ELSEVIER

#### Contents lists available at ScienceDirect

### **Transport Policy**

journal homepage: www.elsevier.com/locate/tranpol



# Subways near the subway: Rail transit and neighborhood catering businesses in Beijing



Siqi Zheng<sup>a</sup>, Xiaoke Hu<sup>a</sup>, Jianghao Wang<sup>b</sup>, Rui Wang<sup>c,\*</sup>

- <sup>a</sup> Hang Lung Center for Real Estate and Department of Construction Management, Tsinghua University, China
- <sup>b</sup> Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, China
- <sup>c</sup> Luskin School of Public Affairs, University of California, Los Angeles, United States

#### ARTICLE INFO

Article history: Received 30 June 2015 Received in revised form 14 March 2016 Accepted 18 March 2016 Available online 7 April 2016

Keywords: Subway Catering Local consumer amenity Beijing

#### ABSTRACT

Beijing has made enormous investment in rail transit since the late 2000s. The rapidly growing subway system greatly improves the accessibility of neighborhoods nearby subway stations and often increases neighborhood population and employment densities, both resulting in a larger market for local retail businesses. While numerous studies have shown that rail transit investment tends to raise property value close to stations, few have provided direct evidence on rail transit's effects on local consumer amenities. Using citywide catering establishment data since 2004 from dianping.com (China's yelp.com), we study the effects of new subway stations on catering openings, diversity and consumer demand in neighborhoods near a subway station opened during 2004–2013. We find that a new subway station positively contributes to the quantity, diversity and consumer demand of nearby food and beverage services. These effects are heterogeneous spatially and in terms of catering types. This study enriches the limited extant empirical evidence on urban rail transit's impact on local economic activities and consumer amenities. In China, where unprecedented rail transit expansion has transformed large cities like Beijing in many ways, our findings can help us better understand how major public investment in cities affects local economy, quality of life, the housing market and related further policy concerns.

© 2016 Elsevier Ltd. All rights reserved.

#### 1. Introduction

Urban rail transit provides convenient access to important destinations in cities, especially for users residing or working within a short distance from a station. Attention has been paid to the connection between rail transit's accessibility benefits and local property values, development activities and job growth (as reviewed in the following section). However, there is a scarcity of knowledge on rail transit's impacts on local consumer amenities. This prevents us from a deeper understanding of the role rail transit investment plays in cities. As pointed out in Glaeser et al. (2001), the rise of the "consumer city" in an era of increasingly mobile firms is inevitable because talents and hence their employers are attracted to cities with more consumer amenities. At the neighborhood level, important consumer amenities such as retail outlets may contribute to community revitalization even it does not dramatically increase labor market outcomes, such as employment, job participation or wage level (Chapple and Jacobus, 2009).

This study attempts to provide a detailed evaluation on the effects of urban rail transit on neighborhood consumer amenities

that are primarily provided by the private sector (rather than publicly supplied amenities such as parks). We apply data during the decade from 2004 to 2013, a time period when Beijing's subway system received hundreds of kilometers of new lines. Such a rapid expansion of subway at a cost of tens of billions of RMB (1 RMB = about 0.15 USD) each year has helped Beijing to meet its increasing travel demand and relieve surface traffic congestion. The subway expansion has also provided enormous accessibility benefits to the city's businesses, especially in neighborhoods nearby the stations. As one would expect, such accessibility benefits should not only reflect in property prices, but also induce more local economic activities that can provide additional local consumer amenities.

We use panel data from station areas (defined by a range of distances from a subway station) where the stations opened during our study period to examine the response of a broad category of catering (food and beverage) business activities to new subway service. The catering industry is selected because it represents perhaps the most important category of local consumer amenity provided largely by the private sector – it not only meets the daily needs of visitors as well as local residents and workers, but also facilitates face-to-face social interactions, even in a fast food

<sup>\*</sup> Corresponding author.

restaurant like a Subway store (an American sandwich chain brand). The differentiated timing of subway construction during this study period allows station areas to serve as both the treatment and control groups in different years for statistical identification. We measure both the supply aspect (number and diversity) and the demand aspect (customer volume) of catering businesses by year in each neighborhood using online data from dianping.com, China's Yelp. We find that subway construction increases the quantity (the number of new catering openings), diversity (the composition of different cuisine genres) and customer volume (proxied by the number of online customer reviews on dianping.com) of catering businesses around stations, with interesting heterogeneities in effects spatially and by catering types.

This paper adds to the limited available evidence on the changes in local consumer amenities following rail transit investments. The rest of the paper is organized as follows: Section 2 reviews relevant literature, Sections 3 and 4 describe data and empirical strategy. Section 5 presents the results, followed by conclusions and discussions in Section 6.

#### 2. Literature review

It is well known that at the city and urban region levels, rail transit investment contributes to productivity gains (see, e.g., Graham, 2007; Haddad et al., 2015). At the local or neighborhood level, much attention has been focused on property values nearby rail transit projects, especially by economists (see literature reviews by Debrezion et al., 2007; Mohammad et al., 2013). Overall, the extensive literature on the property value capitalization of urban rail transit investment, mostly from the US and other industrialized countries, has found generally positive price premiums of properties closer to rail transit stations, although recent development of this literature has pointed to some structural reasons of the wide variations in the estimated property value effects such as land supply elasticity (e.g., Zheng et al., 2014, Sun et al., 2015).

Besides property value impacts, planners have examined categories of broader impacts such as land development and employment. Giuliano (2004) systematically reviews studies of important urban rail systems such as San Francisco's BART, Washington, D.C.'s METRO, and Atlanta's MARTA from the 1970 s to the early 2000 s. She concludes that rail transit's local land use, land value, population, and employment effects are highly localized and uneven across systems and station areas. The variation in local impacts seems to be affected by whether rail transit has a significant impact on local accessibility (e.g., impacts tend to differ between fast-growing and heavily congested areas and areas dominated by automobiles), local zoning and transportation policies, and availability of development subsidies. Similar conclusions appear in more recent studies such as Mejia-Dorantes and Lucas (2014) review of the property value, development density and employment effects of London's Jubilee Line Extension and the Madrid Metrosur.

However, relative to property value and land use effects, there seems to be limited evidence on the changes in local consumer amenities following rail transit investments. The available studies on rail transit and neighborhood amenities primarily focus on local retail activities measured by either the employment or the quantity (density) of retail businesses. The theoretical underpinning of this is obvious. Rail transit decreases transportation costs between its station areas and the rest of city and increases local consumer base as more jobs and housing agglomerate nearby station areas. Increases in both local demand density and neighborhood accessibility result in significantly expanded market potential of local retail businesses. As a result, the number and diversity of local

retail businesses increase as suggested by standard economic theory of firm location choice (Berry, 1967; Stern, 1972; Fischer and Harrington, 1996).

Applying a simultaneous model of census tract population and employment data in 1980 and 1990, Bollinger and Ihlandfeldt (1997) find no impact of Atlanta's MARTA (built between 1979 and 1990) on either commercial or total employment in station areas after controlling for a range of neighborhood land-use and socioeconomic characteristics. In fact, they find that MARTA has altered the composition of employment in station areas in favor of the public sector. In four large California metropolitan areas, Schuetz (2015) examines whether rail and bus rapid transit (BRT) investment stimulates retail employment within a quarter mile from rail and BRT stations built during 1992-2009. She doesnot find any positive effect on retail employment in rail/BRT station neighborhoods. A major explanation for the insignificant retail effect of rail transit provided by Bollinger and Ihlandfeldt (1997) and Schuetz (2015) is the minimal role of rail transit in urban transportation in either the Atlanta region or the large California cities.

However, using panel data of annual census tract retail employment density during 1991-1994 also in Atlanta, Bowes and Ihlanfeldt (2001) estimate a random effects model and find that tracts between one-quarter and one-half mile of a station experience the greatest increase in retail employment density, an effect increases with distance from the city center. Bowes and Ihlanfeldt's findings, while different from those of Bollinger and Ihlandfeldt (1997) and Schuetz (2015), are supported by a recent study in Beijing, China. Among the first to relate rail transit with nearby private consumer amenities in less developed countries, Zheng and Kahn (2013) find that the construction of five subway lines in Beijing during 2003-2009 associates with more chain restaurants (mainly catering to the upper-middle class consumers) nearby subway stations. Their finding, while limited to high-end restaurant businesses, is not surprising because of the more salient role played by rail transit in urban passenger transportation of the dense cities in a developing economy. Compared to the automobile-dominant large cities in the US, rail transit carries a significant share of riders from all socioeconomic classes in Beijing, one of the most motorized cities in China. For example, three out of the four largest California cities have a rail mode split of about 0.5% among commuters during 2005-2009 (Schuetz, 2015), while Beijing's subway carries almost one third of commuters.

#### 3. Background and data

Beijing has experienced massive subway construction since the late 2000 s. From 2007 (all the pre-2007 subways were built before 2004) to 2013, 12 new subway lines were put into use, increasing the total number of subway stations from 70 to 274. By the end of 2013, Beijing's subway reached a total revenue length of 476 km, with a typical daily ridership of more than 10 million. The rapid expansion of Beijing's subway system has been reshaping the intra-city geography of jobs, housing and consumption activities. In this paper, we call a subway station a new station if it was opened in 2007 or later.

Our analysis focuses on the impacts of subway development on neighborhood catering businesses in Beijing during the decade of 2004–2013. We rely on two datasets including: first, when and where a new subway station was opened during the study period; and second, annual catering openings, diversity of catering businesses and the number of online customer reviews in the new subway stations' neighborhoods.

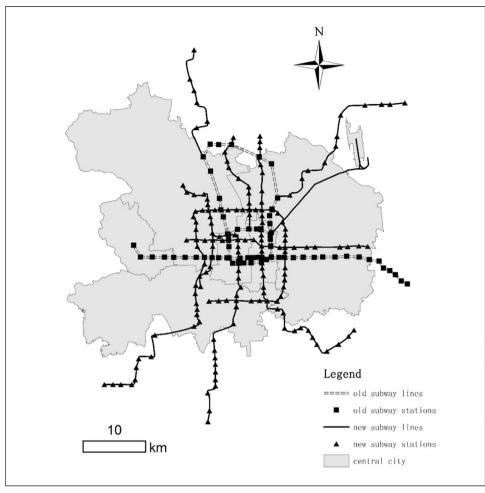


Fig. 1. Beijing's subway system at the end of 2013.

#### 3.1. New subway lines and subway neighborhoods

Fig. 1 shows the Beijing subway system as of December 2013. The triangles and squares represent subway stations. The dashed double lines connecting squares are old subway lines built before 2004. The solid lines connecting triangles are those built during our study period, 2004–2013. It is worth noting that the new subway lines (solid lines) were built at least three years after the old lines (dashed double lines). There may be some systematic differences between the neighborhoods of the old subway stations and those of the new stations, such as the maturity of residential and commercial development. Therefore, we restrict our sample to the 154 new subway station neighborhoods. We define the six central districts (*Xicheng, Dongcheng, Haidian, Chaoyang, Shijingshan, Fengtai*) as the urban area of Beijing. These districts are shown with a darker shade in Fig. 2. Beijing's subway system mainly serves its urban area.

The key spatial unit of analysis in our study, the subway neighborhood, is a buffer around each of the new subway stations. To identify the spatial influence area of a new subway station and how such influence varies when moving a little further away from a station, we construct three buffers for each new station i: a 400-meter (about 0.25 mile) core buffer around the station ( $buffer_{400.i}$ ) and two "donut" buffers including one between 400 and 800 m ( $buffer_{400-800.i}$ , about 0.25–0.5 mile) and the other between 800 and 1200 m ( $buffer_{800-1200.i}$ , about 0.5–0.75 mile) from the station.

We construct a longitudinal dataset of the 154 new subway neighborhoods by year between 2004 and 2013. A key dummy variable,  $connect_{it}$ , indicates whether subway station i at the center

of the neighborhood is opened in year t. This dummy variable takes the value of one starting from the next calendar year after a new station is open (most of the new subway lines opened toward the end of a year, see Appendix A). This is a necessary approximation because of the annual data we use. We have two timeinvariant control variables pop\_density and d\_cbd. pop\_density measures the year 2000 population density in 10,000 people/km<sup>2</sup> of the sub-urban district zone (jiedao) in which a subway neighborhood locates. d\_cbd is a station's distance to city center (Tian'anmen square) to control for unobserved neighborhood characteristics (e.g., proximity to major destinations and local public amenities such as good schools that predominantly concentrate near the city core) and captures the global negative gradients of economic activities from the city center to periphery. We define the area within the Fourth Ring Road as "inner city" and the area outside as "outer city" when examining the heterogeneous effects of subway station opening on catering activities in different regions.

## 3.2. The supply and demand of catering services in subway neighborhoods

Our data of catering establishments comes from *dianpin.com* (the Chinese equivalent to *Yelp*). We obtain the detailed information of each of the 86,456 catering establishments in Beijing from this website, including each business's name, address, date of establishment, number of online reviews by quarter, and cuisine genre. We geocode all catering establishments to construct their quantity and diversity measures by subway neighborhood by year

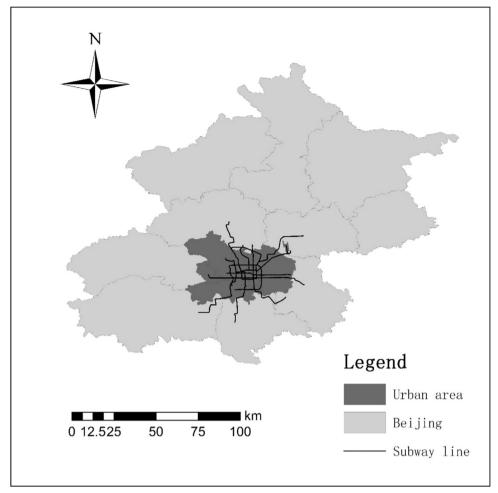


Fig. 2. The urban area of Beijing.

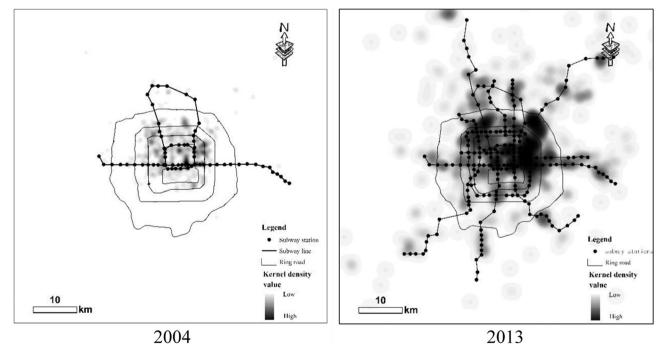


Fig. 3. Spatial distribution of catering establishments.

as described in detail below. Fig. 3 presents the overall spatial distribution of catering establishment density and its relationship with subway service in Beijing in 2004 and 2013. We can observe the emergence of many hotspots of catering businesses along the newly constructed subway lines.

Furthermore, we classify the catering establishments into two subgroups: sit-down establishments and fast food-style ones (such as the American sandwich store "Subway"). Many subway riders may be in a rush so they have a high opportunity cost of time. Instead of ordering and consuming food after sitting down, they likely prefer to stop by a fast food/beverage shop shortly or just to get some take-out food and beverage and continue their journey. On the other hand, some subway riders may come to a station neighborhood with the purpose to spend some time sitting down and enjoying catering services, sometimes with family, friends or business partners. In this case, they usually schedule more time and will be more selective of food and/or beverage shops in a larger geographic area around the subway station. Therefore, we expect that convenience-oriented fast food-style establishments have a higher demand in the immediate adjacency of subway stations. While enjoying the accessibility benefits of subway, sitdown catering services that better facilitate face-to-face interactions may have a flatter demand gradient from a subway station. Fig. 4 graphs the annual shares of fast-food versus sit-down catering establishments in the 0–400 m buffers in the neighborhoods studied, as more and more neighborhoods received subway services over time. There seems to be a trend that more fast-food establishments clustered around subway stations over time.

We use two types of measures – catering openings and online reviews – to quantify the responses of catering industry to a new subway station. The opening of a subway station raises consumer demand in its vicinity due to the improvement in accessibility. For the first quantity measure, we count the annual catering openings in the 0–400 m buffer (openings<sub>400, it</sub>), the 400–800 m buffer (openings<sub>400-800, it</sub>) and the 800–1200 m buffer (openings<sub>800-1200, it</sub>) of each station. We also count the corresponding openings in the two catering subgroups (sit-down establishments and fast foodstyle ones) defined above. We expect the number of new catering businesses (openings) by buffer to be more directly linked to the opening of a new subway station, while the total number of existing catering businesses may be affected by many other historically dependent factors.

However, the number of catering openings does not distinguish large or more popular businesses from the small or less popular businesses. Previous studies (e.g. Bakhshi et al., 2014) have employed online review counts to proxy consumer demand of restaurants. For the second quantity measure, we count the annual online reviews of catering establishments in the *0*–400 *m* buffer (reviews<sub>400, it</sub>), the 400–800 *m* buffer (reviews<sub>400-800, it</sub>) and the

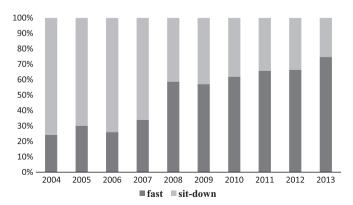


Fig. 4. Shares of fast-food and sit-down catering businesses in  $buffer_{400}$  by year.

800-1200~m buffer ( $reviews_{800-1200, it}$ ). We construct two variables for each buffer: total reviews and reviews per restaurant. The former proxies for the total customer demand in a buffer, while the latter proxies for the average "customer flow intensity" per catering establishment in that buffer.

It is worth noting that not all customers write reviews online after a visit, therefore our review counts can only be regarded as a proxy of customer flows. If the shares of online reviewers among restaurant customers are relatively stable across different buffers, station neighborhoods and over time, such measures are effective proxies for customer demand. Caution should be exercised when interpreting relevant results in our study.

In addition to quantity, the diversity of local catering businesses will likely respond to a new subway station. Adding a subway station increases the market size of local catering businesses because people in a larger area of the city are able to access these local businesses via the accessibility provided by the subway system. This larger market size will bring about a variety of catering businesses to respond to the multiple preferences of potential customers (Schiff, 2014).

We construct two catering cuisine-type-based diversity indices employing the definitions of Shannon index and Gini-Simpson index, both frequently used to measure local species diversity in biostatistics. Because the number of new openings varies significantly across neighborhoods and over time and is sometimes too small (e.g., zero), we use the stock of existing catering establishments in order to obtain more robust measures of catering diversity in each station neighborhood over time.

The first diversity index is based on the Shannon index (Blair, 1996). Its value in station neighborhood i, buffer k and year t is calculated as:

diversity 
$$I_{ikt} = -\sum_{m=1}^{n} \frac{N_{ikmt}}{N_{ikt}} \times \log\left(\frac{N_{ikmt}}{N_{ikt}}\right)$$
 (1)

The second diversity index is the Gini-Simpson index (Simpson, 1949) defined as:

$$diversity2_{ikt} = 1 - \sum_{m=1}^{n} \left(\frac{N_{ikmt}}{N_{ikt}}\right)^{2}$$
(2)

In the above two equations,  $N_{ikt}$  is the total number of catering establishments in neighborhood i, buffer k and year t,  $N_{ikmt}$  is the number of existing catering establishments offering cuisine m in neighborhood i, buffer k and year t. For each of the two diversity indices, we use reported cuisine genre (49 cuisines altogether, see Appendix B) of each establishment in dianping.com. Both indices increase if there are more cuisine types/genres and if the shares of each cuisine type/genre are more even. The use of two diversity indices allows us to validate and test the sensitivity of our results.

The summary statistics of all variables are presented in Table 1. The average numbers of annual catering openings in the 0-400 m, 400-800 m and 800-1200 m buffers are 3.22, 5.52 and 2.92, respectively. The corresponding standard deviations for the three buffers are 5.99, 10.50 and 6.16, suggesting the dispersed distribution of the annual number of catering openings. Within each buffer, the average number of fast food openings is more than twice that of sit-down food openings (2.03 vs. 0.84 in the 0–400 mbuffer, 3.35 vs. 1.45 in the 400–800 m buffer and 1.69 vs. 0.82 in the 800–1200 m buffer). The total count of online reviews per station neighborhood also shows a dispersed distribution - average review counts are 327.01, 414.76 and 209.94 with standard deviations of 857.96, 989.41 and 596.77 in the 0-400 m, 400-800 m and 800-1200 m buffers, respectively. The average count of online reviews per business is higher when located closer to station (18.14 in the 0-400 m buffer, 17.16 in the 400-800 m buffer and 12.62 in

**Table 1**Variable definitions and summary statistics.

Variable		Definition	Obs.	Mean	Std.	Min	Max
openings	0–400 m	The number of catering openings nearby a station, by station-year	1540	3.22	5.99	0	70
	400–800 m		1540	5.52	10.50	0	143
	800–1200 m		1540	2.92	6.16	0	63
openings_sit	0–400 m	Sit-down catering openings	1540	0.84	1.67	0	22
	400–800 m		1540	1.45	2.25	0	27
	800–1200 m		1540	0.82	1.54	0	12
openings_fast	0–400 m	Fast-food catering openings	1540	2.03	4.31	0	39
	400–800 m		1540	3.35	7.49	0	95
	800–1200 m		1540	1.69	4.35	0	48
reviews_total	0–400 m	The total count of online reviews nearby a station, by station-year	1540	327.01	857.96	0	14777
	400–800 m		1540	414.76	989.41	0	16572
	800–1200 m		1540	209.94	596.77	0	8603
reviews_avg	0–400 m	The average count of online reviews per establishment nearby a station, by station-year	1540	18.14	30.41	1	370
	400–800 m		1540	17.16	23.09	1	276
	800–1200 m		1540	12.62	22.24	1	393
diversity1	0–400 m	The Shannon diversity index for 49 cuisines nearby a station, by station-year	1540	1.09	0.98	0	2.93
	400–800 m		1540	1.35	1.03	0	3.10
	800–1200 m		1540	0.96	1.02	0	3.11
diversity2	0–400 m	The Gini-Simpson diversity index for 49 cuisines nearby a station, by station-year	1540	0.48	0.39	0	0.93
	400–800 m		1540	0.56	0.38	0	0.95
	800–1200 m		1540	0.41	0.40	0	0.94
connect		Whether a subway station is in use: YES=1; $NO=0$	1540	0.30	0.46	0	1
pop_d		Average population density in the jiedao where a station locates (10000/km2)	1540	1.39	0.92	0.05	4.14
d_cbd		Distance to city center (m)	1540	8667	4038	1636	19163
suburb		Whether a subway station is located in the outer city: $YES=1$ ; $NO=0$	1540	0.25	0.43	0	1

the 800-1200~m buffer). As for diversity indices, both the Shannon index and the Gini-Simpson index do not have a clear relationship with the distance to subway station (1.09, 3.35 and 1.45 in the 0-400~m, 400-800~m and 800-1200~m buffers for the Shannon index; 0.48, 0.56 and 0.41 in the 0-400~m, 400-800~m and 800-1200~m buffers for the Gini-Simpson index). The average population density by sub-district zone (jiedao) is 13,900/km² in the subway neighborhoods. The average distance of a new subway station to city center is 8.7 km.

#### 4. Empirical strategy

To identify the impact of a new subway station opening on the quantity of catering openings in its neighborhood, we employ a difference-in-differences (DID) strategy by estimating the two-way fixed-effects model below:

$$openings_{ikt} = f(\alpha + \beta_k^{o}connect_{it} + D_i + T_t + \varepsilon_{ikt})$$
(3)

The dependent variable,  $openings_{ikt}$ , measures the number of catering openings in station neighborhood i, buffer k and year t. We include both year fixed effects  $T_t$  and station (neighborhood) fixed effects  $D_i$ .  $connect_{it}$  is a dummy variable that equals one for years after the introduction of station i. The treatment effect,  $\beta_k^o$ , is estimated by subtracting the average time trend of neighborhood catering openings from the average before-after difference in neighborhood catering openings.

The identification of the treatment effect  $\beta_k^o$  in Eq. (3) relies on the assumption that the neighborhood-year observations before receiving a subway station can be considered as the counterfactual of the observations with a subway station. While we have included station fixed effects to control for time-invariant unobservables by station, the DID strategy may still result in a biased estimate of  $\beta_k^o$  if there is any time-variant unobservable associated with the opening of new subway stations. To mitigate the potential problem of time-variant unobservables, we restrict our sample to the station neighborhoods around the new subway stations constructed between 2004 and 2013. These new station neighborhoods should be more homogenous compared to the old subway neighborhoods that have already became well-developed retail hubs.

Within the new subway station neighborhoods, a selection issue may also exist because the neighborhoods receiving subway stations earlier are systematically different from those receiving stations later. For instance, the city government may strategically decide to first build subway lines in the fast-growing areas with anticipated travel demand increase, or instead, in the lagged-behind areas to achieve the goal of more balanced urban development. It will be difficult to measure all the differences between the neighborhoods connected to subway early and late, thus the problem of reversed causality remains a concern. But the below facts about the planning and implementation of Beijing's subway system may alleviate this concern.

A careful read of the Plan for Rapid Rail Transit Network in Beijing approved by the State Council in 1993 suggests the main purpose of building subway lines is to "connect the central city and the suburban area and thus facilitate employment and residential decentralization." As a result, most new subway lines were designed to start from the central city and extend to the urban fringe. So there should be significant variation in urban development maturity along a single subway line. The fact that all stations along a line opened at the same time reduces the possibility that the government just focused on places with stronger demand for retail development when it decided when and where to construct new subway stations. We also find that the 11 new subway lines constructed in our study period exactly followed the original subway construction plan in early 1990s. But the spatial distribution of retail activities in Beijing very likely changed significantly from the early 1990s to the period of study. Our interviews with the local officials in Beijing's urban planning and transportation planning bureaus suggest that the sequence of building subway lines in the last two decades was mainly determined by the transportation investment budget of the municipal government. Therefore, both the "early-determined" subway construction plan and the relatively "exogenous" construction sequence may relieve the concern of the potential reversed causality problem. Furthermore, the fact that we "lag" the neighborhood business activities by one year when relating them to the opening of a subway station (a subway neighborhood is considered as "connected" in the next calendar year after the station opens) also helps mitigate the potential concern of reversed causality.

We employ the negative binomial regression technique to estimate Eq. (3) because a count model fits the data better than a linear regression model when the dependent variable takes the value of dispersed non-negative integers.

In the two-way fixed effects model we include a large number of subway station and year fixed effects, which significantly reduces the degree of freedom in our analyses. So we provide three additional specifications. First, we relax the station fixed effects slightly by replacing them with "subway line segment" fixed effects, and at the same time we include local population density (pop\_density) and the station's distance to CBD (d\_cbd) as two station-specific time-invariant variables. Beijing is a monocentric city with multiple ring roads (the Second to Sixth Ring Roads) from inner city to outer city and the radial subway lines. Therefore, we partition the new subway lines with the ring roads to obtain "subway line segments". An average segment has 4.3 new stations on it. In the second specification, we keep the station fixed effects but replace the year fixed effects with district-specific time trends. In the third specification, we keep all fixed effects but use a linear instead of negative binomial estimator. This allows us see if a random effects model instead of the more restrictive fixed effects DID model can be used. We take logarithm of the dependent variable plus one  $(\log(openings + 1))$  so that we will not lose observations due to the fact that some station-year observations have zero catering opening.

In all the model specifications above, we expect a positive coefficient of *connect* ( $\beta_k^0$ ) because all else being equal, new catering openings should cluster toward a neighborhood with enhanced subway accessibility. By comparing the coefficients of *connect* in different buffers, we can see how the impact of a newly opened subway station on catering openings varies at different distances from the station. We expect that closer buffers gain more catering openings from a new subway station (i.e.,  $\beta_{0-400}^o > \beta_{400-800}^o > \beta_{800-1200}^o$ ).

We apply the same strategy to study the impacts of a new subway station on the number of online reviews as a measure of customer volume and on the catering diversity in its neighborhood. We use the below panel linear regression model:

$$\ln(reviews_{ikt}) = \alpha + \beta_{\nu}^{r} connect_{it} + D_{i} + T_{t} + \varepsilon_{ikt}$$
(4)

reviews $_{ikt}$  measures the total count of online customer reviews and serves as the proxy for the total "customer flow" in station neighborhood i, buffer k and year t. In the fixed effects model we control for station fixed effects  $D_i$  (or alternatively, subway line segment fixed effects). In the random effects model, we replace  $D_i$  with  $X_i$  that includes the logarithms of local population density  $(pop\_density)$  and a station's distance to CBD  $(d\_cbd)$ . The use of the logarithms of the dependent and some independent variables

**Table 2**Effects of a new subway station on nearby catering openings by distance to station.

2A. Negative binomial regression	ıs								
Dep. Var.: openings	(1) 0-400 m	(2) 400-00 m	(3) 800–1200 m	(4) 0-400 m	(5) 400–800 m	(6) 800–1200 m	(7) 0–400 m	(8) 400–800 m	(9) 800–1200 m
connect	0.393*** (0.113)	0.238** (0.0953)	- 1.37e-05 (0.125)	0.262*** (0.0816)	0.322*** (0.0647)	0.250*** (0.0740)	0.0154 (0.0796)	0.0811 (0.0641)	- 0.0591 (0.0760)
$ln(d\_cbd)$	-0.0519 (0.172)	0.0241 (0.159)	0.129 (0.244)	(,	,	(333 2)	(,	,	(**************************************
ln(pop_d)	0.354*** (0.0645)	0.328*** (0.0677)	0.401*** (0.0686)						
constant	- 1.783 (1.721)	- 1.106 (1.553)	- 1.928 (2.379)	-442.7*** (45.44)	-422.3*** (33.43)	-428.0*** (39.18)	- 1.942*** (0.477)	-0.524** (0.209)	-0.371* (0.223)
year dummies	YES	YES	YES	NO	NO	NO	YES	YES	YES
subway station dummies	NO	NO	NO	YES	YES	YES	YES	YES	YES
subway line segment dummies	YES	YES NO	YES NO	NO YES	NO	NO YES	NO NO	NO NO	NO NO
district*year trend N	NO 1540	NO 1540	NO 1540	1540	YES 1540	1540	NO 1540	NO 1540	NO 1540
alpha	0.9414	0.8211	1.6393	0.9414	0.8211	1.6393	0.2705	0.1437	0.0966
LR test (alpha=0)	2005.24	3694.33	2918.09	2005.24	3694.33	2918.09	419.29	374.41	94.86
Prob > chi2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2B. Panel data linear regressions									
Dep. Var.: ln(openings+1)	(1)	(2 0–400 m	2)	(3) (4) 400–800 m		(5)	800-120	(6) 00 m	
Panel setting	fixed effec	t <b>r</b> c	ındom effect	fixed ef	fect	random effect	fixed 6	effect	random effect
connect	0.253*** (0.0694)	0.	299*** ).0733)	0.227** (0.0712	*	0.244*** (0.0745)	0.0865 (0.055	5	0.0905 (0.0587)
$\ln(d\_cbd)$	(******,	_	0.304*** 0.106)	(	•	-0.0904 (0.134)	(	,	0.215 (0.140)
$ln(pop\_d)$			198*** 1.0483)			0.335*** (0.0593)			0.240*** (0.0603)
constant	0.480*** (0.0431)	3.	208*** 0.948)	0.678** (0.0429		1.506 (1.202)	0.453° (0.040		_ 1.452 (1.255)
year dummies	YES		ES	YES		YES	YES		YES
subway station dummies	YES	N		YES		NO	YES		NO
N D	1540	15	540	1,540		1540	1540		1540
R-squared Prob > chi2	0.432 0.1902	Δ.	1902	0.545 0.9305		0.9305	0.412 1.000	n	1.0000
1100 / CIIIZ	0.1302	U.	1302	0.5303		0.5305	1.0000	,	1.0000

Notes: robust standard errors in parentheses; \*\*\*p < 0.01, \*\*p < 0.05, \*p < 0.1; prob > chi2 shows the p-value of Hausman test.

allows us have distributions of variables that are closer to normal and obtain elasticity estimates. We also expect a positive coefficient of connect (  $\beta_k^r$  ) because catering businesses located in a closer buffer likely gain more customers with the opening of a new subway station (i.e.,  $\beta_{0-400}^r > \beta_{400-800}^r > \beta_{800-1200}^r$ ). Alternatively, we use the average count of reviews per business ( $reviews\_avg$ ) to see whether the "customer flow intensity" shows consistent patterns.

For the catering diversity model, we estimate:

$$diversity_{ikt} = \alpha + \beta_k^d connect_{it} + D_i + T_t + \varepsilon_{ikt}$$
 (5)

diversity<sub>ikt</sub> is either the Shannon index or the Gini-Simpson index in station *i*, buffer *k* and year *t*. We expect a positive coefficient of connect ( $\beta_k^d$ ) and  $\beta_{0-400}^d > \beta_{400-800}^d > \beta_{800-1200}^d$  as well. Similar as in Eq. (4), in the random effects model we will use the logarithms of the variables in Xi.

We make two efforts to explore the heterogeneity in the above treatment effects. First, we compare the quantity increases of fast-food and sit-down establishments in the new subway station neighborhoods. We expect that fast-food establishments benefit more and grow faster when a subway station is connected. Whether a new subway neighborhood gains more in fast-food or sit-down establishments is interesting because these two categories of catering services provide different amenities and can have different neighborhood traffic and socio-environmental impacts. Second, we split the observations into those in the inner city (within the Fourth Ring Road, including 116 station neighbor-

hoods) and those in the outer city (outside the Fourth Ring Road, including 38 station neighborhoods). We expect that the subway neighborhoods in the outer city will benefit more because they gain a relatively larger marginal increase in transportation accessibility after being connected to the subway network.

#### 5. Results

Table 2 reports the estimates of new catering openings in subway neighborhoods at different buffer distances using Eq. (3).

Panel 2A estimates three specifications of negative binomial regressions. The first specification, reported in columns (1)–(3), includes year fixed effects, subway line segment fixed effects and two station attribute controls. When a subway station opens, we observe significant rise of new catering businesses in the two inner buffers, and the closest buffer has a larger increase: an increase of 48.1% ( $e^{0.393}-1$ ) in the 0-400 m buffer and an increase of 26.9% ( $e^{0.238}-1$ ) in the 400-800 m buffer. The farthest buffer (800-1200 m) does not have a significant gain in catering establishments. The second specification, reported in columns (4)–(6), includes station fixed effects and district-specific time trends. The coefficient of *connect* is significant in all three buffers, and the 400-800 m buffer has a slightly larger increase than the closest and farthest buffers. The third specification, reported in columns (7)–(9), includes both station fixed effects and year fixed effects. As discussed earlier, the reduced degree of freedom prevents us

**Table 3**Effects of a new subway station on the count of reviews of nearby catering by distance to station.

3A. Count of total reviews						
Dep. Var.: ln(reviews_total+1)	(1)	(2) -400 m	(3)	(4) 0–800 m	(5)	(6) 0–1200 m
Panel setting	fixed effect	random effect	fixed effect	random effect	fixed effect	random effect
connect	0.641***	0.705***	0.499***	0.532***	0.0653	0.0767
1-74 -4.4	(0.155)	(0.160)	(0.141)	(0.145)	(0.149)	(0.153)
$ln(d\_cbd)$		- 1.118*** (0.356)		- 0.429 (0.320)		0.754* (0.408)
ln(pop_d)		0.847***		1.052***		0.780***
m(pop_u)		(0.168)		(0.161)		(0.183)
constant	0.715***	10.76***	1.014***	4.913*	0.566***	- <b>6.127</b> *
	(0.131)	(3.195)	(0.123)	(2.870)	(0.134)	(3.649)
year dummies	YES	YES	YES	YES	YES	YES
subway station dummies	YES	NO	YES	NO	YES	NO
N	1540	1540	1540	1540	1540	1540
R-squared Prob > chi2	0.564 0.0065	0.0065	0.545 0.0818	0.0818	0.531 0.9993	0.9993
Dep. Var.: ln(reviews_avg)	(1) 0-40	(2) 0 m	(3)	(4) -800 m	(5)	(6) 0–1200 m
Panel setting	fixed effect	random effect	fixed effect	randomeffect	fixed effect	random effect
connect	0.285***	0.328***	0.196**	0.225***	0.0337	0.0442
	(0.0893)	(0.0907)	(0.0829)	(0.0843)	(0.0926)	(0.0950)
ln(d_cbd)		-0.646***		-0.348**	, ,	0.428*
		(0.206)		(0.176)		(0.225)
ln(pop_d)		0.497***		0.533***		0.417***
	0.385***	(0.0995) 6.194***	0.521***	(0.0873) 3.666**	0.279***	(0.105) 3.520*
constant	(0.0802)	(1.846)	(0.0735)	(1.585)	(0.0797)	- 3.520 (2.013)
	(0.0002)	(1.040)	(0.0733)	(1.505)	(0.0737)	(2.013)
year dummies	YES	YES	YES	YES	YES	YES
subway station dummies	YES	NO	YES	NO	YES	NO
N B. amusanad	1540	1540	1540	1540	1540	1540
R-squared Prob > chi2	0.435 0.0581	0.0581	0.513 0.0680	0.0680	0.531 0.9976	0.9976
1 100 / CIIIZ	0.0301	0.0361	0.0000	0.0000	0.5570	0.3370

Notes: robust standard errors in parentheses; \*\*\* p < 0.01, \*\* p < 0.05, \* p < 0.1; prob > chi2 shows the p-value of Hausman test.

**Table 4**Effects of a new subway station on nearby catering diversity by distance to station.

4A. Shannon index						
Dep. Var.: diversity1	(1)	(2) -400 m	(3)	(4) 0–800 m	(5)	(6) 0–1200 m
Panel setting	fixed effect	random effect	fixed effect	random effect	fixed effect	random effect
connect	0.242***	0.266***	0.165**	0.180**	0.0393	0.0454
	(0.0690)	(0.0709)	(0.0694)	(0.0704)	(0.0637)	(0.0656)
ln(d_cbd)		<b>-0.401***</b>		<b>-0.125</b>		0.242
		(0.104)		(0.120)		(0.153)
ln(pop_d)		0.247***		0.355***		0.276***
		(0.0546)		(0.0567)		(0.0688)
constant	0.300***	3.896***	0.481***	1.615	0.320***	-1.830
	(0.0417)	(0.926)	(0.0423)	(1.078)	(0.0452)	(1.370)
year dummies	YES	YES	YES	YES	YES	YES
subway station dummies	YES	NO	YES	NO	YES	NO
N	1540	1540	1540	1540	1540	1540
R-squared	0.620		0.632		0.543	
Prob > chi2	0.1164	0.1164	0.5100	0.5100	0.9950	0.9950
4B. Gini-Simpson index						
Dep. Var.: diversity2	(1)	(2)	(3)	(4)	(5)	(6)
	0-	400 m	400	0–800 m	800	0–1200 m
Panel setting	fixed effect	random effect	fixed effect	random effect	fixed effect	random effect
connect	0.0717**	0.0849***	0.0336	0.0442	0.00103	0.00452
			0.0330	0.0 1 12	0.00103	
	(0.0289)	(0.0294)	(0.0296)	(0.0294)	(0.0268)	(0.0272)
ln(d_cbd)		(0.0294) -0.162***		(0.0294) - 0.0453		0.0980
$ln(d\_cbd)$		(0.0294) - 0.162*** (0.0405)		(0.0294) - 0.0453 (0.0416)		0.0980 (0.0603)
, _ ,		(0.0294) -0.162***		(0.0294) - 0.0453		0.0980
, _ ,	(0.0289)	(0.0294) -0.162*** (0.0405) 0.102*** (0.0223)	(0.0296)	(0.0294) -0.0453 (0.0416) 0.132*** (0.0213)	(0.0268)	0.0980 (0.0603) 0.104*** (0.0274)
ln(pop_d)	(0.0289) 0.173***	(0.0294) -0.162*** (0.0405) 0.102*** (0.0223) 1.627***	(0.0296) 0.252***	(0.0294) -0.0453 (0.0416) 0.132*** (0.0213) 0.664*		0.0980 (0.0603) 0.104*** (0.0274) - 0.699
ln(pop_d)	(0.0289)	(0.0294) -0.162*** (0.0405) 0.102*** (0.0223)	(0.0296)	(0.0294) -0.0453 (0.0416) 0.132*** (0.0213)	(0.0268)	0.0980 (0.0603) 0.104*** (0.0274)
ln(pop_d) constant	(0.0289) 0.173*** (0.0187)	(0.0294) -0.162*** (0.0405) 0.102*** (0.0223) 1.627*** (0.361)	(0.0296) 0.252*** (0.0199)	(0.0294) -0.0453 (0.0416) 0.132*** (0.0213) 0.664* (0.375)	(0.0268) 0.171*** (0.0195)	0.0980 (0.0603) 0.104*** (0.0274) - 0.699 (0.539)
In(pop_d)  constant  year dummies	(0.0289)  0.173*** (0.0187)  YES	(0.0294) -0.162*** (0.0405) 0.102*** (0.0223) 1.627*** (0.361)	(0.0296)  0.252*** (0.0199)  YES	(0.0294) -0.0453 (0.0416) 0.132*** (0.0213) 0.664* (0.375)	(0.0268)  0.171*** (0.0195)  YES	0.0980 (0.0603) 0.104*** (0.0274) -0.699 (0.539)
In(pop_d)  constant  year dummies subway station dummies	(0.0289)  0.173*** (0.0187)  YES YES	(0.0294) -0.162*** (0.0405) 0.102*** (0.0223) 1.627*** (0.361) YES	(0.0296)  0.252*** (0.0199)  YES YES	(0.0294) - 0.0453 (0.0416) 0.132*** (0.0213) 0.664* (0.375) YES	(0.0268)  0.171*** (0.0195)  YES YES	0.0980 (0.0603) 0.104*** (0.0274) - 0.699 (0.539) YES NO
ln(pop_d) constant	(0.0289)  0.173*** (0.0187)  YES	(0.0294) -0.162*** (0.0405) 0.102*** (0.0223) 1.627*** (0.361)	(0.0296)  0.252*** (0.0199)  YES	(0.0294) -0.0453 (0.0416) 0.132*** (0.0213) 0.664* (0.375)	(0.0268)  0.171*** (0.0195)  YES	0.0980 (0.0603) 0.104*** (0.0274) - 0.699 (0.539)

Notes: robust standard errors in parentheses; \*\*\* p < 0.01, \*\* p < 0.05, \* p < 0.1; prob > chi2 shows the p-value of Hausman test.

from observing significant subway connection effect.

In Panel 2B we estimate Eq. (3) with a linear regression estimator using  $\ln(openings+1)$  as the dependent variable. In each buffer, we report both fixed effects and random effects results. The p-values (Prob > chi2) of the Hausman tests indicate that the random effects model (in bold) fits the data better. A new subway station neighborhood receives more catering establishments in all three buffers, and closer buffers have larger gains (29.9% for the 0–400 m buffer, 24.4% for the 400–800 m buffer and 9.05% for the 800–1200 m buffer). This pattern holds for the fixed effects results.

Table 3 reports the linear regression estimates of the customer reviews model of Eq. (4). The dependent variables are the total count of reviews in each buffer in Panel 3A and the average count of reviews per catering business in Panel 3B. For each buffer, we report the results of both fixed effects and random effects models. The results with a better data fit (as indicated by the Hausman test) are in bold. After a new subway station is opened, the number of customer reviews (a proxy of customer visits) in the 0–400 m buffer increases most significantly (64.1% for the total review counts in buffer, and 28.5% for the average review counts per business). The 400–800 m buffers also see increased review counts as well (49.9% and 19.6% for total and average review counts, respectively) but in smaller magnitudes compared to the aforementioned 0–400 m buffers. Estimates for the farthest buffers (800–1200 m) suggest no significant change in the count of online customer reviews.

A comparison of results in Tables 2 and 3 suggests that a new

subway station's effects on catering openings and customer visits are mainly within the 800 m radius from the station. While the number of catering businesses likely grow due to a new subway station (based on the results of openings in Table 2), the relative growth of customer visits (proxied by online review counts) is larger in percentage (shown in Table 3A) in the same buffer (0–400 m or 400–800 m). This is confirmed by the results in Table 3B on customer demand intensity per business.

The increased number of and visits to catering businesses following a new subway station is accompanied by the rise in the diversity of choices. Table 4 presents the linear regression estimates of catering diversity based on the 49 cuisine genres in subway neighborhoods using Eq. (5). The dependent variable is the Shannon index in Table 4A and the Gini-Simpson index in Table 4B. The regression results show a consistent pattern that is similar as those found in Tables 2 and 3. The results of both the fixed effects and the random effects models are presented for each buffer, with the results better fitting the data in bold. Following the opening of a new subway station, both the Shannon index and the Gini-Simpson index of neighborhood catering diversity increase in the 0-400 m and the 400-800 m buffers. The closer the buffer to the station, the larger the effect. The Shannon index gains 0.27 (0.18) in the 0–400 m (400–800 m) buffer following a new subway station, or on average about a 24.5% (13.3%) increase in the corresponding buffer category. Assuming equal proportions of each cuisine genre, such an increase in the Shannon index roughly comes from an increase of cuisine genre types from 10 to 19 in the 0-

**Table 5**Effects of a new subway station on catering openings by type and distance to station.

5A. Negative binomial regressions I						
Dep. var.: openings_sit & openings_fast	` '	(2) Fast-food 0–400 m	(3) Sit-down 400–800 m	(4) Fast-food 400–800 m	(5) Sit-down 800–1200 m	(6) Fast-food 800–1200 m
connect	0.307**	0.458***	0.167	0.217**	0.0480	-0.0919
$ln(d\_cbd)$	(0.151) 0.151 (0.225)	(0.120) -0.270 (0.196)	(0.116) -0.0226 (0.185)	(0.107) -0.147 (0.183)	(0.153) -0.120 (0.273)	(0.134) 0.160 (0.288)
$ln(pop\_d)$	0.310*** (0.0827)	0.324***	0.295*** (0.0689)	0.342*** (0.0826)	0.384*** (0.0826)	0.455***
constant	-3.429 (2.334)	- 1.025 (1.940)	-0.761 (1.867)	-0.996 (1.781)	0.661 (2.669)	-4.140 (2.812)
year dummies	YES	YES	YES	YES	YES	YES
subway station dummies	NO	NO	NO	NO	NO	NO
subway line segment dummies	YES	YES	YES	YES	YES	YES
district *year trend	NO	NO	NO	NO	NO	NO
N	1540	1540	1540	1540	1540	1540
Alpha	1.0333	0.9781	0.7887	0.8925	1.5223	1.7918
LR test (alpha=0) Prob > chi2	406.88 0.000	1186.40 0.000	538.90 0.000	2309.04 0.000	880.35 0.000	2657.61 0.000
5B. Negative binomial regressions II						
Dep. var.: openings_sit & openings_fast	. ,	(2) Fast-food 0–400 m	(3) Sit-down 400-800 m	(4) Fast-food 400-800 m	(5) Sit-down 800–1200 m	(6) Fast-food 800–1200 n
connect	0.0994	0.223**	0.0281	0.238***	0.0564	0.184*
constant	(0.114) -44.01 (62.69)	(0.0928) - 755.8*** (54.74)	(0.0907) - 11.84 (42.30)	(0.0803) - 735.3*** (45.21)	(0.104) - 11.04 (51.29)	(0.102) - 743.7*** (52.37)
year dummies	NO	NO	NO	NO	NO	NO
subway station dummies	YES	YES	YES	YES	YES	YES
subway line*ring	NO	NO	NO	NO	NO	NO
district*year trend	YES	YES	YES	YES	YES	YES
N	1540	1540	1540	1540	1540	1540
Alpha	0.3136	0.2488	0.1260	0.2376 391.08	0.07481 6.10	0.1714 86.46
LR test (alpha=0) Prob > chi2	87.77 0.000	165.20 0.000	34.13 0.000	0.000	0.007	0.000
5C. Panel random effects linear regress		0.000	0.000	0.000	0.007	0.000
Dep. var.: openings sit & openings fast		(2)	(3)	(4)	(5)	(6)
DCp. var openings_sit & openings_just	` '	Fast-food 0-400 m		Fast-food 400–800 m	Sit-down	Fast-food 800–1200 m
connect	0.0871* (0.0460)	0.335*** (0.0699)	0.0397 (0.0526)	0.279*** (0.0774)	0.00223 (0.0383)	0.111* (0.0609)
$ln(d\_cbd)$	-0.154** (0.0609)	-0.208** (0.0852)	-0.0797 (0.0835)	-0.0400 (0.106)	0.123 (0.0782)	0.178* (0.0960)
$ln(pop\_d)$	0.103*** (0.0256)	0.125*** (0.0387)	0.178*** (0.0379)	0.243*** (0.0479)	0.146*** (0.0361)	0.172*** (0.0434)
constant	1.734*** (0.554)	2.024*** (0.766)	1.265* (0.748)	0.558 (0.953)	-0.712 (0.702)	- 1.474* (0.861)
year dummies	YES	YES	YES	YES	YES	YES
subway station dummies	NO	NO	NO	NO	NO	NO
N	1540	1540	1540	1540	1540	1540
N Prob > chi2	0.6579	0.6533	0.7079	0.9981	0.9999	1.0000

Notes: robust standard errors in parentheses; \*\*\* p < 0.01, \*\* p < 0.05, \* p < 0.1; prob > chi2 shows the p-value of Hausman test.

400 m buffer and from 10 to 15.5 in the 400–800 m buffer. The results are similar using the Gini-Simpson index, with an average increase of 0.48 (0.56) in the 0–400 m (400–800 m) buffer. That is about a gain of 17.7% in the 0–400 m buffer and 7.9% in the 400–800 m buffer.

It is also interesting to test for the heterogeneity in new subway stations' effects on neighborhood catering businesses. Table 5 reports the different effects of a new station on the openings of the fast-food and sit-down restaurants, because they tend to serve customers with

different time values. Results in Table 5A–B are from negative binomial regressions of fast-food and sit-down restaurant openings. Table 5A controls for year and subway line segment fixed effects, while Table 5B controls for station fixed effects and district-specific trends. Table 5C reports results of linear panel random effects regression models as they fit data better than fixed effects specifications. Table 5's results are parallel to those in Table 2, so we skip detailed descriptions here. Comparing the results of the two types of catering businesses, one can

**Table 6**Effects of a new subway station on catering openings by distance to station: inner city vs. outer city.

Dep. Var.: openings	(1) 0–400 m	(2) 400–800 m	(3) 800–1200 m	(4) 0-400 m	(5) 400-800 m	(6) 800–1200 m	(7) 0–400 m	(8) 400–800 m	(9) 800–1200 m
connect	0.270**	0.180*	-0.0730	0.120	0.212***	0.169**	-0.116	0.00619	-0.132*
suburb	(0.116) - 1.909*** (0.540)	(0.101) 0.328 (0.402)	(0.135) 1.921*** (0.543)	(0.0825)	(0.0662)	(0.0813)	(0.0800)	(0.0673)	(0.0797)
connect*suburb	0.921*** (0.219)	0.391** (0.152)	0.378** (0.164)	0.976*** (0.187)	0.612*** (0.127)	0.332** (0.132)	0.957*** (0.171)	0.454*** (0.109)	0.310*** (0.107)
$ln(d\_cbd)$	-0.0513 (0.169)	0.0159 (0.158)	0.127 (0.243)	(	,	(33.3.)	,	(** ***)	( ,
ln(pop_d)	0.360*** (0.0641)	0.328*** (0.0690)	0.403*** (0.0691)						
constant	-0.527 (1.528)	-1.580 (1.475)	-4.011* (2.226)	-437.2*** (45.01)	-408.3*** (32.69)	-422.0*** (38.60)	-2.624*** (0.461)	-0.807*** (0.193)	-0.544** (0.218)
year dummies subway station dummies	YES NO	YES NO	YES NO	NO YES	NO YES	NO YES	YES YES	YES YES	YES YES
subway line segment dummies	YES	YES	YES	NO	NO	NO	NO	NO	NO
district*year trend	NO	NO	NO	YES	YES	YES	NO	NO	NO
N alpha	1540 0.9315	1540 0.8186	1540 1.6366	1540 0.2951	1540 0.1891	1540 0.1702	1540 0.2598	1540 0.1437	1540 0.0942
LR test (alpha=0) Prob > chi2	1996.50 0.000	3681.40 0.000	2911.61 0.000	456.27 0.000	565.62 0.000	191.86 0.000	405.43 0.000	374.41 0.000	90.08 0.000

Notes: all models use negative binomial regressions; robust standard errors in parentheses; \*\*\* p < 0.01, \*\* p < 0.05, \* p < 0.1.

easily identify the consistent pattern that a new subway station's effect on catering openings is larger and statistically more significant for fast-food restaurants, especially within *800 m* from station. This seems to support the idea that the increase in market demand due to a new subway station comes mainly from rapid transit users who value proximity and quick service.

The spatial heterogeneity is also highlighted in our results. Table 6 reports the effects of a new station on catering openings in the inner city (within the Fourth Ring Road) versus in the outer city (outside the Fourth Ring Road) estimated in the similar way as the negative binomial results in Table 2A. We estimate the difference in effects with a dummy variable *suburb* (indicating outer city station neighborhoods) and its interaction with *connect*. The dummy variable *suburb* is omitted in columns (4)–(9) when station fixed effects are controlled for. The key estimate here is the coefficients of *connect\*suburb*. The positive coefficients of *connect\*suburb* suggest that annual neighborhood catering openings increase more following a new subway station outside the Fourth Ring Road in each buffer category. This is intuitive as one would expect more significant land use changes following a new subway station at the urban fringe.

#### 6. Conclusion

Cities are more likely to be successful if they provide richer public and private consumer amenities. Studying how amenities offered through costly public investment (e.g., the subway) interact with privately provided consumer amenities (e.g., food and beverage services) is important as these public and private amenities together reshape urban spatial structure and contribute to quality of life. Such public and endogenous private amenities deserve the attention of urban policy makers and planners as they not only lead to potential economic development and employment benefits, but also tend to raise property value and increase the risk of gentrification.

This paper enhances the existing literature by studying in detail how local consumer amenities react to urban rail transit development. Using multiyear subway construction history and station neighborhoods' catering business activities in Beijing, we suggest that a new subway station positively contributes to the quantity, diversity and consumer demand of nearby food and beverage services. These effects are larger the closer to a new station, although statistically significant effects are consistently observed only within 800 m from stations. Subway neighborhoods in fringe areas see a much larger increase in catering openings compared to neighborhoods closer to the city center. Moreover, different types of catering businesses change in different ways following the opening of a new subway station in Beijing, with fast food services increasing more than sit-down restaurants following the opening of new stations. The overall findings of this study are consistent with those of Bowes and Ihlanfeldt (2001) and Zheng and Kahn (2013), but with broader aspects examined (such as retail diversity and consumer demand) and more details on effect heterogeneity. Nevertheless, we believe these detailed findings are perhaps more applicable to dense Asian megacities than most cities in the US.

There remain a few limitations in this study. First, while we examine both the supply (quantity and diversity) and demand dimensions of local food and beverage businesses, we leave behind other aspects such as grocery stores and other (dis)amenities that are not provided by the private sector (e.g., crime and traffic). One should factor these other types of transit-induced consumer (dis) amenities when examining consumer willingness to pay and welfare implications of transit investment. Second, while neighborhoods receiving a new subway station during our study period are widely distributed within Beijing, there remains a potential sample selection issue. If the neighborhoods that expect to gain the most in terms of business opportunities are more likely to have subway stations for reasons unobservable to us, then the estimated treatment effects might not apply to every neighborhood in the city. Attention should also be paid to the potential timevarying unobservables that correlate with station opening and affect local retail activities. Finally, the spatial resolution of local population of subway neighborhoods could be further refined to match the geographical boundary of station areas when data are available in the future.

#### Acknowledgements

We thank the editor, two referees and the participants of the 2015 IACP Conference at Chongqing University for valuable

comments. We thank the National Natural Science Foundation of China (No. 71273154, No. 71322307 and No. 71533004) and the Key Project of National Social Science Foundation of China

(No. 13AZD082) for research support. Yangfei Xu provided excellent research assistance.

Appendix A. subway lines as of December 2013

Subway line	Date of opening	Length (km)
Line 1	1971/01/15	31
Line 2	1971/01/15	23.1
Line 13	2002/09/28	40.9
Line Batong	2003/12/27	19
Line 5	2007/10/07	27.6
Line 8	2007/07/19	26.6
Line 10	2008/07/19	57.1
Airport express	2008/07/19	28.1
Line 4	2009/09/28	28.2
Line 15	2010/12/30	31.5
Line Changping	2010/12/30	21.3
Line Daxing	2010/12/30	21.8
Line Fangshan	2010/12/30	24.6
Line Yizhuang	2010/12/30	23.2
Line 9	2011/12/31	16.5
Line 6	2012/12/30	42.8
Line 14	2013/05/05	12.4

Note: the airport express line is excluded from the analysis in this study due to its non-stop service from inner city to airport.

#### Appendix B. catering types and cuisine genre

#### Chinese cuisines

Beijing; Chaoshan; Sichuan; northeastern; guanfu; Guizhou; Hubei; Huaiyang; Hui; Jiangxi; Jiangzhe; Shandong; Mongolia; Shanxi; Shanghai; Taiwan; northwestern; Tibet; Hunan; Xinjiang; Guangdong; Yunnan; Zhejiang; Hong Kong; halal; roast duck; roast fish; farm food; creative food; seafood; crayfish; tea house; abalone; hot pot; incense pot; lamb spine.

#### Western cuisines

Korean; Japan; Southeast Asian; Thai; Indian; Vietnamese; Middle Eastern; Russian; French; Italian.

#### Fast food

Street food; snack; bread or sweets.

#### References

Bakhshi, S., Kanuparthy, P., Gilbert, E., 2014. Demographics, weather and online reviews: a study of restaurant recommendations. In: Proceedings of the 23rd international conference on World wide web. International World Wide Web Conferences Steering Committee. pp. 443–454.

Berry, B.J.L., 1967. Geography of Market Centers and Retail Distribution. Prentice-Hall, Englewood Cliffs, NJ.

Bollinger, C.R., Ihlanfeldt, K.R., 1997. The impact of rapid rail transit on economic development: the case of Atlanta's MARTA. J. Urban Econ. 42 (2), 179–204.

Bowes, D.R., Ihlanfeldt, K.R., 2001. Identifying the impacts of rail transit stations on residential property values. J. Urban Econ. 50 (1), 1–25.

Chapple, K., Jacobus, R., Retail trade as a route to neighborhood revitalization, In: Pindus, Wial and Wolman., (Eds.), Urban and Regional Policy and Its Effects, Volume 2, 2009, Brookings Institute; Washington, DC.

Debrezion, G., Pels, E., Rietveld, P., 2007. The impact of railway stations on residential and commercial property value: a meta-analysis. J. Real. Estate Financ. Econ. 35 (2), 161–180.

Fischer, J.H., Harrington Jr, J.E., 1996. Product variety and firm agglomeration. RAND J. Econ., 281–309.

Glaeser, E.L., Kolko, J., Saiz, A., 2001. Consumer city. J. Econ. Geogr. 1 (1), 27–50.
Graham, D.J., 2007. Agglomeration, productivity and transport investment. J.
Transp. Econ. Policy 41 (3), 317–343.

Giuliano, G., Land use impacts of transportation investments: highway and transit.

In: Hanson and Giuliano, (Eds.), *The Geography of Urban Transportation*, 3rd ed., 2004, Guilford; New York, NY.

Haddad, E.A., Hewings, G.J.D., Porsse, A.A., et al., 2015. The underground economy:
 tracking the higher-order economic impacts of the São Paulo Subway system.
 Transp. Res. Part A: Policy Pract. 73, 18–30.
 Mejia-Dorantes, L., Lucas, K., 2014. Public transport investment and local re-

Mejia-Dorantes, L., Lucas, K., 2014. Public transport investment and local regeneration: a comparison of London's Jubilee Line Extension and the Madrid Metrosur. Transp. Policy 35, 241–252.

Mohammad, S.I., Graham, D.J., Melo, P.C., et al., 2013. A meta-analysis of the impact of rail projects on land and property values. Transp. Res. Part A: Policy Pract. 50, 158–170.

Schuetz J., Do rail transit stations encourage neighborhood retail activity?. Urban Studies, 2015, 52, pp. 2699–2723.

Stern, N., 1972. The optimal size of market areas. J. Econ. Theory 4 (2), 154–173. Sun, W., Zheng, S., Wang, R., 2015. The capitalization of subway access in home value: a repeat-rentals model with supply constraints in Beijing. Transp. Res. Part A: Policy Pract. 80, 104–115.

Zheng, S., Kahn, M.E., 2013. Does government investment in local public goods spur gentrification? Evidence from Beijing. Real Estate Econ. 41 (1), 1–28.

Zheng, S., Sun, W., Wang, R., 2014. Land supply and capitalization of public goods in housing prices: evidence from Beijing. J. Reg. Sci. 54 (4), 550–568.