Political Career Incentives and Infrastructure Investment: Evidence from China*

Zhenhuan Lei† Junlong Aaron Zhou‡

This version: May 12, 2019

Please check the latest version here.

Abstract

Why do politicians who have short tenures have incentives to make long-term investment in infrastructure projects? Earlier research shows that government-led investment in infrastructure can be explained by electoral incentives of politicians in democracy. We argue that infrastructure investment can also be explained by the political career concerns of politicians in a top-down authoritarian system. We test this argument with the case of subway construction in Chinese cities. With both a generalized difference-in-differences design and a fuzzy regression discontinuity design, we demonstrate that obtaining central government's permission to build a subway system helps a mayor get promoted to a higher-level position. We also present evidence that subway construction boosts investment, economic productivity, and fiscal revenue in the city, which in turn helps build up the performance of provincial leaders. Provincial leaders then reward these mayors with higher chances of political promotion.

Keywords: China, Distributive Politics, Subway, Infrastructure, Political Selection

---

*Earlier title: "Seeing Like a Great City: The Political Motivation of Subway Construction in China." We thank Tianyang Xi and Yang Yao for sharing the CCER Officials Dataset. For helpful comments and suggestions, we thank Bruce Bueno de Mesquita, William Godel, Jean Hong, Yue Hou, Ruixue Jia, James Kung, Pierre Landry, Yishuang Li, Xiaobo Lu, Zhaotian Luo, Edmund Malesky, Dan Mattingly, Julia Payson, Adam Przeworski, Pablo Querubin, Peter Rosendorff, Arturas Rozenas, Shanker Satyanath, Peng Shen, Shiran Victoria Shen, Victor Shih, Alastair Smith, Austin Strange, Rocío Titiunik, Erik Wang, Yuhua Wang, Hye Young You, Congyi Zhou, and participants at the 2019 annual conference of Midwest Political Science Association and the Third Quantitative China Studies Seminar at New York University.

†Ph.D. candidate, Department of Politics, New York University. Email: zhenhuan.lei@nyu.edu. Personal website: http://www.reedlei.com. Corresponding author.

‡Ph.D. student, Department of Politics, New York University. Email: jlzhou@nyu.edu.

1 Introduction

Transportation infrastructure such as highways, railways, and subways are vital to urban growth, market integration, improvement of air quality, and job creation (Donaldson, 2018; Duranton & Turner, 2012; Zheng & Kahn, 2013a,b; Banerjee *et al.* , 2012; Baum-Snow & Kahn, 2000; Gendron-Carrier *et al.* , 2018). Given these positive externalities, it is natural to argue that the government should take the lead to invest in transportation infrastructure. However, while city mayors *could* take efforts to build or upgrade urban transportation infrastructure, it is unclear *why* they should do so. More importantly, most transportation infrastructure projects take many years to complete, but city mayors have shorter tenures. Hence, these short-term mayors are unlikely to claim the credit for the provision of transportation infrastructure. Then why should city mayors have incentives to invest in infrastructure? How could infrastructure investment benefit mayors?

These questions are especially important for fast-growing authoritarian countries. These wealthy authoritarian countries have made more investment in physical capital than wealthy democracies (Przeworski *et al.* , 2000, Chapter 3). China is an example in case. Scholars studying the Chinese economy argue that the government investment in key infrastructure has played a crucial role in promoting its economic growth (Lin, 2011).[1] Then what explains the strong incentives of Chinese mayors to make massive government investment in infrastructure? Why would Chinese mayors whose average term is roughly three years (authors' data) embark on long-term infrastructure construction?

We offer a political explanation for these questions. If the construction of infrastructure projects can bring political returns to city mayors, then even short-tenure mayors would have

[1]However, some other scholars point out that much of the infrastructure investment in China does promote the long-term growth and is misplaced in places which do not need those infrastructures (Ouyang & Peng, 2015; Shi & Huang, 2014). Meanwhile, China's focus on the investment in physical capital has crowded out its investment in human capital which may affect China's long-term economic growth (Li *et al.* , 2017).

an incentive to build transportation infrastructure. This is indeed the case in China due to a unique top-down authoritarian system. Unlike most democracies where voters select mayors, provincial leaders appoint city mayors in China. Hence, to get promoted, city mayors need to deliver the policies that are favored by the provincial leadership. The accumulated scholarship of Chinese political selection shows strong evidence that provincial leaders generating higher economic growth rate and more fiscal revenue are more likely to be promoted (Li & Zhou, 2005; Jia *et al.* , 2015; Xu, 2011). As those aspiring city mayors also know these rules for the political promotion of their superiors, they should have a strong incentive to promote economic productivity and fiscal revenue in their cities, which will contribute to the economic performance of provincial leaders. In return, provincial leader will reward those mayors who have produced investment and fiscal revenue with political promotion. Such a political exchange relationship between mayors and the provincial leadership gives city mayors strong incentives to build infrastructure even if the construction will not be completed in his term (it is usually a "he" in China).

To test this argument, we focus on subway construction in Chinese cities. More specifically, we ask why do Chinese mayors have strong incentives to apply to the central government for building a subway system, a typical government-led investment in urban infrastructure among Chinese cities. To answer this question, we collect an original panel data set for Chinese cities from 2003 to 2016. To preview the results, the findings are largely consistent with the theoretical predictions. By employing a generalized difference-in-differences (DID) design, we show that obtaining the subway approval from the central government boosts the mayor's promotion chance by roughly 16.4 percentage points (which is over 40% increase from the baseline promotion rate of 36.8 percent). Using another fuzzy regression discontinuity design (RDD), we obtain even more salient results. Furthermore, once cities start building subway lines, there is indeed a huge surge in the investment level, GDP per capita, and fiscal revenue. However, we do not find evidence to show that subway construction has created more jobs or improved air quality for city residents in the short term.

2 Political Career Incentives and Subway Approval

2.1 Related Literature and Theoretical Framework

Why do politicians have incentives to build infrastructure projects? Earlier research in economics and political science have offered two explanations for these questions. First, infrastructure projects offer rent-seeking opportunities to politicians (Robinson & Torvik, 2005; Keefer & Knack, 2007; Lehne *et al.* , 2018). Second, scholars studying distributive politics have argued that it is in fact the electoral incentives that motivate politicians to make infrastructure investment. The distribution of infrastructure projects can be used to buy the support of voters who will eventually determine the career prospects of politicians (Cadot *et al.* , 2006; Castells & Solé-Ollé, 2005; Voigtländer & Voth, 2014; Hong & Park, 2016).

Building on the second view, we argue that the logic of electoral incentives can also be applied to settings where politicians are not selected via elections. The focus on *electoral* motivation of infrastructure investment makes the theoretical implications unnecessarily narrow. The wisdom embedded in the electorally motivated distribution seems powerless for the understanding of infrastructure or other distributive goods in settings where politicians are not elected. In its essence, the electoral motivation argues that politicians serve for the people who make political selection and promotion. In an electoral system, it is the voters who select their mayors and members of congress. In a non-democratic setting, there will be a smaller group of people doing the same work, such as Politburo of a communist country. Hence, instead of electoral motivation, a broader explanation for politicians' motivation to invest in projects of transportation infrastructure is *selectoral* (Bueno de Mesquita *et al.* , 2003).

In the case of China, the selectorate that city mayors need to please is provincial leaders. Unlike their counterparts in democratic countries, mayors in China are never elected and are not required to be indigenous residents of their cities. In contrast, a group of provincial

leaders such as provincial party secretary and other important members of the provincial Communist Party Committee (e.g., provincial governor) will determine career prospects of city mayors. Moreover, city mayors in China have no control of when or where to run as mayor - these are all decided by provincial leaders. This means that it is impossible for city mayors to strategically choose a more prosperous or familiar city to serve.

Unlike local leaders in democratic settings who may need subordinates to deliver votes, provincial leaders in China do not need people's votes to keep their positions or get promoted. The Chinese political system is described by Bell (2016) as "political meritocracy" because the Chinese government promotes competent officials to more important positions. But what this really means is that provincial leaders with better economic performance are more likely to get promoted (Li & Zhou, 2005; Jia *et al.* , 2015). Hence, provincial leaders need economic performance such as higher economic growth rate and more fiscal revenue to get promoted to higher positions in central government.

This demand of provincial leaders for robust economic performance means that Chinese mayors should have a strong incentive to help the province produce economic growth and fiscal revenue. One possible mechanism is to make large-scale government investment in transportation infrastructure projects such as subways, roads, and bridges. These infrastructure projects can (1) boost economic growth due to the construction investment, (2) generate more land sales revenue which is now the major component of local fiscal revenue in China, and (3) attract more business investment due to an expected more convenient transportation system. After receiving these "porks" offered by city mayors, provincial leaders will reward these mayors who have helped them demonstrate economic performance through the government investment in urban transportation. Hence, we should expect to see that these "helpful" mayors are more likely to get promoted.

While this theoretical account is consistent with the existing research in Chinese politics which demonstrates that cities' economic growth and fiscal revenue size in the *past* increase the promotion chances of their mayors (Yao & Zhang, 2015; Landry *et al.* , 2018; Landry,

2008), the causal mechanism in this paper is different from these earlier works. At the core of these existing works is a signaling model that mayors use their economic performance in the *past* as a signal for their competence. However, our theoretical prediction does not depend on such a signaling mechanism. Instead, the higher promotion chances of mayors obtaining infrastructure projects are due to a "*quid pro quo*" relationship between mayors and their provincial superiors. In such a political exchange relationship, mayors need not to be competent; all they need is to deliver the "porks," that is infrastructure projects in our case, for their selectorate. This could be done by paying extra efforts or using their political wisdom, but is also possible by pulling the right connections that mayors have in Beijing or other ways.

These subtle differences between our theory and the existing research in Chinese politics also lead to different identification strategies. To test whether competence boosts mayors' political promotion chances, one may need to find an exogenous variation in mayors' competence. However, in our case, the empirical challenge is to identify an exogenous variation in "porks" (i.e., infrastructure projects of urban transportation) in the *quid pro quo* relationship between mayors and provincial leaders.

2.2 Subway Approval and the Political Promotion of Mayors

Subway construction in Chinese cities can provide an appropriate setting and the required exogenous variation to test the quid pro quo explanation for the government investment in transportation infrastructure. In Figure 1, we present the administrative process a city government must go through before starting the construction of its first subway line. There are two stages of work. First, the city must propose a plan of subway system (PSS) which includes the layout, financing, expected utilization, and other detailed information of a system of *several* subway lines that the city intends to build in the coming few years. This PSS will be reviewed first by the provincial government, then by the National Development and Reform Commission (NDRC) - a powerful ministry affiliated with the central government,

after that by other related ministries (such as Ministry of Environmental Protection and Ministry of Housing and Rural-Urban Development), and finally approved by a vice Premier or even Premier of the central government.

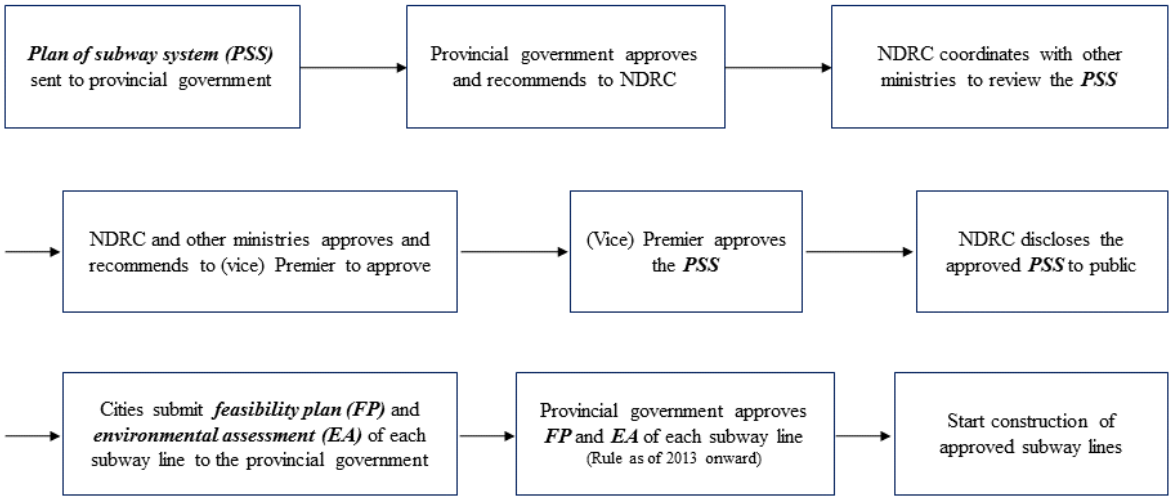

Figure 1: The Administrative Process of Approving Subway Construction in China

This first step of obtaining the approval of PSS (which we will also refer to as "subway approval") by the central government (including the review by NDRC and other ministries as well as the final approval by (vice) Premier) is seen as the most challenging step for a city to realize its "subway dream" and is therefore the focus of this paper. Moreover, the PSS approved by the central government will be disclosed by the NDRC to the public. Hence, we are able to collect a complete list of subway approvals made by the central government. However, we are unable to identify all cities that have applied for a subway but did not obtain the approval from the central government, because such cities are not required to disclose their failed application to the public. The lack of data for failed subway applications is another reason why we focus on the *approval* of PSS.

After obtaining central government's approval of the PSS, the city still cannot start the construction. The city government will still need to submit two other documents to the provincial government for approval: the feasibility plan (FP) and environmental assessment

(EA).[2] Note that FP and EA are prepared for each one of the subway lines that the city will build, whereas PSS is prepared for a system of several subway lines. Although FPs and EAs will be submitted for multiple times, the general understanding is that the approvals of FP and EA are easier to obtain because they are reviewed and approved by the provincial government with which city mayors are more familiar. To our best knowledge, there is not a single case where the construction of subway lines is canceled only because the city has not obtained the FP and EA from the provincial government. Again, this demonstrates that obtaining PSS approval is the key to building subway lines and justifies our focus on the approval of PSS by the central government.

Obtaining subway approvals is mainly the work of city government and its head, namely city mayor. When a city plans to apply for subway approval and build a subway system, the city government will usually establish a new government agency called "the leading team for subway planning and construction" (轨道交通规划与建设领导小组/指挥部). The main function of this "leading team" is to coordinate the efforts of different departments of city government and lead the preparation of the application for subway approval. In most cases, city mayors are the head of this "leading team," who is responsible for obtaining the subway approval and construction.[3] This is the major reason why we focus on city mayors rather

[2]The feasibility plan is a document prepared by the city government to prove that the construction of the subway line is safe, economically sustainable, and feasible (given the city's geographical conditions). The environmental assessment is the report provided by the Bureau of Environmental Protection at the provincial government, informing of the city government that the subway line construction will not incur any significant environmental impact. Before 2013, FPs were submitted to and reviewed by the NDRC. In 2013, the central government decided to delegate the approval of FP to the provincial government (see Chinese media news report here: https://www.yicai.com/news/2706590.html).

[3]For example, city mayor is the head of "leading team" in Dongguan (http://www.dg.gov.cn/007330010/0202/201610/88962478e193413caa7bd6e606f7dcd6.shtml, original government document in Chinese), in Changshu (http://www.sohu.com/a/215820768_707470, news report in Chinese), and in Liuzhou (http://www.liuzhou.gov.cn/xxgk/jcxxgk/zcwj/sjfggz_42671/201808/t20180816_1154085.html, original government document in Chinese).

than city party secretaries (CPS) who are not directly involved in the process of obtaining subway approval or building a subway system. If anything, we should expect that subway approvals do not have a positive influence on CPS's career development since CPS is not responsible for the subway project.

A related concern is whether city mayors can obtain approvals of PSS within their expected tenure. Since Chinese mayors have a short tenure (i.e., roughly 3.5 years), the approval must be quick enough to motivate short-tenure Chinese mayors to apply for and build subway. Indeed, this is the case of subway application in China. On average it takes a mayor 1.4 years to obtain the PSS approval from central government. In the worst case, the longest waiting time for the subway approval was only three years. However, to finish building the first subway line will take another 5.2 years on average. This means that if a mayor aims to finish building a subway line, he should expect to stay in office for 6.6 years to complete both the application and construction. This is much longer than the average tenure of a Chinese mayor. These simple statistics demonstrate that the investment in subway cannot be explained by mayors' motivation to finish building a subway system.

After these mayors obtaining the subway approval from the central government, their cities can start the construction soon, which will boost the investment and economic performance. Moreover, once the news came that the city will build subway lines, more investment and business will pour into the city given the transportation convenience. Subway lines may also generate more land sales revenue for the city (and province) through two possible mechniams: (1) by expanding the size city to the suburb so that the city may sell more land and (2) by increasing the land price near the subway stations. These economic and fiscal benefits will occur almost naturally as the city starts the construction of subway lines. In return, provincial leaders will reward the mayors who have obtained the subway project because subway projects enhance the economic performance of the city and province. Hence, we hypothesize that mayors obtaining approval are more likely to get promoted than other mayors.

3 Data and Identification Strategy

3.1 Data

We have utilized several sources of data to conduct the empirical analysis. To begin, we have collected an original data set of PSS approvals in Chinese cities. To collect this data, we have first relied on the annual reports released by the China Association of Metros to obtain the list of cities with a subway built or under construction. Then we searched on the Internet for each city in this list to obtain the original PSS released by the central government from which we can find out the approval date of the subway plans. We then merge these data with mayors' information from two sources: (1) the CCER Official Dataset[4] which contains the mayors' connection with provincial party secretary, age, and whether a mayor has been promoted or not after the tenure, and (2) Chinese Political Elite Database (CPED) offered by Jiang (2018), which contains rich information on mayors' personal characteristics and previous working experience.

In addition, we construct other city-level variables from another three sources: China City Statistical Yearbook, China Urban Construction Statistical Yearbook, and the air quality index (AQI) from the Ministry of Environmental Protection. The first contains rich information of Chinese cities across years, such as city population, GDP, GDP growth rate, annual fiscal revenue, unemployment rate, and so on. The second one, namely China Urban Construction Statistical Yearbook, records detailed information regarding cities' infrastructure investment level and land sales data. Finally, the AQI is a measure of the city's air quality of the urban area, with a higher number meaning worse air quality.[5] We combine all these datasets above into a city-year panel dataset for empirical tests.

[4]More detailed information about the CCER Official Dataset can be found at Xi *et al.* (2018).

[5]The AQI data from the Ministry of Environmental Protection is only available until 2014. Moreover, the Ministry of Environmental Protection only monitored the air quality of "major cities" in China. Both reasons explain why there are only 1407 observations for the AQI variable.

Table 1: Summary Statistics

Variable	Obs	Mean	Min	Max
Mayor promoted in 3 years	3,556	0.37	0.00	1.00
Mayor connection	3,553	0.01	0.00	1.00
Mayor age	3,549	50.23	33.00	61.00
City population	3,571	417.42	14.19	1591.76
City GDP (billion Yuan)	3,541	138.34	3.18	1954.74
City fiscal revenue (billion Yuan)	3,566	10.78	0.12	313.65
City GDP growth rate (%)	3,557	12.12	-15.95	109.00
Mayor obtaining subway approval	3,571	0.04	0.00	1.00
City investment in infrasturcture per capita (Yuan)	3,559	715.71	0.00	13236.19
City GDP per capita (Yuan)	3,556	32473.41	99.00	467749.00
City land sales revenue per capita (Yuan)	3,566	314.02	0.00	40277.59
City fiscal revenue per capita (Yuan)	3,566	2788.72	70.33	81467.34
City unemployment rate	3,537	0.59	0.00	11.54
Air quality index (AQI)	1,407	84.71	21.01	384.00

Because the central government (mostly NDRC) evaluates cities' PSS based on a government document released in 2003, we focus on the period when subway application is governed by this document, that is 2003 to 2016.[6] Moreover, to facilitate apple-to-apple comparison, we drop few "outlier" cities in our data sample. First, we drop four province-level megacities in China: Beijing, Tianjin, Shanghai, and Chongqing. Their mayors are minister-level appointments and their promotion is completely different from other cities in China. Second, we drop Guangzhou and Shenzhen in Guangdong Province, two so-called "first-tier cities" in China due to their large size and economic prosperity. In addition to their extreme size, these two cities' PSS were approved far earlier than 2003. Hence, their subway approvals are fundamentally different from the approvals of PSS at other cities after 2003.

This leaves us with a city-level panel data of 280 prefecture-level cities[7] spanning from

[6] 2016 is the latest year when we have access to the city mayors' data and city-level data from China Statistical Yearly Book and China Urban Construction Statistical Yearly Book.

[7] There are at least three types of cities in China. The first is province-level municipalities. The party secretaries of these municipalities are usually a member of the politburo, thereby making the administrative level of such cities higher than other cities. There are only four province-level municipalities in China:

2003 to 2016. The descriptive statistics of variables of this dataset can be found in Table 1. Among these 200-plus cities, 32 cities have obtained at least one PSS approval from the central government. Ideally, we want to gather data on citities' application for subway; however, this data is not available to our knowledge. Cities are not required to disclose that they have applied to the central government for building subway before the PSS has been approved. Without the data of cities that have applied for subway, one potential problem is that those have applied for subway may be different in many aspects from other cities which have not applied for subway (and hence have not received any subway approval) and these different features between the two groups of cities may drive the results we report. To overcome this identification challenge and other potential empirical concerns, we utilize a generalized DID design. We introduce this empirical design in the next section.

3.2 Identification Strategy

With the panel structure of the dataset, we mainly exploit the within city variation of PSS approval over time. For these 280 prefecture-level cities, the majority have never received any PSS approval, some have received only one, and some other cities have received multiple PSS over the years because they intend to build additional subway lines (e.g., to expand their earlier subway system). This longitudinal variation of PSS approval over time is the key to the identification strategy based on a generalized DID design. More specifically, we

Beijing, Shanghai, Tianjin, and Chongqing. The second category is what has been studied in this paper: prefecture-level cities. These are cities under the direct control of a province. We do not further distinguish between vice-province-level cities and other prefecture-level cities. The vice-province-level cities are usually capital cities of provinces (with few exceptions such as Shenzhen in Guangdong Province, Xiamen in Fujian Province, etc.) and are also subject to direct control of the provincial government. Yet, to test the robustness of our results, we exclude vice-province-level cities in some additional tests and usually obtain similar results. The final category of cities in China is county-level cities. These county-level cities are usually managed by the nearby prefecture-level cities. We choose prefecture-level cities in China for data analysis because they are by and large comparable to each other, large in number, and have detailed, publicly available data.

use the following equation to empirical analysis.

$$Promotion_{it} = \beta_0 + \beta_1 Approval_{it} + \gamma X_{it-1} + \theta_i + \pi_t + \epsilon_{it} \quad (1)$$

The outcome variable $Promotion_{it}$ is a dummy variable indicating whether or not the mayor of city i has been promoted within three years from year t. We define mayor's promotion by following criteria: the city mayor is promoted to (a) party secretary of a prefecture-level city; or (b) a vice-province-level position (e.g., vice provincial governor, vice minister of a ministry at central government). City mayors cannot be promoted directly to province-level or even higher rank position without first going through (a) or (b). We use a lead of three years because the average tenure of Chinese mayors is roughly three years. Yet, we also use mayors' promotion in one, two, four or five years as robustness checks in Appendix and have found similar results.

Explanatory variable $Approval_{it}$ is a dichotomous indicator equal to one if the mayor of city i and year t has obtained a PSS approval from the central government and to zero if otherwise. One methodological challenge here is that it is hard to identify which mayor can claim the credit of subway approval because city governments do not disclose which mayor started the work on applying for PSS approval. We identify this mayor by following procedures: first, we identify the period of time that the city has been working on the subway approval by finding out when city had applied for PSS (i.e., starting point)[8] and when city got approval (i.e., ending point). Usually during this period of time there was only one mayor and we identify this mayor as the one who applied for and got the PSS approval. In few other cases, there were two different mayors between when the city started to apply for subway and when the city obtained PSS approval. In such situations, we identify the mayor who worked in the city when the approval was made as the one who got approval.[9] These

[8]We use the time when city establishes the subway plan group, gets provincial PSS approval or submits PSS application to central government, whichever is available.

[9]One problem with this coding rule is that the mayor who got the approval may have only worked on

two cases exhaust all situations in our data.

Hence, the within-city variation of this variable $Approval_{it}$ can come from two sources: (a) a mayor is coded as one since the year when he obtains the PSS (but will still be coded as zero before he obtains the PSS); or (b) after the mayor who has obtained the PSS leaves the office, the next mayor will be coded as zero if this new mayor has not obtained another new PSS. One potential selection bias is that only cities without subway are eligible for a subway approval. However, this is not the case because cities can submit multiple applications for PSS even after the first subway approval to, for example, expand the current subway system. In fact, 23 out of 32 cities that have received any subway approval have obtained more than one subway approvals. For those nine cities only obtaining one subway approval, it is mainly because they got their first approval in a very recent year (e.g. 2015 or 2016) and it is possible that they will receive a second subway approval in few years after 2016.

In equation (1), we also include a vector of city-level time-variant control variables X_{it-1}. To avoid post-treatment bias, most control variables included in X_{it-1} are lagged by one year. θ_i and π_t are city and year fixed effect respectively. ϵ_{it} is the stochastic error term. We cluster standard errors at the city level to deal with the serial correlation of mayor promotion in the same city. The focus of the empirical analysis is β_1 which we expect to be positive.

The major challenge of identification is that mayors receiving subway approvals are fundamentally different from other mayors. This selection bias implies that mayors obtaining subway approvals are more likely to be promoted only because they are already more capable and politically connected. To rule out the influence of these selection biases, we have, first, included major sources of selection bias in control variables (X_{it-1}). This includes four groups of control variables. The first group is basic characteristics of this city: population

subway approval for a very short period of time. There are two such special cases: mayors of Dalian in 2009 and Kunming in 2013 were appointed as mayor in the same year as when the subway approval was made. Because their predecessors worked on subway approval for a much longer time than these two mayors, we identify their predecessors as the ones who got approval.

size, GDP size, government fiscal revenue, and economic growth rate. The second group is basic characteristics of mayors such as age, gender, race (i.e., Han or ethnic minority), and number of years in office.

The third group of control variables are measures for mayors' political connections with provincial leaders. We include four measures for connections with provincial party secretary (PPS) and one for connections with provincial governor (PG) since these two provincial leaders are the most powerful politicians at the provincial government who make the decision on political selection and promotion of city mayors. (1) *Birthplace connection with PPS* is a dichotomous variable which is equal to one if the mayor was born in the same prefecture as the PPS, and to zero if otherwise; (2) *alumni connection with PPS* is another dichotomous measure which is equal to one if the mayor went to the same college as the PPS during the same period of time, and to zero if otherwise; (3) *workplace connection with PPS* is a dummy variable which is coded as one if the mayor and the current PPS used to work in the same government unit, and to zero if otherwise; (4) *promotion connection with PPS* is a dummy variable which is equal to one if the mayor was appointed by the current PPS to the current position, and to zero if otherwise; and (5) *promotion connection with PG* is a dummy variable which is equal to one if the mayor was appointed by the current PG to the current position, and to zero if otherwise.[10]

Finally, we also control for a group of measures for mayors' competence and experience. (1) *Education* is a dummy variable which is equal to one if the mayor has obtained a bachelor's degree or above and to zero if otherwise; (2) *local experience* is the number of years the mayor has been working in this city including the years before he becomes the mayor (e.g., as a vice mayor); (3) *province experience* is a dummy variable which is equal to one if the mayor used to work in the provincial government, and to zero if otherwise; (4) *central government experience* is a dummy variable which is equal to one if the mayor used to work in the central

[10]We agree that these are not the *ideal* measures of political connection and competence, but they are, in our view, the *best* measures available.

government, and to zero if otherwise; (5) *SOE experience* is a dummy variable which is equal to one if the mayor used to work in a state-owned enterprise (SOE), and to zero if otherwise; (6) *university experience* is a dummy variable which is equal to one if the mayor used to work in a university, and to zero if otherwise; (7) *youth league experience* is a dummy variable which is equal to one if the mayor used to work in China Communist Youth League, and to zero if otherwise; and (8) *finance experience* is a dummy variable which is equal to one if the mayor used to work in a financial institution, and to zero if otherwise.

In addition to including these four groups of control variables, we also check the "parallel trends assumption" of a generalized DID design in Section 4.1. Because we cannot control for *all* sources of observed or unobserved selection bias, we need to ensure that cities who later obtain subway approvals are not more likely to get promoted even before the subway application is approved. This should alleviate our concern that subway approval has only picked up the difference among mayors before the PSS is approved. Similarly, we should *not* expect to see that after the mayor who has obtained the approval leaves the office, the promotion chances of this city's later mayors are higher due to the subway approval. This is because while these later mayors work in a city which has obtained subway approval, the credit goes to that earlier mayor who obtained the approval. By showing that before the subway approval and after the mayor obtaining subway approval leaves the office, city mayors' promotion chances are similar (i.e., "parallel trends assumption"), we should be less concerned about selection biases involved.

Finally, to further demonstrate the robustness of the results, we employ a fuzzy regression discontinuity design. Because the central government requires that only cities with more than 3 million population are eligible for subway approval, we utilize this feature and essentially compare cities with just above and below 3 million residents. By utilizing this "quasi-exogenous" variation in subway approval allocation, we show in section 4.2 that our results are still robust.

4 Does Subway Approval Lead to Promotion?

4.1 Baseline Results with a Generalized DID Design

We present the baseline results based on a generalized DID design (specified in equation (1)) in Table 2. We use mayor's political promotion in three years as the measure for the outcome variable, which is equal to one if the mayor is promoted within three years and to zero if otherwise. In column (1), we only include whether the city mayor obtained the approval of PSS from the central government as the explanatory variable together with city and year fixed effects. In column (2), we add several mayor characteristics into then regression, including mayor's age, gender, ethnicity, and number of years in the current position. Next, we further include the five measures for mayors' political connections in column (3) and mayors' education and experience in column (4) to exclude the alternative explanation that politically connected or competent mayors are more likely to obtain subway projects and are at the same time more likely to get promoted. Then, we include major city-level variables in column (5), such as city population, city GDP size, city government fiscal revenue, and city GDP growth rate, to control for the selection bias that more populous, prosperous, and fiscally robust cities are more likely to obtain the approval of PSS from the central government and at the same time mayors of such cities are most likely to get promoted. As shown in all these specifications, a mayor obtaining the approval of PSS is more likely to be promoted in three years than other mayors by 13.3-16.4 percentage points (which is over 40% increase from the baseline promotion rate of 36.8%).

Finally, in column (6), we exclude "vice-province level" cities (副省级城市). Mayors from these cities are "vice governor/minister" level appointees (副省级干部) whose promotion rules are different from mayors in other prefecture-level cities. Moreover, because the political promotion of mayors from "vice-province level" cities are ultimately determined by the central government, the *quid pro quo* relationship between these mayors "vice-province level" cities and provincial leaders should be weaker. After excluding these "vice-province

level" cities, we find that the effect of subway approval on the promotion chances of mayors become even more salient.

Table 2: Subway Approval and Mayors' Promotion

	Mayor promoted in three years					
	(1)	(2)	(3)	(4)	(5)	(6)
Subway approval	0.133**	0.142**	0.146**	0.157**	0.164**	0.262***
	(0.067)	(0.067)	(0.067)	(0.068)	(0.071)	(0.087)
City FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Mayor basic characteristics		✓	✓	✓	✓	✓
Mayor connections			✓	✓	✓	✓
Mayor competence				✓	✓	✓
City characteristics					✓	✓
Exclude vice-province cities						✓
Outcome variable mean	0.367	0.370	0.370	0.369	0.368	0.379
Observations	3843	3516	3509	3481	3230	3074

Notes: Standard errors clustered at the city level are reported in parentheses. Control variables: (1) mayor basic characteristics include mayor's age, gender, ethnicity, and number of years in office; (2) mayor connections include mayor's connections with PPS based on birthplace, alumni, workplace and promotion connections and mayor's promotion connection with PG; (3) mayor competence include mayor's education level and mayor's previous working experience in this city government, the provincial government, the central government, state-owned enterprises, university, the Communist Youth League, and financial sector; (4) city characteristics include city's population size, GDP size, GDP growth rate, and fiscal revenue in the previous year. The Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. FE = fixed effects.

We perform several additional tests to check the robustness of the results presented in Table 2. First, we test if our results are sensitive to alternative measures for outcome variable. In Appendix Table A1, we use mayor's promotion in one, two, four, and five years as outcome variables and find similar results. Next, we replace year fixed effects by province-year fixed effects to allow the heterogeneous effects of time in different provinces. We report these results in Appendix Table A2 and they are similar to those reported in Table 2. We also use alternative measures for the economic performance of the city to avoid the influence of extreme values or misreported economic data by city governments. We do

so by (1) taking the average of city-level variables in the previous three years, (2) replacing GDP growth by nighttime light intensity, and (3) decomposing GDP growth into the growth of different sectors. The results are contained in Appendix Table A3 and the coefficient of subway approval is still positive and significant. Furthermore, we conduct a placebo test in Appendix Table A4 where we show that subway approval does not affect the promotion of earlier mayors. As another placebo test, we examine in Appendix Table A5 whether subway approval increases the promotion chances of CPS. As discussed in Section 2.2, because CPS is not responsible for subway application or construction, we should not observe increased promotion chances for CPS due to subway approval. And indeed, in Appendix Table A5 we see that subway approvals are not correlated with the promotion chances of CPS. Finally, we report in Appendix Table A6 that the effect of subway approval on mayor's promotion does not depend on the location of cities or the political connections of city mayors. However, we do find that the effect of subway approval on promotion tends to be larger among smaller cities with fewer than 4 million residents.

One critical assumption of empirical tests employing a generalized DID design is that the treatment group and control group are similar before the former receives the treatment, also known as "parallel trends assumption." To test this assumption and to show the effect of subway approval on mayor's promotion more clearly, we employ a more flexible model specification below:

$$Promotion_{it} = \sum_{\gamma=-5}^{+5} \beta_\gamma Approval_{i(t+\gamma)} + \sum_{\delta=+1}^{+4} \alpha_\delta Leave_{i(t+\delta)} + \omega X_{it-1} + \theta_i + \pi_t + \epsilon_{it} \qquad (2)$$

where $Approval_{i(t+\gamma)}$ is a set of dummy variables indicating whether the city i has obtained the subway approval at time $t+\gamma$. Hence, γ is a more flexible approach to measure the effect of subway approval before the city has obtained the approval (i.e., $\gamma > 0$) and when

the city has obtained the subway approval for γ years (i.e., $\gamma < 0$).[11] To capture the effect of subway approval on mayor promotion after the mayor obtaining the subway approval leaves office, we introduce a series of dummy variables $Leave_{i(t+\delta)}$ to indicate that this is the δth year after the mayor who obtains the subway approval has left office. The parallel trends assumption requires that all β_γ (when $\gamma > 0$) and α_δ are not significantly different from zero. This means that before the city receives its subway approval and after the mayor who obtains the subway approval leaves the office, political promotion rate of city mayors are not significantly different between cities later obtaining subway approvals and other cities which have never received any subway approval. The positive and significant effect of subway approval on mayor's promotion will only appear during the period when the PSS is approved and the mayor who obtains the subway approval is still in office (i.e., $\beta_\gamma > 0$ when $\gamma < 0$).

Figure 2 contains the results of this test.[12] It shows that before the city obtains the subway approval (-1 to -5 on the X-axis), the promotion chances of mayors in cities that have later obtained the subway approval are not significantly different from that of mayors in other cities which have never received the subway approval. However, once the city receives subway approval, its mayor has a substantial and significant higher chance of getting promoted. This effect continues to hold as long as the mayor obtaining the subway approval is still in office. After this mayor (who obtains subway approval) leaves the office (i.e., +1 to +4 on X-axis), the effect of subway approval on mayor's promotion again drops to the level around zero, which shows that later mayors who have not worked on the subway approval would not enjoy a higher chance of promotion, even though this city has obtained subway approval thank to the efforts of an earlier mayor. Taken together, results contained in Figure 2 demonstrate that our generalized DID design is valid.

[11] For example, $Approval_{i(t-2)} = 1$ means year t is two years after the year when city i gets its first subway approval, and $Approval_{i(t+2)} = 1$ means year t is two year prior to that.

[12] We report regression results in Appendix Table A7.

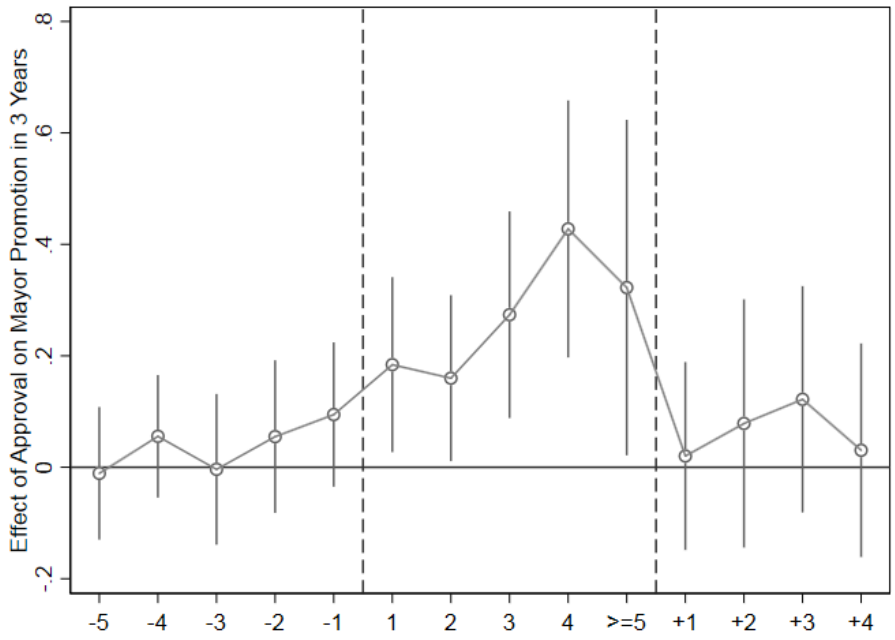

Figure 2: **Dynamic Effects of Subway Approvals on Mayor Promotion**. Each circle indicates a point estimate for the effect of subway approval and the vertical bars are the 90% confidence intervals. Negative numbers on the horizontal axis refer to years before the city receives subway approval. Numbers without positive or negative signs on the horizontal axis indicate the years when the mayor obtaining subway approval still remains in office. Positive numbers on the horizontal axis are the years after the mayor obtaining the subway approval leave office.

4.2 A Fuzzy Regression Discontinuity Design

To further limit the influences of various selection bias that we cannot control for in the generalized DID model shown in Table 2, we employ a fuzzy RDD as a robustness check in this subsection. In 2003, the central government of China required all cities with the intention to build the subway that they should satisfy the following four criteria to be eligible for building a subway: (1) the city annual fiscal revenue must exceed 10 billion Yuan; (2) the city GDP must reach 100 billion Yuan; (3) the number of permanent residents of city must be more than 3 million people; (4) the expected number of users of a subway line will be more than 30 thousand people per hour during rush hours. In other words, while cities satisfying these four requirements are not guaranteed to have subways, those not reaching them are not eligible to have subways. These requirements imposed by the central government allow us

to compare the cities just above and below these thresholds and identify a treatment effect that is induced by a quasi-exogenous variation.

We use the third requirement of the city population to construct the fuzzy RDD setup. Before proceed to our analysis, let us first explain why the other three requirements cannot be employed as a "discontinuity" in the fuzzy RDD. To begin, most obviously, the fourth requirement cannot be used because only cities applying for building subways will be required to calculate the expected number of users and those cities not applying for building subways will not have this statistic. Second, it is widely known that local governments in China have been manipulating key economic statistics such as city GDP and fiscal revenue (Wallace, 2016). Chances are that the manipulation of these economic statistics will result in non-random distribution of cities' GDP and fiscal revenue. Hence, cities' GDP and fiscal revenue may not be qualified running variables.

We conduct the McCrary test (McCrary, 2008) to examine if the density of distribution is smooth around three candidate running variables: GDP around 100 billion Yuan, fiscal revenue around 10 billion Yuan, and permanent residents around 3 million.[13] We report McCrary tests of all three candidate discontinuities in Table 3. The numerical result for McCrary test of population line around 3 million shows that there is no significant density jump around 3 million people. However, there is a very clear density jump of cities just below and above GDP of 100 billion Yuan (P-value is roughly 0). The case of fiscal revenue line of 10 billion Yuan is more complex. Although the McCrary test does not reject the null that there is no density jump around cities with roughly 10 billion Yuan of fiscal revenue, the statistical power of this rejection is very weak (p-value = 0.117). These results show that GDP and fiscal revenue are not appropriate running variables because there is clear sorting around the cutoffs.

[13]Because it will usually take roughly 1-2 years for the central government to approve a subway application, we lag three candidate running variables by two years to ensure that the city has reached the thresholds when it applies for the subway.

Table 3: McCrary Tests of Candidate Discontinuities

Candidate Discontinuity	McCarry Test (P-value)
Permanent resident = 3 million	-0.722 (0.471)
Fiscal revenue = 10 billion Yuan	-1.569 (0.117)
GDP = 100 billion Yuan	3.624 (0.000)

Another test for a valid "discontinuity" is that pre-treatment covariates should be balanced above and below the cutoff. Following Cattaneo *et al.* (2018), we test if pre-treatment covariates are balanced by identifying the "jump" around the cutoff, namely cities with 3 million residents. We apply this test to all control variables used in Table 2, which include mayor's basic characteristics, connections, and working experience and city's economic performance in the previous year. By definition, these variables are *pre*-treatment covariates because the values of these variables are realized before the city receives subway approval. These results are reported in Figure 3. From this figure, we see that most pre-treatment variables are balanced.[14]

In addition to sorting and imbalance, another concern for population RDD is compound treatment (Eggers *et al.* , 2018); that is, the same population cutoff could be used for different policies. This could be the case if, for instance, multiple social programs utilize the same population cutoff as criterion. We believe this should not be a major concern to our identification strategy. To our best knowledge, 3 million population is only the criterion for building subways in Chinese cities. In particular, 3 million population is not a cutoff for mayors' wage rate and cities' fiscal transfer size in China, both of which are popular sources

[14]The only exception is mayor's alumni connection with PPS, which presents a significant jump around the cutoff. As discussed in Lee & Lemieux (2010), sporadic "imbalances" could happen by chance if the researcher studies a large number of pre-treatment covariates. Moreover, the magnitude of imbalance in the case of mayor's alumni connection with PPS is not large.

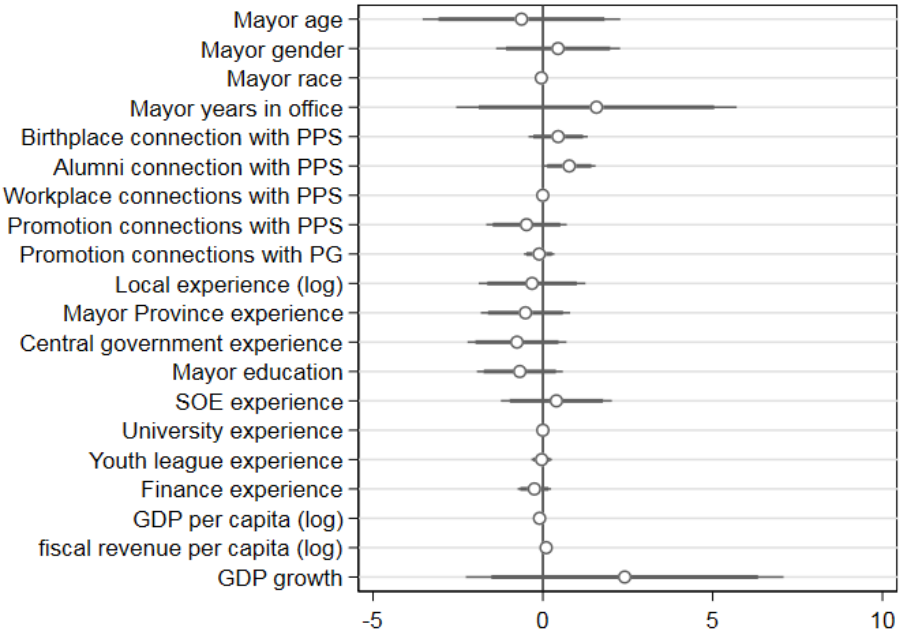

Figure 3: **Balance Test on Pre-treatment Covariates** Each circle is a point estimate and horizontal bars are the 90% and 95% confidence intervals.

of compound treatment in population RDDs in European countries (Eggers *et al.* , 2018).

With this fuzzy RDD setting based on the population requirement to build the subway, we are able to identify the effects of subway approval with a quasi-exogenous variation. We conduct the fuzzy RDD in an instrumental variable (IV) framework, i.e., a parametric two-stage least squares (2SLS) regressions. We assume that the promotion of city mayors take the following functional form:

$$Promotion_{it} = \beta_0 + \beta_1 \widehat{Approval}_{it} + \tau f(Z_{i,t-2}, Pop_{i,t-2}) + \gamma X_{it-1} + \theta_i + \pi_t + \epsilon_{it} \qquad (3)$$

where $Z_{i,t-2}$ is the IV, i.e., whether or not the city is home to more than 3 million residents; hence, $f(Z_{i,t-2}, Pop_{i,t-2})$ is a function of running variable (i.e., the population size of the city) and its interaction with $Z_{i,t-2}$. The first stage of the 2SLS model is specified by the

following equation:

$$Approval_{it} = \alpha_0 + \alpha_1 Z_{i,t-2} + \lambda f(Z_{i,t-2}, Pop_{i,t-2}) + \mu X_{it-1} + \theta_i + \pi_t + \epsilon_{it} \quad (4)$$

where the subway approval is predicted by the IV, i.e., whether the city had 3 million people two years ago. One potential problem of this parametric strategy of fuzzy RDD is that we have to assume a functional form of $f(Z_{i,t-2}, Pop_{i,t-2})$ and the results are dependent on this functional form assumption. To overcome this problem, we follow Gelman & Imbens (2018) and include first- and second-order polynomials of running variables and their interactions with IV to relax the assumption of functional form.

Moreover, we also include city fixed effect θ_i and year fixed effect π_t. Controlling for these fixed effects in a fuzzy RDD setting can protect our results against the potential bias that observations on either side of the 300 million population cutoff are correlated in unobserved ways (Jacob & Lefgren, 2004; Holbein & Hillygus, 2016; Dague *et al.* , 2017). In fact, because we use variations that some cities move across the cutoff over years (i.e. the population was below the cutoff years ago, but it went beyond the cutoff later), we need to include fixed effects to capture this variation.

Finally, we select the optimal bandwidth for fuzzy RDD by following Imbens & Kalyanaraman (2012). The optimal bandwidth is roughly 1 million people around the cutoff 3 million people. That is, we have mainly focused on city-year observations with a population ranging from 1.87 to 4.13 million. While the bandwidth may seem large, these are merely medium-sized Chinese cities (the median population size for a Chinese city is 3.6 million). Considering the standard deviation of city population size is 2.4 million, our choice of bandwidth (based on Imbens & Kalyanaraman (2012)) of 1.13 million people is not very large. To further alleviate the concern that results are sensitive to the choice of bandwidth, we also conduct analysis with both narrower and wider bandwidths than 1.13 million people as robustness check. The results are generally consistent with those reported in this section.

Finally, we only focus on cities with GDP larger than 100 billion Yuan and fiscal revenue larger than 10 billion Yuan, because these are the cities eligible to apply for the subway.

The results based on the aforementioned fuzzy RDD are reported in Table 4. First, we check if the results are susceptible to the weak instrument problem. We report the first-stage results in Panel B and the F Statistics of the first stage at the bottom of Table 4. The results of the first stage show that our instrumental variable is strongly associated with the subway approval and F statistics of all specifications are safely larger than 10, the conventional cutoff to identify a weak instrument (Staiger & Stock, 1997).

Table 4: Subway Approval and Mayors' Promotion: Fuzzy RDD

Panel A (second stage):	Mayor promoted in three years					
	(1)	(2)	(3)	(4)	(5)	(6)
Subway approval	0.868	0.620*	0.656**	0.770**	0.621**	0.541*
	(0.659)	(0.340)	(0.265)	(0.335)	(0.305)	(0.315)
Population	-0.012	-0.012*	-0.013*	-0.015	-0.013	-0.015
	(0.008)	(0.007)	(0.007)	(0.009)	(0.009)	(0.015)
Population × IV	0.022*	0.023*	0.024**	0.024	0.019	0.026
	(0.012)	(0.012)	(0.012)	(0.015)	(0.015)	(0.018)
Panel B (first stage):	Subway approval					
	(7)	(8)	(9)	(10)	(11)	(12)
IV (Population >3 million)	0.612*	0.739***	0.754***	0.879***	0.835***	0.820***
	(0.363)	(0.270)	(0.273)	(0.210)	(0.215)	(0.186)
City FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Mayor basic characteristics		✓	✓	✓	✓	✓
Mayor connections			✓	✓	✓	✓
Mayor competence				✓	✓	✓
City characteristics					✓	✓
Exclude vice-province cities						✓
Observations	162	159	159	157	157	147
Cragg-Donald F statistic	27.479	35.910	33.506	40.845	33.888	32.567

Notes: Bandwidth is 1.13 million people, which is selected based on the optimal bandwidth methods by Imbens & Kalyanaraman (2012). Standard errors clustered at the city level are reported in parentheses. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. FE = fixed effects. IV = instrumental variable (i.e., population >3 million).

The second-stage results are reported in Panel A. We add control variables in the same order as we did in Table 2. The results across all specifications in general show that subway approval from the central government increases the promotion chances of mayors obtaining subway approvals. One may also notice that the coefficients reported in Table 4 are much larger than those in Table 2 and wonder what may explain the gap between the different results in these two tables. We believe that this difference between generalized DID design and fuzzy RDD is mainly due to the different data samples utilized in these two statistical tests. Most notably, the sample for the fuzzy RDD is only medium-size cities with a population around 3 million, while the generalized DID design is based on all city-year observations. It is possible that the effect of subway approval is much larger among these cities around 3 million residents and the fuzzy RDD captures this local average treatment effect (Angrist & Pischke, 2008). To demonstrate this empirically, we estimate the heterogeneous effect of subway approval in cities with different population size in Appendix Table A6. We indeed find that subway approval is more helpful for mayor's promotion in smaller cities with fewer than 4 million residents than other larger cities.

We perform several robustness checks to our results presented in Table 4. First, we investigate whether the results are robust to alternative functional forms. In Appendix Table A8, we further add the quadratic term of running variable and its interaction with instrumental variable and obtain similar results. Furthermore, the coefficients of these two additional quadratic terms are mostly small and insignificant, showing that the influence of higher terms of running variable is not strong. Next, we perform another test showing that our fuzzy RDD results reported in Table 4 are robust to alternative outcome measures for mayor promotion, namely mayor promoted in one, two, four or five years, in Appendix Tables A9 to A12. Finally, we show in Figure 4 that results are robust to alternative bandwidth choices.

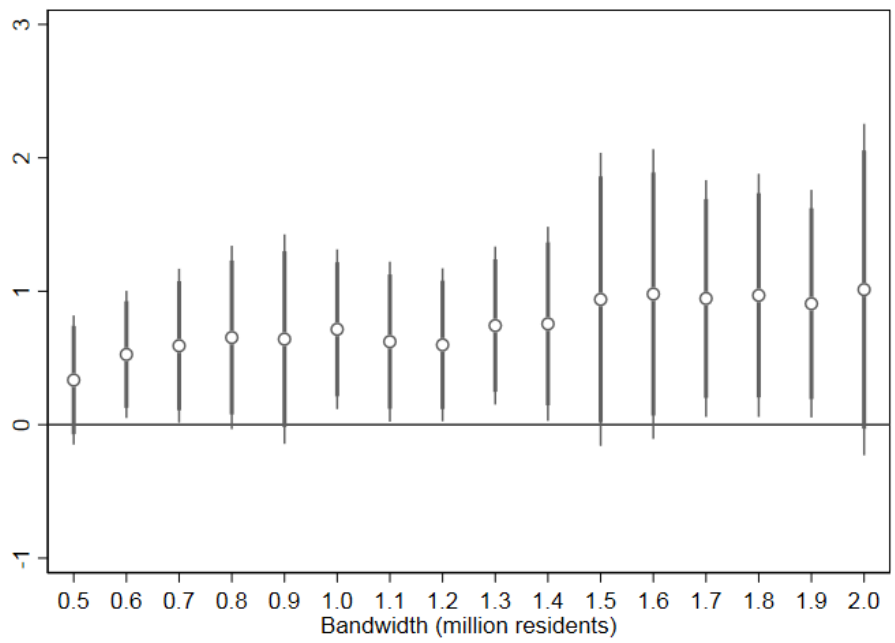

Figure 4: **Fuzzy Regression Discontinuity Design Results with Alternative Bandwidth Choices**. Each circle indicates a point estimate for the effect of subway approval on the promotion of city mayors in three years. Vertical bars are the 90% and 95% confidence intervals. All baseline control variables and two-way fixed effects are included.

5 Mechanism and Alternative Explanations

Results reported in the previous section show that subway approval enhances the promotion chances of Chinese mayors. In this section, we first present some suggestive evidence showing the mechanism behind this finding is that subway approval boosts investment, economic productivity, and fiscal revenue. Then we consider several potential rivalry explanations and demonstrate that they are unlikely to be the causal mechanism behind the findings reported in Section 4.

5.1 Investment and Economic Performance

The first plausible mechanism, as argued earlier, is that after obtaining the subway approval, the city can start making a massive amount of investment in infrastructure and can generate more fiscal revenue by selling land along with the (planned) subway lines. These investment

and fiscal revenue will last for a very long time, meaning that the subway approval can boost the economic performance of the city in the long run. Based on a *quid pro quo* relationship between provincial leaders and mayors, provincial leaders will reward mayors who have helped the province generate growth and fiscal revenue, because they help provincial officials demonstrate their economic performance to the central government.

To test this mechanism, we use the following model specification:

$$Y_{it} = \sum_{\sigma \geq -4, \sigma \neq +1}^{+5} \beta_\sigma Approval_{i(t+\sigma)} + \omega X_{it-1} + \theta_i + \pi_t + \epsilon_{it} \tag{5}$$

where $Approval_{it}$ is a dummy variable indicating whether city i at year t gets its first subway approval. Moreover, similar to equation (2), we add in total nine different dummy variables $Approval_{i(t+\sigma)}$ to check if subway approval's effect only starts to appear after the city has obtained the approval and before that the "parallel trends assumption" holds. We leave the year right before approval year (i.e. when $\sigma = +1$) as our baseline. Therefore, when "parallel trends assumption" holds, no difference between baseline year and other years prior to approval years should be observed, and difference should appear after approval year. Y_{it} in equation (1) is the mechanism outcome variable we will test.

First, the results for infrastructure investment and economic performance of the city are reported in Panel A and B of Figure 5. We use fixed asset investment in infrastructure per capita (Yuan) and GDP per capita (Yuan), respectively, as measures. In both cases, there is no significant difference between cities which later have obtained the subway approval and other cities before the year when the city has obtained its first subway approval (≤ -5 to -2 on the X-axis). Yet, once the city has obtained a subway approval (0 to ≥ 4 on the X-axis), its investment in infrastructure per capita sees a jump of roughly 500 Yuan (per capita) and its GDP per capita also becomes significantly larger than other cities. Both effects become even larger as time goes by, probably because more construction projects of subway lines start and there will be more investment then.

Next, we test whether the subway approval will lead to an increase of government fiscal

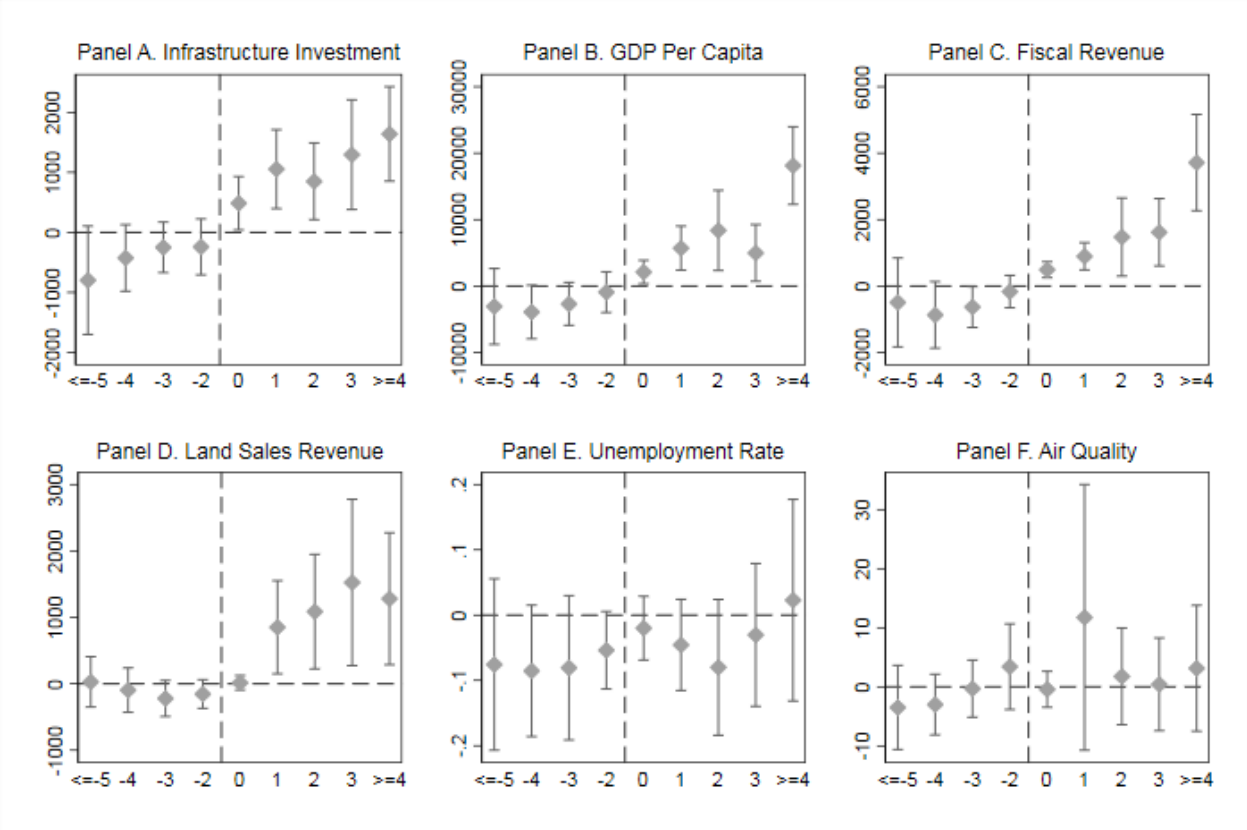

Figure 5: **Dynamic Effects of Subway Approvals on Economic Performance, Unemployment Rate and Air Quality**. Each circle indicates a point estimate. Vertical bars are 95% confidence intervals. The dummy variable indicating one-year prior treatment status is omitted from the regression, so that post-treatment and pre-treatment effects are relative to the period immediately prior to the treatment.

revenue. In Panel C of Figure 5, we report the dynamic effects of subway approval on total government fiscal revenue per capita. We see that the government fiscal revenue becomes larger in cities with subway approval after they obtain the subway approval; yet, before these cities obtain their first subway approval, their fiscal revenue size are similar to other cities, showing that cities which have never received subway approval are an appropriate control group.

Then we focus more closely on a specific type of fiscal revenue for city governments, namely the land sales revenue. Since all the land in China is owned by the state, only the government can sell the land.[15] The sales of land has recently become a major source of

[15]In fact, one cannot buy the property rights of land in China. She can only buy the using rights of land.

government revenue for many city governments in China, given that most tax revenue is taken by the central and provincial government. In Panel D of Figure 5, we test if subway approval increases the land sales revenue for a city government. We use the land sales revenue per capita (Yuan/person) as the measure for outcome variable. As shown in Panel D, since the second year after the city obtains its subway approval, the land sales revenue for this city government becomes significantly larger than other cities which have never received subway approvals. Similar to other Panels in Figure 5, we do not find this difference of land sales revenue before cities receiving subway approval, showing that the increase of land sales revenue is due the subway approval.

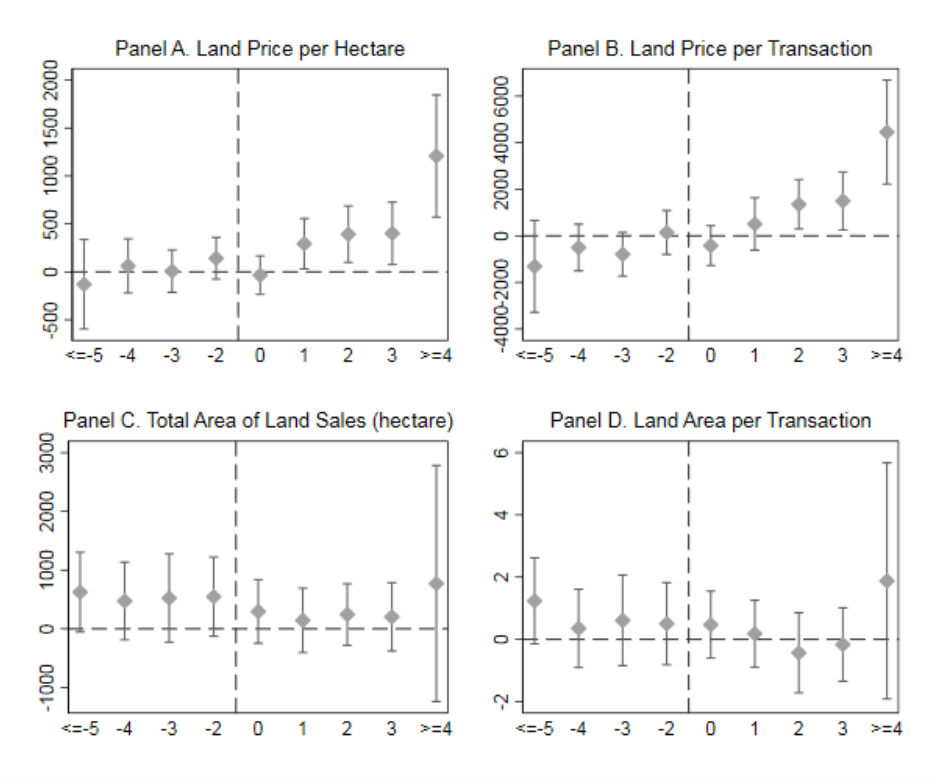

Figure 6: **Dynamic Effects of Subway Approvals on Land Price and Area for Government Land Sales**. Each circle indicates a point estimate. Vertical bars are 95% confidence intervals. The dummy variable indicating one-year prior treatment status is omitted from the regression, so that post-treatment and pre-treatment effects are relative to the period immediately prior to the treatment.

Only governments can grant and sell these using rights of land in China.

Given the importance of land sales revenue for local governments in China, we further differentiate two different channels for the higher land sales revenue in cities building subway lines. First, it is possible that land around subway lines will be sold at a *higher price*. Or alternatively, subway lines extend the urban periphery of the city so that the city can sell *a larger area* of land. To test these these two arguments, we further collect land sales data for city governments from the China Land and Resources Statistical Yearbook, which contains the land sales data for all prefecture-level cities from 2004 to 2016. To measure land price, we use (1) land price per hectare of land sold by the city government and (2) land price per transaction, which is the total land sales revenue divided by the number of land sales transactions made by the city government. In Panels A and B of Figure 6, we use these two measures for land price as outcome variables. In both panels, we see that land price increase after the city obtains the subway approval, showing that subway construction allows the city government to generate more land sales revenue by selling land at a higher price.

Next, we test if subway approval allows a city government to sell larger area of land. We use total area of land sales (hectare) and land area per transaction (i.e., total area of land sales divided by the number of land sales transactions made by the city government) as outcome variables. We show these results in Panels C and D of Figure 6. Contrary to the expectation, we have not found that cities which have obtained the subway approval sell more land than other cities. Hence, from the evidence presented in Figure 6, we conclude that subway construction boosts land sales revenue for a city government because it allows the city government to sell land at a higher price.

Taken together, these results demonstrate that mayors obtaining the subway approval help both the city and province generate more investment and fiscal revenue. More importantly, as the benefits of investment, economic growth, and fiscal revenue continue to grow in the long run, mayors who have obtained the subway approval should enjoy a higher chance of receiving political promotion.

5.2 Welfare Improvement

A competing explanation is that subway approval can bring in some welfare improvement to city residents which explains the higher promotion chances of city mayors who have obtained the subway approval. While this explanation is unlikely to hold because subway lines are most likely to bring in welfare improvement after they start operating, we consider two potential welfare improvements before the operation of subway. One possibility is that cities may need to hire more workers to design and build subway lines and new business will develop along the planned subway lines. Hence, the city will have a lower unemployment rate. The other possibility is that once residents learn that their city is going to build a subway, they are going to delay their plans to buy cars and will use public transportation more. This may lead to better air quality. We test these two conjectures in Panels E and F of Figure 5. However, we have not found evidence to support these two conjectures. Cities obtaining subway approval are not fundamentally different from other cities in terms of unemployment rate and air quality either before or after these cities receive subway approvals.

Furthermore, one may argue that while subway approvals do not produce short-term welfare improvement for residents (i.e., before the operation), approvals pave the pay for the long-term welfare improvement for residents once the subway lines start to operate. Hence, these mayors are rewarded with political promotion for these long-term welfare improvements. Following this explanation, one can think of the promotion as the current value of future welfare improvement once the subway starts to operate. Although we are unable to entirely exclude this alternative explanation due to data limitation, we believe this explanation is unlikely to hold. One implication of this view is that among all city mayors who have obtained subway approvals, those who are promoted must have produced more welfare improvement for their residents once the subway lines start to operate than those who are not promoted. To test this implication, we collect the subway ridership data from the China Association of Metros, which organization has released detailed subway ridership data since 2013. We use the number of subway riders to measure the future welfare improvement be-

Table 5: Mayor Promotion and Subway Ridership Intensity

	Promoted mayor	Not promoted mayor	Mean difference
Subway ridership intensity	5.22	6.69	-1.47
	(2.62)	(4.58)	(1.79)
Subway ridership intensity >7	0.38	0.46	-0.09
	(0.52)	(0.52)	(0.23)

Notes: Unit for subway ridership intensity is thousand people per day and kilometer. Standard errors are reported in parentheses. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

cause taking the subway (and any consequences of this action) is the most direct way of welfare improvement that subway lines can bring to city residents.

More specifically, we measure the subway ridership by "ridership intensity," defined as the number of subway riders per day and kilometer (thousand people / day and kilometer). The central government has required that subway lines should obtain the ridership intensity of 7,000 people/day and kilometer even for subway lines that just start to operate. Hence, we test if cities with mayors who have obtained subway approval and are promoted have larger ridership intensity and are more likely to reach the threshold set by the central government.

We show these results in Table 5. Because the China Association of Metros only reports rider statistics for cities that have already operated a subway system, we only have ridership measures for 22 cities.[16] With this small sample of cities, we only report a T test of the difference of ridership intensity between promoted mayors and non-promoted mayors. Yet, we have not found that promoted mayors' cities will have more subway riders or are more likely to pass the central government's requirement, if not fewer riders or less likely to reach 7,000 riders per day and kilometer. These results demonstrate that the promotion of mayors obtaining subway approval may not be due to the potential welfare improvement brought by the subway in the long run.

[16]We exclude major cities such as Beijing, Shanghai, Guangzhou, and Shenzhen for the reasons we have mentioned in Section 3.

5.3 Strategic Appointment of Mayors

Another interpretation of the findings is that provincial leaders appoint politically connected mayors to cities which will obtain subway approval soon. Because provincial leaders may have friends at the central government and can help them know which cities will obtain subway approvals in advance, they can replace city mayors with their loyal subordinates just before the central government announcing the subway approval, letting their loyal subordinates to take the credit of obtaining subway approval. If this is the case, the results are driven by political loyalty and internal information rather than a political exchange relationship between mayors and provincial leaders

Yet, we believe that this argument cannot explain our findings. First, we have controlled for four different measures of mayors' political connections with PPS and promotion connection with PG. With these control variables in both generalized DID design and fuzzy RDD, we argue that political connections with provincial leaders are unlikely to be the main explanation for our results.

Second, to further alleviate the concern for the strategic appointment of politically connected mayors just before the announcement of subway approvals, we test if city mayors are more likely to be replaced in one, two, and three year(s) before the city is granted with a subway approval. The results are presented in Table 6. The outcome variable is a dummy variable, which is equal to one if the mayor is replaced in a given city and year and is zero otherwise. The three explanatory variables are indicators of one, two, or three years before the city obtains a subway approval. From Table 6, we see that the turnover rate of mayors is not significantly different in years shortly before the city obtains subway approval than other time. This demonstrates that there is no systematic pattern that provincial leaders replace city mayors shortly before a city obtains subway approval.

Table 6: Is There Strategic Appointment of Mayors before Subway Approvals?

	Turnover of City Mayor		
	(1)	(2)	(3)
1 year before subway approval	0.023		
	(0.055)		
2 years before subway approval		-0.047	
		(0.055)	
3 years before subway approval			0.092
			(0.066)
City FE	✓	✓	✓
Year FE	✓	✓	✓
Outcome variable mean	0.285	0.285	0.285
Observations	3592	3591	3590

Notes: Standard errors clustered at the city level are reported in parentheses. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. FE = fixed effects.

5.4 Corruption

Finally, we consider if the results are driven by a corruption mechanism that subway construction breeds rent-seeking opportunities and mayors may trade money for promotion. To test if this mechanism can explain our results, we need a measure for the corruption level of mayors. While to our knowledge there is not a widely accepted measure for corruption of Chinese mayors, we build a proxy for corrupt mayors by the following procedures. First, we identify all mayors who were investigated due to corruption from two data sources: (1) the CCER Official Database which covers the corruption investigation against mayors from 2012 to 2016, and (2) the Chinese Political Elite Database (CPED) from Jiang (2018), which contains anti-corruption investigations of city mayors from 2000 to 2015. By merging these two datasets, we compile a full record of all anti-corruption investigations of city mayors from 2003 to 2016.[17] With this new dataset, we identify a mayor as a "corrupt mayor" if he has ever been investigated for corruption until 2016.

[17]We drop years 2000 to 2002 because our empirical analysis uses subway data from 2003 to 2016.

Since our goal is only to check whether the corruption mechanism invalidates the political exchange mechanism, we only test if our main results are robust to the inclusion of this measure for corrupt mayors. We present the results in Table 7. In column (1), we only include subway approval and corrupt mayor as explanatory variables together with city and year fixed effects. Then in column (2) we further add mayor and city control variables which we have used in Table 2. In both columns, we see that the coefficients of subway approval are similar to those reported in Table 2, while we also see that corrupt mayors enjoy a slightly higher chance of promotion. One major problem of using corruption investigation as the measure for corruption is that corruption investigations are controlled by provincial government and the provincial leaders may choose not to investigate loyal mayors who are politically connected with the provincial leadership. To alleviate the concern of selective investigation, we present in column (3) the results based on city-year observations after 2012 when the central government centralized the power of anti-corruption investigations. In other words, provincial leaders had fewer opportunities to influence the investigations against corrupt mayors after 2012. Again, we find that subway approval is still robust in this smaller sample of years after 2012.

6 Conclusion

Why do city mayors have incentives to make investment in transportation infrastructure such as subway lines? While economists have offered plenty of evidence that subway lines bring in many benefits to city residents, it is unclear why and how city mayors could benefit from them. Following earlier literature treating infrastructure projects as distributive goods, this paper argues that the electoral motivation for infrastructure project distribution can also be applied to settings where city mayors are not elected. We study the incentives of Chinese mayors to build a subway system in their cities. Instead of delivering porks for voters, city mayors in China still have strong incentives to build subway lines because by doing so, they

Table 7: Testing Corruption Mechanism

	Mayor promoted in three years		
	(1)	(2)	(3)
Subway approval	0.133**	0.164**	0.153*
	(0.067)	(0.071)	(0.082)
Corrupt mayor	0.059	0.049	-0.025
	(0.048)	(0.052)	(0.130)
City FE	✓	✓	✓
Year FE	✓	✓	✓
City and mayor controls		✓	✓
Years after 2012 only			✓
Outcome variable mean	0.367	0.367	0.288
Observations	3843	3230	1270

Notes: Standard errors clustered at the city level are reported in parentheses. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. FE = fixed effects.

can deliver investment, economic performance, and fiscal revenue for their political patrons, that is provincial leaders.

With an original panel data of Chinese cities from 2003 to 2016, we have shown that mayors obtaining the approval from the central government to build subway have an additional 16.4 percentage points larger chance of getting promoted based on the whole sample. By focusing only on cities with a population around 3 million and employing a fuzzy RDD, we have found that mayors with subway approvals can have an additional 62.1 percentage points of chance to get promoted for these medium-size cities. The major explanations for these promotion bonuses are the increase of investment, economic performance, and fiscal revenue. However, subway construction does not create more jobs or improve air quality for city residents in the short term.

These results explain why the Chinese mayors have the incentives to work on subway construction even if they are unlikely to the ones to take the credit of completing the construction. We show that the Chinese political system rewards the mayor who has obtained the subway approval even without starting to build any subway line. And this task of ob-

taining the subway approval can be accomplished within one mayor term. However, there are potential problems of this incentive structure. While we have not tested in this paper, city mayors may apply for subway without clearly understanding whether the city really needs a subway line if the approval alone can win mayors a promotion bonus. Meanwhile, subway construction may incur a huge cost which may become huge burden for future city governments. This can lead to massive amounts of city debts which the future mayors and city residents need to pay for. The mayor applying for subway in the first place may walk away with a promotion and leave the heavy city debts to his successors. Hence, we expect future research can further explore fiscal and welfare implications for politically motivated investment in urban infrastructure.

References

Angrist, Joshua D, & Pischke, Jörn-Steffen. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Banerjee, Abhijit, Duflo, Esther, & Qian, Nancy. 2012. On the Road: Access to Transportation Infrastructure and Economic Growth in China. *NBER Working Paper No. 17897*. Available at https://www.nber.org/papers/w17897.

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

Bell, Daniel A. 2016. *The China Model: Political Meritocracy and the Limits of Democracy*. Princeton University Press.

Bueno de Mesquita, Bruce, Smith, Alastair, Morrow, James D, & Siverson, Randolph M. 2003. *The Logic of Political Survival*. MIT press.

Cadot, Olivier, Röller, Lars-Hendrik, & Stephan, Andreas. 2006. Contribution to Productivity or Pork Barrel? The Two Faces of Infrastructure Investment. *Journal of Public Economics*, **90**(6-7), 1133–1153.

Castells, Antoni, & Solé-Ollé, Albert. 2005. The Regional Allocation of Infrastructure Investment: The Role of Equity, Efficiency and Political Factors. *European Economic Review*, **49**(5), 1165–1205.

Cattaneo, Matias D, Idrobo, Nicolás, & Titiunik, Rocío. 2018. *A Practical Introduction to Regression Discontinuity Designs: Volume I*.

Dague, Laura, DeLeire, Thomas, & Leininger, Lindsey. 2017. The Effect of Public Insurance Coverage for Childless Adults on Labor Supply. *American Economic Journal: Economic Policy*, **9**(2), 124–54.

Donaldson, Dave. 2018. Railroads of the Raj: Estimating the Impact of Transportation Infrastructure. *American Economic Review*, **108**(4-5), 899–934.

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

Eggers, Andrew C, Freier, Ronny, Grembi, Veronica, & Nannicini, Tommaso. 2018. Regression Discontinuity Designs Based on Population Thresholds: Pitfalls and Solutions. *American Journal of Political Science*, **62**(1), 210–229.

Gelman, Andrew, & Imbens, Guido. 2018. Why High-Order Polynomials Should not be Used in Regression Discontinuity Designs. *Journal of Business & Economic Statistics*, 1–10.

Gendron-Carrier, Nicolas, Gonzalez-Navarro, Marco, Polloni, Stefano, & Turner, Matthew A. 2018. Subways and Urban Air Pollution. *NBER Working Paper No. 24183*. https://www.nber.org/papers/w24183.

Holbein, John B, & Hillygus, D Sunshine. 2016. Making Young Voters: The Impact of Preregistration on Youth Turnout. *American Journal of Political Science*, **60**(2), 364–382.

Hong, Ji Yeon, & Park, Sunkyoung. 2016. Factories for Votes? How Authoritarian Leaders Gain Popular Support Using Targeted Industrial Policy. *British Journal of Political Science*, **46**(3), 501–527.

Imbens, Guido, & Kalyanaraman, Karthik. 2012. Optimal Bandwidth Choice for the Regression Discontinuity Estimator. *The Review of Economic Studies*, **79**(3), 933–959.

Jacob, Brian A, & Lefgren, Lars. 2004. Remedial Education and Student Achievement: A Regression-Discontinuity Analysis. *Review of Economics and Statistics*, **86**(1), 226–244.

Jia, Ruixue, Kudamatsu, Masayuki, & Seim, David. 2015. Political Selection in China: The Complementary Roles of Connections and Performance. *Journal of the European Economic Association*, **13**(4), 631–668.

Jiang, Junyan. 2018. Making Bureaucracy Work: Patronage Networks, Performance Incentives, and Economic Development in China. *American Journal of Political Science*, **62**(4), 982–999.

Keefer, Philip, & Knack, Stephen. 2007. Boondoggles, Rent-seeking, and Political Checks and Balances: Public Investment under Unaccountable Governments. *The Review of Economics and Statistics*, **89**(3), 566–572.

Landry, Pierre F., Lü, Xiaobo, & Duan, Haiyan. 2018. Does Performance Matter? Evalu-

ating Political Selection Along the Chinese Administrative Ladder. *Comparative Political Studies*, **51**(8), 1074–1105.

Landry, Pierre Francois. 2008. *Decentralized Authoritarianism in China: The Communist Party's Control of Local Elites in the Post-Mao Era*. Vol. 1. Cambridge University Press New York.

Lee, David S, & Lemieux, Thomas. 2010. Regression Discontinuity Designs in Economics. *Journal of economic literature*, **48**(2), 281–355.

Lehne, Jonathan, Shapiro, Jacob N, & Eynde, Oliver Vanden. 2018. Building Connections: Political Corruption and Road Construction in India. *Journal of Development Economics*, **131**, 62–78.

Li, Hongbin, & Zhou, Li-An. 2005. Political Turnover and Economic Performance: the Incentive Role of Personnel Control in China. *Journal of Public Economics*, **89**(9-10), 1743–1762.

Li, Hongbin, Loyalka, Prashant, Rozelle, Scott, & Wu, Binzhen. 2017. Human Capital and China's Future Growth. *Journal of Economic Perspectives*, **31**(1), 25–48.

Lin, Justin Yifu. 2011. *Demystifying the Chinese Economy*. Cambridge University Press.

McCrary, Justin. 2008. Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test. *Journal of Econometrics*, **142**(2), 698–714.

Ouyang, Min, & Peng, Yulei. 2015. The Treatment-Effect Estimation: A Case Study of the 2008 Economic Stimulus Package of China. *Journal of Econometrics*, **188**(2), 545–557.

Przeworski, Adam, Alvarez, Michael E, Cheibub, Jose Antonio, & Limongi, Fernando. 2000. *Democracy and Development: Political Institutions and Well-being in the World, 1950-1990*. Vol. 3. Cambridge University Press.

Robinson, James A, & Torvik, Ragnar. 2005. White Elephants. *Journal of Public Economics*, **89**(2-3), 197–210.

Shi, Hao, & Huang, Shaoqing. 2014. How Much Infrastructure is Too Much? A New Approach and Evidence from China. *World Development*, **56**, 272–286.

Staiger, Douglas, & Stock, James H. 1997. Instrumental Variables Regression with Weak Instruments. *Econometrica*, **65**(3), 557.

Voigtländer, Nico, & Voth, Hans-Joachim. 2014. Highway to Hitler. *NBER Working Paper*. Available at https://doi.org/10.3386/w20150.

Wallace, Jeremy L. 2016. Juking the Stats? Authoritarian Information Problems in China. *British Journal of Political Science*, **46**(1), 11–29.

Xi, Tianyang, Yao, Yang, & Zhang, Muyang. 2018. Capability and Opportunism: Evidence from City Officials in China. *Journal of Comparative Economics*.

Xu, Chenggang. 2011. The Fundamental Institutions of China's Reforms and Development. *Journal of Economic Literature*, **49**(4), 1076–1151.

Yao, Yang, & Zhang, Muyang. 2015. Subnational Leaders and Economic Growth: Evidence from Chinese Cities. *Journal of Economic Growth*, **20**(4), 405–436.

Zheng, Siqi, & Kahn, Matthew E. 2013a. China's Bullet Trains Facilitate Market Integration and Mitigate the Cost of Megacity Growth. *Proceedings of the National Academy of Sciences*, **110**(14), E1248–E1253.

Zheng, Siqi, & Kahn, Matthew E. 2013b. Understanding China's Urban Pollution Dynamics. *Journal of Economic Literature*, **51**(3), 731–72.

Online Appendix for
Political Career Incentives and Infrastructure Investment: Evidence from China

Contents of Online Appendix

Table A1: Subway Approval and Mayors' Promotion with Flexible Outcome Measures

	Mayor promoted in X years (X=1, 2, 4 or 5)					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: promotion in 1 year						
Subway approval	0.077**	0.077**	0.079**	0.085**	0.095**	0.133***
	(0.035)	(0.035)	(0.036)	(0.035)	(0.038)	(0.047)
Outcome variable mean	0.144	0.148	0.147	0.147	0.147	0.152
Observations	3851	3516	3509	3481	3230	3074
Panel B: promotion in 2 years						
Subway approval	0.115**	0.122**	0.125**	0.134**	0.144**	0.221***
	(0.054)	(0.053)	(0.053)	(0.054)	(0.058)	(0.065)
Outcome variable mean	0.262	0.266	0.266	0.266	0.265	0.273
Observations	3846	3516	3509	3481	3230	3074
Panel C: promotion in 4 years						
Subway approval	0.130*	0.131*	0.130*	0.137*	0.140*	0.240***
	(0.071)	(0.071)	(0.070)	(0.072)	(0.074)	(0.092)
Outcome variable mean	0.432	0.435	0.435	0.435	0.433	0.446
Observations	3841	3516	3509	3481	3230	3074
Panel D: promotion in 5 years						
Subway approval	0.130*	0.124*	0.119	0.126*	0.133*	0.222**
	(0.075)	(0.074)	(0.073)	(0.075)	(0.078)	(0.096)
Outcome variable mean	0.470	0.470	0.470	0.470	0.469	0.483
Observations	3841	3516	3509	3481	3230	3074
City FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Mayor basic characteristics		✓	✓	✓	✓	✓
Mayor connections			✓	✓	✓	✓
Mayor competence				✓	✓	✓
City characteristics					✓	✓
Exclude vice-province cities						✓

Notes: This table shows that the main results are robust to alternative outcome measures. To do so, we use mayor's promotion in one, two, four or five years as outcome variable and obtain similar results as those reported in Table 2. Control variables: (1) mayor basic characteristics include mayor's age, gender, ethnicity, and number of years in office; (2) mayor connections include mayor's connections with PPS based on birthplace, alumni, workplace and promotion connections and mayor's promotion connection with PG; (3) mayor competence include mayor's education level and mayor's previous working experience in this city government, the provincial government, the central government, state-owned enterprises, university, the Communist Youth League, and financial sector; (4) city characteristics include city's population size, GDP size, GDP growth rate, and fiscal revenue in the previous year. Standard errors clustered at the city level are reported in parentheses. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. FE = fixed effects.

Table A2: Subway Approval and Mayor's Promotion with Province-Year Fixed Effects

	Mayor promoted in three years					
	(1)	(2)	(3)	(4)	(5)	(6)
Subway approval	0.107*	0.110*	0.110*	0.119*	0.118*	0.200**
	(0.062)	(0.061)	(0.061)	(0.064)	(0.069)	(0.084)
City FE	✓	✓	✓	✓	✓	✓
Province-year FE	✓	✓	✓	✓	✓	✓
Mayor basic characteristics		✓	✓	✓	✓	✓
Mayor connections			✓	✓	✓	✓
Mayor competence				✓	✓	✓
City characteristics					✓	✓
Exclude vice-province cities						✓
Outcome variable mean	0.367	0.370	0.370	0.369	0.368	0.379
Observations	3843	3516	3509	3481	3230	3074

Notes: This table shows that the main results are robust to the inclusion of province-year fixed effects. Province-year fixed effects capture the differentiated impact of year dummies in different provinces. To do so, we replace year fixed effects by province-year fixed effects and obtain similar results as those reported in Table 2. Control variables: (1) mayor basic characteristics include mayor's age, gender, ethnicity, and number of years in office; (2) mayor connections include mayor's connections with PPS based on birthplace, alumni, workplace and promotion connections and mayor's promotion connection with PG; (3) mayor competence include mayor's education level and mayor's previous working experience in this city government, the provincial government, the central government, state-owned enterprises, university, the Communist Youth League, and financial sector; (4) city characteristics include city's population size, GDP size, GDP growth rate, and fiscal revenue in the previous year. Standard errors clustered at the city level are reported in parentheses. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. FE = fixed effects.

Table A3: Subway Approval and Mayor's Promotion with Alternative Measures for Cities' Economic Performance

	Mayor promoted in three years					
	(1)	(2)	(3)	(4)	(5)	(6)
Subway approval	0.166**	0.245**	0.252**	0.360***	0.264***	0.352***
	(0.080)	(0.096)	(0.102)	(0.096)	(0.099)	(0.103)
Population (3-year mean)	0.251	0.301	0.161	0.260	0.282	0.378
	(0.215)	(0.224)	(0.318)	(0.333)	(0.269)	(0.286)
GDP (3-year mean)	0.056	0.055	0.042	0.039	-0.010	-0.014
	(0.094)	(0.096)	(0.133)	(0.135)	(0.115)	(0.117)
Fiscal revenue (3-year mean)	-0.097	-0.095	-0.183	-0.181	-0.197	-0.194
	(0.141)	(0.144)	(0.176)	(0.178)	(0.165)	(0.168)
GDP growth (3-year mean)	0.004	0.004				
	(0.005)	(0.005)				
Nighttime light intensity			-0.082	-0.096		
			(0.107)	(0.108)		
GDP growth: agriculture (3-year mean)					-0.014*	-0.015*
					(0.008)	(0.008)
GDP growth: industry (3-year mean)					0.001	0.001
					(0.005)	(0.005)
GDP growth: service (3-year mean)					0.011	0.012
					(0.008)	(0.008)
City FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Mayor basic characteristics	✓	✓	✓	✓	✓	✓
Mayor connections	✓	✓	✓	✓	✓	✓
Mayor competence	✓	✓	✓	✓	✓	✓
Exclude vice-province cities		✓		✓		✓
Observations	2719	2587	1975	1879	2231	2123

Notes: This table shows that the main results are robust to the inclusion of alternative measures for cities' economic performance. In columns (1) and (2), we use the average of cities' population size, GDP size, GDP growth rate, and fiscal revenue in the previous three years as control variables. The average of city-level variables in the previous three years is less sensitive to the influence of extreme values. Next, in columns (3) and (4) we replace GDP growth rate by nighttime light intensity, which is measured by Log average nighttime brightness per square km. The data come from Jiang (2018), where the author merges the nighttime light intensity data from National Oceanic and Atmospheric Administration of the United States with the panel data for Chinese cities. Because many Chinese cities manipulate their economic data, we use nighttime light intensity to have a less biased measure for cities' economic performance. Finally, in columns (5) and (6), we decompose economic growth into the growth in three sectors and include them as control variables. Across all these specifications, we obtain similar results as those reported in Table 2. Control variables: (1) mayor basic characteristics include mayor's age, gender, ethnicity, and number of years in office; (2) mayor connections include mayor's connections with PPS based on birthplace, alumni, workplace and promotion connections and mayor's promotion connection with PG; (3) mayor competence include mayor's education level and mayor's previous working experience in this city government, the provincial government, the central government, state-owned enterprises, university, the Communist Youth League, and financial sector. Standard errors clustered at the city level are reported in parentheses. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. FE = fixed effects.

Table A4: Placebo Test: Promotion before the Subway Approval

	Mayor's promotion in X years before approval				
	(1) 1 yr	(2) 2 yrs	(3) 3 yrs	(4) 4 yrs	(5) 5 yrs
Subway approval	0.048	-0.032	-0.116*	0.020	0.035
	(0.046)	(0.043)	(0.070)	(0.061)	(0.068)
City FE	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
Outcome variable mean	0.373	0.389	0.403	0.412	0.412
Observations	3565	3287	3008	2729	2450

Notes: This table shows that mayors obtaining the approval of PSS are not associated with a higher chances of political promotion for mayors before the subway approval. Standard errors clustered at the province and year level are reported in parentheses. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. FE = fixed effects.

Table A5: Placebo Test: Subway Approval and the Promotion of City Party Secretary

	City party secretary promoted in three years					
	(1)	(2)	(3)	(4)	(5)	(6)
Subway approval	0.013	0.013	0.014	0.023	0.028	-0.016
	(0.039)	(0.041)	(0.041)	(0.043)	(0.047)	(0.036)
City FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
CPS basic characteristics		✓	✓	✓	✓	✓
CPS connections			✓	✓	✓	✓
CPS competence				✓	✓	✓
City characteristics					✓	✓
Exclude vice-province cities						✓
Observations	3832	3494	3491	3455	3213	3061

Notes: This table shows that subway approval does not increase the promotion chance for city party secretary (CPS). This result is consistent with our theory because CPS is not responsible for applying for subways in China. When a city applies for subway approvals, the city government usually establishes a new government agency called "the leading task group for subway planning and construction" (轨道交通规划与建设领导小组/指挥部) and city mayor is *always* the head of this organization based on our limited knowledge. Hence, it is safe to assume that subways are pork provided by city mayors to provincial leaders in most cases, and so, we should not observe increased promotion chances for CPS due to subway approvals. Control variables: (1) CPS basic characteristics include CPS's age, gender, ethnicity, and number of years in office; (2) CPS connections include CPS's connections with PPS based on birthplace, alumni, workplace and promotion connections and CPS's promotion connection with PG; (3) CPS competence include CPS's education level and CPS's previous working experience in this city government, the provincial government, the central government, state-owned enterprises, university, the Communist Youth League, and financial sector; (4) city characteristics include city's population size, GDP size, GDP growth rate, and fiscal revenue in the previous year. Standard errors clustered at the city level are reported in parentheses. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. FE = fixed effects. CPS = city party secretary.

Table A6: Heterogeneous Effects

	Mayor promoted in three years							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Subway plan	0.164**	0.186**	0.189*	0.160**	0.169**	0.146*	0.166**	0.362**
	(0.076)	(0.085)	(0.103)	(0.073)	(0.074)	(0.086)	(0.079)	(0.165)
Subway plan × western province	-0.001							
	(0.182)							
Subway plan × mid province		-0.099						
		(0.144)						
Subway plan × eastern province			-0.056					
			(0.140)					
Subway plan × alumni connection (PPS)				-0.141				
				(0.200)				
Subway plan × workplace connection (PPS)					-0.276			
					(0.192)			
Subway plan × promotion connection (PPS)						-0.037		
						(0.095)		
Subway plan × promotion connection (PG)							-0.037	
							(0.138)	
Subway plan × 𝟙{4 mi <pop<6 mi}								-0.099
								(0.222)
Subway plan × 𝟙{6 mi <pop<8 mi}								-0.207
								(0.184)
Subway plan × 𝟙{8 mi <pop}								-0.291
								(0.205)
City FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Controls used in column (5) of Table 2	✓	✓	✓	✓	✓	✓	✓	✓
Observations	3230	3230	3230	3235	3235	3232	2990	3230

Notes: Standard errors clustered at the city level are reported in parentheses. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. FE = fixed effects. mi = million.

Table A7: Test Parallel Trends Assumption

	Mayor promoted in three years	
	(1)	(2)
To be approved in 5 years	-0.075	-0.011
	(0.074)	(0.072)
To be approved in 4 years	0.032	0.056
	(0.074)	(0.067)
To be approved in 3 years	0.003	-0.004
	(0.080)	(0.082)
To be approved in 2 years	0.059	0.055
	(0.080)	(0.083)
To be approved in 1 year	0.094	0.095
	(0.075)	(0.078)
Approved for 1 year	0.154*	0.184*
	(0.092)	(0.095)
Approved for 2 years	0.108	0.160*
	(0.092)	(0.090)
Approved for 3 years	0.225**	0.274**
	(0.111)	(0.112)
Approved for 4 years	0.375***	0.428***
	(0.122)	(0.140)
Approved for 5 or more years	0.300**	0.322*
	(0.152)	(0.182)
Mayor left for 1 year	0.086	0.020
	(0.097)	(0.102)
Mayor left for 2 years	0.065	0.079
	(0.125)	(0.135)
Mayor left for 3 years	0.113	0.122
	(0.114)	(0.123)
Mayor left for 4 years	0.032	0.031
	(0.099)	(0.116)
City FE	✓	✓
Year FE	✓	✓
Baseline controls		✓
Observations	3843	3230

Notes: This table reports the regression results for Figure 2. Standard errors clustered at the city level are reported in parentheses. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. FE = fixed effects.

Table A8: Fuzzy Regression Discontinuity Design: Check Functional Form

Panel A (second stage):	Mayor promoted in three years					
	(1)	(2)	(3)	(4)	(5)	(6)
Subway approval	0.840	0.696*	0.782**	0.999**	0.930**	1.096**
	(0.593)	(0.358)	(0.330)	(0.408)	(0.395)	(0.510)
Population	-0.027**	-0.034***	-0.042***	-0.057***	-0.055**	-0.079***
	(0.012)	(0.011)	(0.011)	(0.022)	(0.022)	(0.028)
Population2	-0.000	-0.000*	-0.000**	-0.001**	-0.001**	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Population * IV	0.055***	0.059***	0.066***	0.084***	0.074***	0.091***
	(0.018)	(0.017)	(0.018)	(0.028)	(0.028)	(0.025)
Population2 * IV	0.000	0.000	0.000	0.000	0.000	0.001*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Panel B (first stage):	Subway approval					
	(7)	(8)	(9)	(10)	(11)	(12)
IV (Population >3 million)	0.840	0.777***	0.775***	0.884***	0.852***	0.783***
	(0.593)	(0.289)	(0.279)	(0.218)	(0.239)	(0.229)
City FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Mayor basic characteristics		✓	✓	✓	✓	✓
Mayor connections			✓	✓	✓	✓
Mayor competence				✓	✓	✓
City characteristics					✓	✓
Exclude vice-province cities						✓
Observations	24.359	30.110	27.158	29.712	26.157	21.095
Cragg-Donald F statistic	162	159	159	157	157	147

Notes: We check the functional form of our parametric fuzzy RDD by including quadratic term of running variable and its interaction with instrumental variable. The results are robust to this alternative functional form. Furthermore, the coefficients of these two additional quadratic terms are mostly small and insignificant, showing that the influence of higher terms of running variable is not strong. Standard errors clustered at the province and year level are reported in parentheses. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. FE=fixed effects. IV=instrumental variable (i.e., population > 3 million).

Table A9: Fuzzy RDD with Promotion in 1 year as Outcome Variable

	Mayor promoted in one year					
	(1)	(2)	(3)	(4)	(5)	(6)
Subway approval	0.621	0.543	0.610*	0.540**	0.504**	0.757**
	(0.480)	(0.366)	(0.327)	(0.237)	(0.224)	(0.382)
Population	-0.025*	-0.025**	-0.028**	-0.023***	-0.025***	-0.061***
	(0.014)	(0.013)	(0.012)	(0.008)	(0.009)	(0.023)
Population * IV	0.054***	0.053***	0.058***	0.060***	0.057***	0.083***
	(0.020)	(0.019)	(0.017)	(0.017)	(0.016)	(0.021)
City FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Mayor basic characteristics		✓	✓	✓	✓	✓
Mayor connections			✓	✓	✓	✓
Mayor competence				✓	✓	✓
City characteristics					✓	✓
Exclude vice-province cities						✓
Observations	154	151	151	149	149	139

Notes: This table shows that the results are robust to using mayor promoted in one year as outcome variable. Standard errors clustered at the city level are reported in parentheses. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. FE=fixed effects. IV=instrumental variable (i.e., population $>$ 3 million).

Table A10: Fuzzy RDD with Promotion in 2 years as Outcome Variable

	Mayor promoted in two years					
	(1)	(2)	(3)	(4)	(5)	(6)
Subway approval	0.614*	0.426***	0.440***	0.550***	0.515***	0.727***
	(0.355)	(0.152)	(0.145)	(0.145)	(0.169)	(0.245)
Population	-0.010***	-0.011**	-0.013**	-0.007	-0.007	-0.041**
	(0.004)	(0.005)	(0.005)	(0.010)	(0.011)	(0.018)
Population * IV	0.021***	0.025***	0.027***	0.022	0.021	0.051***
	(0.008)	(0.009)	(0.009)	(0.015)	(0.015)	(0.017)
City FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Mayor basic characteristics		✓	✓	✓	✓	✓
Mayor connections			✓	✓	✓	✓
Mayor competence				✓	✓	✓
City characteristics					✓	✓
Exclude vice-province cities						✓
Observations	120	119	119	117	117	108

Notes: This table shows that the results are robust to using mayor promoted in two years as outcome variable. Standard errors clustered at the city level are reported in parentheses. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. FE=fixed effects. IV=instrumental variable (i.e., population > 3 million).

Table A11: Fuzzy RDD with Promotion in 4 years as Outcome Variable

	Mayor promoted in four years					
	(1)	(2)	(3)	(4)	(5)	(6)
Subway approval	0.627	0.520**	0.636***	0.627***	0.479***	0.579***
	(0.440)	(0.237)	(0.245)	(0.186)	(0.135)	(0.169)
Population	-0.014	-0.013	-0.014	-0.011	-0.012	-0.033**
	(0.012)	(0.011)	(0.009)	(0.009)	(0.009)	(0.014)
Population * IV	0.035**	0.032*	0.040***	0.048***	0.035***	0.050***
	(0.017)	(0.019)	(0.014)	(0.013)	(0.013)	(0.012)
City FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Mayor basic characteristics		✓	✓	✓	✓	✓
Mayor connections			✓	✓	✓	✓
Mayor competence				✓	✓	✓
City characteristics					✓	✓
Exclude vice-province cities						✓
Observations	90	89	89	87	87	82

Notes: This table shows that the results are robust to using mayor promoted in four years as outcome variable. Standard errors clustered at the city level are reported in parentheses. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. FE=fixed effects. IV=instrumental variable (i.e., population $>$ 3 million).

Table A12: Fuzzy RDD with Promotion in 5 years as Outcome Variable

	Mayor promoted in five years					
	(1)	(2)	(3)	(4)	(5)	(6)
Subway approval	0.585	0.460*	0.414**	0.493**	0.409**	0.411*
	(0.487)	(0.242)	(0.200)	(0.219)	(0.181)	(0.215)
Population	-0.011	-0.010	-0.010	-0.010	-0.008	-0.016
	(0.009)	(0.008)	(0.007)	(0.010)	(0.009)	(0.013)
Population * IV	0.022*	0.020*	0.020*	0.024*	0.019	0.030**
	(0.012)	(0.012)	(0.011)	(0.014)	(0.013)	(0.012)
City FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Mayor basic characteristics		✓	✓	✓	✓	✓
Mayor connections			✓	✓	✓	✓
Mayor competence				✓	✓	✓
City characteristics					✓	✓
Exclude vice-province cities						✓
Observations	135	132	132	130	130	120

Notes: This table shows that the results are robust to using mayor promoted in five years as outcome variable. Standard errors clustered at the city level are reported in parentheses. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. FE=fixed effects. IV=instrumental variable (i.e., population > 3 million).