet in reality, we simplify the estimation framework by treating them as three nested events (stations opening *after 2003, after 2008* and *after 2009*).

3 Theoretical Framework and Estimation

3.1 Theoretical framework

Economic researchers, planners and policymakers often cite long travel time, low mobility as problems that can interfere with household living convenience---the main claims of advocacy groups for improving transport infrastructure. However, household living convenience is not the only potential benefit of improved transport infrastructure. Transport investments may also lead to enhancements in local safety, pollution and the environmental and aesthetic appeal of station areas. A full evaluation of place-based investment decisions must capture all of these potential impacts. But rather than investigating each outcome separately, one can use land developers' bid decisions to identify their revealed preferences using the hedonic methods. The working assumption here is that any shift in the desirability of a land parcel—along different dimensions—would be reflected in equilibrium land prices.

We develop a simple theoretical model to support our empirical strategy [7]. In light of recent literature, we assume that the utility of homeowner i living in community j depends on local transport accessibility A_j, amenities Z_j, and other consumption C_i: $U_{ij} = U_i(A_j , Z_j , C_i)$. The household has income (I_i) and faces the budget constraint $C_i \leq I_i - T_j - P_j$, where T_j represents property taxes and P_j is the housing price. Local public goods accessibility depends on tax revenues if it is provided by the community j, $A_j = A(T_j)$.

We consider first the household location decision with predetermined capital spending in transport infrastructure. To maximize utility, a household chooses the community to maximize utility, conditional upon housing prices, and public goods accessibility (i.e. rail stations): $U (A(T_j), Z_j , I_i - T_j - P_j)$. This simplified theorem yields the household's bid for housing in community j as an implicit function of local amenities and taxes: $f_{ij} = f_i(Z_j, T_j)$. Holding prices, amenities, and all other factors constant, community j will provide higher utility than any alternative community if $P_j < f_{ij}$. The household's WTP (willingness-to-pay) for a marginal increase in T_j in its chosen district can be written as:

$$\partial f_i(Z_j, T_j)/\partial T_j = (\partial f/\partial C)^{-1}[A(T_j)\cdot(\partial U/\partial A)] - 1 \qquad (1)$$

The WTP will be positive if the marginal product of tax revenues multiplied by the marginal utility of transport accessibility (in brackets) exceeds the marginal utility of consumption. Under the equilibrium assumption, the housing price in community j, $P_j = P*(Z_j, T_j)$, equals the bid of the marginal consumer, who must be indifferent

[7] See the seminal work by Rosen (1974), and see a recent survey by Gibbs and Pryce (2012) about future directions for housing economics. We adapt the classic scenario to China (Zheng et al. 2009).

between this community and other alternatives. As such P_j will respond positively to increases in T_j when the transport investment was lower than the preferred level of the marginal household.

In China, homeowners do not pay property tax (Zheng and Kahn 2008; Wu et al. 2013). Thus, household budget constraint can be simplified as $C_i \leq I_i - P_j$. In addition, as mentioned above, the urban land supply and the provision of public facilities like rail stations are highly controlled by the central government, rather than by local communities. Thus, the local amenity capitalization effects are expected to be more straightforward than in US and European cities with property taxes (Gyourko et al. 1999), because real estate developers implicitly purchase local public goods through buying land parcels. This is especially the case in the context of transport improvements, where positive effects on land prices capture benefits from better accessibility[8].

3.2 Econometric models

Using a rich geographically-coded dataset, this study estimates the effects of increased station proximity on land prices in Beijing[9]. The baseline equation for our analysis is expressed as follows[10]:

$$Ln\,Price_{ilt} = \beta_0 + \sum_{j=1}^{3} \beta_j Lndist_{it} + \sum_{t=1}^{3} \beta_t Y_t + \beta_k X_{ilk} + f_l + \varepsilon \qquad (2)$$

Where $Price_{ilt}$ represents the price of vacant residential or commercial land

[8] It is important to note that this paper does not provide a unified framework for linking sorting with institutional roles (Gibbs and Pryce 2012).

[9] Recall that our analysis does not attempt to account for the impact of financial changes on the real estate markets (Deng and Liu 2009; Jackson and Orr 2011).

[10] In this study, we have tried estimating flexible-form models with Box–Cox transformation but could not reject a strong log–log relationship between land prices and key explanatory variables.

parcel i located at area l in the period t; $dist_{it}$ is the distance to the nearest station; X_{ilk} is a matrix of land structural and localized characteristics; Y_t presents the time trend effects; f_l indicates area-specific fixed effect; ε is a random error term[11]. Other Greek letters are parameters to be estimated.

This traditional cross-sectional approach is highly successful at capturing long-run relationships between land prices and rail access, but may not recover the impact of increased station proximity on local prices before and after a change in transport improvement policy. To explicitly account for this, we adopt a conceptually more attractive approach. By focusing on what happens after the transport improvement, in places affected and unaffected by the change, we can more reliably assess the new rail transit's impact on local land prices.

To achieve this, we need data on land price changes and rail station access changes. In contrast to the systemic repeated sales data and limited transport infrastructure changes in the developed countries, it is easy to observe an opposite scenario in China: an emerging land market system since the 1990s and the rapid urban rail transit development. The first data requirement is met by using a 1999-2009 cross-sectional land parcel transaction data. The second data requirement is easier to meet because of the recent dramatic changes in public transport infrastructure in Beijing. The supply of new rail transit stations increased over time---two subway lines were opened in 2003, four lines were opened around 2008 and another eight lines were planned to open after 2009. These improvements will lead to the increased

[11] Standard errors are clustered at the zone level to allow for heteroscedasticity and spatial-temporal correlation in the error structure within zones.

proximity to stations for a series of subset of land parcels in our data set after 2003, after 2008, and after 2009 respectively. This means that we can, in principle, estimate the increased station proximity effect in the multi-nested treatment scenarios[12]. The outcome regression equation becomes:

$$Ln\,\mathrm{Pr}\,ice_{iljt} = \beta_0 + \sum_{j=1}^{3}\beta_j Treatment_j + \sum_{t=1}^{3}\beta_t Period_t + \sum_{j=1}^{3}\sum_{t=1}^{3}\beta_{jt} Treatment_j * Period_t + \beta_k X_{ilk} + f_l + \varepsilon \quad (3)$$

In this equation, $Treatment_j$ refers to a specific treatment group (e.g. $(station \geq 2003)$, $(station \geq 2008)$, $(station \geq 2009)$). $Period_t$ is a set of "policy-on" time dummy variables ($(1999 \leq year < 2003)$, $(2003 \leq year < 2008)$, $(year \geq 2008)$). The coefficients β_{jt} then show the various treatment effects ($Treatment_j * Period_t$) in different periods[13]. Table 2 summarized the underlying meanings and expected signs of these treatment effects.

The rationale behind this multiple intervention research design is that, it allows us to test for heterogeneous new rail transit's impacts on Beijing's land market along several dimensions. First, it is expected that these estimates are significantly positive in corresponding periods. For example, the interactions between $(station \geq 2003 * 2003 \leq year < 2008)$ and $(station \geq 2008 * year \geq 2008)$ should be significantly positive and show the opening effect for stations in 2003 and in 2008

[12] In the presence of nested treatment groups, our study's estimates provide new insights about each treatment effect conditional on the subsequent treatment scenarios. One major concern is to test whether there are spillover effects among treatment groups when adding all of them into one model specification. As a robustness check, we have tried to add each treatment group subsequently in different model specifications, but the difference between their coefficients won't tell anything about the spillover effect because the sum up value of the treatment coefficients remains the same as when adding all of them into one model specification.

[13] β_{j1} represent a set of baseline categories ($Treatment_j * Period_1$) that are omitted in the estimating result tables.

respectively. A second dimension is captured by estimates of $(station \geq 2003 * year \geq 2008)$ and $(station \geq 2008 * 2003 \leq year < 2008)$. These two coefficients allow us to test post-opening effect for stations in 2003 and pre-opening effect for station in 2008 respectively. The expected signs largely depend on the price growth trends during 2003 and 2008 versus after 2008. If the price growth trends after 2008 are greater than that during 2003 and 2008, then their estimates would be less positive and insignificant. A third dimension is to examine the net planning effect[14] for stations opening after 2009 relating to different land market periods. It is expected that there would be positive signs associated with estimates of $(station \geq 2009 * 2003 \leq year < 2008)$ and $(station \geq 2009 * year \geq 2008)$.

There are at least three limitations to the models presented above. The first limitation is the common time-trend assumption. In general, one would expect observed and unobserved characteristics to be evolved with the transport improvement. This is particularly the case if we do not have access to repeated observations for the same parcel over time and therefore cannot apply panel-data methods to control for fixed-over-time omitted variables. To address this concern, this paper does provide an extremely rich data set which allows us to mitigate this problem (at least partially) by controlling for a wide range of parcel and location characteristics. Despite of this effort, our results might still underestimate the rail access effect if price adjustment process is long before or after the building of new subway lines, or might overestimate the amenity benefits if other local externalities at

[14] Note that the net planning effect includes a combination of the potential negative construction effect and the positive anticipation effect for planned stations (Knaap et al, 2001).

station areas evolve with the increased rail access. This problem is not unique here. Ideally, one could control for a number of things (i.e. crime, shops, cafes, travel time) change together as a result of the stations opening if those detailed data is accessible. However, to the best of our knowledge, there are no publicly available data sources in which we can merge systemic information on localized changes with detailed data on land price and demographic characteristics. When one is reading the results, it is important to keep in mind that our measures might capture the additional impact of variation at the local areas.

Second, many location factors associated with rail stations would have interaction effects on land prices for reasons other than the benefits of increased station proximities due to the building of new railway lines. For example, stations located near employment-centre could offer more job opportunities and other amenities that might provide additional land values, whereas increasing proximity to station areas with high crime levels may actually decrease the benefits of transport accessibility on land values. To help identify such interaction effects, the model specification can be written as:

$$\begin{aligned}\mathrm{Ln}Price_{\mathrm{iljt}} = \beta_0 + \sum_{\mathrm{j=1}}^{3} \beta_{\mathrm{j}} Treatment_{\mathrm{j}} + \sum_{\mathrm{t=1}}^{3} \beta_{\mathrm{t}} Period_{\mathrm{t}} \\ + \sum_{\mathrm{j=1}}^{3} \sum_{\mathrm{t=1}}^{3} (\beta_{\mathrm{jt}} + \beta_{\mathrm{k}}^{'} \mathrm{X}_{\mathrm{ilk}})\, Treatment_{\mathrm{j}} * Period_{\mathrm{t}} + \beta_{\mathrm{k}} X_{\mathrm{ilk}} + f_{\mathrm{l}} + \varepsilon \quad (4)\end{aligned}$$

Finally, another potential source of bias may arise from the conducted timing of the data. It is worth noting that when the new metro lines and stations were being constructed, accessibility for residents at station areas might be in fact lowered by localized congestion---which could lead to lower land values before the opening of new stations. When new stations were opened, the changes in land values would reflect not just the observable accessibility benefits, but also the disappearance of the noise or congestion effects at the station areas. Despite these limitations, we believe that our analysis provides some "food-for-thoughts" about real estate consequences of transport improvements in China.

4 Results

4.1 Balance of "treated" and "control" places

As a first step towards valuing rail access, it is worthwhile to do the "balancing" tests to see if treated places would be significantly different from the control places[15] in terms of the observable pre-treatment demographic characteristics.

Columns (1)–(6) of Table 3 present regressions of initial residential land prices, educational attainment, population density, public housing rent ratio, employment accessibility and other demographic variables measured in the pre-treatment period (before 2003), on the indicators for treatment variables using 2km distance band[16]. The upper panel of Table 3 shows estimates using the residential land parcel sample. The lower panel shows coefficients using the commercial land parcel sample.

For the most part, there is a near-zero and insignificant coefficient in all these

[15] Intuitively, the control group is places that have never been within a 2km radius of a rail transit station. For a more nuisance assessment, we have also used the propensity score matching techniques to select the control group without 2km distance band of a rail station based on local demographic characteristics, and the results are virtually similar.

[16] Due to the lack of census panel data, we cannot measure demographics dynamics in treated places relative to observationally identical control places as a result of transport improvements.

regressions, suggesting that treated and control places do not have markedly different pre-treatment characteristics. We do find marginally significant coefficients for the employment accessibility, however, the magnitude of such coefficients are very small. Although not reported, repeating this exercise for 1km and 4km distance-band treatment scenarios tends to improve the balancing conditions in terms of pre-treatment characteristics. This gives us more confidence about the reliability of the results. In any case, we will adjust the estimates further for differences in characteristics, in the regression estimates that follow in the next part of this section.

4.2 Baseline estimates

Tables 4 and 5 show baseline regression estimates of the model in equation (3) using the same three-period, parcel-level land data. The only variation in station distance is between periods, in places affected by the rail access changes. So, any measured effects of station distances on land prices occur only through parcel-station distance reductions due to the building of new rail transit lines after 2003.

Column (1) in both tables shows estimates that include proximity effects for parcels that are beyond the 4km distance band, treatment dummies, fixed effects, general time effects [17], but no additional controls. In the first treatment group $(station \geq 2003)$, the opening effect of stations in 2003 on the residential land prices is found significantly positive when treated with all distance bands [18]. In particular,

[17] To further control the spatial-temporal effect, we also include the interactions between time trends and parcels in each treatment group that only meet the first treatment selection criteria---parcels that experienced distance reductions to the closet stations(Treatment Criteria 1* Time); and interactions between time trends and parcels in each treatment group that only meet the second treatment selection criteria---the parcel-station distance is now within the distance bands(Treatment Criteria 2* Time).

[18] Recall that the distance bands are cumulative, which make the interpretation of the results more

parcels that are now within 2km from a station have a significantly higher price premium compared to other distance bands. These results suggest that residential land parcels that are very close to stations might be affected by negative externalities, but those at an intermediate spatial range are beyond potential negative externality effects and benefit more from increased rail access. There are no statistically significant post-opening impacts from distance reductions to parcels that are beyond 1km, 2km, and 4km spatial contours from new stations in 2003.

Estimates from the second treatment group $(station \geq 2008)$ are qualitatively similar to those reported in the first treatment group $(station \geq 2003)$. But the price premium paid for being closer to a station opening in 2008 is much larger than that of newly-opened station in 2003. This is expected because more new stations were opened in 2008 than in 2003, resulting in obvious parcel-station distance reductions. The pre-opening effects for stations in 2008 are positively significant when treated with the 1km, 2km and 4km distance-bands[19]. Continuing to discuss the results in Column (1), we next focus on the third treatment group $(station \geq 2009)$. This treatment group highlights the net planning effects for stations opening after 2009. As expected, we find that prices rise significantly in areas affected by planned stations opening after 2009 when treated with both of the $(2003 \leq year < 2008)$ period and the

straightforwardly. For example, for residential parcels we find a positive effect within 1km of a station. But the next band is within 2km of a station and the results show a stronger positive effect. This result implies that the proximity impact of rail stations is determined by the mix of properties within 1 km and between 1km and 2km. Of course, researchers can further disaggregate the distance band selection into the 0.5km range, or choose to define the bands as 0 to 0.5km, 0.5km to 1km, 1km to 2km, and 2km to 4km. Recall that our purpose here is to shed light on the importance of considering the distributional proximity impacts of rail stations on land prices over space.

[19] Note that treatment dummies have insignificant signs, which can help explain the pre-opening effect of station in 2008 is not caused by the price-growing trends in the treated places.

$(year \geq 2008)$ period. However, we find that the price premiums associated with the 2km distance band are higher during the period after 2008 than that of during 2003 and 2008. This result confirms the possibility that the under-constructed rail transit plans are observed by the developers and increasingly capitalized into land prices when closing to their completion times. Again, it is noteworthy that the data limits our transport improvement analysis to changes that occurred within about 3 years of the new rail transit development. The estimates might underestimate the whole effect of transport accessibility when the price-lag adjustment process is long before or after the opening of new lines[20], or might overestimate the benefits if negative externalities at station areas evolve with the improved transport accessibility.

For mega-cities like Beijing, part of the increased station proximity effects could be attributed to the spatial effects, like differences in price trends in the central city and suburbs. In Column (2), we estimate the same specification but augmented with a set of spatial measures by allowing the interactions between the time trend and distance to CBD, and by allowing for time trends interacted with the distance to the Olympic park, and the distance to "bedroom" communities[21]. The rationale behind this is that, during the study period, there was a boom in land price growth in Beijing. Although not shown in the table, this is confirmed by the significantly negative coefficient on the distance to CBD and its interactive terms with the time trend. The key finding here is that whilst the price growth trend effect matters, the increased

[20] See McDonald and Osuji (1995) and McMillen and McDonald (2004) for a detailed discussion.

[21] The estimated coefficients of these interaction terms are not reported. The results remain robust by controlling the interactions between time trends and distance-to-stations.

station proximity effects are still robust and contribute to significantly higher residential land prices.

In Columns (3) and (4), we control for a wide range of land structural and location-specific characteristics (documented in Appendix Table 1). About 45% of the variation in the log of land prices is explained by transport improvement models. This compares favourably to previous hedonic literature in China. After controlling for the full set of localized characteristics and adjusting for different temporal-spatial trends in column (4), we find that the opening effects of station in 2003, on average, are valued at around 0.61%, 1.96%, 1.25% of residential land prices at affected areas (within 1km, 2km, 4km respectively). The opening effects of station in 2008, on average, are valued at about 3.75%, 4.20%, 2.02% of the prices of affected residential land parcels (within 1km, 2km, 4km respectively). The positive and significant signs associated with the pre-opening effect for station in 2008 show that the potential increased station proximity effect is capitalised into local land prices (within 1km, 2km and 4km). In terms of the net planning effect, prices rise by about 0.23%, 0.58%, 0.48% on average when treated with the time period during 2003 and 2008 (using 1km, 2km, 4km distance band respectively); and prices rise by around 3.01%, 3.79%, 3.51% on average when treated with the time period after 2008 (using 1km, 2km, 4km distance band respectively). These results suggest that land prices would rise more than proportionately with station proximity and there is a bit of a non-linear price elasticity effect going from those within 4km to those within 1km of a station[22].

[22] Note that land parcels located more than 4 km away from a new station might also benefit from the improvements in rail access and would be far enough from the localized congestion nuisances at the station areas.

Switching to the commercial land parcel sample in Table 5, we find quite similar qualitative patterns with the residential land parcel sample results. As for the quantitative nature, we find that station proximity impacts on commercial land prices are slightly lower than those on residential land prices, possibly because that the commercial land market is thinner than the residential land market. We also find that these proximity impacts decay in a non-linear distance trend. For example, the most affected places for commercial land parcels are those that are now within 1km station area. This finding is in line with the expectation that commercial land parcels would accrue greater benefit than residential land parcels at a closer distance range from a station---by gathering large population flows and high demand for commercial activities.

To further explore whether the observed differences in proximity impacts on residential and commercial land prices are statistically significant, the Chow statistical test (Chow 1960) is conducted. The null hypotheses are: the set of coefficients for the treatment effects on the commercial parcels and the corresponding set of coefficients for the treatment effects on the residential parcels are not significantly different from each other. In the results not reported, all null hypotheses are rejected at the 5% significance level. This provides strong evidence of the spatial heterogeneity in the proximity impacts of rail stations across residential and commercial land markets.

4.3 Robustness

To test the robustness of main findings, we now examine how sensitive the

---

We have tested this hypothesis and find little evidence to support this claim.

baseline results are to changes in data samples and econometric specifications. The first sensitivity analysis is to adjust spatial selections in the land parcel sample. As the Beijing urbanized area is so large, it may have a large influence on the baseline estimates. We therefore, in model specifications of Table 6 report results that only include the land parcel sample located within the central city (within the 3rd ring road) and within the suburb (within the 5th ring road) subsequently. The results, reported in Columns 1-4 of Table 6, generally mirror that of the baseline estimates, suggesting that the spatial trimming of parcel sample does not significantly affect the new rail transit's impact[23].

Next, we consider whether there are significant spillover effects within and across residential/commercial treatment groups. Such test helps to gauge the robustness of the results more fully. Following the Irwin and Bockstael (2001), the *within-group* spillover effects are measured by the interactions between distance from parcels in $Treatment_j$ (but not belong to the $Treatment_{j+1}$) to parcels in $Treatment_{j+1}$, and its corresponding treatment effect ($Treatment_j$*$Year_t$). The results in Columns 1-2 of Table 7 show that the estimated spillover effect coefficients are small in magnitudes and insignificant for both residential and commercial parcel sample[24].

The *cross-group* spillover effects are calculated through the interactions of the distance between all treated commercial land parcels[25] and residential land parcels in

---

[23] Note that while the qualitative nature of the results is relatively robust, the estimated coefficients of treatment effects within the central city have lower magnitudes than those within the suburbs. In the results not reported, we also find that the opening and planning effects are more pronounced in station areas that look to be getting several new lines rather than just one.

[24] There are no significant spillover effects within groups when using the 1km and 4km distance bands.

[25] Note that we have also interacted the residential land parcels in each treatment group with both of treated and

each treatment group. Estimates from column (3) in Table 7 show that most of the residential treatment effect variables are reassuringly quite robust to the potential spillover impacts from nearby commercial land parcels. The only two exceptions are associated with the treated residential land parcels receiving distance reductions to stations after 2009. The small magnitudes of their coefficients affirm the possibility that residential land parcels could gain slightly positive spillover effects from adjacent commercial land parcels when treated with planned station areas.

Finally, reliance on estimates of amenity benefits for the average sample effect in a metropolitan area would mask rail access values to parcels in particular places. Thus we now turn to the results with interaction terms estimated by using Eq.(4). In Table 8, the interactions between treatment effect variables and local residents' median educational attainment level show that residential land price premiums are valued greater for being close to a station in high- than in low-educational attainment areas. Meanwhile, the commercial land prices are found to be valued higher when treated in high- than in low-educational attainment places, possibly because the larger consumption capability for well-educated residents gentrifies the value they attach to rail access. Estimates from the treatment effect variables and employment accessibility interactions show that the effect of increased station proximity is more valuable in places with higher- than lower-employment accessibility. We find no significant evidence on the interactions between treatment effect variables and crime rates.

---

control commercial land parcels. Because the estimating results are not significant, they were dropped from the table.

Beyond these interaction variables, what other local amenities and disamenities that might have significant complementarities with rail stations were overlooked? The list could be very long, including climate, social capital, architectural factors, and other forces of local heterogeneity that are unlikely to be observed by the econometrician. Thus, our results sheds light on the importance of valuing rail access, not just in terms of controlling its structural characteristics, but how those characteristics interact with local demographics.

6 Conclusions

Public investments on transport infrastructure has been and will continue to be critical parts of government budgets, yet little is known about the real estate consequences of these investments, especially in transitional economy countries. This paper uses unique data and innovative methods to assess the effects from transport improvements on vacant residential and commercial land prices at a micro-geographic scale. The transport improvements we consider entailed the opening and planning of new stations in Beijing from 2003 to 2012, so we can look at what happened to land prices when geographical distances to the nearest station were reduced.

Our results yield three important insights. First, residential and commercial land parcels receiving increased station proximity have experienced appreciable price premiums. For example, we find the treatment effects of 2.02% or more on residential land prices in places affected by the opening of new stations in 2008, whilst such effects are slightly less for commercial land prices. Second, the results highlight the importance of land price changes in the planned station areas. This finding add to the growing body of literature that suggests that government plans in rail transit constructions do matter in influencing property values even before they are in service. Finally, we find some evidence that these proximity effects on land prices are not distributed evenly over space and local demographics.

These results are important because they show the direct evidence about how land prices respond to rail access changes, and because they provide a sound empirical basis for planners to consider wider gentrification effects that might be affected by infrastructure investments. Within the "new urbanism" process, transport-oriented development strategies were designed to regenerate depressed areas and reduce congestion in downtown areas. Classic examples of this include Boston's Big Dig, Chicago's Midway line, Los Angeles's Bay Area subway line, Toronto's Spadina Subway line and London's Jubilee and DLR lines. Given the massive investments in rail infrastructure, empirical answers are scarce on whether public investments and private investments are complements that spur the emerging land markets in developing countries. Our findings offer a good support for this argument by demonstrating that the same "game" plays out in Beijing, where transport infrastructure investments have been shown to stimulate spatially targeted land markets.

Reference

Adams, D., and Tiesdell, S. (2013). *Shaping Places: Urban Planning, Design and Development.* London: Routledge.

Bajic, V., (1983). The effects of a new subway line on housing prices in metropolitan Toronto. *Urban Studies*, *20*(2),147–158.

Baum-Snow N. (2010). Changes in transportation infrastructure and commuting patterns in US metropolitan areas, 1960-2000. *American Economic Review*, *100*(2), 378–382.

Baum- Snow, N., and M. E. Kahn. 2005. Effects of urban rail transit expansions: Evidence from sixteen cities from 1970– 2000. *Brookings-Wharton papers on urban affairs,* vol. 6, ed. G. Burtless and J. R. Pack, 147– 206. Washington, DC: Brookings Institution Press.

Baum-Snow, N., Kahn, M.E. (2000). The Effects of New Public Projects to Expand Urban Rail Transit, *Journal of Public Economics, 77*(2), 241–263.

Billings, S. B. (2011). Estimating the value of a new transit option. *Regional Science and Urban Economics, 41*(6), 525–536.

Bowes, D., and Ihlanfeldt, K. (2001) Identifying the impacts of rail transit stations on residential property values. *Journal of Urban Economics, 50*(1), 1–25.

Cai, H., Henderson, J.V., Zhang, Q.(2009). China's land market auctions: Evidence of corruption. *NBER Working Paper, No. 15067.*

Cheshire, P.C., and Sheppard S. (2004). Capitalising the value of free schools: the impact of supply characteristics and uncertainty. *Economic Journal, 114*(499),397–424.

Cheshire, P.C., and Sheppard, S. (1995). On the price of land and the value of amenities. *Economica, 62*(258), 247–267.

Chow, G.C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, *28*(3), 591–605.

Davis, F.W., (1970). Proximity to a rapid transit station as a factor in residential property values. *The Appraisal Journal*, *38*(4), 554–572.

Deng, Y., and Liu, P., (2009). Mortgage prepayment and default behavior with embedded forward contract risks in China’s housing market. *Journal of Real Estate Finance and Economics, 38*(3), 214–240.

Ding, W., Zheng, S., Guo, X. (2010). Value of access to jobs and amenities: Evidence for new residential properties in Beijing. *Tsinghua Science and Technology*, *15*(5), 595–603.

Duranton, G. and Turner M.A. (2012). Urban Growth and Transportation. *Review of Economic Studies, 79*(4), 1407–1440.

Gibbons, S. (2004). The costs of urban property crime. *Economic Journal, 114*(499), 441–463.

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

Gibbons, S., and Machin, S. (2008). Valuing school quality, better transport and lower crime: Evidence from house prices. *Oxford Review of Economic Policy*, *24*(1), 99–119.

Gibbons, S., Lyytikäinen,T., Overman, H. G., Sanchis-Guarner,R. (2012). New road infrastructure: The effects on firms. *SERC Discussion Papers, SERCDP00117*. Spatial Economics Research Centre (SERC), London School of Economics and Political Sciences, London, UK.

Gibb, K., Pryce, G. (2012). New directions in housing economics: Introduction. *Journal of Property Research, 29*(4), 271–279.

Gunn, H. (2000). *An introduction to the valuation of travel time savings and losses,* Chapter 26 in: D.A. Hensher, K.J. Button (Eds.), Handbook of Transport Modelling, Elsevier Science.

Gyourko, J., Kahn, M.E., Tracy, J. (1999). *Quality of life and the environment.* In: Cheshire, P.,Mills, E. (Eds.), Handbook of Regional and Urban Economics, vol. 3. North-Holland.

Irwin, E .G., and Bockstael, N. E. (2001). Interacting agents, spatial externalities, and endogenous evolution of residential land use pattern. *Journal of Economic Geography*, *2*(1), 31–54.

Kahn, M. E. (2007). Gentrification trends in new transit-oriented communities: Evidence from 14 cities that expanded and built rail transit systems. *Real Estate Economics*, *35*(2), 155–182.

Knaap, G.J., Ding,C., Hopkins, L.D. (2001). Do plans matter? Effects of light rail plans on land values in station areas. *Journal of Planning Education and Research*, *21*(1), 32–39.

Jackson, C., and Orr, A. (2011). Real estate stock selection and attribute preferences. Journal of Property Research, *28* (4), 317–339.

McMillen, D.P. (2001). Nonparametric employment subcenter identification. *Journal of Urban Economics, 50*(3), 448–473.

McMillen, D.P., McDonald, J. (2004). Reaction of house prices to a new rapid transit line: Chicago's midway line 1983-1999. *Real Estate Economics, 32*(3), 463–486.

McDonald, J.F., Osuji, C.I. (1995). The effect of anticipated transportation improvement on residential land values. *Regional Science and Urban Economics 25*(3), 261–278.

Wang,Y.P., Murie, A. (2000). Social and spatial implications of housing reform in China, *International Journal of Urban and Regional Research, 24*(2), 397–417.

World Bank. (2008). *World Bank Transport Business Strategy 2008-2012.* Washington DC: World Bank.

Wu, W., and Zhang, W. (2009). PLS Path Model Building: A multivariate approach to land price studies-A case study in Beijing. *Progress in Natural Science, 19*(11), 1643–1649.

Wu, W., Zhang, W., Jin, F. and Deng, Y. (2009). Spatio-temporal analysis of urban spatial interaction in globalizing China: A case study of Beijing-Shanghai corridor. *Chinese Geographical Science, 19*(2), 126–134.

Wu, W., Zhang, W., Dong, G. (2013). Determinant of residential location choice in a transitional housing market: Evidence based on micro survey from Beijing. *Habitat International*, *39*(3), 16–24.

Wu, W. (2014). Does public investment improve homeowners' happiness? New evidence based on micro surveys in Beijing. *Urban Studies, 51*(1), 75–92.

Wu, W., Dong, G. (2014). Valuing "green" amenities in a spatial context. *Journal of Regional Science,* in press.

Zheng, S., and Kahn, M.E. (2008). Land and residential property markets in a booming economy: New evidence from Beijing. *Journal of Urban Economics*, *63*(2), 743–757.

Zheng, S., Fu, Y., and Liu, H. (2009). Demand for urban quality of living in China: Evidence from cross-city land rent growth. *Journal of Real Estate Finance and Economics, 38*(3), 194–221.

Zheng, S., Cao, J., Kahn, M.E, Sun, C. (2013). Real estate valuation and cross-boundary air pollution externalities: Evidence from Chinese cities. *Journal of Real Estate Finance and Economics*, *46*, 1–17.

Zheng, S., Kahn, M.E. (2013) Does Government Investment in Local Public Goods Spur Gentrification? Evidence from Beijing. *Real Estate Economics*, *41*(1), 1–28.

Table list

Table 1 New rail transit constructions in the urbanized area of Beijing

Line	Start by (year)	Open by (year)	Cost (RMB billions)	Length (kilometre)	Station (number)
13	2000	2003	6.6	40.5	16
Batong	2001	2003	3.4	19	13
4	2004	2008	15.2	28	24
5	2003	2008	11.9	27.6	23
10A	2004	2008	12.8	24.6	22
8A	2005	2008	2.5	15.8	4
Daxing	2008	2010	6.0	22	12
Yizhuang	2008	2011	11.0	23.2	14
8B	2009	2012	10.1	17	11
6	2007	2012	18.2	39	30
7	2009	2012	15.1	24	21
9	2007	2012	8.8	16.4	13
10B	2007	2012	18.5	32.9	23
15A	2009	2012	18.1	20.2	13

Notes.---The information on the rail transit lines that have not been completed yet may be changed. See the updated information on the Beijing Municipal Committee of Transport's official website http://www.bjjtw.gov.cn/

Table 2 Underlying meanings and expected signs of the treatment effects

Treatment effects	Underlying Meaning	Expected signs
$(station \geq 2003) * (2008 > year \geq 2003)$	Opening effect of stations in 2003	+
$(station \geq 2003) * (year \geq 2008)$	Post-opening effect of stations in 2003	+/-
$(station \geq 2008) * (2008 > year \geq 2003)$	Pre-opening effect of stations in 2008	+/-
$(station \geq 2008) * (year \geq 2008)$	Opening effect of stations in 2008	+
$(station \geq 2009) * (2008 > year \geq 2003)$	Net planning effect of stations after 2009	+
$(station \geq 2009) * (year \geq 2008)$	Net planning effect of stations after 2009	+

Table 3 "Balancing" test results

	Land Price (1)	Education (2)	Public Housing (3)	Density (4)	Heritage (5)	Employment (6)
Panel A: Residential land sample						
2km_$(station \geq 2003)$	0.173	0.013	-0.038	-0.972	0.015	-0.013
	(0.935)	(0.121)	(-0.950)	(-1.358)	(0.938)	(-3.250)
2km_$(station \geq 2008)$	-0.235	0.163	-0.005	1.398	-0.039	-0.024
	(-1.343)	(1.598)	(-0.132)	(1.431)	(-1.560)	(-6.001)
2km_$(station \geq 2009)$	-0.072	0.019	-0.019	0.198	-0.010	-0.004
	(-1.075)	(0.487)	(-1.267)	(0.759)	(-1.429)	(-2.021)
Constant	0.304	1.011	-0.402	-3.016	-0.106	-0.272
	(0.749)	(4.284)	(-4.568)	(-2.811)	(-3.029)	(-6.727)
Fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1054	1054	1054	1054	1054	1054
Adjusted R-squared	0.416	0.34	0.207	0.222	0.153	0.577
Panel B: Commercial land sample						
2km_$(station \geq 2003)$	0.323	0.242	-0.028	-1.373	0.057	-0.023
	(0.441)	(0.804)	(-0.252)	(-0.454)	(0.838)	(-1.769)
2km_$(station \geq 2008)$	-0.312	0.117	0.097	3.737	-0.143	-0.056
	(-0.429)	(0.391)	(0.875)	(1.217)	(-1.607)	(-4.308)
2km_$(station \geq 2009)$	0.681	0.303	-0.145	-4.506	0.081	-0.039
	(0.799)	(0.866)	(-1.124)	(-1.253)	(1.025)	(-2.610)
Constant	1.357	0.428	-0.536	-5.312	-0.263	-0.190
	(1.182)	(0.909)	(-3.045)	(-2.006)	(-2.430)	(-9.048)
Fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	466	466	466	466	466	466
Adjusted R-squared	0.277	0.278	0.199	0.278	0.148	0.538

Notes.--- Each column reports estimates of the balancing tests from a separate regression. The dependent variable for each regression is listed in the first row of the table (initial average land prices, educational attainment, public housing rent ratio, population density, heritage building percentage, employment accessibility), as described in the text. The independent variables are the treatment group. Results in the top panel are from residential land parcel sample. Results in the bottom panel are from commercial land parcel sample. t-statistics in parentheses, clustered on zone unit.

Table 4 Baseline estimates of rail transit's effect on residential land parcel sample

Distance band	Variables	Model 1	Model 2	Model 3	Model 4
1 km	$(station \geq 2003) * (2008 > year \geq 2003)$	0.664	0.642	0.621	0.611
		(1.829)	(1.778)	(1.876)	(1.746)
	$(station \geq 2003) * (year \geq 2008)$	0.383	0.196	0.183	0.199
		(0.834)	(0.422)	(0.416)	(0.449)
	$(station \geq 2008) * (2008 > year \geq 2003)$	0.653	0.592	0.584	0.575
		(0.351)	(0.333)	(0.312)	(0.319)
	$(station \geq 2008) * (year \geq 2008)$	4.532	4.218	3.992	3.750
		(4.263)	(3.957)	(3.580)	(3.378)
	$(station \geq 2009) * (2008 > year \geq 2003)$	0.298	0.256	0.242	0.239
		(1.776)	(1.631)	(1.779)	(2.025)
	$(station \geq 2009) * (year \geq 2008)$	3.276	3.134	3.188	3.009
		(2.884)	(2.796)	(3.107)	(3.067)
2 km	$(station \geq 2003) * (2008 > year \geq 2003)$	2.337	2.148	2.026	1.968
		(2.452)	(2.201)	(2.034)	(1.977)
	$(station \geq 2003) * (year \geq 2008)$	1.121	1.799	1.206	1.053
		(1.045)	(1.598)	(1.193)	(1.020)
	$(station \geq 2008) * (2008 > year \geq 2003)$	1.695	1.675	1.509	1.281
		(2.042)	(2.204)	(2.219)	(2.100)
	$(station \geq 2008) * (year \geq 2008)$	4.661	4.427	4.221	4.206
		(4.099)	(3.900)	(3.765)	(3.862)
	$(station \geq 2009) * (2008 > year \geq 2003)$	0.699	0.634	0.601	0.584
		(3.344)	(3.268)	(3.284)	(3.281)
	$(station \geq 2009) * (year \geq 2008)$	4.206	4.052	4.001	3.799
		(3.456)	(3.289)	(3.143)	(2.954)
4 km	$(station \geq 2003) * (2008 > year \geq 2003)$	1.664	1.459	1.361	1.259
		(2.956)	(2.727)	(2.638)	(2.596)
	$(station \geq 2003) * (year \geq 2008)$	1.992	1.750	1.518	1.332
		(1.515)	(1.345)	(1.248)	(1.213)
	$(station \geq 2008) * (2008 > year \geq 2003)$	1.449	1.225	0.912	0.941
		(1.723)	(1.690)	(1.737)	(1.860)
	$(station \geq 2008) * (year \geq 2008)$	2.589	2.297	2.129	2.023
		(2.631)	(2.441)	(2.234)	(2.168)
	$(station \geq 2009) * (2008 > year \geq 2003)$	0.538	0.496	0.485	0.481
		(1.724)	(1.664)	(1.792)	(1.979)
	$(station \geq 2009) * (year \geq 2008)$	4.179	4.156	4.043	3.511
		(4.053)	(4.194)	(4.092)	(4.388)
Distance to CBD*Trends		No	Yes	No	Yes
Distance to Stations>4KM		Yes	Yes	Yes	Yes
Parcel Characteristics		No	No	Yes	Yes
Treatment dummies		Yes	Yes	Yes	Yes
Time dummies		Yes	Yes	Yes	Yes
Treatment Criteria 1*Time		Yes	Yes	Yes	Yes

Treatment Criteria 2*Time	Yes	Yes	Yes	Yes
Distance to OlympicPark*Time	No	Yes	No	Yes
Distance to New Residential Area$_i$*Time	No	Yes	No	Yes
Station-distance*Time	No	Yes	No	Yes
Fixed effect	Yes	Yes	Yes	Yes
Location-specific characteristics	No	No	Yes	Yes
Observations	2,343	2,343	2,343	2,343
Adjusted R-squared	0.384	0.393	0.437	0.456

Notes.---Dependent variable is log residential land price. Data is the disaggregated parcel-level data for three periods: pre-2003, 2003-2007 and after. Coefficients are×100. The baseline omitted category is *Treatment$_j$***Period$_1$(pre-2003)*. Regressions include control variables detailed in Appendix Table 1. t-statistics in parentheses, clustered on zone unit.

Table 5 Baseline estimates of rail transit's effect on commercial land parcel sample

Distance band	Variables	Model 1	Model 2	Model 3	Model 4
1 km	$(station \geq 2003) * (2008 > year \geq 2003)$	0.889	0.763	0.662	0.625
		(2.102)	(2.079)	(2.181)	(2.097)
	$(station \geq 2003) * (year \geq 2008)$	0.805	0.621	0.556	0.511
		(1.214)	(0.944)	(0.832)	(0.768)
	$(station \geq 2008) * (2008 > year \geq 2003)$	1.605	1.469	1.185	0.871
		(2.439)	(2.652)	(2.319)	(1.819)
	$(station \geq 2008) * (year \geq 2008)$	2.788	2.376	2.128	1.963
		(3.584)	(3.337)	(2.653)	(3.035)
	$(station \geq 2009) * (2008 > year \geq 2003)$	0.795	0.668	0.622	0.582
		(1.944)	(1.663)	(1.709)	(1.921)
	$(station \geq 2009) * (year \geq 2008)$	1.463	1.298	1.165	1.081
		(1.701)	(1.728)	(1.686)	(1.941)
2 km	$(station \geq 2003) * (2008 > year \geq 2003)$	0.736	0.687	0.675	0.622
		(1.669)	(1.789)	(1.843)	(1.891)
	$(station \geq 2003) * (year \geq 2008)$	0.753	0.501	0.499	0.389
		(0.506)	(0.334)	(0.341)	(0.268)
	$(station \geq 2008) * (2008 > year \geq 2003)$	1.359	1.135	0.986	0.766
		(1.810)	(1.736)	(1.680)	(1.662)
	$(station \geq 2008) * (year \geq 2008)$	1.913	1.616	1.567	1.449
		(2.142)	(1.973)	(2.040)	(2.153)
	$(station \geq 2009) * (2008 > year \geq 2003)$	0.706	0.616	0.588	0.516
		(1.709)	(1.735)	(1.861)	(1.823)
	$(station \geq 2009) * (year \geq 2008)$	1.493	1.346	1.211	1.014
		(1.868)	(1.775)	(1.670)	(1.684)
4 km	$(station \geq 2003) * (2008 > year \geq 2003)$	0.646	0.627	0.552	0.396
		(1.755)	(1.923)	(1.890)	(1.692)
	$(station \geq 2003) * (year \geq 2008)$	1.268	1.007	0.939	0.604
		(0.856)	(0.679)	(0.644)	(0.586)
	$(station \geq 2008) * (2008 > year \geq 2003)$	1.251	1.208	1.024	0.876
		(1.700)	(1.808)	(1.769)	(1.708)
	$(station \geq 2008) * (year \geq 2008)$	2.494	2.111	1.988	1.834
		(2.269)	(2.359)	(2.513)	(2.327)
	$(station \geq 2009) * (2008 > year \geq 2003)$	0.885	0.756	0.688	0.582
		(1.667)	(1.662)	(1.707)	(1.813)
	$(station \geq 2009) * (year \geq 2008)$	1.651	1.489	1.439	1.322
		(1.705)	(1.686)	(1.725)	(1.737)
Distance to CBD*Trends		No	Yes	No	Yes
Distance to Stations>4KM		Yes	Yes	Yes	Yes
Parcel Characteristics		No	No	Yes	Yes
Treatment dummies		Yes	Yes	Yes	Yes
Time dummies		Yes	Yes	Yes	Yes
Treatment Criteria 1*Time		Yes	Yes	Yes	Yes

Treatment Criteria 2*Time	Yes	Yes	Yes	Yes
Distance to OlympicPark*Time	No	Yes	No	Yes
Distance to New Residential Area$_i$*Time	No	Yes	No	Yes
Station-distance*Time	No	Yes	No	Yes
Fixed effect	Yes	Yes	Yes	Yes
Location-specific characteristics	No	No	Yes	Yes
Observations	1,341	1,341	1,341	1,341
Adjusted R-squared	0.297	0.331	0.365	0.388

Notes.---Dependent variable is log commercial land price. See other notes in Table 4.

Table 6 Regression estimates of rail transit's effect on selected sample, sensitivity analysis

Distance band	Variables	Residential land parcel sample		Commercial land parcel sample	
		(1)	(2)	(3)	(4)
1 km	$(station \geq 2003) * (2008 > year \geq 2003)$	0.565	0.597	0.568	0.592
		(1.652)	(1.860)	(1.656)	(1.935)
	$(station \geq 2003) * (year \geq 2008)$	0.093	0.135	0.377	0.425
		(0.178)	(0.288)	(0.475)	(0.613)
	$(station \geq 2008) * (2008 > year \geq 2003)$	0.328	0.551	0.798	0.889
		(0.818)	(1.662)	(1.659)	(1.912)
	$(station \geq 2008) * (year \geq 2008)$	2.851	3.031	1.392	1.556
		(2.021)	(2.403)	(1.891)	(2.542)
	$(station \geq 2009) * (2008 > year \geq 2003)$	0.216	0.225	0.501	0.614
		(1.649)	(1.844)	(1.176)	(1.878)
	$(station \geq 2009) * (year \geq 2008)$	1.949	2.352	0.981	1.016
		(1.633)	(2.277)	(1.657)	(1.648)
2 km	$(station \geq 2003) * (2008 > year \geq 2003)$	1.763	1.835	0.579	0.605
		(1.676)	(1.829)	(1.662)	(1.790)
	$(station \geq 2003) * (year \geq 2008)$	0.825	1.027	0.172	0.225
		(0.621)	(0.973)	(0.089)	(0.142)
	$(station \geq 2008) * (2008 > year \geq 2003)$	1.146	1.162	0.628	0.791
		(1.654)	(1.793)	(1.244)	(1.750)
	$(station \geq 2008) * (year \geq 2008)$	2.202	2.911	1.185	1.295
		(1.661)	(2.281)	(1.676)	(1.986)
	$(station \geq 2009) * (2008 > year \geq 2003)$	0.456	0.512	0.578	0.603
		(2.151)	(2.653)	(1.656)	(2.003)
	$(station \geq 2009) * (year \geq 2008)$	2.271	2.862	1.552	1.068
		(1.706)	(2.273)	(1.685)	(1.687)
4 km	$(station \geq 2003) * (2008 > year \geq 2003)$	0.981	1.148	0.278	0.423
		(1.654)	(2.199)	(0.921)	(1.652)
	$(station \geq 2003) * (year \geq 2008)$	0.967	1.201	0.212	0.278
		(0.729)	(1.055)	(0.113)	(0.168)
	$(station \geq 2008) * (2008 > year \geq 2003)$	0.695	0.852	0.732	0.813
		(1.221)	(1.661)	(1.386)	(1.886)
	$(station \geq 2008) * (year \geq 2008)$	1.793	1.916	1.663	1.735
		(1.732)	(1.912)	(1.691)	(2.070)
	$(station \geq 2009) * (2008 > year \geq 2003)$	0.381	0.458	0.481	0.545
		(1.180)	(1.665)	(1.033)	(1.548)
	$(station \geq 2009) * (year \geq 2008)$	1.889	2.878	1.026	1.211
		(2.373)	(4.389)	(1.177)	(1.821)
Distance to CBD*Trends		Yes	Yes	Yes	Yes
Distance to Stations>4 KM		Yes	Yes	Yes	Yes
Parcel Characteristics		Yes	Yes	Yes	Yes
Treatment dummies		Yes	Yes	Yes	Yes
Time dummies		Yes	Yes	Yes	Yes

Treatment Criteria 1*Time	Yes	Yes	Yes	Yes
Treatment Criteria 2*Time	Yes	Yes	Yes	Yes
Distance to Olympic Park*Time	Yes	Yes	Yes	Yes
Distance to New Residential Area$_i$*Time	Yes	Yes	Yes	Yes
Station-distance*Time	Yes	Yes	Yes	Yes
Fixed effect	Yes	Yes	Yes	Yes
Location-specific characteristics	Yes	Yes	Yes	Yes
Observations	1181	1826	707	1036
Adjusted R-squared	0.389	0.431	0.322	0.346

Notes.---The dependent variable is the log of land prices. This table reports the estimates using spatially selected data samples. Columns 1-2 are based on the residential land parcel sample within the central city and suburb respectively. Columns 3-4 are based on the commercial land parcel sample within the central city and suburb respectively. Coefficients are×100. t-statistics in parentheses, clustered on zone unit. See text for details.

Table 7 Regression estimates of spatial spillover effects, sensitivity analysis

Variables	(1)	(2)	(3)
$Dist*(station \geq 2003) * (2008 > year \geq 2003)$	0.011	0.020	-0.021
	(1.222)	(0.153)	(0.375)
$Dist*(station \geq 2003) * (year \geq 2008)$	0.033	0.080	-0.095
	(1.031)	(0.320)	(0.429)
$Dist*(station \geq 2008) * (2008 > year \geq 2003)$	0.020	0.010	-0.024
	(1.001)	(0.250)	(0.381)
$Dist*(station \geq 2008) * (year \geq 2008)$	0.040	0.010	-0.029
	(0.801)	(0.166)	(1.223)
$Dist*(station \geq 2009) * (2008 > year \geq 2003)$	0.040	0.010	-0.011
	(1.000)	(0.250)	(1.911)
$Dist*(station \geq 2009) * (year \geq 2008)$	-0.080	0.010	-0.021
	(-1.143)	(0.142)	(2.131)
Distance to CBD*Trends	Yes	Yes	Yes
Distance to Stations>4KM	Yes	Yes	Yes
Parcel Characteristics	Yes	Yes	Yes
Treatment dummies	Yes	Yes	Yes
Time dummies	Yes	Yes	Yes
Treatment Criteria 1* Time	Yes	Yes	Yes
Treatment Criteria 2* Time	Yes	Yes	Yes
Distance to Olympic Park* Time	Yes	Yes	Yes
Distance to New Residential Area$_i$ * Time	Yes	Yes	Yes
Station-distance*Time	Yes	Yes	Yes
Fixed effect	Yes	Yes	Yes
Location-specific characteristics	Yes	Yes	Yes
Observations	2343	1341	2343
Adjusted R-squared	0.428	0.278	0.439

Notes.---This table reports the estimates of spillover effects. The within-group spillover effects estimates are shown on model specification 1 and 2 based on residential and commercial land parcel sample respectively. The sample sizes are the same as the baseline resulting tables. Estimates of cross-group spillover effects from commercial parcels to residential parcels are shown on specification 3. In specifications 1-2, *Dist* represents a series of distance (in kilometre) interactions between parcels in the subsequent treatment group and parcels in the prior treatment group, as described more details in the text. In specification 3, *Dist* means the interactions of the distance (in kilometre) between treated commercial parcels and treated residential parcels with each residential treatment effect. All specifications are based on treated parcels that experienced distance reductions and the outcome distance to the nearest stations are now within the 2km distance band. t-statistics in parentheses, clustered on zone unit.

Table 8 Regression estimates of interaction effects, sensitivity analysis

Distance band	Variables	Residential land parcel sample			Commercial land parcel sample		
		Educational attainment	Employment accessibility	Crime	Educational attainment	Employment accessibility	Crime
1km	$(station \geq 2003) * (2008 > year \geq 2003)$	0.138	0.064	-0.007	0.296	0.085	-0.016
		(1.816)	(1.685)	(-0.121)	(2.176)	(1.667)	(-0.262)
	$(station \geq 2003) * (year \geq 2008)$	0.937	0.239	-0.201	0.054	0.293	-0.225
		(1.583)	(1.067)	(-1.142)	(0.185)	(1.296)	(-1.271)
	$(station \geq 2008) * (2008 > year \geq 2003)$	0.191	0.005	-0.026	0.197	0.058	-0.009
		(1.073)	(0.172)	(-0.473)	(1.225)	(0.925)	(-0.148)
	$(station \geq 2008) * (year \geq 2008)$	2.071	4.837	-0.782	2.268	2.676	-0.036
		(3.277)	(2.072)	(-1.367)	(2.187)	(1.988)	(-0.165)
	$(station \geq 2009) * (2008 > year \geq 2003)$	0.032	0.028	-0.016	0.131	0.075	-0.053
		(0.481)	(0.622)	(-0.941)	(0.824)	(1.019)	(-1.104)
	$(station \geq 2009) * (year \geq 2008)$	0.148	1.156	-0.063	2.082	0.966	-0.145
		(0.534)	(1.883)	(-0.488)	(2.511)	(1.845)	(-0.879)
2km	$(station \geq 2003) * (2008 > year \geq 2003)$	2.114	0.105	-0.004	0.266	0.378	-0.166
		(4.161)	(2.283)	(-0.058)	(1.750)	(1.979)	(-0.933)
	$(station \geq 2003) * (year \geq 2008)$	0.191	0.325	-0.292	0.292	0.712	-0.809
		(1.073)	(1.109)	(-0.598)	(0.861)	(0.698)	(-0.967)
	$(station \geq 2008) * (2008 > year \geq 2003)$	0.281	0.026	-0.028	0.208	0.942	-0.322
		(1.965)	(0.116)	(-0.444)	(0.504)	(0.661)	(-0.578)
	$(station \geq 2008) * (year \geq 2008)$	3.007	4.660	-0.353	1.751	2.115	-2.793
		(1.755)	(2.149)	(-1.587)	(1.689)	(1.779)	(-1.623)
	$(station \geq 2009) * (2008 > year \geq 2003)$	0.068	0.022	-0.012	0.568	1.213	-0.211
		(1.243)	(0.688)	(-0.800)	(1.303)	(0.719)	(-1.148)
	$(station \geq 2009) * (year \geq 2008)$	2.382	2.521	-0.061	1.979	1.185	-0.389
		(4.436)	(2.942)	(-0.457)	(2.213)	(1.787)	(-1.154)

4km	$(station \geq 2003) * (2008 > year \geq 2003)$	0.462	0.646	-0.039	0.332	0.236	-0.163
		(3.756)	(1.737)	(-0.582)	(1.677)	(2.165)	(-0.896)
	$(station \geq 2003) * (year \geq 2008)$	0.672	0.316	-0.202	0.311	0.642	-2.751
		(1.566)	(1.295)	(-1.270)	(0.816)	(0.633)	(-1.597)
	$(station \geq 2008) * (2008 > year \geq 2003)$	0.186	0.070	-0.027	0.356	0.901	-0.935
		(1.420)	(0.731)	(-0.519)	(0.866)	(0.632)	(-0.962)
	$(station \geq 2008) * (year \geq 2008)$	0.969	3.046	-0.218	1.071	1.439	-1.782
		(2.612)	(1.765)	(-0.965)	(2.052)	(2.129)	(-1.129)
	$(station \geq 2009) * (2008 > year \geq 2003)$	0.039	0.044	-0.012	0.255	1.162	-0.144
		(0.283)	(0.201)	(-0.185)	(0.593)	(0.689)	(-1.321)
	$(station \geq 2009) * (year \geq 2008)$	2.064	1.636	-0.106	0.654	3.395	-0.456
		(4.291)	(3.752)	(-0.404)	(2.003)	(2.066)	(-0.889)
Observations			2343			1341	

Notes.---This matrix table can be viewed as two parts with respect to the residential and commercial land parcel sample respectively. The sample sizes are the same as baseline results. Each part of the table reports the estimates of the interactions between treatment effect variables and educational attainment, employment accessibility, crime rates from one single regression. Coefficients are×100. The regressions shown in the table also include a full set of controls. t-statistics in parentheses, clustered on zone unit.

Figure list

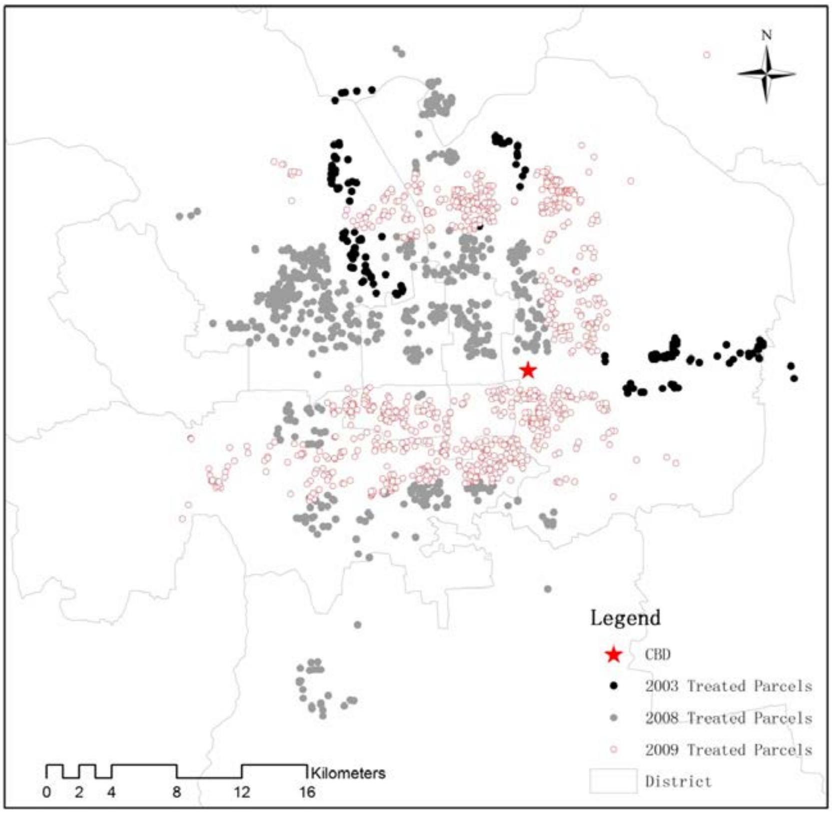

Figure 1 Spatial distributions of treated residential land parcels

Notes.---"2009 Treated Parcels" refer to the parcels in the $Treatment_3$ (station $\geq$ 2009); In comparison to the $Treatment_3$, "2008 Treated Parcels" are the additional parcels that belong to the $Treatment_2$ (station $\geq$ 2008). In comparison to the $Treatment_2$, "2003 Treated Parcels" are the additional parcels that belong to the $Treatment_1$ (station $\geq$ 2003). All treated parcels are selected using the 2km distance band.

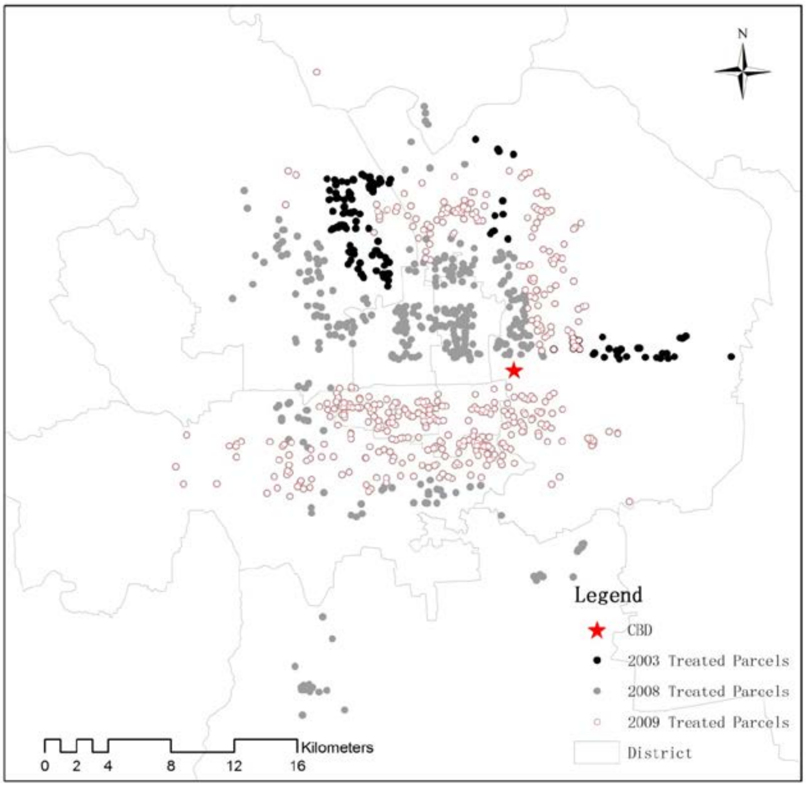

Figure 2 Spatial distributions of treated commercial land parcels

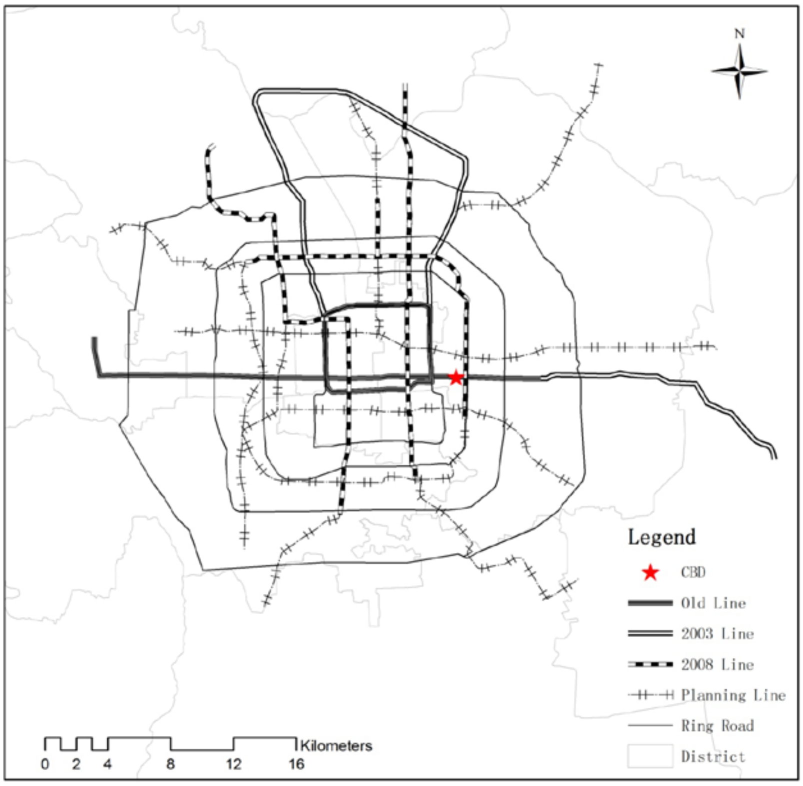

Figure 3 Rail transit network in the Beijing urbanized area

Appendix A

Appendix Table 1 Descriptive statistics of variables

Variables	Definition	Residential land sample Mean/ (Std.Dev)	Commercial land sample Mean/ (Std.Dev)
Dependent Variable			
Land Price	Ln (Land parcels' leasing price per square meter (RMB/sq.meter))	7.45(1.08)	7.76(1.42)
Locational-specific Variables			
CBD	Ln (Distance between a land parcel and CBD (meters))	9.03 (0.64)	8.85(0.75)
Land parcel size	Ln (The area of a land parcel (m^2))	9.06 (1.34)	7.59(1.78)
Park	Ln (Distance to the nearest park (meters))	7.77 (0.72)	7.61(0.81)
River	Indicator of proximity to rivers (<500 meter)	0.18 (0.38)	0.11(0.31)
Air quality	Indicator of Air pollution index to each parcel	1.93 (0.87)	1.99(0.88)
Bus	Ln (Distance to the nearest bus stop (meters))	6.03(0.82)	6.12 (1.06)
Expressways	Ln (Distance to the nearest expressway (meters))	6.43(1.14)	6.36 (0.98)
School	Ln (Distance to the nearest middle school*school rank)	25.01 (5.68)	24.34(6.34)
Employment	Indicator of employment accessibility (see McMillen (2001))	0.04(0.05)	0.06(0.07)
Density	Population density per zone (1,000 people per km^2)	2.37 (3.35)	2.76(4.35)
Heritage	Ratio of heritage architectures built before 1949per zone (%)	0.03(0.09)	0.07(0.14)
Education Attainment	Median resident educational attainment in each zone:1=middle school or lower;2=high school;3=university;4=post graduate	1.715(0.508)	1.91(0.46)
Crime	Number of crimes per 1000 people per zone	5.335(6.655)	4.08(5.15)
Public Housing	Percentage of people renting public housing per zone	0.31(0.20)	0.33(0.21)