Subways and Road Congestion

Yizhen Gu, Chang Jiang, Junfu Zhang and Ben Zou*

Abstract

We study whether subways alleviate road congestion by examining 45 subway line launches in China and using detailed data on road speed. Our difference-in-differences estimation finds that in the first year after a subway line is launched, rush-hour speed on nearby roads increases by about 4%. The effect is most prominent in initially-congested roads and declines over distance to the new subway line. Evidence on road speed is corroborated with substitution patterns among modes of transportation. Using auxiliary data from Beijing, we calculate that the time savings for each automobile or bus commute from faster speed is worth 0.1 USD.

Keywords: Subway, congestion, public transit
JEL Classification: R41, R42, L92

*Gu: Institute for Economic and Social Research, Jinan University; email: yizhengu@jnu.edu.cn. Jiang: Maps Open Platform Department, Baidu, Inc.; email: jiangchang@baidu.com. Zhang: Department of Economics, Clark University; email: juzhang@clarku.edu. Zou: Department of Economics, Michigan State University; email: benzou@msu.edu. We would like to thank Seema Jayachandran and four anonymous referees for extremely helpful comments. We thank Richard Arnott, Victor Couture, Mark Jacobsen, Ajin Lee, Matt Turner, Cliff Winston, and participants at seminars and conferences at Brookings-Tsinghua Center, Fudan University, Harvard University, Huazhong University of Science and Technology, Jinan University, MIT, Michigan State University, University of Chicago, and Singapore Management University for helpful comments. We are grateful to the people at the Baidu Maps Open Platform department, in particular Tianqi Liu and Xiangmin Li, for providing and helping us with the data. We have benefited from outstanding research assistance by Hui Hu, Wenwei Peng, Zixin Wei and Kai Wu. All views expressed and errors are our own.

1 Introduction

Traffic congestion is a major challenge facing many cities around the world. The problem is particularly acute in large cities in the developing world that have recently experienced rapid increases in population and car ownership. Recent studies have found that many developing-country cities are among the world's most congested (e.g., Reed & Kidd, 2009; TomTom, 2019).

Various policies have been implemented aiming to reduce traffic congestion. Demand-side policies include congestion pricing (Small et al., 2007) and restrictions on automobile ownership and usage (e.g., Li, 2018; Davis, 2008; Gu et al., 2017). On the supply side, urban rail transit systems (henceforth subways) are considered an effective way to reduce congestion, because they have large capacity and do not require much surface land. By 2014, 171 cities worldwide had a subway system in operation (Gendron-Carrier et al., 2018). Cities in developing countries account for a large share of recent subway construction (Gonzalez-Navarro & Turner, 2018). In China alone, total subway length increased from less than 400 kilometers in 4 cities in 2001 to about 5,000 kilometers in 30 cities by 2017 (Cui et al., 2018).

Due to data limitations, empirical evidence on the congestion-reducing effect of subways is scarce. First, subway line openings are infrequent. An analysis covering multiple subway lines often requires data from a number of cities over many years, which are hard to come by. Second, standard datasets in socioeconomic studies lack good measures of road speed. Speed can be backed out from travel diaries available in many household transportation surveys (e.g., Couture et al., 2018). However, these surveys are infrequent, have limited sample sizes, and suffer from substantial measurement errors. Traffic data can also be derived from traffic cameras, wired loop detectors, or GPS-enabled bus and taxi fleets (e.g., Anderson, 2014). But municipal governments usually administer these data separately, making it difficult to compare across cities.

This paper takes advantage of the recent massive construction of subways across Chinese cities. We focus on the 45 new subway lines (including 7 extensions to existing lines) launched between August 2016 and December 2017 across 25 cities. We use a novel source of traffic speed data from China's leading provider of digital map services, which collects and processes real-time traffic information from user-generated data. The company offers a free digital map application on smartphones and other mobile devices. With the location service turned on, mobile devices remit bits of location data every few seconds to the company's database. With these data the company generates real-time speed information at the road segment level. The speed information is then displayed on the digital map and can be used to generate real-time optimal travel routes.

We obtain hourly speed data between August 1, 2016 and January 31, 2018 for a sample of road segments near subway lines. Road segments in the sample are those inside 2.5-kilometer buffer zones of selected 5-kilometer segments of existing, new, and planned subway lines. After applying several restrictions, our baseline sample includes more than 20,000 road segments that cover all 42 cities in Mainland China that had an existing or planned subway system by the end of 2017. The spatial and temporal dimensions of our data are a significant improvement on existing studies that typically look at a shorter period of time for a much smaller number of subway lines.

Big data at the granular level allow us to adopt a saturated empirical model. Log hourly speed is first regressed on a set of fully interacted indicators for road segment, day of the week, and hour of the day. The residual from this regression captures the deviation from the "usual" traffic speed. It is then averaged at the weekly level and used as the outcome variable in a difference-in-differences (DID) estimation: Road segments in cities that had existing or planned subway lines but did not have a new line launched during the sample period are randomly assigned to each treated line to serve as the control group. The control road segments are given a fake opening date that is the same as that of the corresponding treated line. All 45 treated lines and their controls are then stacked according to the week relative to opening. This stacked DID approach helps control for seasonality and macro trends common to all cities.

Perhaps the biggest challenge to identification is that most of the openings are at a time of the year, Chinese New Year, when traffic eases up anyway in major cities, combined with the fact that our control cities are systematically smaller. We address this by allowing for different seasonality for larger cities. We find a similar effect when we restrict the analysis to the small number of openings that occur at other times of the year. The results are also similar from a two-way fixed effects model using only treated cities.

We first investigate whether subways have *any* effect on road speed by focusing on weekday rush hours and road segments that are likely to be directly affected. Directly-affected road segments are those that a person needs to travel through if one chooses to drive instead of taking the subway. They are identified by using the route-planning function of the digital map platform.

Inspection of the coefficients in the pre-treatment period suggests that we cannot reject the parallel trend assumption. The launch of a new subway line has an immediate positive effect on the speed in directly affected road segments. In the first week after a line opens, speed in those road segments increases by about 2.5% relative to that in the control segments. The effect increases to about 5% in the 6^{th} week before it declines and stabilizes between 2% and 3%. We track up to 48 weeks after the line opening; the effect remains stable and statistically significant by then. The average effect in the post-treatment period is between 3.6% and 4.4%.

We conduct a host of additional robustness checks. Evidence from a placebo test with fake opening dates suggests that the result is unlikely to be driven by confounding trends. Results are similar with various cuts of the sample or time period. A difference-in-discontinuity design also suggests an immediate jump in road speed around the time of the subway line opening.

We then investigate the effects of a new subway line on different types of road segments. For 23 out of the 25 treated cities, we are able to extract road segments near an existing or planned subway line. We include all sampled road segments in these treated cities to study how the effect spreads through the urban transportation network. We find substantial heterogeneity in the congestion-relieving effect of a new subway line. The effect is larger for directly affected road segments, for initially more congested roads, suggesting that marginal switchers from road vehicles to subways are more likely to be those who experienced high levels of road congestion (Anderson, 2014). The effect declines quickly with the distance to the new subway line. There is suggestive evidence of a network effect of the subway system: Road segments that are far away from the new subway line but close to an existing subway line also experience substantial increases in speed. We also expand the analysis sample to include non-rush hours and find that the effect is larger in rush hours, especially for road segments that have the same direction as the flow of the traffic.

To corroborate our finding on road speed, we study the substitution patterns among different modes of transportation using micro data from household transportation surveys in Beijing. We find that improved access to the subway is associated with increases in subway trips and decreases in bus trips, car trips, and annual vehicle kilometers traveled of private cars.

We build a conceptual framework of transportation mode choices. The framework leads to a formula for the welfare impact of subways that relies only on observable quantities and prices. The welfare impact can be decomposed into three components: (1) welfare gains from savings in travel time by road vehicles (including automobiles and buses), (2) welfare gains for those who switch to the subway from road vehicles, and (3) changes in government spending on public transit.

The main empirical result of the paper, the effect of subways on road speed, helps us estimate the first component of the welfare expression. We do so, focusing on commutes in Beijing. Our estimate suggests that Beijing's subway increases rush-hour average road speed by 3 percent. Using additional data on volumes of ridership, average length of commutes, and average wages, we calculate that the time saved for each automobile or bus commute from reduced congestion is worth 0.1 USD. We need to rely on stronger assumptions and correlational evidence to estimate the other two terms. We find that each of these terms is large in magnitude, but they offset each other. If accurate, this would suggest that the benefit of Beijing's subway system exceeds its cost. Note that other potential benefits of the subway, such as the reduction in air pollution (Chen & Whalley, 2012; Gendron-Carrier et al., 2018) and car accidents are left out of our framework and calculation.

Congestion-reduction is one of the most cited reasons by proponents of subways. Using a panel of U.S. cities, Winston & Langer (2006) find that longer rail transit mileage is associated with lower congestion costs. However, empirical evidence on the causal effect is limited. Yang et al. (2018) find that the city-level "congestion index" drops sharply following 6 subway line openings in Beijing. Yet the event study design does not purge out potential time trends. Other studies find significant reductions in air pollution upon the launch of the subway (Chen & Whalley, 2012; Gendron-Carrier et al., 2018; Li et al., 2019), which *implies* reduced traffic. Our paper provides direct evidence on this relationship.

There has been a long debate in the literature of urban economics on the value of public transit, mostly in the setting of developed countries (Voith, 1991; Baum-Snow & Kahn, 2005; Winston & Maheshri, 2007; Parry & Small, 2009). Some studies have used temporary interruptions to the public transit system, such as strikes, to evaluate the benefits (e.g., Anderson, 2014; Adler & van Ommeren, 2016). However, halting existing systems is likely to induce very different responses compared with the introduction or expansion of services. In addition, because public transits in developing countries face very different constraints, evidence from developed countries may not immediately apply.

Although building additional road capacity is a natural response to traffic congestion and is widely adopted by policymakers, there is debate on whether the supply-side approach is effective. The "fundamental law" of highway congestion (Downs, 1962, 2000; Hsu & Zhang, 2014) suggests that the elasticity of vehicle kilometers traveled (VKT) with respect to lane kilometers is one, thus adding road capacity does not reduce congestion. Using a panel of U.S. cities, Duranton & Turner (2011) find empirical evidence that the fundamental law also applies to urban roads; the supply of public transit does not reduce VKT and is unlikely to alleviate traffic congestion. In contrast, our finding suggests that building subways—an alternative way to add road capacity—does reduce congestion in Chinese cities, at least in the short run.

This paper also contributes to a new and growing strand of literature that takes advantage of user-generated big data in urban transportation. Some recent studies have collected data on millions of simulated trips using the "usual" traffic speed provided by Google Maps (Akbar & Duranton, 2017; Akbar et al., 2018). Kreindler (2018) designed a new smartphone application that collected precise GPS coordinates over 100,000 commuter trips in an experimental setting in Bangalore. Our collaboration with a large digital map provider allows us to analyze real-time speed data from a large set of road segments over a relatively long period of time.

The remainder of the paper is structured as follows. Section 2 describes the institutional background, data sources and sample constructions. Section 3 presents the main empirical analyses. Section 4 documents supporting evidence from mode substitution using household travel data. Section 5 briefly discusses the welfare impact. Section 6 concludes.

2 Background, Data and Sample

2.1 Subways in Chinese Cities

There has been a large boom in subway construction across Chinese cities in the past two decades. Figure 1 shows the growth of subway length and ridership in China. In 2001, only four cities in Mainland China—Beijing, Shanghai, Guangzhou and Tianjin—had a subway system. The combined length of all subway lines was below 400 kilometers. By the end of 2017, 30 cities had a total of 4,476 kilometers of subway lines, and 12 other cities had their first subway lines under construction. Subway ridership increased accordingly, from just under 1 billion in 2001 to about 16 billion in 2017.

Massive subway construction is a response to the rapid growth in population and car ownership in China's major cities. The overall urbanization rate increased from 35% in 2000 to 58% in 2017. Much of the increased urban population concentrated in large cities. Rapid increase in car ownership has made major Chinese cities among the most congested and polluted in the world. Many city governments regard subways as the essential infrastructure to reduce congestion and pollution. Billions of dollars have been invested in building and expanding subway systems. Despite the large amounts of investment, there has been little empirical evidence on whether these subway lines achieved the stated goal of reducing congestion.

2.2 Data on Road Speed

The company we work with, Baidu Maps, is China's leading provider of digital maps and online navigation services. It runs a smartphone application that is similar to the one provided by Google Maps or Apple Maps. With the application installed and location service turned on, mobile devices remit bits of location information every few seconds to the company's data center. Mobile devices also transmit information on acceleration, rotation, and angle, which helps distinguish devices in vehicles and those carried by pedestrians or cyclists. The location of the device is matched with a digital map of roads. Speed can be calculated with multiple location records from the same device and the time lapse in between. The application has 280 million monthly active users in China. With large amounts of data flowing into the company's server every second, it is able to compute real-time speed at very fine geographic levels. Real-time speed is displayed on the digital map and can be used to calculate optimal routes for different modes of transportation.[1]

We obtain hourly average speed at the road segment level between August 1, 2016 and January 31, 2018.[2] A road segment is a short stretch of road. In our baseline sample, the median length of a road segment is about 50 meters. For each road segment we know the name of the road, direction of the segment (the two directions of a stretch of a two-way road are two separate road segments), and a set of coordinates indicating its location. From these coordinates we calculate the length of the segment and its geographic relation to the subway and other urban features. Road segments are classified into five hierarchical categories: highways, urban expressways, arterial streets, sub-arterial streets, and local streets.

The speed data also come with a "congestion index," which is defined as time needed to travel through the road segment under the current speed relative to time needed under traffic-free speed. The traffic-free speed is the average speed on the same road segment between midnight and 5 AM. For example, half of the traffic-free speed corresponds to a congestion index of 2.

[1] Online Appendix Figure A.1 shows a screenshot of the digital map with color-coded road speed. Online Appendix A.1 presents additional facts about the data.

[2] Except for the periods between September 1 and September 10, and between October 10 and November 30, 2016, for which the original data were no longer available. For each day we have hourly speed between 7AM and 7PM.

2.3 Data Sample

The ideal data would be a random sample of road segments in the city combined with an over-sampling of those near new subway lines. Such a sample could give us both a big-enough sample size in regions that are most likely to be affected and an overall representative sample. However, resources allocated to help us extract and prepare the data were too limited for such a sampling procedure.

Instead, we extract road segments in the neighborhoods of selected subway lines. We first select segments of new, existing, and planned subway lines across all 42 cities with an existing or planned subway system, and then extract all road segments in the neighborhood of these subway line segments. This approach comes at a cost. We do not have data on segments further from the subway, so we cannot assess effects on the city's overall traffic during rush hour.

2.4 Subway Lines and Road Segments in the Sample

Our data include subway lines that were launched between August 1, 2016 and December 31, 2017. There were 45 such lines (including 7 extensions to existing lines) across 25 cities. Table 1 Panel A lists these "treated lines." The 25 cities (henceforth the "treated cities") include China's largest metropolises and provincial capitals. The table lists the official opening date for each new subway line. It is worth noting that a disproportionately large share of the lines opened towards the end of the calendar year. 25 out of the 45 subway lines were launched in December. There were 7 new line openings on Dec 28, 2016 alone, and 6 on Dec 28, 2017. There are reasons to believe that the official opening dates are not randomly determined. The clustering of opening dates can be correlated with potential confounding factors. Indeed, the end of the calendar year is also the start of China's holiday season. Chinese New Year (CNY), the nation's most important holiday, follows a lunar calendar and usually falls between mid-January and mid-February. With economic activity on a slower pace, it is possible that the roads will be less congested even without the launch of new subway lines. Thus it is important to purge out seasonality in traffic patterns.

For 23 out of the 25 treated cities, our data also include an existing or planned subway line. Changes in speed on road segments near these existing or planned subway lines shed light on how the effect spreads along the road and subway network, which we study in Section 3.3. These lines are chosen such that they are at least 3 kilometers away from the treated subway line, and have a comparable distance to the downtown. Table 1 Panel B lists these lines. Among the 25 such lines there are 22 existing and 3 planned. The opening dates of these lines are, in general, at least one year apart from those of the treated lines in the same city. So they should not generate confounding effects.

In addition to the 25 treated cities, there are 17 "control cities" that have existing or planned subway systems but did not have a new line launched during the sample period. Road segments from these cities serve as the control sample. In each of these control cities, we pick the latest line completed; if the city does not have a completed line, we choose the first line to be built. Table 1 Panel C lists these lines.

We choose a 5-kilometer stretch of each subway line listed in Table 1. We then create a 2.5-kilometer buffer zone on both sides of the segment, and extract all road segments that lie within the buffer zone.[3] Because most subways are designed to alleviate congestion in urban centers, in general, we pick subway segments close to the downtown. In cities with multiple subway lines included in the sample, we adjust the positions of the segments such that their buffer zones do not overlap, so that we can maximize the sample size. These buffer zones include 1.3 million unique road segments. Most of these segments are tiny streets and do not have speed information. Our raw data have about 1.8 billion hourly speed observations from more than 350,000 unique road segments.

2.5 Baseline Sample

2.5.1 Road Segments Directly Affected by the New Subway Lines

Our sample includes non-random patches of a city's roads. The average effect on sample road segments does not have an intuitive interpretation. Because subway lines are built in places with heavy traffic and aim to alleviate congestion on nearby roads, as a first pass, it is interesting to investigate whether subways have *any* effect on road congestion. Distance to the subway is arguably related to how much a road segment is affected, but it is an imperfect measure. Whether the subway is effective in diverting traffic from certain roads depends on the substitutability between subway trips and traffic through these roads.

The trip-planning feature of the digital map application can be used to identify road segments that are close substitutes to, and are directly affected by the new subway line. These segments are those en route if one chooses to drive instead of taking the subway. Specifically, we first divide the buffer zone around each new subway line into 1km-by-1km grids. Between any pair of grids, we find the best public transit route (or routes, as the digital map service sometimes recommends several alternative routes). We save the collection of pairs between which the best public transit routes involve the newly built subway line. Then for each pair of grids in this collection, we find the best route(s) for driving. Road segments in these optimal driving routes are regarded as those directly affected by the new subway line. We repeat the same process under the typical traffic conditions for weekday morning and afternoon rush hours. This procedure yields 8,275 road segments that are most likely to be directly affected by the new subway lines. Online Appendix A.1 provides additional illustrations on how directly-affected roads are selected.

[3] We picked these buffer zones before seeing the speed data.

2.5.2 Assigning Control Segments to Treated Lines

In the baseline sample, there are 45 groups of treated road segments. Each group represents road segments directly affected by a new subway line. Our identification strategy is a stacked DID specification, where road segments from the 17 control cities serve as controls. We randomly divide control segments into 45 equal-sized subsamples and assign them to each treated group. For control segments, we include all road segments in the buffer zone, not only those that are directly affected.

We create a variable for each treated line that indicates the time relative to the opening date. The opening date of the treated line is assigned to the corresponding control road segments as their "fake" opening date. Road segments near each treated subway line and the corresponding control segments form a "group" of comparison. These groups are then stacked according to the time relative to the opening date. We control the group-by-time-elative-to-opening fixed effects, which restricts the comparison between the treated and the control to be within the same *calendar* time. The inclusion of road segments from control cities and the use of the stacked DID approach eliminates macro trends and seasonality that are common to all cities.

2.5.3 Further Restrictions on the Sample

We impose several additional restrictions on the sample. We first drop weekends and national holidays. In the baseline we keep only morning rush hours between 7 AM and 9 AM and evening rush hours between 5 PM and 7 PM.[4]

We drop local streets due to concerns over the quality of speed information on those narrow, less travelled streets. We notice some local streets are actually in restricted-access areas such as paved paths in a park or driveways in a gated residential community. Many local streets have a substantial portion of missing values, presumably because no devices have passed through during the period of time. Although we could assign traffic-free speed to these missing values, we choose to be on the conservative side.[5]

For each group of comparison we include observations that are up to 48 weeks post opening. Although our sample expands over 18 months, the longest time we can track a new subway is about 1 year after its opening. This is due to two reasons. First, the first large batch of subway

[4]Baidu Maps uses the same definition of rush hours. Using the data we have, Appendix Figure A.4 shows that peak traffic takes place between 7 and 10 AM in the morning and between 5 and 7 PM in the evening. The discrepancy between our data and Baidu Maps' definition is probably due to the fact that our data have a much smaller scope and are not representative of the cities in our sample. As a robustness check, we also estimate the effect on road speed in non-rush hours in Section 3.3.

[5]About 76% of our road segments are local streets. Local streets may play an important role in alleviating congestion on the main roads. Akbar & Duranton (2017) find that even at times when main roads are very congested, many local streets are still relatively smooth. They suggest that the existence of small local streets essentially puts an upper bound on traffic congestion on the main roads. We leave the substitution patterns between main and minor roads to future research.

openings took place in December 2016, about a year before the end of our sample period. Second, 12 of the 25 treated cities have multiple line openings during the sample period; the gap between two openings is typically about one year. Tracking the effect up to 48 weeks guarantees that we do not include effects from further expansions of the subway system.

We also restrict the baseline sample up to 6 weeks prior to the opening. Subway construction itself could affect traffic conditions on nearby roads, thus making pre-trends uninformative about the nature of traffic conditions on these roads. We take advantage of an engineering fact of subway construction. Once the construction is mostly finished, a new subway line requires some time testing the hardware and software systems. For the subway lines we study here, the testing period usually runs 2 to 3 months before opening to the public. So traffic during the 6 weeks prior to opening can be assumed to be unaffected by the construction.[6] This relatively short pre-treatment period is also chosen because in many cases the pre-treatment is in the post-treatment period of another new subway line in the same city. We perform robustness checks to validate this assumption. In particular, the inclusion of flexible time trends and a longer pre-treatment period also yield quantitatively similar results.

An additional 10% of the road segments are dropped from the sample due to missing values during the sample period. The baseline sample is a perfectly balanced panel of road segments.

The official opening of a subway line usually involves a ceremony with the presence of government officials and media. In order to make sure that the system is ready for carrying real passengers, in some cases there were "test rides" before the line officially opens to the public. The existence of those test ride periods may generate a spurious positive trend on road speed in the pre-treatment period, and bias our DID estimate downward. We find detailed project progresses for each subway line from various sources.[7] We treat these test ride periods differently depending on whether they are contiguous to the official opening date. If the official opening date immediately follows the test ride period, we re-define the opening date as the first day of the test ride. If the test ride ends several days prior to the official opening date, we drop the test ride period from the sample.

[6]The construction of a subway takes several major steps. The first step involves digging tunnels. Modern shield tunneling technology can advance tunnels underneath the roads. Yet it still requires operations on the ground, such as lifting the tunnel boring machine up and down, trucking rocks and pebbles and pumping water from the underground. These operations are likely to affect traffic. Once tunnels are dug, tracks are laid and electrified. Usually one year prior to the line opening, scale test cars are run on the electrified tracks. In the meantime, subway stations are under construction. All these constructions are likely to affect traffic on nearby roads. Usually, major constructions are completed 6 months before opening. Then subway trains not carrying passengers are tested. In the last few months before opening, final touches on the system may still be in progress. The final approval for public operation comes after a panel of specialists inspects the system, which usually takes place weeks ahead of the official opening. Appendix A.4 shows examples of the detailed administrative and engineering processes from two subway lines in the sample.

[7]Appendix Table A.1 lists the periods for test rides of the treated subway lines. Links to sources are also provided.

2.5.4 Summary Statistics

Table 2 reports the summary statistics of the sample. Panel A compares city-level characteristics between treated and control cities. The average treated city is more than twice as large (by population) and about 10% richer (by GDP per capita) than the average control city.

Panel B presents the speed and congestion index by the type of road and treatment status in the baseline sample. Arterial and sub-arterial streets account for the majority of road segments in the sample (local streets are dropped). Treated and control cities have similar compositions of road types. This is reassuring because it shows that although the control cities are in general much smaller and less prosperous than the treated cities, the road networks of the studied areas are not substantially different. The average speed is slightly above 30 kilometers per hour during rush hours. Except for highways, which we have a small number of segments, the average congestion indices are surprisingly similar across all types of roads, at around 1.7. This suggests that in equilibrium, there is little room for arbitrage by taking a detour on smaller roads. Average speed and congestion index for each road type are also similar between the treated and control road segments.

2.5.5 Weekly Average Residual Log Speed

In order to reduce dimensionality and computational burden, we first partial out fixed effects in road speed and group the residuals at the weekly level before embarking on the baseline model.[8] We first run the following regression:

$$\ln speed_{lt} = \lambda_{l,dow,h} + \varepsilon_{lt}, \tag{1}$$

where $\lambda_{l,dow,h}$ is the complete set of fixed effects, each of which indicates a specific road segment (l) in a given hour of the day (h) on a given day of the week (dow). The residual from this regression, $\ln \tilde{speed}_{lt} = \hat{\varepsilon}_{lt}$, measures the log point deviation of hourly speed from the "usual" speed—the average speed on the same road segment, in the same hour of the day, on the same day of the week. Such a saturated model is not possible without the high-frequency data at the granular level. The adjusted R-squared of this regression is 0.67, suggesting that a large portion of the variation in speed can be explained by the "usual" traffic pattern. Note that $\ln \tilde{speed}_{lt}$ may still include time trends and seasonality. The standard deviation of the residual is 0.24, indicating that seasonality or trends in speed remains non-negligible.

Hourly residual log speed is then averaged at the weekly level.

$$\ln \tilde{speed}_{lw} = \frac{1}{N_w} \sum_{t \in w} \ln \tilde{speed}_{lt},$$

[8] The results are almost identical when we run the regressions in one step. Appendix Table B.5 replicates the baseline with one-step regressions. The small differences are due to different sizes across weeks: due to days without observations and national holidays, some weeks may have a smaller number of hourly observations, so the weights for some observations are different between the one-step and the two-step estimations.

where w is the week *relative* to opening. Week 0 starts from the day the subway line was launched. N_w is the number of hours sampled in the week. $\ln\tilde{speed}_{lw}$ should be interpreted as the weekly average log point deviation from the road segment's "usual" speed.

3 Effects of Subway on Road Speed

3.1 Baseline Specifications

We estimate the effect of subway on nearby road speed first using a stacked DID model with 45 subway line openings. Each of the 45 components of the DID model compares the change in residual log speed in treated road segments before and after the launch of a new subway line with the contemporaneous change in speed in control road segments. The empirical model can be written as:

$$\ln\tilde{speed}_{lgw} = \sum_{w=\underline{w},\ w\neq -1}^{\bar{w}} \beta_w \cdot T_{lg} \cdot d_{gw} + \lambda_l + \lambda_{gw} + \gamma_t \cdot d_t \cdot \mathbf{X}_c + \varepsilon_{lgw}, \tag{2}$$

where g indicates a group of road segments that includes the directly-affected road segments near a treated subway line and their randomly assigned road segments from control cities. T_{lg} is a binary variable indicating treatment status of the road segment. d_{gw} is a binary variable that takes value 1 when an observation in group g is w weeks away from the launch of the new subway line, with $w = 0$ indicating the week of launch. $[\underline{w}, \bar{w}]$ is the extent of time periods included in the sample, where $\underline{w} = -6$ and $\bar{w} = 47$. λ_l is the road segment fixed effect. λ_{gw} is the fixed effect of group g in the w^{th} week relative to the opening. d_t is binary variable that indicates *calendar* week t; $\mathbf{X}_c$ is a vector of city-level characteristics. Note that d_t is colinear with λ_{gw}, $\mathbf{X}_c$ is colinear with λ_l, so d_t and $\mathbf{X}_c$ are not separately included in the regression.[9]

We choose the week prior to the launch, $w = -1$, as the base for comparison and impose β_{-1} to be 0. All β_w with $w \geq 0$ should be interpreted as the percent increase in speed in week w as a result of a new subway line opening, relative to the week before opening. Because the outcome variable is the logarithm of speed relative to the usual level in the same road segment, the coefficient can also be interpreted as the percent of time saved to travel through the road segment. We expect $\beta_w > 0$ for $w \geq 0$ if subways are effective in alleviating road congestion. Standard errors are clustered at the group level.

The fairly saturated model of Equation 2 is robust to many common concerns over identification. Differences in levels of speed between the treated and control groups do not matter because road segment fixed effects purge out any level differences. Because in each case, time is realigned

[9]There are 45 groups, which is not a small number. We also experiment with the 95% confidence intervals from the block wild bootstrap procedure (Cameron et al., 2008). Results are almost identical. Randomly assigning the control segments into different groups could generate spatial and temporal correlations as nearby segments will appear as controls in different weeks for different cities. Clustering the standard errors at the city level could address this problem. Standard errors are also almost identical if they are two-way clustered at the city and the group levels.

to indicate the week relative to opening, λ_{gw} captures the macro trend and seasonality common to both treated and control road segments. Key to identification is the usual parallel trend assumption, which we can assess by testing whether $\beta_w = 0$ for $w < 0$, both individually and jointly.

Probably the biggest challenge to identification is the *differential seasonality* in traffic speed between treated and control cities. Differential seasonality means that traffic speed in some particular months is systematically different between the treated and control cities, even though there may not be a difference in overall trends. Three facts make this a valid concern. First, Table 1 shows that most (25 out of 45) of our treated lines opened in December. Second, January is China's holiday season and economic activity slows down. Appendix Figure A.5 shows that road speed in control cities increased by about 20% relative to the usual level in weeks around the 2017 Chinese New Year (CNY), which landed on January 28$^{\text{th}}$. Thus the launching dates of many subway lines coincide with a period when road speed start to increase. Third, Table 2 Panel A shows that treated and control cities differ substantially in size and economic performance. Larger and richer cities attract more migrant workers, and thus are likely to experience a larger hike in traffic speed around the CNY, as migrant workers return to their hometowns. With many lines launched right before the CNY, differential seasonality may confound our result, and it cannot be ruled out by a parallel pre-trend.

To account for potential differential seasonality, the term $\gamma_t \cdot d_t \cdot \boldsymbol{X}_c$ allows the residual speed in each *calendar* week to differ linearly with observable city-level characteristics, $\mathbf{X}_c$. In the regression $\mathbf{X}_c$ includes log population and log GDP per capita. Admittedly, this approach is imperfect because (1) we are only able to control for a limited number of city-level characteristics and (2) the relationship may not be linear. Online Appendix B.1 presents additional analyses to support this approach. In particular, Figure B.2 shows that results are robust to the inclusion of higher-order polynomials of city-level characteristics. Figure B.3 plots the estimates of γ_t, which captures how road speed systematically differs by city characteristics in each calendar week. It shows that the role of differential seasonality is limited, it is only relevant in the weeks around the CNYs. Results are robust when these weeks are excluded.

The issue of differential seasonality arises because the treated and control cities are significantly different along some dimensions. As an alternative approach to address this issue, we estimate a standard two-way fixed effects model using observations only from treated units:

$$\ln \tilde{speed}_{lt} = \sum_{w=\underline{w},\ w\neq -1}^{\bar{w}} \beta_w \cdot T_l \cdot d_{lt}^w + \lambda_l + \tau_t + \gamma_t \cdot d_t \cdot \boldsymbol{X}_c + \varepsilon_{lt}, \tag{3}$$

where the binary variable d_{lt}^w takes value 1 when segment l in calendar week t is w weeks away from the launch. τ_t is the calendar week fixed effect and captures the average seasonality or time trend in treated cities. $\gamma_t \cdot d_t \cdot \boldsymbol{X}_c$ is controlled to guard against the possibility that cities with larger seasonal fluctuations are more likely to have new lines launched around the time when seasonality is most pronounced (e.g., December). Because every unit is eventually treated, τ_t and β_w are separately estimated from the variation in the timing of launch. Pre-treatment periods of the units that are treated later serve as controls for the post-treatment periods of the units that are treated earlier. With most of the treated lines launched in December, it is a concern that our sample may not have enough variation to credibly identify the two time-varying parameters (Abraham & Sun, 2018; Goodman-Bacon, 2019). We estimate Equation 3 to corroborate the results from the stacked DID model.

Panel A of Figure 2 shows $\hat{\beta}_w$'s and the corresponding 95% confidence intervals from the stacked DID model in Equation 2. There is no evidence of differential pre-treatment trends. Road speed increases by about 2.5% in the week of opening and keeps rising in the first month. The effect reaches its maximum in the 6$^{\text{th}}$ week at around 5% before declining and stabilizing at between 3% and 4%. We still see a sizable effect 48 weeks after the launch. The remaining panels report results from estimating different versions of the two-way fixed effects model in Equation 3. Panel B includes only the treated lines and does not adjust for seasonality. Overall, we still see a positive and significant effect throughout the sample period, although the magnitude of the effect is somewhat smaller and the coefficients associated with later weeks are less precisely estimated. Panel C adds the adjustment for differential seasonality to account for the possibility that cities with larger seasonality are more likely to have lines launched in December; results are almost identical.

Week-by-week estimates in Equations 2 and 3 can be seen as "dynamic" specifications. We estimate the average effect in "static" specifications by replacing weekly dummies with a post-treatment indicator. In Equation 2, $\sum_{w=\underline{w},\ w\neq -1}^{\bar{w}} \beta_w \cdot T_{lg} \cdot d_{gw}$ is replaced by $\beta \cdot T_{lg} \cdot Post_{gw}$; in Equation 3, $\sum_{w=\underline{w},\ w\neq -1}^{\bar{w}} \beta_w \cdot T_l \cdot d_{lt}^w$ is replaced by $\beta \cdot T_l \cdot Post_{lt}$. Table 3 reports results from specifications corresponding to the three panels in Figure 2. In the stacked DID specification, the estimated average effect is 4.4%. The two-way fixed effects models, whose results are reported in Columns 2 and 3, lead to a somewhat smaller effect at 3.6%. We use the stacked DID specification for the remaining analyses in the paper.

A few remarks are in order about endogeneity. Subway lines are likely to be built in areas with congested traffic. But this should not be a threat to the internal validity as long as these areas do not exhibit different *trends* in traffic congestion. Depending on the initial levels of road congestion, the magnitude of the subway's effect on alleviating congestion may be different. To the extent that subway lines are likely to be built in otherwise congested areas, the congestion-relieving effect we find here is expected to be greater than the hypothetical case where subway lines are randomly placed. We believe that the effect of randomly-placed subway lines is a parameter that bears little policy relevance. We do, however, present evidence in Section 3.3 on how the effects differ by the road segment's initial level of congestion and its geographical location relative to the subway.

Because a subway takes many years to build, its opening is widely reported and largely anticipated. But urban traveling is not something one can easily arbitrage over time — one needs to get to work every workday. It could be a concern though, with the expectation of a subway line opening, people who have a higher idiosyncratic value for public transit are more likely to move to places near the line before it opens. This may have mixed impacts on our estimate. On the one hand, one may imagine that these people are more likely to ride the subway, so the estimation out of this selected population is likely to be larger than the effect on the whole population. On the other hand, these people are arguably less likely to own a car and more likely to travel by bus. To the extent that the marginal negative effect of a bus ride on road congestion is smaller than a car trip, diverting bus rides to subway trips is not as effective in reducing traffic congestion as diverting the same number of car rides. So our result may underestimate the average effect on the whole population. Nevertheless, using individual-level data from household travel surveys, Section 4 shows that restricting the sample to households that had not moved in the past 5 years does not affect the substitution patterns between modes of transportation, suggesting that the endogenous selection of residents is unlikely to be a major concern here.

Duranton & Turner (2011) find that over a period of 10 years, VKT increases proportionately to road lane kilometers, which is consistent with the prediction of the fundamental law of road congestion—building more roads does not reduce traffic congestion. They also find that the provision of public transportation does not affect VKT. Our results suggest otherwise, at least in the short run. Increases in aggregate VKT may come from increases in driving by current residents or from increased migration. Existing residents likely respond to the change in transportation infrastructure quickly. The fact that we find a persistent positive effect one year after a subway line opens basically rules out that demand for driving from existing residents explains the fundamental law. We cannot rule out the possibility that in the longer run, an improved public transit system will attract migration, increasing VKT and congestion on nearby roads. However, in a recent study, Gendron-Carrier et al. (2018) find that the opening of a subway system has a persistent and stable effect on reducing air pollution for up to 3 years. Using a panel dataset that tracks expansions of subway and road systems in U.S. cities from 1991 to 2014, Pang & Shen (2019) find that a 10 percent expansion in subway length reduces traffic on interstate highways by 0.8 percent. These findings imply that the effect on road congestion may persist for a relatively long period of time.

3.2 Robustness Checks

3.2.1 Confounding Trends

To alleviate the concern that our results are driven by some confounding time trends, we test whether the timing of the effect coincides with the timing of subway line openings. We expand the sample to up to 48 weeks before and after the subway line opening and repeat the same specification as in Column 1 of Table 3, each with a placebo subway opening week between 48 weeks prior and 48 weeks posterior to the actual opening week. The Wald statistics associated with the treat×post variable from these regressions are then plotted against the placebo opening week relative to the actual week of opening. If the results are truly due to the subway line opening instead of some confounding trends, the largest Wald statistic should be found at around the actual week of the line opening. Figure 3 confirms that this is the case.

Some other events may take place at the same time of the launch of a subway line. Road blocks due to construction may be removed; shops may open near the subway station; road patterns may change; buses may be re-routed. Admittedly, we cannot partial out impacts from all these possible events. However, as long as the timing of these events does not synchronize with the subway opening, the null effect in the pre-period and the placebo results in Figure 3 suggest that they are unlikely to drive our result. The spatial pattern of these confounding effects are likely to be different from the effects of the new subway line. Therefore, we can also devise ways to separately distinguish these effects. We discuss more evidence on this front in Section 3.3.[10]

3.2.2 Longer Pre-treatment Periods

We restrict the sample to be within 6 weeks prior to a subway line opening due to the concern that subway construction may directly affect road congestion. Within the 6-week window of the pre-treatment period, all major constructions should have concluded and traffic on the ground should be back to "normal." Nevertheless, 6 weeks is a short period of time, so we experiment with including 12, 24, and 48 weeks prior to the subway launch. Columns 1 to 3 of Table 4 show that the estimates are similar across different lengths in the pre-treatment periods. Figure 4 presents the week-by-week estimates up to 48 weeks pre-treatment. The pre-treatment effects all hover around 0 and there is an immediate and significant effect right at the time of the subway line opening.

The issue of pre-treatment trends can be circumvented when we focus on the effect in a narrow neighborhood around the date of launch. We estimate difference-in-discontinuity models by including a flexible time trend with up to the 5^{th} polynomial, separately for the treated and control segments, and separately for the pre- and post-periods. The coefficient associated with the treated × post variable thus indicates the discontinuous change in road speed at the time of subway launch relative to that in the control segments. Columns 4 to 6 of Table 4 report that the discontinuous effect is salient across different specifications. The most saturated model yields an effect of around 2.3%, which is close to the first-week effect as shown in Figure 4 and various panels of Figure 2 .

[10]For example, Online Appendix B.4 shows that conditional on the distance to the subway line, road segments that are closer to a subway station do not seem to have a slower traffic speed after the launch. This could rule out the effects due to shop openings and bus re-routing, provided that such events bring additional traffic to road segments near subway stations. In general, Section 3.3 shows various pieces of evidence that the effect is indeed larger in segments that we would expect to be more affected. Confounding factors are unlikely to be consistent with all these patterns.

3.2.3 Subgroups of Subway Lines by Time of Opening

Our data cover 18 months between August 2016 and January 2018. The opening dates of the treated lines spread across this period. Therefore, although we restrict the observations from each treated line to be within 6 weeks before and 48 weeks after the opening, the number of post-opening periods in the sample differs across treated lines. In the dynamic specification in Equation 2, each β_w is estimated from a different mix of treated lines. In the static specification the treated $\times$ post variable is identified from a mix of treated lines and periods. Different composition complicates the interpretation of our results.

We can circumvent this problem by restricting the sample period such that all β_w's are estimated from all subway lines, although this would substantially reduce the sample size. For lines launched in December 2017, we are able to track up to 4 weeks since launch. Most new subway line openings in 2016 were in December. Although our data start from August 2016, there is a gap between early October through November. In Table 5 Column 1, we narrow the sample period to 6 weeks prior to and 3 weeks posterior to the subway line opening so that the coefficient of treated$\times$post is estimated from the same set of weeks relative to line opening for most subway lines; this results in a significant effect of 3.6%.

To avoid the problem of differential composition yet still be able to include a reasonably-long sample period, we divide the treated lines into three sub-samples: 21 lines launched before January 31st, 2017; 9 lines launched between February 1st and November 30th, 2017; and 15 lines launched after December 1st, 2017. We include the maximum length of period that all the lines in a specific subsample can cover.[11] We run separate estimations for each sub-sample. Because there are small numbers of lines in each group, we report the 95% confidence intervals from the block wild bootstrap (Cameron et al., 2008).

The dynamic effects of the subway line opening for each group are shown in Figure 5. For all three subsamples we see a statistically significant increase in road speed right after the opening of a subway line. For the first and second subsamples, for which there are several months in the post-period, the effect hovers between 4% and 5%. For the last subsample, the effect is about 2.5% in the first four weeks. In general the coefficients in the pre-treatment period are around zero, except for the earliest few weeks in the estimation for the last subsample, where the coefficients are estimated less precisely. We speculate that the small sample size and the small number of cities in each subsample makes it hard to separately identify β_w and γ_t.[12] Results from the static model are shown in Columns 2 to 4 of Table 5. The average effect is about 5.0%, 3.9%, and 2.2%, respectively,

[11] For the earlier group, we include 6 pre-treatment weeks and 48 post-treatment weeks. Note that not every line in this subgroup has an observable pre-period up to 6 weeks. For the middle group, we include 20 weeks in the pre-period and 20 weeks in the post-period. For the later group, we include 48 weeks in the pre-period and 3 weeks in the post-period including the week of opening.

[12] Appendix B.7 shows results from an alternative specification where the $\gamma_t \cdot d_t \cdot X_c$ terms are not controlled for. We find the pre-period coefficients are more precisely estimated and are closer to zero.

for the three subsamples.

3.3 Heterogeneous Effects by Road Segments Characteristics

The treated road segments in the baseline are those directly affected by the new subway lines. These are the likely group of road segments for which the congestion-relieving effect of a subway is the largest. However, an arguably more interesting question is how much a subway reduces the *overall* road congestion and how the effect differs across different road segments in the urban transportation network. There is at least one constraint and one caveat to properly answer this question. The constraint is that we do not have speed data on all, or a random sample of, the road segments in a city. The caveat is that the effect of a subway line on the overall congestion in the city depends on the size of the city, location of the subway, and the layout of the existing road network. A couple of subway lines may be sufficient for a medium-sized city, while China's largest cities, such as Beijing and Shanghai, already have over a dozen subway lines. The marginal effect of a subway line on the overall road congestion may be different by line, but this difference is less informative without knowing the existing traffic patterns and transit infrastructure in the city.

In light of these two observations, we rephrase the question to how the effects differ on road segments with different characteristics. These characteristics include the segment's geographic location relative to the subway line, its initial congestion level, and its position in the broader urban transportation network. These heterogeneities could speak to the effect of a typical subway line in a hypothetical road network. We expand our treated sample to include all road segments in treated cities, and run the following regression:[13]

$$\ln \tilde{speed}_{lgw} = \sum_{k \in K} \beta_k \cdot T_{lg} \cdot Post_{gw} \cdot z_{lk} + \lambda_l + \lambda_{kgw} + \gamma_t \cdot d_t \cdot \mathbf{X}_c + \varepsilon_{lgw}. \tag{4}$$

K is a set of complete partition of road segment characteristics. Binary variable z_{lk} is equal to one if segment l has characteristic k. λ_{kgw} is the segment-characteristic-by-group-by-week-to-opening fixed effects. This is a set of more saturated fixed effects than those in the baseline. It allows time trends in traffic speed to differ by road characteristics. Each β_k indicates the speed-enhancing effect of the subway on road segments in treated cities with a certain characteristic, compared with road segments with the same characteristic in control cities.[14] For each set of segment characteristics, each block of Figure 6 plots β_k's and their associated 95% confidence intervals from estimating

[13] We include all road segments in cities with treated subway lines, except for local streets, in the treated sample. Because in some cities there were multiple subway line openings during the sample period, in order to avoid double counting, we group subway lines that opened within the same city-month and consolidate comparing groups. In each consolidated group, the treatment time is assigned as the first opening date of the treated lines. We consolidate 45 line openings into 35 groups.

[14] In some cases we are unable to produce the exact same characteristics in the control sample. Because some subway lines in control cities are still under construction, we cannot identify "directly affected" road segments using the trip planning function on the online digital map. In these cases we define directly affected segments as those within 1 kilometer away from the proposed subway line. In each control city we only extract one buffer zone. So in the set of results where we investigate how the effect spreads across the road network, we control for λ_{gw} instead of λ_{kgw}.

Equation 4. The figure also reports the magnitude of the coefficient and its standard error, as well as the average congestion index for road segments with the corresponding characteristic.

We first focus on nearby roads and include road segments that are within the 2.5-kilometer buffer zones from treated subway lines. The first estimate in Figure 6 shows that on average the launch of a new subway line increases nearby road speed by 3.1%. The second block of coefficients is from a regression where we divide the nearby road segments into those directly affected by the treated subway lines (those in the baseline regressions) and those not directly affected. The effect for the former group is 1.8 percentage points larger than the latter group, a statistically significant difference. The third block of coefficients show that the speed-enhancing effect of a subway declines quickly by distance. Speed on road segments that are within 1 kilometer of the treated subway lines increases by 4.8%, while the effect on those between 1 and 2.5 kilometers is 2.2%. The next regression divides the nearby road segments into those largely parallel to the new subway and those largely orthogonal to it. The effect is somewhat stronger in parallel roads.

Next we divide road segments into halves based on whether the average of pre-treatment congestion index is below or above the sample median (1.7). The congestion-relieving effect is much larger in initially more congested road segments. Speed on these road segments increases by 4% upon a subway line opening, while the effect on less congested roads is less than 2%. Next we divide the sample by road type. Because there are not many highways and urban expressways, we combine these two. The effect is strongest in arterial roads. It is intuitive because most of the subway lines are parallel with, and sometimes directly underneath, arterial urban roads.

Recall in the baseline sample that road speed is derived from four hours in a day, two in the morning (7AM-9AM) and two in the evening (5PM-7PM). We divide the observations into morning and evening rush hours, and categorize road segments-by-rush hour into those "with" or "against" the traffic. A segment is "with" traffic in morning rush hours if its average congestion index in morning rush hours is larger than that in evening rush hours. If a segment is "with" traffic in the morning, it is considered "against" traffic in the evening, and vice versa. Average congestion index is 1.89 for road segments that are with traffic and 1.54 for those against traffic. The effect is larger in road segments that are with traffic. In the next regression we also include observations from off-peak hours. Traffic is 20 percent worse in rush hours and the effect is twice as large.

Two conclusions are immediately evident from these heterogeneous regressions. First, the congestion-relieving effect is larger in more congested road segments. This is consistent with the hypothesis that travelers who experience high congestion levels are more likely to switch to public transit (Anderson, 2014). Second, the effect declines quickly by distance to the treated subway line, which suggests that access is an important constraint to using public transit.

Proponents of subways argue that subway lines exhibit a network effect. The addition of a new subway line could increase ridership in existing subway lines, reducing traffic congestion on road segments that are close to existing lines even if they are far away from the new subway line.

The final block in Figure 6 examines this hypothesis. In the regression, we include road segments in buffer zones of existing subway lines in treated cities (listed in Table 1 Panel B). Road segments in treated cities are divided into four mutually-exclusive groups. The first group includes those within 1 kilometer from the new subway line ($d_1 \in [0,1]$). The second group includes those between 1 and 2.5 kilometers from the new subway line ($d_1 \in (1,2.5]$). The remaining two groups are more than 2.5 kilometers from the new subway line. While those in the third group are within 1 kilometer of an existing subway line ($d_1 \in (2.5,\infty)$, $d_2 \in [0,1]$), those in the fourth group are farther away ($d_1 \in (2.5,\infty)$, $d_2 \in (1,\infty)$). If there is a network effect in the subway system, we expect to find a meaningful speed increase in the latter two groups, and the effect on the third group to be larger than that on the fourth group. Our findings confirm these hypotheses. The effect is 2.1% for the third group and 1.6% for the fourth group.

4 Substitution Patterns among Modes of Transportation

Higher speed suggests less traffic on the road. In this section we use individual-level travel data from Beijing to show substitution patterns between subway and other modes of transportation. Due to data limitations, the results presented in this section are largely descriptive and should be interpreted as correlations rather than causalities.

Individual travel records are from one-day travel diaries, which are part of the *Household Travel Surveys* in Beijing (BTI, 2010,2015). We obtain two rounds of surveys in 2010 and 2015. During this period 16 new subway lines and extensions were launched in Beijing. The smallest identifiable geographic level is a Transportation Analysis Zone (TAZ).[15] For each sample year, we measure a TAZ's access to the subway by its average distance to the nearest subway station.[16] The travel diaries record detailed modes of transportation used in each trip. We drop respondents who did not have a trip record or only had walking trips. We count the number of trips that use bus, subway, and car. Notice that an individual can use multiple modes of transportation in a day's trip.

Table 6 reports the correlations between TAZs' distance to a subway and residents' choice of transportation modes. All models control for TAZ and year fixed effects, so the coefficient should be interpreted as the correlation between the log change in distance to a subway and the change in the number of trips that use a certain mode of transportation. The surveys are repeated cross-sectional and do not track the same households across different rounds, so we control for detailed household and individual characteristics to account for changes in the demographic composition

[15]TAZs are small geographic areas. In 2010, Beijing is divided into about 2,000 TAZs. Within the 4$^{\text{th}}$ Ring Road, where the density of subway lines is the highest, the median TAZ covers an area of 1.3 square kilometers and contains about 10,000 residents.

[16]We calculate a TAZ's average distance to the nearest subway station in the following steps. First, we divide the city into 500m$\times$500m grids. Second, we calculate the distance between the center of each grid and the nearest subway station. Third, we calculate the TAZ's average distance to the nearest subway station as the weighted average of grids' distances to their nearest subway stations, where the intersected area between the TAZ and the grid is used as weight. We repeat the same procedure separately for subway systems in 2010 and 2015.

in a TAZ.

Panel A shows that on average, each commuter has 0.47 subway trips, 0.5 bus trips, and 0.56 car trips per day. Improved access to subway is positively associated with subway trips and negatively with bus and car trips. Estimates in Columns 1-3 suggest that when the TAZ's distance to the nearest subway station declines by 70% (log distance declines by 1), the number of subway trips per commuter increases by 0.21, and the number of bus and car trips per commuter decreases by 0.085 and 0.062, respectively. The coefficients imply that about 1 additional subway trip is associated with 0.4 fewer bus trips and 0.3 fewer car trips.

The *Survey* also has detailed information on car ownership and usage; in particular, it records the mileage driven on cars in the past year. Although the mileage variable is based on recollection and likely has plenty of measurement error, we use it to check whether car usage drops in association with improved access to a subway. We sum over VKT from all the cars a household owns, plus one to that sum, and take the natural logarithm, so households that do not own a car are also included in the sample. Column 4 shows that, in contrast to what the fundamental law of road congestion predicts, improved access to a subway is indeed associated with less driving. The elasticity of VKT on distance to subway is about 0.7. Column 5 estimates the model with the level of VKT as the outcome variable. As the log distance to the nearest subway station decreases by 1, average VKT per household declines by 632 km, which is about 15% of the sample average.

Households that value subway trips more may move to neighborhoods that are expected to have improved access to a subway. Although it is hard to credibly identify causal effects using the current data, we can alleviate concerns over false correlation by focusing on the non-movers. Panel B includes only households that had not moved between 2009 and 2014. The results are essentially unchanged.

To investigate the driving behavior in more detail, Panel C further restricts the sample to households that had a car in 2010.[17] For this smaller sample, access to a subway is associated with increased subway trips and reduced bus trips, some evidence of reduced car trips, but no evidence of smaller mileage on cars. This result suggests that for households that already own a car, driving behavior does not change significantly in association with improved access to a subway. Therefore, the reduction in overall driving may come from reduced car purchases.

[17] The 2015 *Survey* asks the age of each vehicle. We include all households in the 2010 *Survey* that owned a car and households in the 2015 *Survey* that owned a car made in or before 2010. This imputation is not perfect. Households that owned a car in 2010 but replaced it with a new car between 2010 and 2015 will be mistakenly excluded from the sample. First-time car owners after 2010 who bought a second-hand car made before 2010 will be mistakenly included in the sample. The lower share of car trips in this sample compared with those in Panel A and Panel B is probably due to the fact that the sample here is more skewed towards respondents in the 2010 survey.

5 Welfare Analysis

5.1 Conceptual Framework

In this section we build a conceptual framework of transportation mode choices, which helps link our empirical finding to welfare implications. Consider a city with N commuters, each inelastically demanding one trip. Indirect utility of a commute for individual i using transportation mode m is written as:

$$V_{im} = b + A_m - c_m + \xi_{im}. \tag{5}$$

The individual can choose between two modes of transportation, road vehicle (denoted as a, including cars and buses) or subway (denoted as s), $m \in \{a, s\}$. b is the benefit from the trip, which is assumed to be a constant. A_m is the amenity value of mode m. Amenities include things like comfort and privacy. c_m is the *private* cost of each commute in mode m. Specifically:

$$c_m = \begin{cases} c(N_a), & \text{if } m = a \\ r, & \text{if } m = s. \end{cases} \tag{6}$$

The private cost of a commute by road vehicle has two components. The first component is the fixed cost of taking such a trip. For commutes by car, this includes costs of fuel, parking, and depreciation of the vehicle; for commutes by bus, it is the fare. The second component is the time cost, which is an increasing function of the number of total road vehicle commutes, $c'(.) > 0$. $N_a \cdot c'(.)$ captures the negative externality one additional trip imposes on all other trips on the road. r is the cost of one subway commute, which is the sum of the subway fare and the time cost. We assume there is no congestion in the subway system, so the cost of a subway trip does not depend on the number of riders.

In addition to the private cost, the government spends τ_s on each subway trip, which includes construction costs as well as subsidies to cover operating costs. Assuming the subway fare plus government subsidies equal total cost, the *social* cost of each subway trip is $r + \tau_s$. Similarly, the government spends τ_a on an average commute by road vehicle, which includes construction and maintenance costs of the road infrastructure, as well as subsidies to bus trips.

We denote $u_m = b + A_m - c_m$, which is the same for all commutes given the mode of transportation. ξ_{im} captures commuter i's idiosyncratic preference for mode m. We assume ξ_{im} follows *i.i.d.* distribution across individuals. $\xi_{ia} - \xi_{is}$ captures individual i's preference for road vehicle relative to subway.

Without subway, r can be seen as infinitely large and everyone uses road vehicle for commute, $N_a = N$. The aggregate social welfare is:

$$W_{ns} = N \cdot (b + A_a - c(N)) + \sum_i \xi_{ia} - \tau_a \cdot N, \tag{7}$$

where the subscript *ns* indicates the case with "no subway." When a subway is available, we assume a non-empty subset of individuals will switch to the subway. The person who is indifferent between the two modes of transportation has preference such that $\xi_{ia} - \xi_{is} = c(N_a) - r + A_s - A_a = \bar{\Delta\xi}$. Commuters who have $\xi_{ia} - \xi_{is} > \bar{\Delta\xi}$ will remain traveling by road vehicle. We denote this group of commuters as Ω_a. Commuters who have $\xi_{ia} - \xi_{is} \leq \bar{\Delta\xi}$, denoted as Ω_s, switch to the subway. The aggregate social welfare is:

$$W_s = [N_a(b + A_a - c(N_a)) + \sum_{i \in \Omega_a} \xi_{ia}] + [N_s(b + A_s - r) + \sum_{i \in \Omega_s} \xi_{is}] - (\tau_s \cdot N_s + \tau_a \cdot N_a). \quad (8)$$

Change in social welfare due to the introduction of a subway is therefore:

$$\Delta W = \underbrace{[c(N) - c(N_a)] \cdot N_a}_{\text{stayers's gain}} + \underbrace{[(c(N) - r) \cdot N_s - (A_a - A_s) \cdot N_s - \sum_{i \in \Omega_{is}} (\xi_{ia} - \xi_{is})]}_{\text{switchers's gain}} + \underbrace{[\tau_a - \tau_s] \cdot N_s}_{\text{change in govt spending}}. \quad (9)$$

The first term is the welfare gain accrued to those who remain commuting by road vehicle (the "stayers"). The introduction of a subway benefits this group by increasing road speed and reducing travel time. The second term is the welfare gain for those who switch to the subway (the "switchers"); this is the sum of the difference in trip cost, the difference in amenities, and the difference in idiosyncratic preferences for the two modes. The third term is the change in government spending.

Figure 7 illustrates the model. Commuters are ranked according to their idiosyncratic preference along the horizontal axis. From left to right people have increasing preferences for road vehicle over subway. When there is no subway, everyone travels by road vehicle; their utility is represented by $u_a + \xi_{ia}$. When a subway becomes available, some are diverted and road speed increases, the utility from a commute by road vehicle shifts up by $\Delta u_a = c(N) - c(N_a)$, shown in the graph by the dashed curve. Utility from a subway trip is represented by the curve $u_s + \xi_{is}$.[18]

The empirical result found in this paper speaks directly to the first term of Equation 9.[19] The welfare gain for the stayers is represented by the shaded area A in Figure 7. It is the product of $c(N) - c(N_a)$—the value of time saved in each commute by road vehicle due to faster road speed (Δu_a)—and the number of remaining commutes by road vehicle.

[18]Several assumptions are made to keep the model tractable. Demand for trips is assumed to be inelastic. In fact, additional trips could be induced, generating additional welfare gains, or deadweight loss if the private gains from those marginal trips are smaller than government subsidies. Utility from completing a trip is assumed to be the same, while in reality different trips have different purposes and values. Relaxing these assumptions complicates the analysis and could change our conclusion substantively.

[19]Note that $c(N) - c(N_a)$ is related to the concept of "road technology"—how road speed changes as a function of traffic volume. Suppose the substitution pattern between subway and car trips follows the results in Table 6, Panel B: 1 subway trip replaces 0.315 car trips. Also assume the supply of bus services does not change. Beijing had 1.4 billion car commutes and 1.1 billion subway trips in 2016. Without the subway system, there would have been 1.7 billion car commutes. We show later that our empirical result suggests that Beijing's subway system increases city-wide average speed by 3%. This implies an elasticity of road speed with regard to the number of car trips of -0.14.

5.2 Quantifying the Welfare Gain from Beijing's Subway

We quantify the first term of Equation 9 using data from Beijing. We choose Beijing as the example not only because it provides the best data available, but also because it has one of the most extensive and complete subway networks among Chinese cities. Most additional data used for this exercise are from the *2016 Beijing Transportation Annual Report (BTAR)* published by the municipal government (BTI, 2016). Online Appendix C.1 lists detailed sources of data used in the back-of-the-envelope calculations. In line with our empirical setting, we focus on commuting trips and assume they all take place during rush hours.

The monetized value of time saved in each commute by road vehicle, $c(N) - c(N_a)$, can be calculated with the time cost of each commute, the increase in road speed (our main empirical result), and the monetized value of time. In 2016, the population weighted average distance to the subway is about 1 kilometer in Beijing. At this distance, we estimate that the road speed increases by 3%.[20,21] The average one-way commute costs 56 minutes (BTI, 2016). A 3% increase in speed implies a 1.68-minute decrease in commuting time. Average annual wage in Beijing was 92,456 *yuan*. Assuming 2,000 hours a year for a full-time job, this translates into a wage of 0.77 *yuan* per minute. The monetary value of commute time is typically assumed to be half of the wage rate (Parry & Small, 2009; Anderson, 2014), so the time savings for each commute from reduced congestion is worth 0.65 *yuan* (0.1 USD). In 2016, there were 1.4 billion car commutes and 1.1 billion bus commutes in Beijing (BTI, 2016). This translates into a welfare increase of 1.6 billion *yuan* per year (249 million USD) due to faster commutes.[22]

The other two terms in Equation 9 are not directly related to our empirical results. Quantifying them requires additional data, stronger assumptions, and correlational evidence. Additional cautions are called for when interpreting these calculations. We delegate details to Online Appendix C.2. We find that each of these terms is much larger than the value of time saved from faster road speed, but according to our calculations they largely offset each other. If accurate, this would suggest that, when only considering the cost and benefit of commuting trips, the benefit of Beijing's subway system seems to exceed its cost.

[20] The population-weighted average distance to the nearest subway line is calculated as the average distance of TAZ to the nearest subway line weighted by population. We run the following regression: $\ln \tilde{speed}_{lgw} = \beta_0 \cdot T_{lg} \cdot Post_{gw} + \beta_1 \cdot T_{lg} \cdot Post_{gw} \cdot \ln dist_l + \lambda_l + \lambda_{gw} + \iota \cdot \lambda_{gw} \cdot \ln dist_l + \gamma_t \cdot d_t \cdot \mathbf{X}_c + \varepsilon_{lgw}$. The differences in this specification from the baseline are the inclusion of terms (1) $\ln dist_l$ and (2) $Post_{gw} \cdot \ln dist_l$. $\ln dist_l$ is the log distance to the treated (for segments in treated cities) or the control (for segments in control cities) subway line. $Post_{gw} \cdot \ln dist_l$ accounts for potential heterogeneous trends for road segments with different distances to the subway line. β_0 indicates the effect at 1 kilometer away from the subway, which is estimated to be 0.03 with a standard error of 0.0068.

[21] We assume the estimated effect on average road segment applies to the average commute. In exploiting the heterogeneous effects, we show that the effect is larger in more congested roads. To the extent that more congested roads are also more traveled, our estimate of the speed increase on the average road segment is an underestimate of the gains to an average commute.

[22] 0.77/2 (unit time value)×1.68 (time saved each commute)×(1.4+1.1) billion (number of commutes).

6 Conclusions

This paper studies the effect of subways on road congestion in the setting of the rapid expansions of subway networks across Chinese cities. We use user-generated, real-time big data to measure road speed, which is new to the literature. We find that the launch of a new subway line immediately and significantly increases speed on nearby roads. On average the speed on road segments that are close substitutes to the new subway line increases by about 4% over the first year following the line opening. The effect is concentrated in road segments that were initially congested and declines quickly with distance to the new subway line. There is also suggestive evidence of spillover effects along the subway network.

We present a conceptual framework of transportation mode choices. Through the lens of the model, we show how our empirical result is related to the welfare calculation. Using data from Beijing, we estimate that the time savings for each commute by road vehicle from faster speed is worth 0.1 USD.

Many other potential benefits, such as reduction in air and noise pollution and traffic accidents, are not considered in this paper. In fact, existing studies suggest that health benefits from reduced air pollution account for a substantial share of the construction and operation costs of the subway (Chen & Whalley, 2012; Gendron-Carrier et al., 2018; Li et al., 2019). Our calculations are also based on short-run results. In the long run, the existence of an extensive subway system could change location choices of residents and firms, encourage migration, and facilitate economic agglomeration (e.g., Tsivanidis, 2018; Heblich et al., 2018; Gonzalez-Navarro & Turner, 2018).

References

Abraham, S., & Sun, L. (2018). *Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects* (Tech. Rep.).

Adler, M. W., & van Ommeren, J. N. (2016). Does public transit reduce car travel externalities? Quasi-natural experiments' evidence from transit strikes. *Journal of Urban Economics*, *92*, 106-119.

Akbar, P. A., Couture, V., Duranton, G., & Storeygard, A. (2018). *Mobility and congestion in urban India.*

Akbar, P. A., & Duranton, G. (2017). *Measuring the Cost of Congestion in a Highly Congested City: Bogotá* (Tech. Rep.).

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

Baum-Snow, N., & Kahn, M. E. (2005). Effects of urban rail transit expansions: Evidence from sixteen cities, 1970-2000. *Brookings-Wharton Papers on Urban Affairs*, 147-206.

BTI. (2010,2015). *Household travel surveys*. Beijing Transportation Institute.

BTI. (2016). *Beijing Transportation Annual Report*. Beijing Transport Institute.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, *90*(3), 414-427.

Chen, Y., & Whalley, A. (2012). Green infrastructure: The effects of urban rail transit on air quality. *American Economic Journal: Economic Policy*, *4*(1), 58–97.

Couture, V., Duranton, G., & Turner, M. A. (2018). Speed. *Review of Economics and Statistics*, *100*(4), 725-739.

Cui, X., Jia, W., Xiao, S., & Mao, B. (2018). *Report on the Development of Urban Rail Transit Operation in China* (Tech. Rep.).

Davis, L. W. (2008). The effect of driving restrictions on air quality in Mexico City. *Journal of Political Economy*, *116*(1), 38–81.

Downs, A. (1962). The law of peak-hour expressway congestion. *Traffic Quarterly*, *16*(3), 1962-1967.

Downs, A. (2000). *Stuck in Traffic: Coping with Peak-hour Traffic Congestion*. Brookings Institution Press.

Duranton, G., & Turner, M. A. (2011). The fundamental law of road congestion: Evidence from US cities. *American Economic Review*, *101*(6), 2616–52.

Er, L. (2014). *Beijing Subway Starts New Fare Schedule Today, Saying Goodbye to the Era of 2-yuan Flat Rate.* Retrieved from https://www.bbc.com/zhongwen/simp/china/2014/12/141228_beijing_subway_new_price

Gendron-Carrier, N., Gonzalez-Navarro, M., Polloni, S., & Turner, M. A. (2018). *Subways and urban air pollution* (Tech. Rep.). National Bureau of Economic Research.

Gonzalez-Navarro, M., & Turner, M. A. (2018). Subways and Urban Growth: Evidence from Earth. *Journal of Urban Economics*, *108*, 85-106.

Goodman-Bacon, A. (2019). *Difference-in-Differences with Variation in Treatment Timing* (Tech. Rep.). Vanderbilt University.

Gu, Y., Deakin, E., & Long, Y. (2017). The effects of driving restrictions on travel behavior evidence from Beijing. *Journal of Urban Economics*, *102*, 106–122.

Heblich, S., Redding, S. J., & Sturm, D. M. (2018). *The Making of the Modern Metropolis: Evidence from London* (Tech. Rep.). Working paper, Princeton University.

Hsu, W.-T., & Zhang, H. (2014). The fundamental law of highway congestion revisited: Evidence from national expressways in Japan. *Journal of Urban Economics*, *81*, 65–76.

Imbens, G., & Lemieux, T. (2008). Regression discontinuity design: A guide to practice. *Journal of Econometrics*, *142*(2), 615-635.

Kreindler, G. (2018). *The Welfare Effect of Road Congestion Pricing: Experimental Evidence and Equilibrium Implications* (Tech. Rep.). Working Paper, MIT.

Lee, D., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, *48*, 281-355.

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

Li, S., Liu, Y., Purevjav, A.-O., & Yang, L. (2019). Does subway expansion improve air quality? *Journal of Environmental Economics and Management*.

NBS. (2010-2017). *Statistical Yearbook of Chinese Cities*. National Bureau of Statistics.

Pang, J., & Shen, S. (2019). *Do Subways Reduce Congestion? Evidence from US Cities.* (Tech. Rep.). Working paper, Syracuse University.

Parry, I. W. H., & Small, K. A. (2009). Should urban transit subsidies be reduced? *American Economic Review*, *99*(3), 700-724.

Reed, T., & Kidd, J. (2009). *Global Traffic Scorecard* (Tech. Rep.). INRIX Research.

Small, K. A., Verhoef, E. T., & Lindsey, R. (2007). *The Economics of Urban Transportation*. Routledge.

TomTom. (2019). *TomTom Traffic Index 2019* (Tech. Rep.). TomTom Inc.

Tsivanidis, N. (2018). *Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogota's TransMilenio* (Tech. Rep.). Working paper, University of Chicago.

Voith, R. (1991). The long-run elasticity of demand for commuter rail transportation. *Journal of Urban Economics*, *30*(3), 360–372.

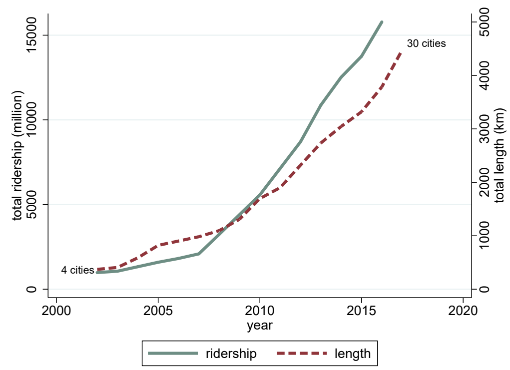
Winston, C., & Langer, A. (2006). The effect of government highway spending on road users' congestion costs. *Journal of Urban Economics*, *60*(3), 463–483.

Winston, C., & Maheshri, V. (2007). On the social desirability of urban rail transit systems. *Journal of Urban Economics*, *62*(2), 362–382.

Yang, J., Chen, S., Qin, P., Lu, F., & Liu, A. A. (2018). The effect of subway expansions on vehicle congestion: Evidence from Beijing. *Journal of Environmental Economics and Management*, *88*, 114–133.

Figures and Tables

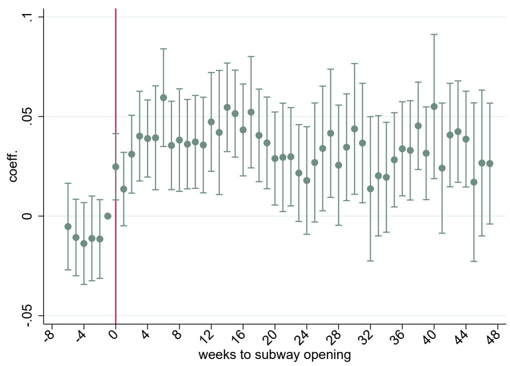
Figure 1: Subway Construction in China



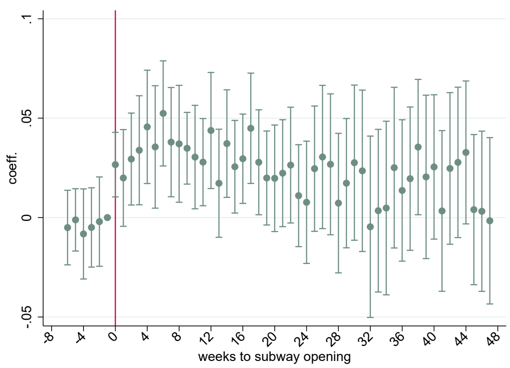
Note: Data from *Statistical Yearbooks of Chinese Cities*, published annually by China's National Bureau of Statistics (NBS, 2010-2017).

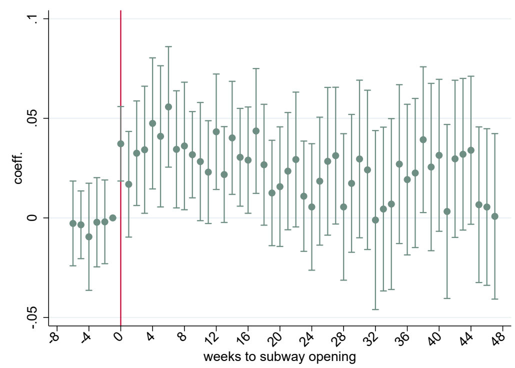
Figure 2: Dynamic Effects of Subway Launches

Panel A: stacked DID



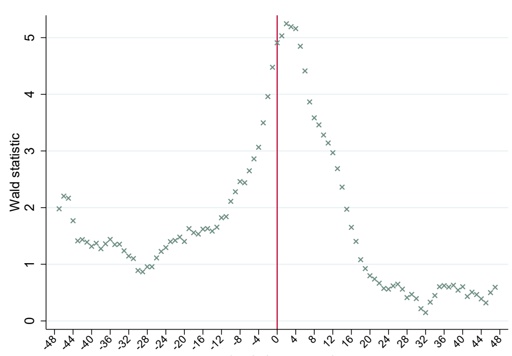
Panel B: Two-way FE, not accounting for seasonality	**Panel C: Two-way FE, accounting for seasonality**





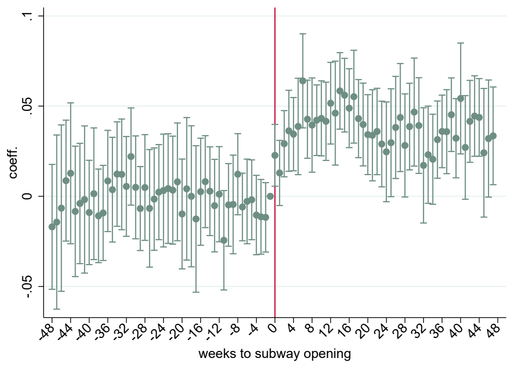
Note: Panels A is estimated using Equation 2. Panels B and C are estimated using Equation 3, using only observations from treated cities. Panels A and C control for differential seasonality by allowing week-specific effects to differ by city characteristics. See Section 3.1 for details.

Figure 3: Wald Statistics from Estimates with Placebo Opening Dates



Note: Each dot is a Wald statistic from the t-test of the key coefficient in the stacked DID specification where a placebo week of subway opening is used. The x-axis indicates the placebo week of opening relative to the actual week of opening.

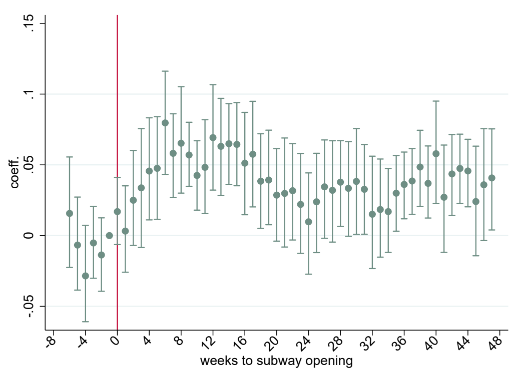
Figure 4: Longer Pre-treatment Periods



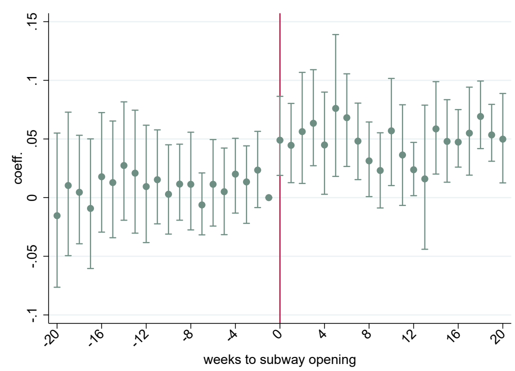
Note: The model is estimated using the stacked DID specification. Up to 48 weeks in the pre-treatment period and 48 weeks in the post-treatment periods are included.

Figure 5: Lines Launched in Different Periods

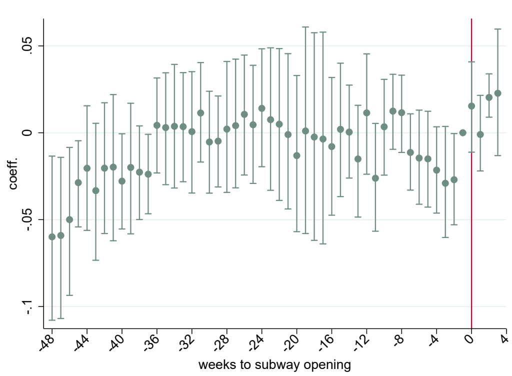
Panel A: Lines launched before Jan 31, 2017



Panel B: Lines launched between Feb 1 and Nov 30, 2017

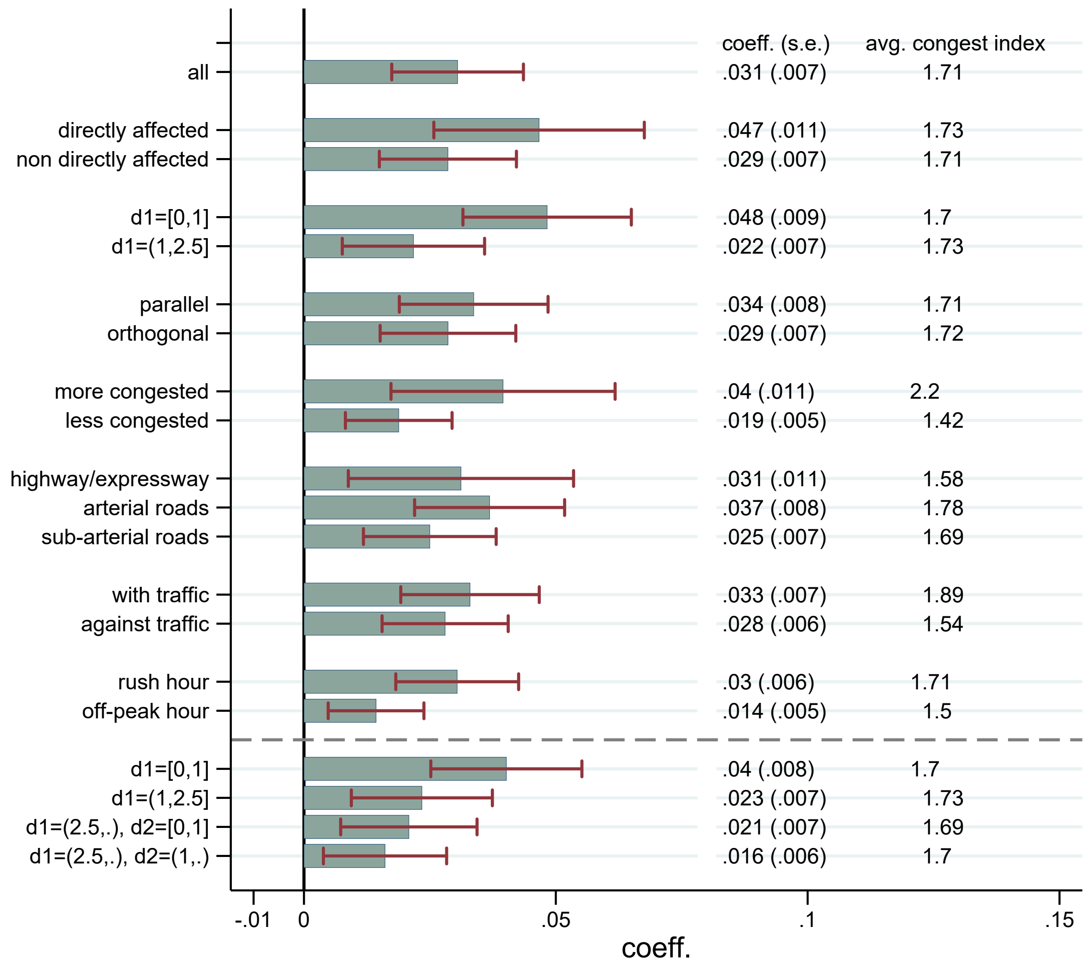


Panel C: Lines launched after Dec 1, 2017



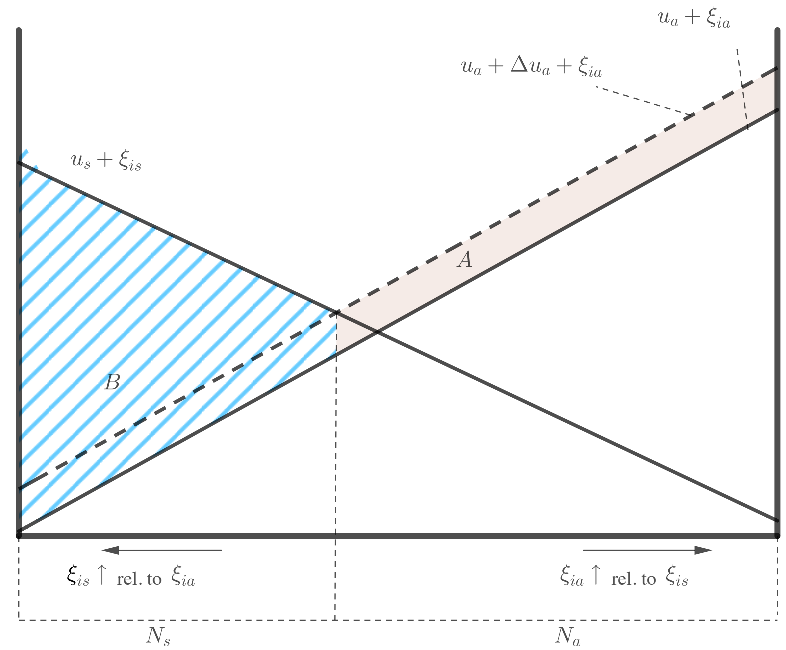
Note: 95% confidence intervals are from 499 repetitions of wild block bootstrapping (Cameron et al., 2008).

Figure 6: Heterogeneous Effects by Road Segment Characteristics



Note: Each group of bars report $\hat{\beta}_k$'s from estimating Equation 4. Except for the last block, the treated road segments include those within 2.5 kilometers from the treated subway line. The length of the bar indicates the magnitude of the coefficient. The range bar indicates the 95% confidence interval. $\hat{\beta}_k$ and its corresponding standard error, and average congestion index for the group of road segments are also reported next to each bar. In the third-to-last block, each observation is split into weekly average in morning rush hours and weekly average in evening rush hours. Road segment-by-morning/evening indicator fixed effects are also controlled for. In the second-to-last block, weekly average in non-rush-hour (9AM-5PM) speed is also included, road segment-by-rush hour indicator fixed effects are also controlled for. In the last group, treated road segments also include those within 2.5 kilometers from an existing or planned subway line. See Table 1 Panel B for these lines. The last group is estimated from a specification that replaces λ_{kgw} with λ_{gw}. $d1$ represents the road segment's distance to the treated subway line and $d2$ represents the road segment's distance to the nearest existing subway line.

Figure 7: Illustration of the Welfare Impact



Note: $\Delta u_a = c(N) - c(N_a)$ is the gain for a commute by road vehicle from reduced congestion. The shaded area A indicates the welfare gain to those who continue to commute by road vehicle. Hashed area B indicates the welfare gain to those who switch to subway. See Section 5.1 for details.

Table 1: Subway Lines in the Sample

City	Line(s)	Open Date	City	Line(s)	Open Date
Panel A: New Subway Lines Opened between August 1, 2016 and December 31, 2017					
Beijing	16	12/31/16	Nanchang	2	8/18/17
Beijing	Xijiao	12/30/17	Nanjing	4	1/8/17
Changchun	1	6/30/17	Nanjing	S3	12/6/17
Chengdu	4 (2^{nd} phase east), 4 (2^{nd} phase west)	6/2/17	Nanning	1	12/28/16
Chengdu	10	9/6/17	Nanning	2	12/28/17
Chengdu	7	12/6/17	Qingdao	3 (2^{nd} phase)	12/18/16
Chongqing	Airport	12/28/16	Qingdao	2	12/10/17
Chongqing	5,10	12/28/17	Shanghai	9	12/30/17
Dalian	1	6/8/17	Shenzhen	7,9	10/28/16
Foshan	Guang-Fo	12/28/16	Suzhou	2 (2^{nd} phase)	9/24/16
Fuzhou	1 (2^{nd} phase north)	1/6/17	Suzhou	4	4/15/17
Guangzhou	6 (2^{nd} phase), 7 (1^{st} phase)	12/28/16	Tianjin	6	8/6/16
Guangzhou	9,13	12/28/17	Wuhan	6, Airport	12/28/16
Guiyang	1	12/28/17	Wuhan	8,Yangluo	12/26/17
Hangzhou	2 (1^{st} phase northwest)	7/3/17	Xi'an	3	11/8/16
Harbin	3	1/26/17	Xiamen	1 (1^{st} phase)	12/31/17
Hefei	1	12/26/16	Zhengzhou	2	8/19/16
Hefei	2	12/26/17	Zhengzhou	1 (2^{nd} phase), Suburban	1/12/17
Kunming	3	8/29/17			
Panel B: Exisiting or Planned Subway Lines in Treated Cities					
Beijing	15	12/30/10	Nanjing	1	9/3/05
Changchun	3	6/30/11	Qingdao	3	12/16/15
Changsha	2	4/29/14	Shanghai	10	2010
Chengdu	1	9/27/10	Shanghai	11 extensions	4/26/16
Chongqing	1	3/18/11	Shenzhen	5	6/22/11
Dalian	2	5/22/15	Shenzhen	11	6/28/16
Fuzhou	5 (1^{st} phase)	2021	Suzhou	1	4/28/12
Guangzhou	6	12/28/13	Tianjin	1	12/28/84
Guiyang	1 (old town)	2018	Wuhan	3	12/28/15
Hangzhou	4 (1^{st} phase)	2/2/15	Xi'an	1	9/15/13
Harbin	1	9/26/13	Xiamen	2	2019
Kunming	6	6/28/12	Zhengzhou	1 (1^{st} phase)	12/28/13
Nanchang	1	12/26/15			
Panel C: Exisiting or Planned Subway Lines in Control Cities					
Changsha	2	4/29/14	Shaoxing	1	2022
Changzhou	1	2019	Shenyang	1	9/27/10
Dongguan	1	2022	Shijiazhuang	3	2022
Hohhot	1 (1^{st} phase)	2020	Taiyuan	2	2020
Jinan	R2	9/30/21	Urumqi	3	2021
Lanzhou	1	2018	Wuhu	2	2020
Luoyang	2	2022	Wuxi	1	7/1/14
Nantong	1	2021	Xuzhou	1	2019
Ningbo	2	9/26/15			

Note: Opening dates of subway lines are from pages on Baidu Baike, Wikipedia, and various news sources.

Table 2: Summary Statistics of the Baseline Sample

Panel A: City level characteristics

	treated cities ($N = 25$)				control cities ($N = 17$)			
	mean	p25	median	p75	mean	p25	median	p75
population (million)	8.68	3.98	5.54	11.04	3.81	2.66	3.54	4.11
GDP per capita (*yuan*)	105,597	82,082	105,417	126,364	95,674	71,120	94,402	109,106

Panel B: Road speed and congestion

	road segments in treated cities				road segments in control cities			
	# of obs.	# of unique segments	avg. speed (km/h)	avg. congest. index	# of obs.	# of unique segments	avg. speed (km/h)	avg. congest. index
all road segments	284,301	8,342	31.76	1.69	358,152	12,088	31.48	1.7
highways	1,353	36	82.9	1.18	1,685	56	60.83	1.61
urban expressways	5,773	219	50.09	1.77	7,921	269	53.18	1.59
arterial streets	117,321	3,287	33.14	1.76	135,256	4,715	34.21	1.71
sub-arterial streets	159,854	4,800	29.66	1.65	213,290	7048	28.71	1.69

Note: Data in Panel A are from the *2017 Statistical Yearbook of Chinese Cities*. In Panel B, each observation is a road segment-by-week-to-opening. Segments in the baseline regression sample are included. Treated road segments are those directly affected by the new subway lines. Week-to-opening is between 6 weeks before and 48 weeks after line opening.

Table 3: Baseline Estimates

	(1)	(2)	(3)
treated×post	0.044	0.036	0.036
	(0.008)	(0.009)	(0.010)
model	stack DID	two-way FE	two-way FE
group-by-week-to-open FE	✓		
road segment FE	✓	✓	✓
calendar week FE		✓	✓
adj. for differential seasonality	✓		✓
incl. segments from control cities	✓		
N	642,453	284,301	284,301

Note: The dependent variable is weekly average residual speed. Standard errors are in parentheses, clustered at the group level in Column 1, clustered at the subway line level in Columns 2 and 3.

Table 4: Varying the Length of the Pre-Period

	Length of pre-periods			**Discontinuity around the launch**		
	# of weeks before launch			treat-specific time trend poly.		
	12	24	48	linear	up to 3rd	up to 5th
	(1)	(2)	(3)	(4)	(5)	(6)
treated×post	0.043	0.039	0.039	0.042	0.028	0.023
	(0.007)	(0.008)	(0.008)	(0.009)	(0.012)	(0.013)
range of weeks rel. to opening	[-12,47]	[-24,47]	[-48,47]	[-48,47]	[-48,47]	[-48,47]
poly. of time trends	none	none	none	linear	up to 3rd	up to 5th
N	726,727	915,041	1,174,944	1,174,944	1,174,944	1,174,944

Note: The dependent variable is a segment's weekly average log residual speed. All models are estimated using the stacked DID model and include all 45 treated lines. Polynomial of flexible time trends indicate the degree of the polynomials for treatment-specific pre- and post- time trends. Standard errors are in parentheses, clustered at the group level.

Table 5: Subsamples by Time of Launch

	(1)	(2)	(3)	(4)
	all lines	before 1/31/17	2/1-11/30/17	after 12/1/17
treated×post	0.036	0.050	0.039	0.022
	(0.009)	[0.030,0.069]	[0.021,0.055]	[-0.008,0.051]
range of weeks rel. to opening	[-6,3]	[-6,47]	[-20,20]	[-48,3]
# of treated lines	45	21	9	15
N	179,741	408,546	143,567	445,626

Note: The dependent variable is a segment's weekly average log residual speed. All models are estimated using the stacked DID model. In Column 1, standard errors are in parentheses, clustered at the group level. In Columns 2 to 4, numbers in brackets are 95% confidence intervals from 499 repetitions of wild block bootstrapping.

Table 6: Individual Transportation Mode Choices and Household VKT

Panel A: All households					
	(1)	(2)	(3)	(4)	(5)
	subway	bus	car	log VKT	VKT
log dist. to subway station	-0.209	0.085	0.062	0.686	620
	(0.026)	(0.017)	(0.014)	(0.191)	(326)
mean dep. var	0.472	0.499	0.558	2.964	4302
N	63,710	63,710	63,710	49,500	49,500
Panel B: Households not moved between 2009 and 2014					
	(1)	(2)	(3)	(4)	(5)
	subway	bus	car	log VKT	VKT
log dist. to subway station	-0.203	0.072	0.064	0.696	632
	(0.025)	(0.016)	(0.015)	(0.194)	(333)
mean dep. var	0.463	0.504	0.561	2.948	4289
N	60,591	60,591	60,591	47,305	47,305
Panel C: Households not moved and had a car since 2009					
	(1)	(2)	(3)	(4)	(5)
	subway	bus	car	log VKT	VKT
log dist. to subway station	-0.467	0.156	0.043	-0.050	4.412
	(0.121)	(0.067)	(0.064)	(0.088)	(940)
mean dep. var	0.577	0.519	0.458	9.337	14,075
N	12,482	12,482	12,482	10,195	10,195
hhd. chars.	✓	✓	✓	✓	✓
ind. chars.	✓	✓	✓		
TAZ FE	✓	✓	✓	✓	✓
year FE	✓	✓	✓	✓	✓

Note: The key explanatory variable is TAZ's log average distance to the nearest subway station. See text for details about how this average distance is calculated. Data are from the Beijing's *Household Travel Surveys* in 2010 and 2015 (BTI, 2010,2015). Each observation is an individual in Columns 1-3. The outcome variable is the number of trips in the corresponding mode of transportation on the day of the survey, taken from the *One-day Travel Diaries* in the *Survey*. Individuals in the sample include those who have trip records on the day of survey. Trips that use only walking are excluded. An observation is a household in Columns 4 and 5. The outcome variable in Column 4 is the log total kilometers driven in the past year from all the cars owned by the household. Total vehicle-kilometer travelled (VKT) is added by one before taking logs, so households that do not own a car are also included in the sample. The outcome variable in Column 5 is the level of VKT. Household characteristics (hhd. chars.) include indicators for household income brackets, home ownership, house type (commercial apartment, work unit dormitory, low-income housing, etc.), whether having children under age 5, and household size. Individual characteristics (ind. chars.) include gender, age, indicators for educational levels, industry and occupation. All models include TAZ and year fixed effects. Standard errors are clustered at the TAZ level.

Online Appendix for "Subways and Road Congestion" (Not for Publication)

Authors: Yizhen Gu, Chang Jiang, Junfu Zhang and Ben Zou

A More Details on Data and Sample

A.1 Selection of Subway Lines and Road Segments

Figure A.1 shows a snapshot of the digital map on the Baidu Maps website (http://maps.baidu.com/). Color-coded roads indicate current speed. We digitize the subway network in all 42 cities in the sample. Figure A.2 shows the selection of road segments using Guangzhou as an example. The grey lines are surface roads. Red lines are existing subway lines by August 1, 2016. Green lines are new subway lines launched between August 1, 2016 and December 31, 2017. Dots indicate subway stations. Yellow areas are the buffer zones around the selected subway line segments. All road segments within the buffer zones are in our dataset.

Figure A.3 illustrates the procedure of picking the "directly affected" road segments. These are road segments that a person needs to travel through if one chooses to drive instead of taking the public transit, while the public transit trip would involve using the new subway line. First, the buffer zone around each new subway line is divided into 1km-by-1km grids. Between any pair of grids, we find the best public transit route (or routes, as the digital map service usually recommends several alternative routes). We save the collection of pairs between which the best public transit routes involve the newly launched subway line. Then for each pair of grids in this collection, we find the best route(s) for driving. In the case of Figure A.3, the digital map suggests two alternative routes. We take road segments in these routes as being directly affected by the new subway line. The same process is repeated under the typical traffic conditions for weekday morning and afternoon rush hours.

A.2 Traffic Congestion by Hour

We define rush hours as 7-9 AM in the morning and 5-7 PM in the evening. These are the same definitions adopted by Baidu Maps, which determines the rush hours by examining the usual traffic pattern across Chinese cities. Using our sample data, we plot the average hourly congestion index for weekdays on the third week in November, 2017 in Figure A.4.[1] Morning traffic peaks at 7-10 AM and evening traffic peaks at 5-7 PM. Evening rush hours are more congested than morning rush hours.

A.3 Seasonality

We inspect seasonality in traffic speed in control cities. To obtain adjusted log speed in each day, hourly log speed in each road segment is first regressed on a full set of indicators with segment, hour of the day, and day of the week indicators fully interacted. Rush hours are included. Weekends and national holidays are excluded. Residual log speed is averaged within each day. The average log speed over the sample period is then added back.

Figure A.5 presents the adjusted log rush-hour speed. There is no clear overall trend over the sample period. Speed increased after the 2017 New Year and reached the apex around the time of the Chinese New Year (CNY, on January 28th). The daily average rush-hour speed in the days around the CNY was about 20% higher than the average in the sample period. The spike in speed near the 2017 CNY lasted for a little more than one month, from early January to mid-February. There is also a decline in road speed between late August and late September, by about 8% at its peak, probably due to the fact that new academic year usually starts in early September and parents drive children to school.

A.4 Example Timelines of Subway Construction

Beijing Line 16

- 5/4/2011 — Construction plan was approved by the Beijing municipal government.
- 10/26/2012 — Study on environmental impacts was completed. One station was removed from the plan.
- 4/25/2013 — Designs for the subway stations were on display to the public.
- 12/10/2013 — Second round of evaluation on the environmental impacts was published. The evaluation was released to the public for comments.
- 12/12/2013 — Construction officially started.
- 8/16/2016 — Scale-test cars were run on completed tracks.
- 9/5/2016 — Tracks were electrified.
- 9/20/2016 – Passenger trains were tested without passengers on board.

[1] There is no major holidays or events in or around this week. We pick it to represent the "typical" traffic pattern. Also note that our dataset has hourly traffic speed in 12 hours every day between 7AM and 7PM.

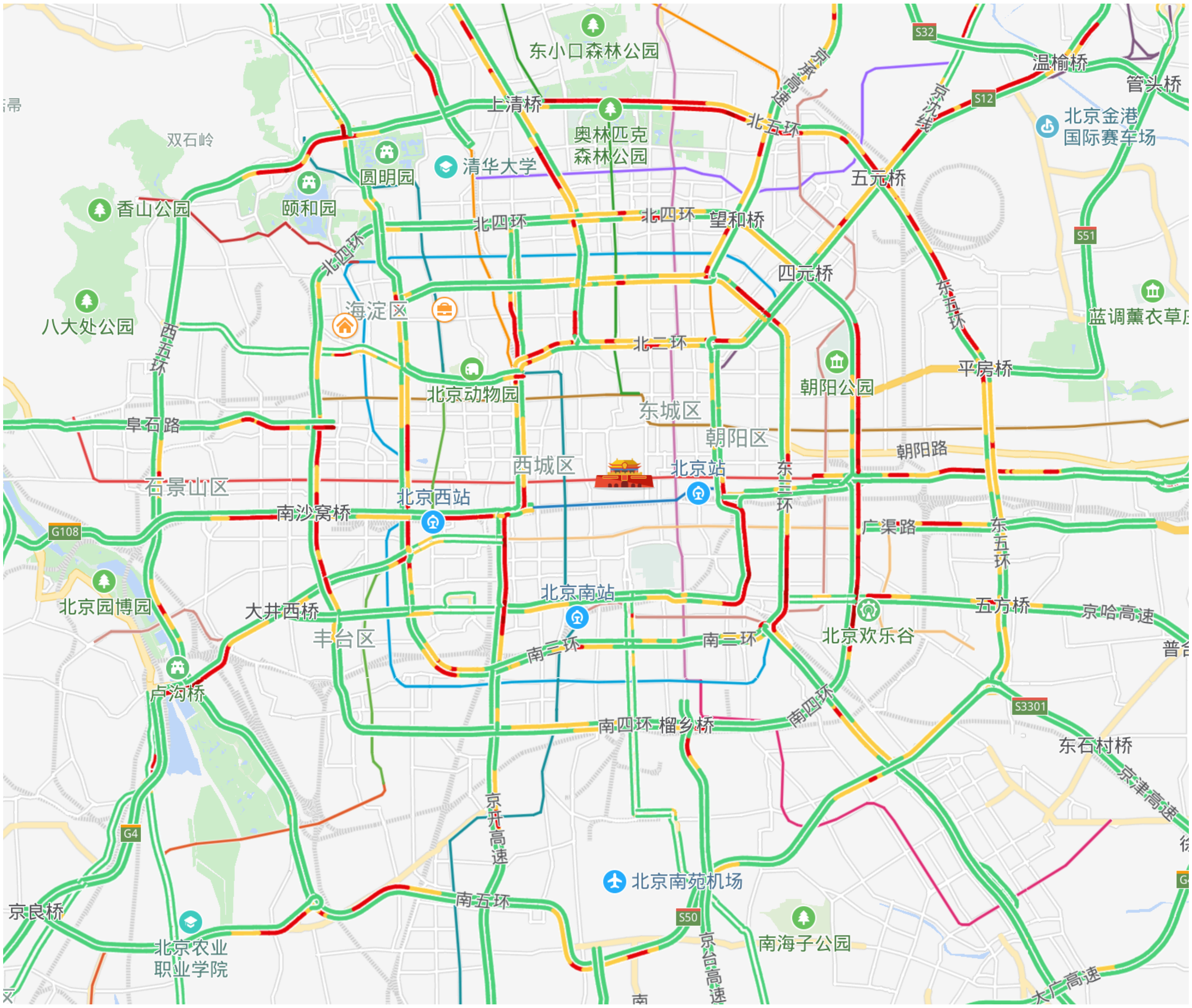
- 12/31/2016 — Line opened officially.

Suzhou Line 4

- 6/18/2012 — Evaluation of environmental impacts was approved by the Ministry of Environmental Protection.
- 6/19/2012 — Plans of land use along the subway line were approved by the Ministry of Land and Resources.
- 9/27/2012 — Construction officially started.
- 9/12/2013 — First tunnel boring shield machine was deployed.
- 12/18/2014 — Tracks were laid.
- 8/16/2015 — Tunnels were completed.
- 12/18/2015 — Tracks were completed.
- 4/9/2016 — Tracks were electrified.
- 12/28/2016 — Test rides were run with trains carrying no passengers.
- 3/13/2017 — The subway line passed the final examination by a panel of experts and was accepted as suitable for operation.
- 3/20/2017-3/24/2017 — Free tickets were handed out to residents for test rides.
- 3/25/2017-3/31/2017 — Test rides with passengers.
- 4/15/2017 — Line opened officially.

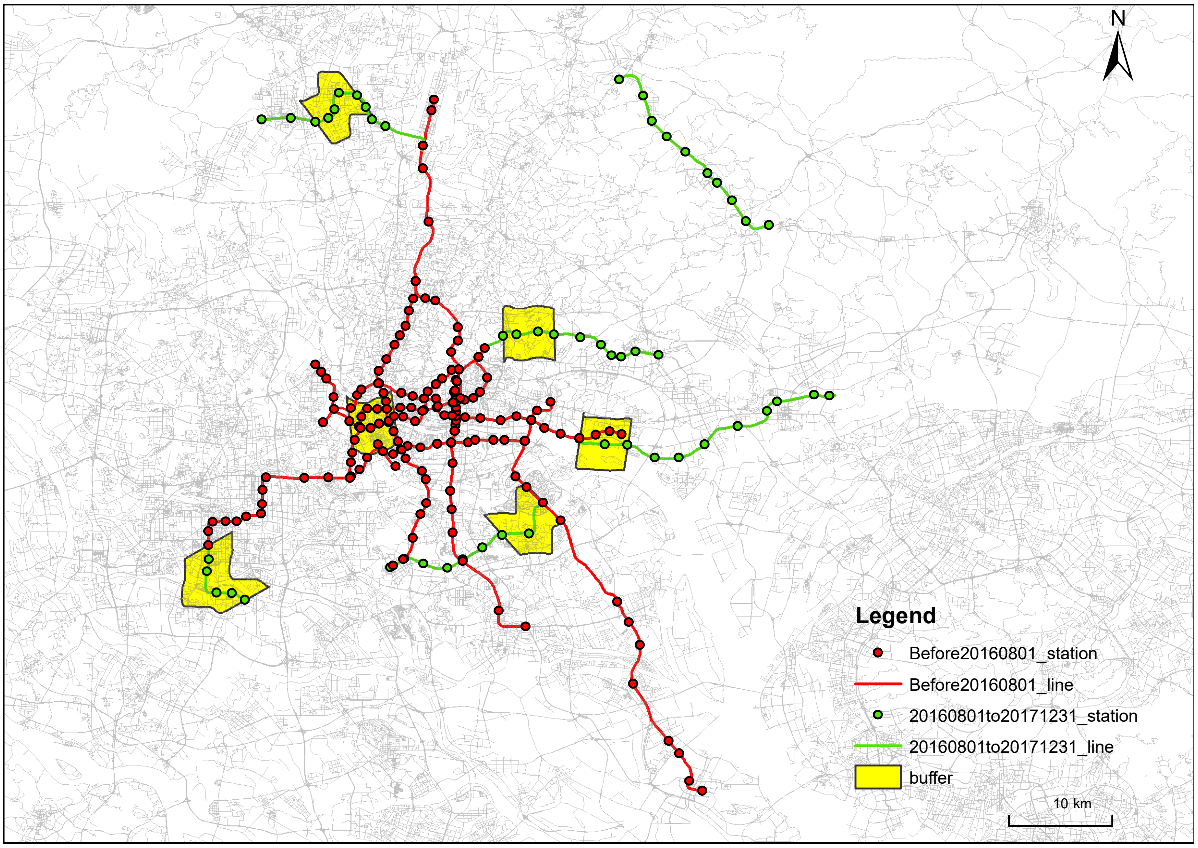
Table A.1 lists days of test ride for each new subway line.

Figure A.1: A Snapshot of the Digital Map with Real-time Traffic Congestion



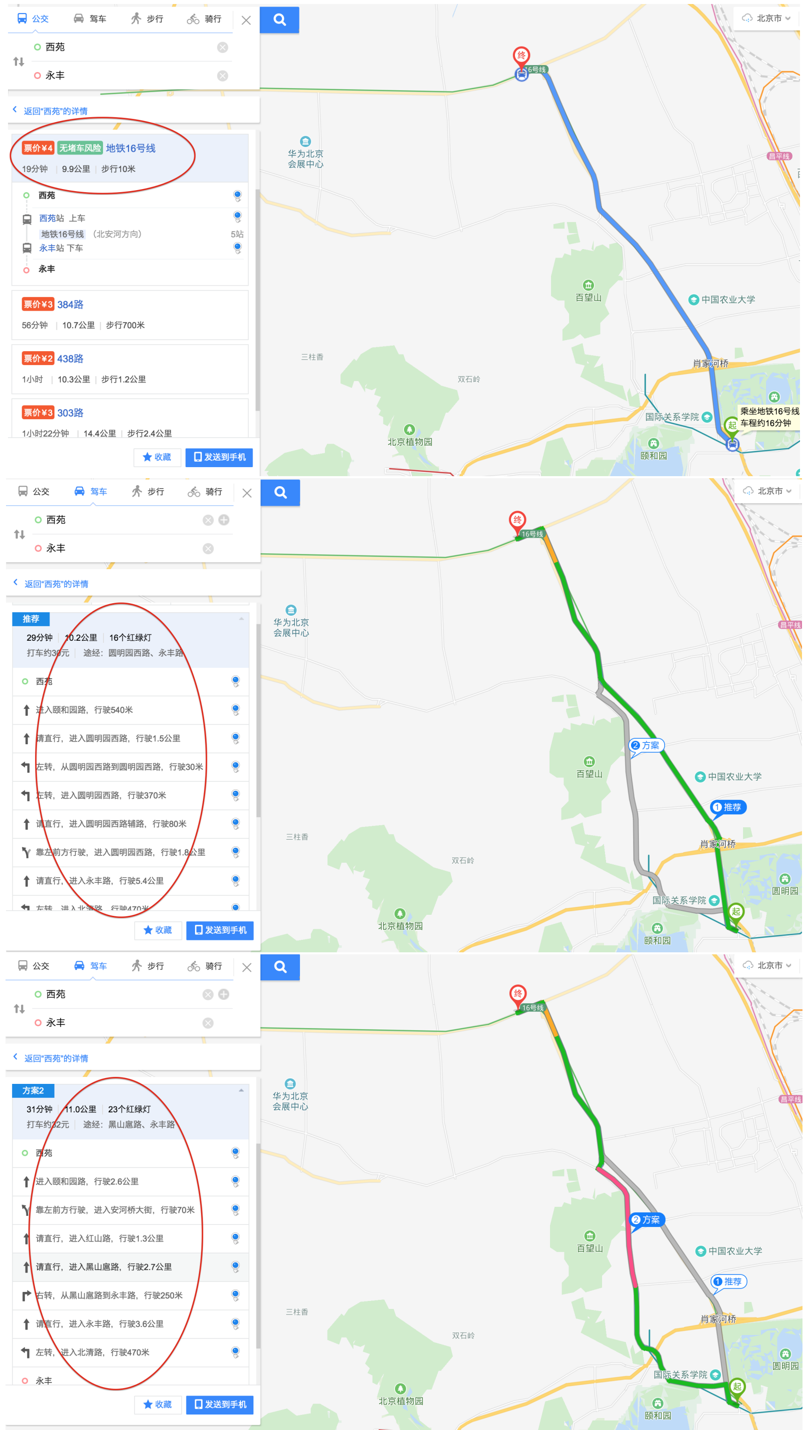
Note: Screenshot from the web version of our data provider's digital map. It shows the color-coded roads in Beijing, at 8:30 AM on Monday, September 3, 2018. Red indicates the road is congested; yellow indicates heavy traffic; green indicates light or smooth traffic.
Source: Baidu Maps website (maps.baidu.com).

Figure A.2: Illustration of Road Segment Selection



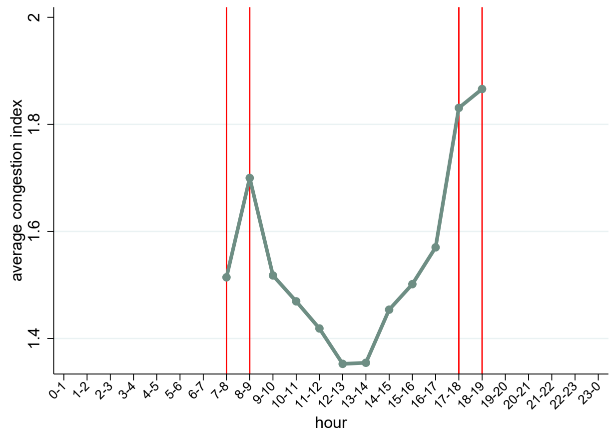
Note: The map illustrates the city of Guangzhou. Roads are shown as grey lines, which are from OpenStreetMap. Colored lines and dots indicate subway lines and stations, which are digitized manually from Baidu Maps. Yellow areas are buffer zones around new and existing subway lines.
Sources: Road networks are from OpenStreetMap (https://www.openstreetmap.org/). Subway lines and stations, as well as buffer zones are from authors' calculations.

Figure A.3: Road Segments Directly Affected by the Subway



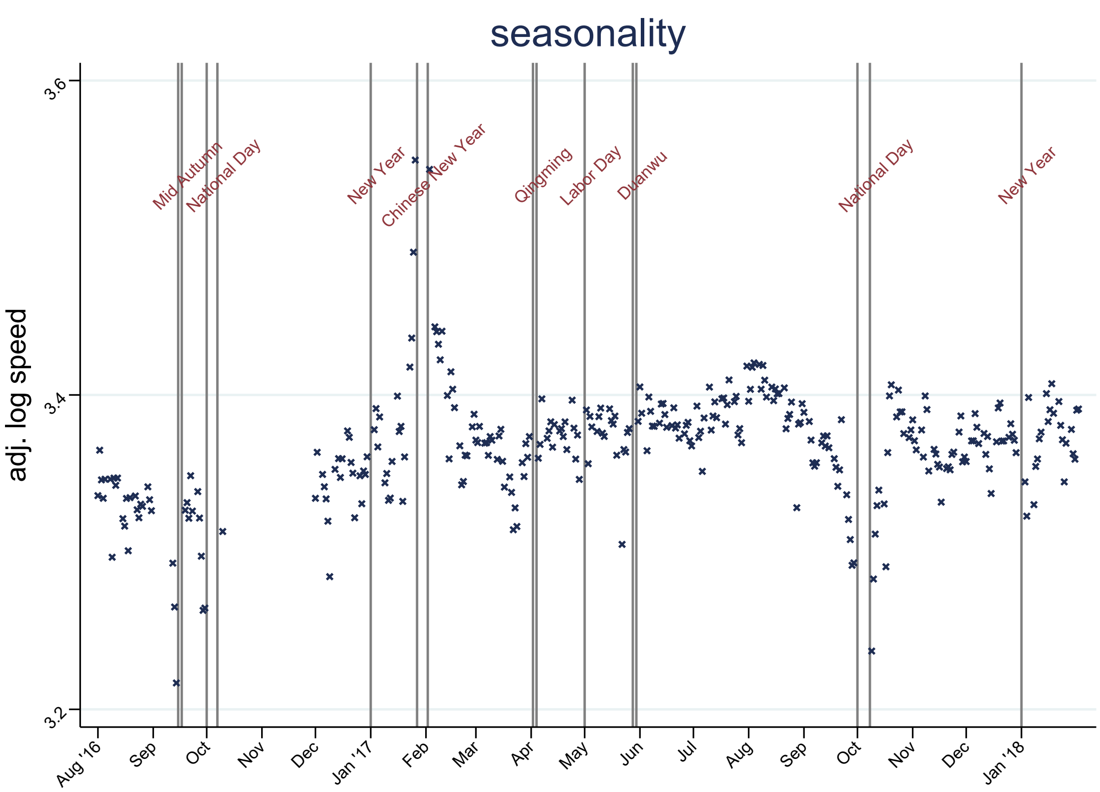
Note: Screenshots from data provider’s website. The upper panel shows a trip for which taking the subway is the best public transit option. The middle panel shows the best driving route as a substitute for the subway trip. The lower panel indicates the alternative driving routes. The left part of the webpage indicates the names of the roads involved in the route, which we use to extract directly-affected segments in our data sample.
Source: Baidu Maps website (maps.baidu.com).

Figure A.4: Congestion Index by Hour



Note: Congestion index is the ratio between travel time under current speed and that under traffic-free speed, which is the average speed in the same road segment between 12AM and 5AM. We have hourly data for the sample road segments between 7AM and 7PM. Rush hours are 7-9AM and 5-7PM. Plotted in the graph are hourly congestion indices on the weekdays in the third week of November, 2017.

Figure A.5: Seasonality in Traffic in Control Cities



Note: Average daily log rush hour speed in sample road segments from 17 control cities are plotted. Speed is adjusted for by a full set of road segment-by-day of the week-by-hour of the day fixed effects. Weekends and national holidays are excluded.

Table A.1: Official Opening Dates and Test Ride Periods

City	Line	Open Date	Test Ride Dates	Sources
Shanghai	9	12/30/17		
Beijing	16	12/31/16		
Beijing	Xijiao	12/30/17	2/28/18	1*
Nanjing	S3	12/6/17		
Nanjing	4	1/8/17		
Nanning	1	12/28/16		
Nanning	2	12/28/17		
Xiamen	1 (1st phase)	12/31/17	10/6/17-10/11/17	1,2
Harbin	3	1/26/17		
Dalian	1	6/8/17	6/7/17	1
Tianjin	6	8/6/16	7/4/16-7/19/16	1
Guangzhou	Guang-Fo, 7 (1st phase), 6 (2nd phase)	12/28/16		
Guangzhou	9, 13	12/28/17		
Chengdu	4 (2nd phase E), 4 (2nd phase W)	6/2/17		
Chengdu	7	12/6/17		
Chengdu	10	9/6/17		
Hefei	1	12/26/16		
Hefei	2	12/26/17	12/6/17-12/10/17	1
Nanchang	2	8/18/17		
Kunming	3	8/29/17		
Wuhan	6, Airport	12/28/16		
Wuhan	8, Yangluo	12/26/17		
Shenzhen	7, 9	10/28/16		
Fuzhou	1 (2nd phase N)	1/6/17	12/25/16-1/3/17	1,2
Suzhou	4	4/15/17	3/25/17-3/31/17	1
Suzhou	2 (2nd phase)	9/24/16		
Xi'an	3	11/8/16		
Guiyan	1	12/28/17		
Zhengzhou	2	8/19/16	8/10/16-8/19/16	1,2
Zhengzhou	1 (2nd phase), Suburban	1/12/17		
Chongqing	5	12/28/17	9/30/17-11/10/17	1
Chongqing	10	12/28/17		
Chongqing	Airport	12/28/16		
Changchun	1	6/30/17	6/25/17-6/29/17	1,2
Qingdao	3 (2nd phase)	12/18/16	12/7/16-12/9/16	1
Qingdao	2	12/10/17	12/3/17-12/5/17	1
Hangzhou	2 (1st phase NW)	7/3/17		

* Beijing's Xijiao Line opened on 12/30/17. There was an accident on 1/1/18, which rendered the line to be closed until 2/28/18.

Note: Test ride dates of new subway lines are from pages on Baidu Baike, Wikipedia, and various news sources. Hyperlinks to these sources are provided in the column titled "Sources."

B Additional Empirical Results

B.1 Additional Discussions on Differential Seasonality

We present evidence of differential seasonal effect in traffic speed between treated and control cities. The divergence in seasonality mostly appear around the 2017 CNY. We show that the way we treat differential seasonality in the baseline specification, by allowing calendar time effects to differ linearly along the dimensions of log population and log GDP per capita, can remove the impact of differential seasonality.

Figure B.1 shows how accounting for differential seasonality matters for the estimates. Panel A estimates Equation 2 without accounting for differential seasonality ($\gamma_t \cdot d_t \cdot \mathbf{X}_c$ not controlled for). It shows the effect is hump-shaped in the first few months following the launch of a new subway line. Panels B to E add different accounts of differential seasonality, with log population (Panel B), log GDP per capita (Panel C), both (Panel D, which is the same as Figure 2 Panel A), and both as well as their quadratic terms (Panel E) interacted with a full set of calendar week indicators. The hump shape becomes less salient, but the overall effect remains largely unchanged.

Figure B.2 summarizes the estimated coefficients from these five alternative specifications. Confidence intervals are omitted for a clean presentation. The estimated effects from different models largely track well with one another. The only discrepancy takes place at the hump-shaped part in the first 20 weeks after the line opening. Accounting for differential seasonality makes the hump shape less salient.

The different result from a model that takes differential seasonality into consideration comes almost solely from larger seasonal fluctuations in the weeks around the 2017 CNY. Figure B.3 plots the coefficients (γ_t's) associated with the interactive terms between city characteristics ($\mathbf{X}_c$) and calendar week dummies (d_t's) in Equation 2. In Panel A, $\mathbf{X}_c$ is log population; in Panel B, $\mathbf{X}_c$ is log GDP per capita. In both cases, most of the $\hat{\gamma}_t$'s hover around 0 except for a few episodes. The first is when we have a small number of observations in the sample between August and October in 2016. This is because when we restrict the sample to be up to 6 weeks to the line opening, only 5 lines from 5 cities have observations that are before September 2016. The estimates are not precise and fluctuate substantially. The second episode is in the weeks between the New Year and the CNY in 2017. When log population is used to account for differential seasonality, the peak coefficient is around 0.05. The average treated city has a population of 8.68 million while the average control city has a population of 3.81 million. This translates into a 4 percentage points difference due to differential seasonality (0.05 * (ln(8.68) - ln(3.81))). According to the estimates in Panel B, differences in log GDP per capita between treated and control cities accounts for up to 1 percentage point adjustment at the peak (0.1 * (ln(105509) - ln(95674))). In Panel A, there is also some evidence of differential seasonality in the last week of January, 2018, as it approaches the 2018 CNY (February 16th). In Panel B, there is also a sharp but brief spike in the week of the National Day (October

1) in 2017. For the rest of the sample period, differential seasonality is not a serious concern.

We are only able to control for a limited number of city-level characteristics. This is primarily because of two reasons. First, we are only able to observe a limited number of city characteristics relevant for urban transportation. Second, including more characteristics and interact them with calendar week indicators runs the risk of over-controlling. The characteristic-specific time trend will eventually soak out time variation and make estimates of key coefficients imprecise.

There is evidence that our approach to account for differential seasonality, although not exhaustive, is adequate. Almost all the effects due to differential seasonality is concentrated around the 2017 CNY. Other city-level characteristics may account for differential seasonality in a way that log population and log GDP per capita do not pick up, but the *timing* when such differential seasonality matters is less likely to differ. We still find a positive and statistically significant effect, although somewhat smaller, when observations from January and February of 2017 are excluded. In addition, by using only treated cities, the two-way fixed effects model circumvents the problem of differential seasonality. Figure 2 Panels B and C show that the results are similar.

Table B.1 summarizes weekly estimates by estimating static DID models with different accounts for seasonality. The estimated effect ranges between 3.9% and 4.7%. Column 5 adds quadratic terms of log population and log GDP per capita, and the estimated effect is 3.9%. Column 6 includes up to the cubic terms, and the estimated effect comes back to 4.4%.

Whether differential seasonality, after our adjustment, remains to have an impact on our results can be gauged using a set of placebo tests. Suppose a subway line was launched in December 2017, we can estimate a placebo effect with a fake launch date that is one year before the actual launch date. If differential seasonality does affect our result and the $\gamma_t \cdot d_t \cdot \mathbf{X}_c$ term is inadequate to account for it, we would find some positive and significant effect in the placebo test. Similarly, we can construct a placebo test with lines launched in December 2016 and set fake launch dates that are one year after the actual launch dates.

One problem with this placebo test is that many treated cities in our sample have multiple lines launched during the sample period, often with launching dates that are one year apart. A positive result from the placebo test could be confounded with the network effect of subway. We therefore focus on "singleton" line launches: cities that did not have multiple line openings that are several months apart. These singleton launches allow us to use a placebo launch date without worrying about network effect from the other line launches in the same city.[2]

[2]There are 23 such singleton lines. We divide them into two groups. The first group includes those launched *before* April 30th, 2017 and there was no additional line launched in the same city before January 31, 2018. For these lines we use the placebo opening date on December 31, 2017. These lines include Fuzhou Line 1 (2nd phase north), Harbin Line 3, Shenzhen Lines 7 and 9, Suzhou Lines 2 (2nd phase) and 4, Tianjin Line 6, Xi'an Line 3, Zhengzhou Lines 1 (2nd phase) and 2, and the Suburban Line. The second group includes those launched *after* April 30th, 2017 and there were no other lines launched in the same city between August 1, 2016 and April 30, 2017. For these lines we use the placebo opening date on December 31, 2016. These lines include Changchun Line 1, Chengdu Lines 4 (2nd phase east and west), 10, and 7, Dalian Line 1, Guiyang Line 1, Hangzhou Line 2 (1st phase northwest), Kunming Line 3, Nanchang Line 2, Shanghai Line 9, and Xiamen Line 1 (1st phase).

Panels A and B of Figure B.4 show that for both groups of singleton lines, we still find a positive and significant effect on road speed upon the actual launching dates. Panels C and D show that the road speed does not change around the placebo launching dates.

B.2 Event Study

The simple event study approach essentially compares the outcome before and after the policy change. One can control for flexible time trends and identify a discontinuous change in outcome in the neighborhood of policy change. This revised approach is sometimes also called the regression discontinuity design using time as the running variable (Imbens & Lemieux, 2008; Lee & Lemieux, 2010). One crucial limitation of such an empirical design is that it cannot distinguish the treatment effect of the policy change from other time-varying confounding factors, including seasonality and macroeconomic trends. In our case, most of the subway lines were launched right before China's holiday season, and our data show that road congestion tends to ease with reduced economic activity. A naive event study will mistakenly attribute the seasonal increase in speed to the causal effect of subway openings.

Appendix Figure B.5 illustrates this point. From the baseline sample, we plot weekly average residual log speed separately for the treated and control road segments. The weeks are realigned relative to the week of subway line opening so we can take averages across all groups. For each week relative to the subway line opening, the averages are first taken within each subway line, and then taken across 45 lines. The red dots show the weekly averages of residual log speed from treated road segments, while the teal crosses show those from control road segments. The dashed lines are Lowess non-parametric smooth fits, separately for the treated (in red) and the control (in teal), and separately for before and after subway line openings.

Speed on treated road segments shows a clear jump by about 5% upon the time of a subway line opening. However, this could be misleading because the pattern could also be driven by seasonality. Speed on control road segments exhibits a similar, albeit smaller, jump of about 1% around the time of a subway line opening in *treated* cities. A comparison of the non-parametric smoothing lines for the treated and control road segments suggests that there is indeed a speed-enhancing effect of subway line opening, but the true effect is smaller than what a simple event study using only the treated segments would suggest.

A formal regression discontinuity design where the running variable is time relative to the subway opening can be specified as follows:

$$\ln \tilde{speed}_{lw} = \beta_0 + \beta_1 \cdot post_{lw} + \beta_2 \cdot post_{lw} \cdot w + f(w) + \varepsilon_{lw}, \tag{B.1}$$

where $\ln \tilde{speed}_{lw}$ is the residual log speed. w is the week relative to the subway opening, with $w = 0$ for the week when the subway line opens. $post_{lw}$ is a binary variable which takes value 1 for weeks on or after the subway opening. $f(w)$ is a flexible function of w, which is taken up to the 5$^{\text{th}}$ polynomial of w. ε_{lw} is the error term. β_1 captures the discontinuity in levels of average residual log speed as the subway line opens. β_2 captures the possible trend break after the subway line opening. We vary the bandwidth of weeks included in the regression. Up to 48 weeks before and 48 weeks after the subway line opening are used. We split the sample into treated-only and control-only sub-samples and estimate Equation B.1 separately on each sub-sample.

Panel A of Appendix Table B.2 presents the results from the regression discontinuity design on the treated road segments. Column 1 uses the baseline sample. It includes a linear time trend, $f(w) = w$. The result shows that the average speed on treated road segments increases by 6.4% after the subway line opens. Panel B presents the regression-discontinuity results from the control road segments. With the same specification as in the corresponding column in Panel A, it finds an "effect" of about 1.3%. Therefore, an event study approach relying only on the treated lines over-estimates the true effect. It is worth noting that subtracting the estimated coefficient in Panel B from that in Panel A (which yields a sort of "difference-in-differences" estimate) leads to an effect of 5.1%, which is similar to Panel A of Figure 2 (Note that even this approach generates an over-estimation because it does not account for seasonality). As higher-order polynomials are added in Columns 2 and 3, the estimates for both treated and control segments become smaller. In particular, the coefficients in Panel B become only marginally significant. The differences between the coefficients in the two panels remain fairly close to those in the stacked DID result before adjusting for differential seasonality. Column 4 further allows a linear time trend break after the line opening. We do not find different trend breaks for the treated and control lines. Columns 5-8 expand the sample period to 48 weeks before and 48 weeks after the subway line opening. For both treated and control lines we obtain slightly larger effects. Nevertheless, throughout these columns, the effect for the treated segments is larger than that for control segments by about 4-5 percentage points.

B.3 Effect by Individual Lines

Do subways differ in their effects in alleviating congestion? We estimate the effect separately for each new subway line using the stacked DID specification:

$$\ln\tilde{speed}_{lgw} = \beta_g \cdot T_{lg} \cdot Post_{lw} + \lambda_l + \lambda_w + \gamma_t \cdot d_t \cdot \mathbf{X}_c + \varepsilon_{lgw}, \tag{B.2}$$

Each regression includes road segments near *one* treated subway line and their randomly matched control segments (i.e., each regression includes one group g). The standard errors are adjusted for temporal (up to 20 periods) and spatial (up to 50 kilometers) autocorrelations.

In Appendix Figure B.6, the graph on the left reports the results of these estimates. Because new subway lines opened at different times during the sample period, with dynamic effects over time, the binary variable $Post_{lw}$ captures different periods relative to the opening dates for different lines. To avoid attributing dynamic effects to heterogeneous effects, we also estimate versions of the model in which we restrict the sample to be between 6 weeks before and 4 weeks after the opening. The graph on the right reports the results. In most cases there is a positive effect and most estimates are bounded between -10% and 15%.

It is interesting though that the heterogeneity in effects does not fall along the lines of usual suspicion. It does not seem that subway lines in larger, more populous cities have a larger effect, nor does it appear that new lines have a larger effect than extensions, or vice versa. It remains interesting for future research to determine what characteristics of a subway line and its surrounding environment are important for its effectiveness.

B.4 Effect by the Distance to Subway Station

Section 3.2 of the paper briefly discusses the challenge of isolating the true effect of subway from the effect of "subway opening" when there are other events taking place contemporaneous to the launch of a new subway line. These events include, but are not restricted to, opening of new shops, rerouting of buses, and changing of traffic patterns on surface roads. Some of these confounding events can be separately identified because their effects have a different *geographic* pattern from the effect from subway itself. In particular, events like shop openings and bus reroutings are likely to take place near subway stations. They bring new traffic to road segments near subway stations, but are less likely to affect traffic in road segments farther away from the stations but are nevertheless close to the subway line. Conditional on the distance to the subway *line*, road segments closer to a subway *station* are expected to have a smaller increase in road speed.

Table B.3 presents the heterogeneous effects by road segment's distance to the treated subway line and its stations. Columns 1 and 2 show that the effect declines by distance to the treated line and the treated stations. Column 3 shows that conditional on the distance to the treated line, road segments closer to a subway station indeed experience a smaller increase in speed. But the four-way interactive term is not statistically significant. Column 4 nonparametrically divides road segments into three groups: (1) those within 1 kilometer from the treated line and within 500 meters from a treated station; (2) those within 1 kilometer from the treated line but more than 500 meters away from a treated station; and (3) those more than 1 kilometer away from the treated line. Notice that on average segments in group (2) also have a longer distance from the treated line than those in in group (1), so caution is needed when interpreting the results. We find that the effect on road segments closer to the subway station is somewhat larger. Road segments that are close to the subway line but a bit farther away from subway stations also experience substantial increases in road speed. They are less likely to be affected by contemporaneous events like shop opening or bus rerouting. And the magnitude of the coefficient is similar to those in the baseline.

B.5 Mode Substitution from a Panel of Cities

Section 4 of the paper presents descriptive evidence of transportation mode substitution following subway expansion from household-level travel data in Beijing. Here we show supporting evidence from a panel of cities.

Information on transportation infrastructure and usage at the city level is from *Statistical Yearbooks of Chinese Cities*, compiled and published by China's National Bureau of Statistics (NBS, 2010-2017). These yearbooks report key statistics on urban transportation. Although variables included in these publications vary by year and city, annual subway ridership, bus ridership, and car ownership are available for most major cities in recent years. We compile a panel of 36 cities between 2010 and 2017. We investigate the correlations between subway ridership and volumes of alternative modes of transportation while controlling for year and city fixed effects. Column 1 of Table B.4 shows a clear substitution between subway and bus trips: As the number of subway trips increases by 100, the number of bus trips declines by 22. Increases in subway ridership also reduce the service mileage of buses. Admittedly, the substitution between subway and bus ridership is not necessarily a result of individual choices. It could be due to changes in the city's provision of public transportation. That is, when a new subway line opens, the city government may reduce bus services on routes that overlap with the new subway services. It is possible that with the expansion of the subway system, bus services are re-routed mainly to solve the last-mile problem of transporting passengers to subway stations. Column 2 shows that the total bus mileage reduces by 5 kilometers for every additional 100 trips by subway, although the coefficient is not statistically significant at the conventional level. Column 3 shows some evidence that increases in subway trips are associated with a reduced number of registered civilian-use cars. The coefficient is not statistically significant. The magnitude of the coefficient nevertheless indicates that 5,000 subway trips a year translates into one fewer car. This suggests a moderate substitution between subway trips and car ownership.

B.6 Estimating the Baseline Results in One Step

Throughout the paper, we adopt a two-step estimation approach. Log speed at the hourly level is first regressed on the road segment-by-hour of the day-by-day of the week fixed effect ($\lambda_{l,dow,h}$ in Equation 1). Residual from that regression is averaged by week relative to subway line opening. The average residual log speed is then used as the dependent variable in the baseline regression in Equations 2 and 3.

The choice of the two-step approach is primarily for reducing computational burden. Once weekly average residual log speed is obtained, sample size is reduced by about 20 fold (4 hours a day, 5 days a week), thus second-step estimations in various specifications are made much easier.

If we run the regression in one-step, the specifications would be, with a slight abuse of notation, as follows:

$$\ln speed_{lghd} = \sum_{w=\underline{w},\ w\neq -1}^{\bar{w}} \beta_w \cdot T_{lg} \cdot d_{gw} + \lambda_{l,dow,h} + \lambda_{gw} + \gamma_t \cdot d_t \cdot \mathbf{X}_c + \varepsilon_{lghd}, \tag{B.3}$$

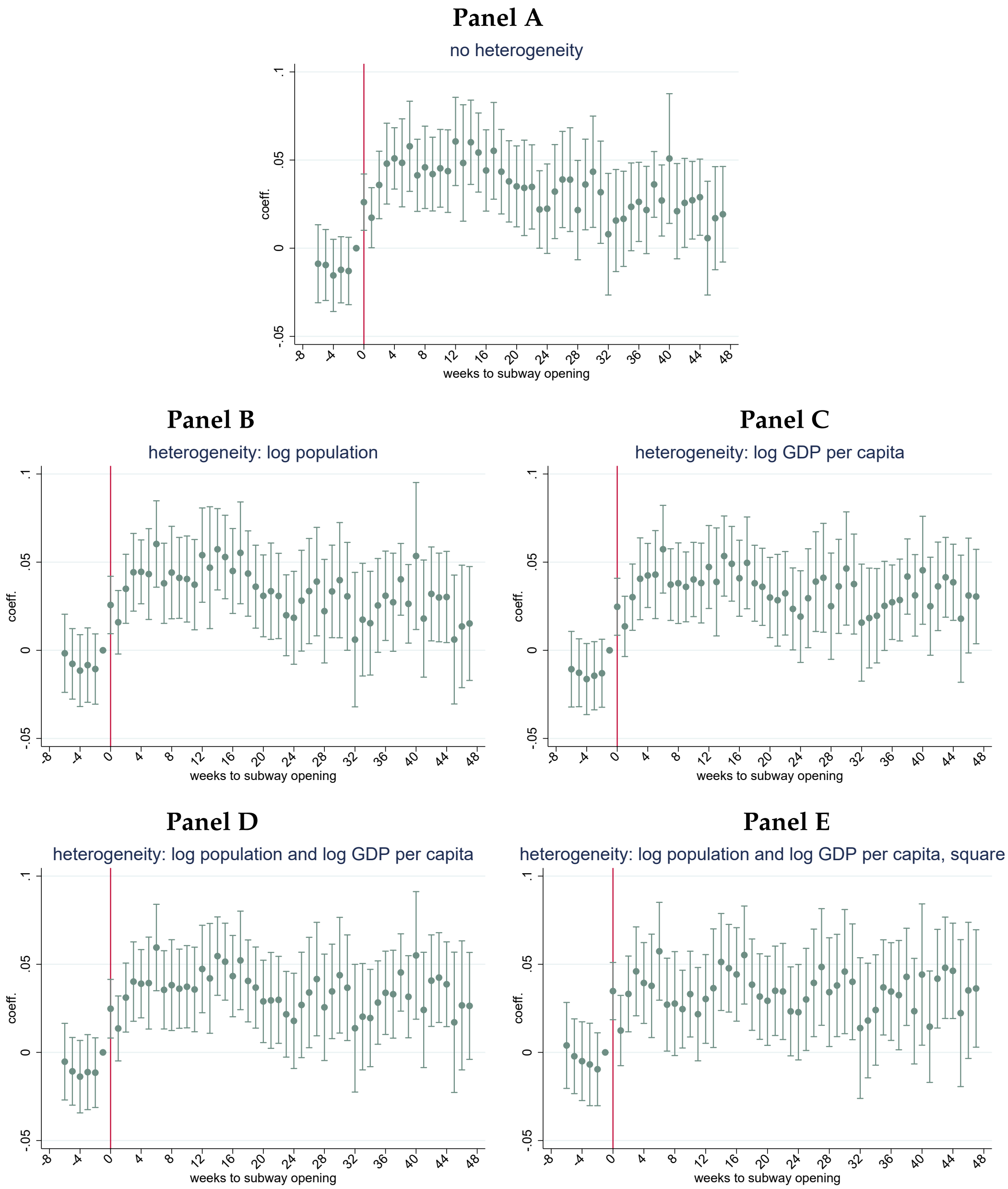
$$\ln speed_{lhd} = \sum_{w=\underline{w},\ w\neq -1}^{\bar{w}} \beta_w \cdot T_l \cdot d_{lt}^{w} + \lambda_{l,dow,h} + \tau_t + \gamma_t \cdot d_t \cdot \mathbf{X}_c + \varepsilon_{lhd}, \tag{B.4}$$

where h indicates hour and d indicates day. w in Equation B.3 is the week relative to subway line opening, and t in Equation B.4 is the calendar week. $\lambda_{l,dow,h}$ is the road segment-by-day of the week-by-hour of the day fixed effect.

The two-step approach is largely inconsequential, but the results will be slightly different from the one-step specification for two reasons. First, the two-step approach takes average residual hourly log speed at the weekly level, then treat each weekly observation equally. However, due to national holidays, missing data in several weeks in 2016, and exclusion of test ride periods, not all weeks have 20 hours. These differences may generate some discrepancies in the estimate. Second, the two-step approach uses an estimated variable as the dependent variable in the second step, the standard error in the second step is incorrect.

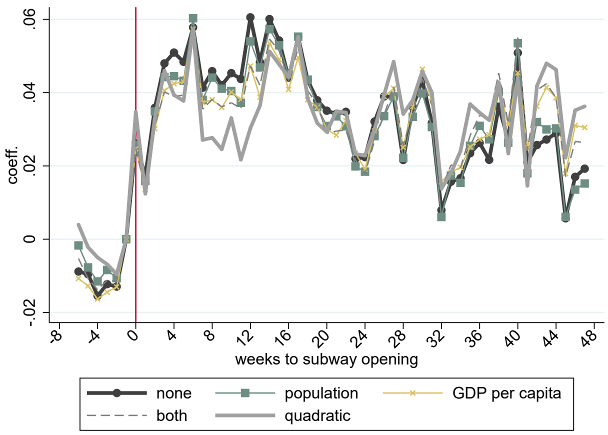
We replicate the baseline Table 3 using the one-step specifications. The results are reported in Table B.5. Overall we obtain very similar results. Coefficients and standard errors are slightly different in some cases, but the differences are small and do not affect our conclusions.

Figure B.1: Accounting for Differential Seasonality



Note: Each panel is from the estimation of a version of Equation 2. Panel A does not account for differential seasonality across cities. Panel B includes log population and its interaction with calendar week dummies. Panel C includes log GDP per capita and its interaction with calendar week dummies. Panel D includes both log population and log GDP per capita and their interactions with calendar week dummies. Panel E adds quadratic terms.

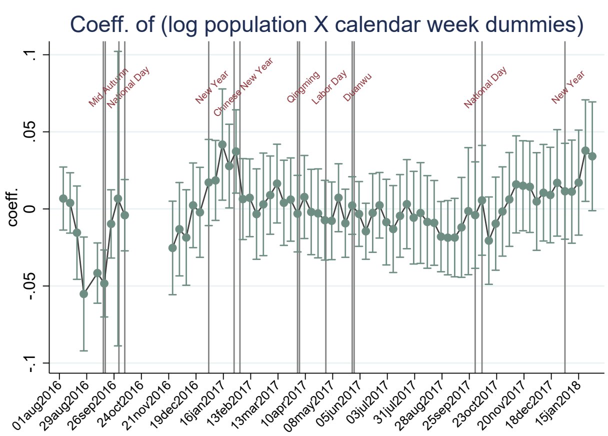
Figure B.2: How Much Does Differential Seasonality Matter?



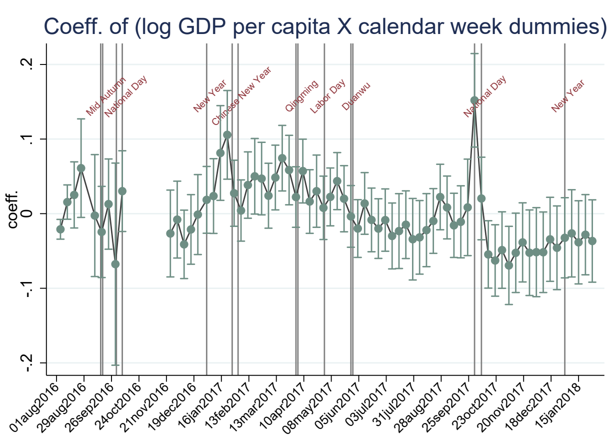
Note: This graph plots coefficients from 5 estimations as shown in Figure B.1.

Figure B.3: Visualizing Seasonality by City Characteristics

Panel A

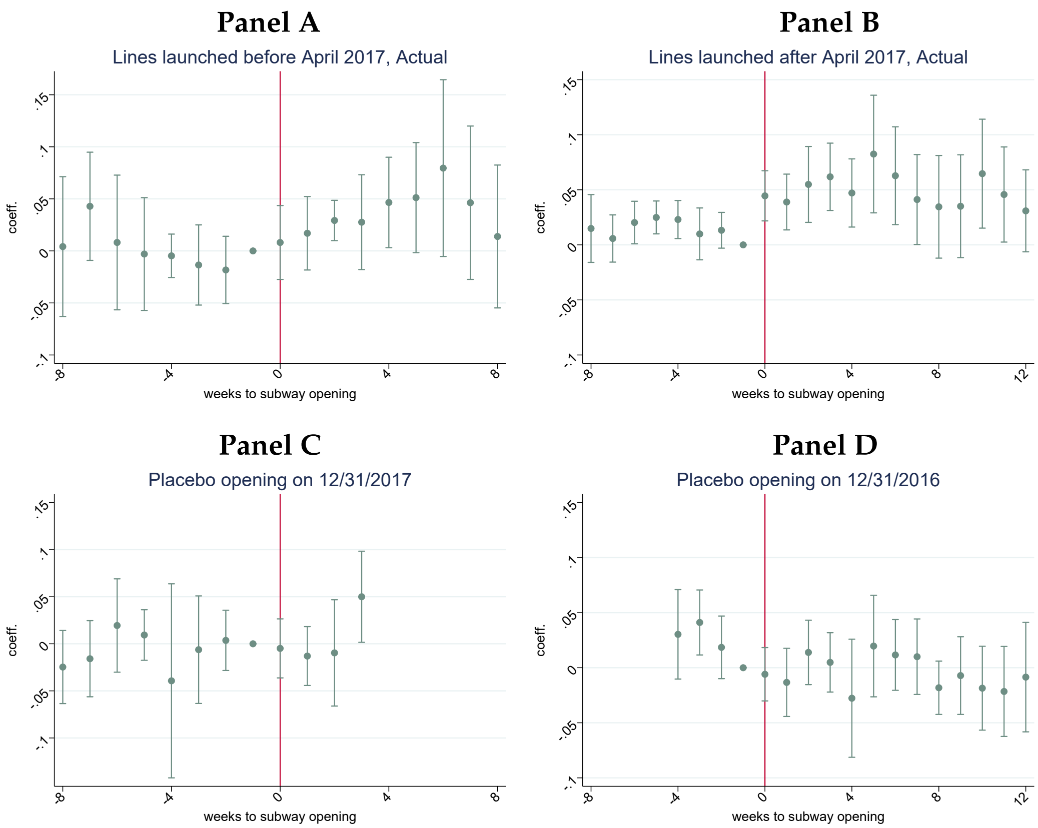


Panel B



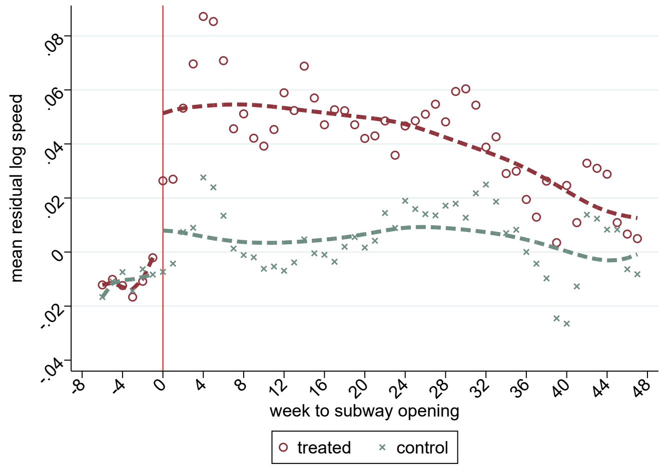
Note: $\hat{\gamma}_t$'s and the corresponding 95% confidence intervals from estimating Equation 2 are reported. $\mathbf{X}_c$ includes log population in Panel A and log GDP per capita in Panel B.

Figure B.4: Placebo Tests for Singleton Launches



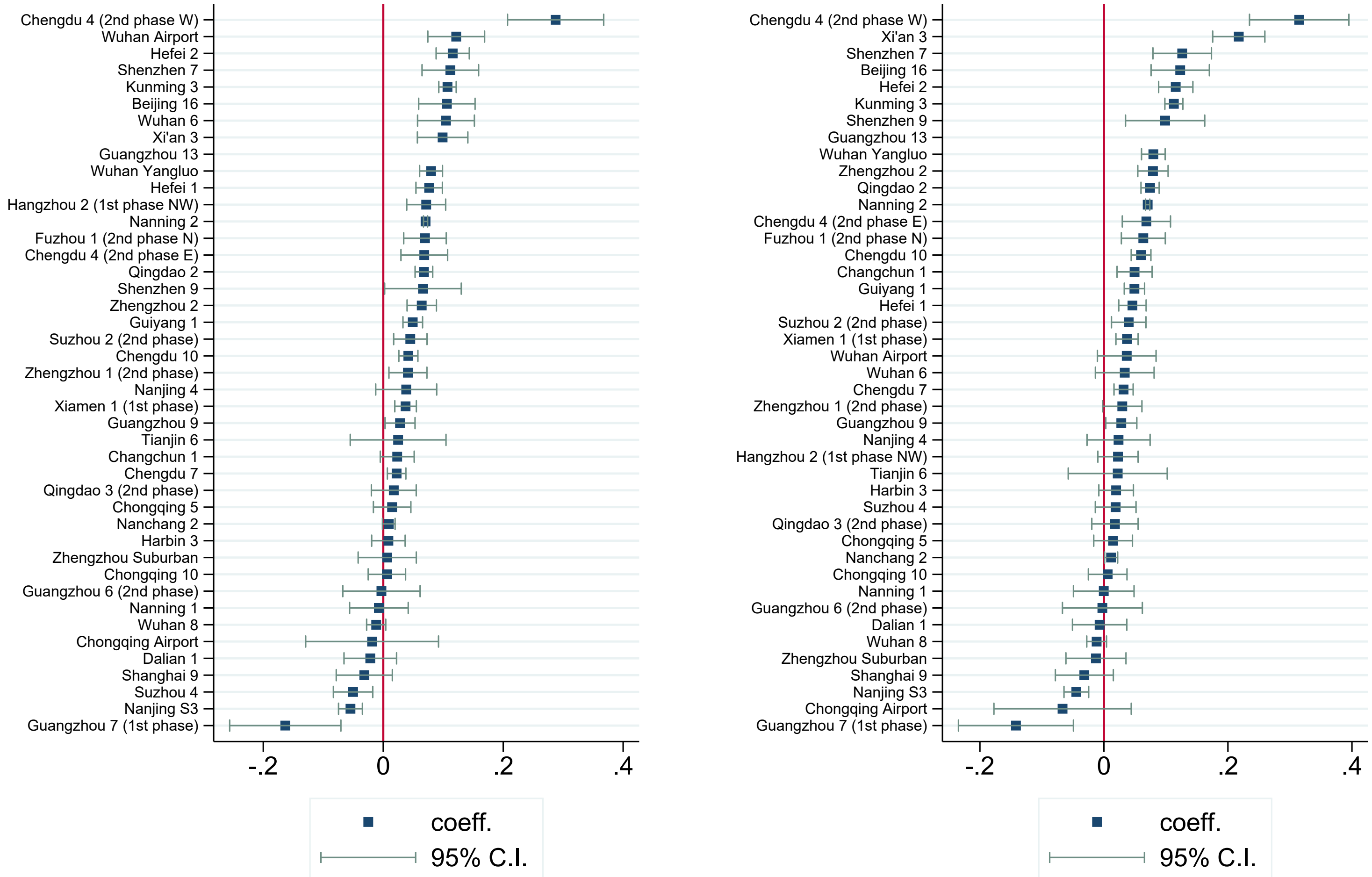
Note: The sample in Panels A and C includes subway lines that were launched before Apr 30, 2017 and there was no additional line launched before Jan 31, 2018 in the same city. Panel C shows the placebo effect where the placebo launch date is December 31, 2017. The sample in Panels B and D includes subway lines that were launched after Apr 30, 2017 and there was no line launched between Aug 1, 2016 and Apr 30, 2017 in the same city. Panel D shows the placebo effect where the placebo launch date is December 31, 2016. All models account for potential differential seasonality. Vertical bars represent 95% confidence intervals. Standard errors are clustered at the group level.

Figure B.5: Event Study



Note: Each red dot represents the average weekly residual log speed in treated road segments. Each teal cross represents the corresponding value from control road segments. The averages are first taken at the group-treatment status-week level, then taken across all groups.

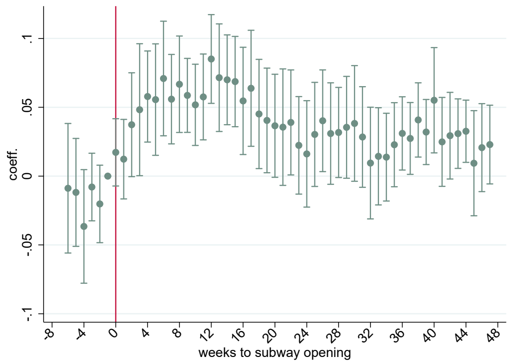
Figure B.6: Case-specific Estimates



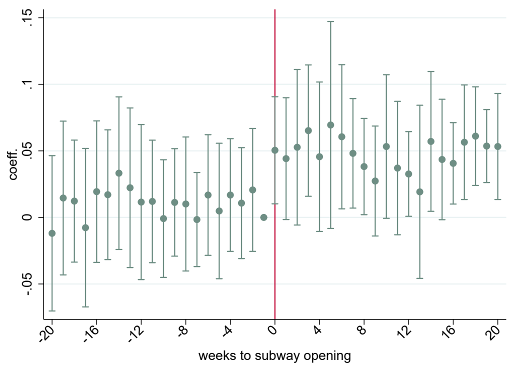
Note: Graph on the left: case-specific estimates with standard errors clustered at the subway line level, weeks relative to opening between -6 and 47 are included in the sample. Graph on the right: weeks relative to opening between -6 and 3 are included in the sample. Model specification is as in Equation B.2. Standard errors are adjusted for spatial and serial correlations. Guang-Fo line in Foshan has few treated road segments extracted. Xijiao Line in Beijing was launched on 12/30/2017, but was closed on 1/4/2018 due to an accident and was not reopened until 3/1/2018, which was after the end of our sample period. These two lines are excluded from case-specific estimations.

Figure B.7: Lines Launched in Different Periods — Not Controlling for Differential Seasonality

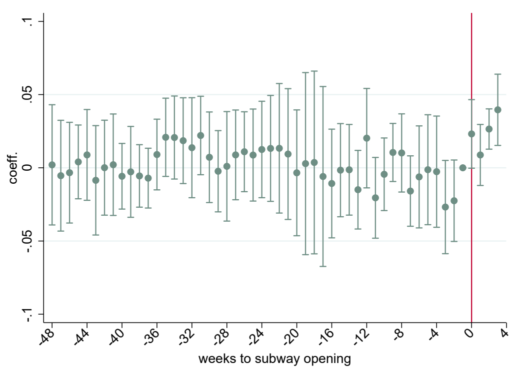
Panel A: Lines launched before Jan 31, 2017



Panel B: Lines launched between Feb 1 and Nov 30, 2017



Panel C: Lines launched after Dec 1, 2017



Lines by time of opening	Panel A	Panel B	Panel C
treated×post	0.052	0.038	0.023
	(0.011)	(0.011)	(0.015)
range of week rel. to opening	[-6,47]	[-20,20]	[-48,3]
# of lines	21	9	15

Note: Differential seasonality is not accounted for. Standard errors are clustered at the group level.

Table B.1: Accounting for Differential Seasonality

	(1)	(2)	(3)	(4)	(5)	(6)
treated×post	0.047	0.044	0.046	0.044	0.039	0.044
	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)
segment and group-by-week-to-open FE	✓	✓	✓	✓	✓	✓
calendar week × city characteristics						
log population		✓		✓	✓	✓
log GDP per capita			✓	✓	✓	✓
quadratic terms					✓	✓
cubic terms						✓

Note: Stacked difference-in-differences models with different accounts for differential seasonality. Standard errors are clustered at the group level.

Table B.2: Regression Discontinuity Using Time as the Running Variable

Panel A: treated	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
post	0.064	0.046	0.044	0.055	0.070	0.062	0.067	0.068
	(0.012)	(0.012)	(0.010)	(0.013)	(0.015)	(0.013)	(0.011)	(0.012)
post×weeks to open				0.023				0.003
				(0.011)				(0.003)
order of polynomial	1	3	5	5	1	3	5	5
weeks in sample	[-6,47]	[-6,47]	[-6,47]	[-6,47]	[-48,47]	[-48,47]	[-48,47]	[-48,47]
N	284,301	284,301	284,301	284,301	492,996	492,996	492,996	492,996
Panel B: control	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
post	0.013	0.009	0.009	0.017	0.020	0.010	0.021	0.022
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)
post×weeks to open				0.017				0.002
				(0.005)				(0.001)
order of polynomial	1	3	5	5	1	3	5	5
weeks in sample	[-6,47]	[-6,47]	[-6,47]	[-6,47]	[-48,47]	[-48,47]	[-48,47]	[-48,47]
N	358,152	358,152	358,152	358,152	681,948	681,948	681,948	681,948

Note: The dependent variable is the residual log speed. Standard errors are clustered at the subway line level.

Table B.3: Heterogeneous Effects by Distance to Subway Station

	(1)	(2)	(3)	(4)
treated×post	0.046	0.040	0.049	
	(0.009)	(0.009)	(0.010)	
treated×post×				
dist to line	-0.013		-0.018	
	(0.004)		(0.007)	
dist to station		-0.010	-0.002	
		(0.005)	(0.011)	
dist to line×dist to station			0.003	
			(0.006)	
dist to line < 1 km				0.056
& dist to station < 500 m				(0.009)
dist to line < 1 km				0.041
& dist to station > 500 m				(0.009)
dist to line between 1 km and 2.5 km				0.017
				(0.007)
λ_{lg}	✓	✓	✓	✓
λ_{gw}	✓	✓	✓	✓
λ_{kgw}				✓
$d_t \cdot \mathbf{X}_c$	✓	✓	✓	✓

Note: Standard errors are clustered at the group level.

Table B.4: Subway and Other Modes of Transportation in a Panel of Cities

	(1) bus ridership per resident (trips/resident)	(2) bus mileage per 100 residents (kilometers)	(3) # of cars per 10,000 residents (count/10,000 residents)
subway ridership per resident	-0.218	-5.312	-2.203
	(0.103)	(10.488)	(3.471)
city FE	✓	✓	✓
year FE	✓	✓	✓
mean dependent variable	264	9,623	2,123
N	529	249	426
N of cities	36	36	36

Note: Data are from *Statistical Yearbooks of Chinese Cities* (NBS, 2010-2017), published annually by the National Bureau of Statistics. Standard errors are clustered at the city level.

Table B.5: One-Step Estimates

	(1)	(2)	(3)
treated×post	0.044	0.033	0.033
	(0.008)	(0.009)	(0.010)
model	stack DID	two-way FE	two-way FE
group-by-week-to-open FE	✓		
segment-day of week-hour of day FE	✓	✓	✓
calendar week FE		✓	✓
adj. for differential seasonality	✓		✓
incl. segments from control cities	✓		
N	11,931,441	5,292,057	5,292,057

Note: Standard errors are clustered at the group level.

C Additional Details on Welfare Calculations

C.1 Sources of Data

Numbers of Commuting Trips—Numbers of commuting trips by mode are from the *2016 Beijing Transportation Annual Report* (BTAR) (BTI, 2016).[3] The 2016 BTAR reports 5.7 million commutes by car, 4.53 million by subway, and 4.52 million by bus on an average working day. Assuming 250 working days in a year, the annual number of commutes is 1.4 billion by car, 1.1 billion by subway, and 1.1 billion by bus.

Time Cost of Each Commute—The 2016 BTAR reports that the average commuting time is 56 minutes. We assume this number applies to all commutes regardless of mode of transportation.

Time Value— The *2016 Statistical Yearbook of Beijing* (SYB) (NBS, 2010-2017) reports that the average annual salary is 92,456 *yuan*. Assuming a full-time worker works 2,000 hours per year, this translates into an hourly wage of 46 *yuan*, or 0.77 *yuan* per minute. Monetary value for time spend on a trip is assumed to be half that value, or 0.39 *yuan*.

Pecuniary Cost of Each Trip—The BTAR reports the a single subway trip costs 4.7 *yuan* on average and a single bus trip costs 1.5 *yuan* (there is no peak pricing in Beijing's transit system in 2016). For car commutes, the BTAR reports that the average distance is 13.2 kilometers. We cannot find a reliable estimate of the operational costs of a car in Beijing. American Automobile Association estimates that cars cost 0.62 USD per mile.[4] This translates to 33 *yuan* per car trip.[5] A taxi trip of

[3] Available for download at http://www.bjtrc.org.cn/List/index/cid/7.html, last accessed on 10/25/2019.

[4] https://newsroom.aaa.com/2013/04/cost-of-owning-and-operating-vehicle-in-u-s-increases-nearly-two-percent-according-to-aaas-2013-your-driving-costs-study-archive/. Last accessed on 10/25/2019.

[5] While the cost of living in China is in general lower than that in the United States, cost of owning and operating a car is not necessarily cheaper, especially in Beijing. Cars, especially those imported, are sold at higher prices in China. Gas price is regularly 20-30% higher, while insurance and maintenance is less expensive due to lower labor cost. The calculation also does not take into consideration the fact that carpooling is more common in China than in the United

the same distance costs about 45 *yuan*, which arguably is a slightly more expensive option than driving one's own car (adding extra to pay for the service, less the cost of parking).

Mode Substitution— Table 6 Panel B shows that bus and car trips are diverted to subway trips. Among the diverted trips, 53% are bus trips and 47% are car trips. Panel B also indicates evidence of crowd-in to subway trips. Because the model assumes inelastic demand for commute, the welfare calculation here ignores welfare gains from these trips.

Government Spending on Public Transit— The cost of subway consists of construction and operational costs. Yang et al. (2018) estimate that the construction cost of 1 kilometer of subway is 598 million *yuan* (92 million USD) in Beijing. By the end of 2016, the system had a total length of 600 kilometers of track. With an annual discount rate of 3%, the annualized construction cost of the system is 10.79 billion *yuan*. On the operational side, it is reported that under the current fare schedule, government subsidizes 50% of each subway ride (Er, 2014). The average fare for a subway ride is 4.7 *yuan*. The Beijing subway had 3.78 billion total rides in 2016 (including both commuting and non-commuting trips), so the average government spending, including both annualized construction cost and subsidies in operational cost, is 7.48 *yuan* per ride.[6] This gives an overall subsidy rate for a subway ride of 62%. Each bus trip is also subsidized by 62% (Er, 2014) (including only operational costs). The average fare for a bus trip is 1.5 *yuan*. So the average subsidy per bus ride is 2.45 *yuan*.

Data sources and key parameter values are summarized in Table C.1.

C.2 Additional Welfare Calculations

Here we present the back-of-the-envelope calculations of the second and third terms of Equation 9. The quantification uses additional data, stronger assumptions, and non-causal correlations, and the results should be interpreted with additional caution.

The second term—the welfare gain for the switchers—is represented by the hashed area B in Figure 7. Among the switchers, the marginal traveler is indifferent between road vehicle and subway; her welfare gain is Δu_a. The individual with the strongest preference for subway benefits more, as illustrated in Figure 7 by the gap between $u_s + \xi_{is}$ and $u_a + \xi_{ia}$ on the left $y-$axis. Thus the second term can be approximated by the average of these two extreme cases multiplied by the number of subway commutes.

The maximal gain among the switchers is governed by ξ_{im} and is unknown to the researcher.[7] To gauge the size of that area, we make an additional assumption: If the costs are the same between subway and road vehicle commutes, there will be some positive number of individuals who choose the subway. That is, $\max_i\{\xi_{is} - \xi_{ia}\} \geq A_a - A_s$. With this assumption, the second

States.

[6]10.79 billion *yuan*/3.78 billion rides + 4.7 *yuan*. Assuming fare plus subsidy equal total cost.

[7]Intuitively, ξ_{im} governs the elasticity of demand for trips in each mode of transportation.

term is bounded from below by $[(c(N)-r)+(c(N)-c(N_a))]\cdot N_s/2$. Equation 9 becomes:

$$\Delta W \geq [c(N)-c(N_a)]\cdot N_a + [(c(N)-r)+(c(N)-c(N_a))]\cdot N_s/2 + [\tau_a - \tau_s]\cdot N_s. \tag{C.1}$$

Assuming all commutes cost the same amount of time, to calculate $(c(N)-r)$ we only need to focus on pecuniary costs. The average subway fare of a single trip in Beijing was about 4.7 *yuan*, $r = 4.7$. $c(N)$ is the weighted average of private cost per bus trip and cost per car trip. One bus trip costs 1.5 *yuan* and one car trip costs 33 *yuan*. 53% of the diverted trips are from bus trips and 47% are from car trips. So $c(N)$ is 16.31 *yuan*. From the empirical estimation, $(c(N)-c(N_a))$ is 0.65 *yuan*. In 2016, Beijing's subways carried 1.1 billion commutes, i.e., $N_s = 1.1$ billion. So the second term amounts to about 6.74 billion *yuan*.[8]

The third term is the aggregate difference in government subsidies between bus and subway commutes. Each subway commute is subsidized by 7.48 *yuan*. So the total subsidy to subway commutes is 8.2 billion *yuan*. We assume there is no subsidy or tax on car commutes.[9] Diverting bus to subway reduces subsidies to bus commutes. 53% of the subway trips were diverted from buses, and each bus trip is subsidized by 2.45 *yuan*, which amounts to 1.4 billion *yuan* (2.45×1.1 billion×0.53). The net subsidy in commuting trips is therefore 6.8 billion *yuan* per year. These rough calculations indicate that each of the latter two terms are larger in magnitude than the first term, but they largely cancel each other.

Cautions are called for when interpreting these numbers. The largest welfare gain comes from the switchers, on whose preferences we have to make additional assumptions. Correlational instead of causal evidence is used. Costs and subsidies are from rough estimates. As discussed in the main text, many sources of benefits and costs are also omitted from the framework.

[8] The second term is $[(c(N)-r)+(c(N)-c(N_a))]\cdot N_s/2$, $c(N)-r = 16.31-4.7 = 11.61$, $c(N)-c(N_a) = 0.65$, $N_s = 1.1$ billion.

[9] There is no congestion pricing or toll on Beijing's urban roads. We do not consider the implicit taxes on car trips by various restrictions imposed by the government on car travels.

Table C.1: Parameters for Welfare Calculation

variable	notation	value	source
Panel A: Numbers of commuting trips			
annual # of subway commutes	N_s	1.1 billion	BTAR
annual # of bus commutes	part of N_a	1.1 billion	BTAR
annual # of car commutes	part of N_a	1.4 billion	BTAR
Panel B: Time value of each commute			
time cost per **commute**	-	56 min	BTAR
average annual wage	-	92,456 *yuan*	SYCC
time cost per trip (monetized)	part of c_m	22 *yuan*	AC
Panel C: Pecuniary cost of each commute			
single-trip subway fare	part of c_s	4.7 *yuan*	BTAR
single-trip bus fare	part of c_a	1.5 *yuan*	BTAR
cost of a single-trip car ride	part of c_a	33 *yuan*	AC
Panel D: Mode substitution			
# of bus rides diverted per subway trip	-	0.53	Table 6
# of car trips diverted per subway trip	-	0.47	Table 6
Panel E: Subsidies to public transits			
subsidy per subway trip	τ_s	7.48 *yuan*	AC
subsidy per bus trip	τ_a	2.45 *yuan*	AC

Note: Explanation of abbreviations: AC—authors' calculations; BTAR–*2016 Beijing Transportation Annual Report* (BTI, 2016); SYCC—*2016 Statistical Yearbook of Chinese Cities (NBS, 2010-2017).*