Environment for Development

Discussion Paper Series December 2018 ■ EfD DP 18-20

# The Effects of Urban Rail Transit on Air Quality:

## *New Evidence from Multiple Chinese Cities*

**Lunyu Xie, Joshua Linn, and Haoshen Yan**

The Environment for Development (EfD) initiative is an environmental economics program focused on international research collaboration, policy advice, and academic training. Financial support is provided by the Swedish International Development Cooperation Agency (Sida). Learn more at www.efdinitiative.org or contact info@efdinitiative.org.

**Central America**
Research Program in Economics and Environment for Development in Central America Tropical Agricultural Research and Higher Education Center (CATIE)

**Chile**
Research Nucleus on Environmental and Natural Resource Economics (NENRE)
Universidad de Concepción

**China**
Environmental Economics Program in China (EEPC)
Peking University

**Colombia**
The Research Group on Environmental, Natural Resource and Applied Economics Studies (REES-CEDE), Universidad de los Andes, Colombia

**Ethiopia**
Environment and Climate Research Center (ECRC)
Ethiopian Development Research Institute (EDRI)

**India**
Centre for Research on the Economics of Climate, Food, Energy, and Environment, (CECFEE), at Indian Statistical Institute, New Delhi, India

**Kenya**
School of Economics
University of Nairobi

**South Africa**
Environmental Economics Policy Research Unit (EPRU)
University of Cape Town

**Sweden**
Environmental Economics Unit
University of Gothenburg

**Tanzania**
Environment for Development Tanzania
University of Dar es Salaam

**USA (Washington, DC)**
Resources for the Future (RFF)

**Vietnam**
University of Economics
Ho Chi Minh City, Vietnam

The Environment for Development (EfD) initiative is an environmental economics program focused on international research collaboration, policy advice, and academic training. Financial support is provided by the Swedish International Development Cooperation Agency (Sida). Learn more at www.efdinitiative.org or contact info@efdinitiative.org.

## Contents

# The Effects of Urban Rail Transit on Air Quality: New Evidence from Multiple Chinese Cities

Lunyu Xie, Joshua Linn, Haosheng Yan[1]

## Abstract

Automobiles are a major contributor to pervasive urban air quality problems. Partly to reduce these air quality problems, China has invested heavily in urban subway systems. This paper provides the first comprehensive estimates of the effects of these investments on urban air quality. The analysis uses a unique data set that combines hourly air quality data, daily meteorological data, and characteristics of cities for all major subway projects between 2013 and 2014 in China. Based on a Discontinuity Based Ordinary Least Squares (DB-OLS) model, we find that a subway project reduced air pollution by 14% on average; it alleviated more pollution in the daytime, especially during non-rush hours; the effects are smaller in the cities with higher income and lower population density; and, the more existing subway stations, the weaker the effect of opening an additional subway line on air quality.

**Keywords:** urban rail transit, air quality, China

**JEL Classification:** L92, Q53, R41

[1] Lunyu Xie (corresponding author) is an Associate Professor at Renmin University of China, 59 Zhongguancun Avenue, Haidian District, Beijing, 100872, China. Email address: lunyuxie@ruc.edu.cn. Joshua Linn is an Associate Professor at the University of Maryland and a Senior Fellow at Resources for the Future, Washington, DC. Haosheng Yan is a Ph.D. candidate at Renmin University of China.

# The Effects of Urban Rail Transit on Air Quality: New Evidence from Multiple Chinese Cities

## 1. Introduction

Rising traffic congestion and urban air pollution have unfortunately accompanied rapid economic growth in many major cities in developing countries. The adverse human health effects of automobile tailpipe emissions include cardiopulmonary diseases, respiratory infections, lung cancers, infant mortality, and childhood asthma (EPA 2004, Chay and Greenstone 2003, Currie and Neidell 2005, Neidell 2004). The World Health Organization (WHO) estimated that urban air pollution accounts for 6.4 million years of life lost worldwide annually (Cohen et al. 2004). Automobiles are a particularly important contributor to urban air quality problems, due to their emissions of small particulates ($PM_{2.5}$), carbon monoxide (CO), and nitrogen oxides ($NO_x$) as well as other pollutants. Besides, traffic congestion tends to worsen air pollution in urban areas (Gibson and Carnovale 2015; Chen et al. 2015). Partly because of heavy traffic and congestion, cities in China are facing serious air pollution problems. In 2013, China's Ministry of Environmental Protection announced that only 3 of 74 large Chinese cities met official air quality standards. The 71 cities, including Beijing, that do not meet the standards have an aggregate population of roughly 503 million. In Beijing, the number of automobiles increased by a factor of five between 2001 and 2013 (Beijing Transportation Commission, 2014). In 2013, the average congestion time on weekdays was around two hours (Beijing Traffic Analysis Report, 2013). The concentration of $PM_{2.5}$ was eight times greater than the 2013 WHO guidelines (BMEPU, 2014). It is estimated that vehicle emissions account for about 20 percent of Beijing's $PM_{2.5}$ pollution (BMEPU, 2014).

To reduce traffic congestion and air quality problems caused by automobiles, many countries have turned to policies that either raise the cost of owning and using automobiles, or reduce the costs of alternative transportation modes. For example, many cities restrict or tax the use of vehicles in certain areas and times of day, either by taxing entry into the city center at certain times of day, or outright banning certain vehicles from entering the city center on certain days. In addition, many cities ration new license plates and limit the growth of automobile ownership. Whereas the driving restrictions have had mixed effects on air quality (e.g., Davis 2008, Cao et al. 2014, Sun et al. 2014, Viard and Fu 2015), ownership restrictions have been more effective (Yang et al. 2017). However, the costs of both types of policies have been high compared to the benefits (Blackman et al. 2018).

Besides these vehicle policies, many cities have invested heavily in urban subway systems. A total of 48 lines in 18 Chinese cities have been built or extended since 2013. China's urban rail transit system is projected to have 6000 kilometers of track by 2020.[2] By investing $30 billion, Beijing tripled the size of its subway system between 2003 and 2011. Beijing's subway system expanded from 114 kilometers in 2003 to 371 kilometers in 2011, and the number of stations increased from 70 to 219. Worldwide, about 155 million passengers per day in over 116 cities used urban rail transit in 2006, and those numbers are growing (Hester and Harrison 2009). Since the beginning of 2015, 57 cities opened or extended metro systems, tram and light rail systems, with many of those openings occurring in developing countries.[3]

By making it more convenient for commuters to use the subway, expanding subway systems reduces the time cost of using this transit option, thus reducing driving and the associated traffic congestion and air pollution (Kain 1968, Vickrey 1969, Chen and Whalley 2012). The benefits of these investments depend on the responsiveness of vehicle use to reduced time costs of riding the subway, as well as the relative emissions of vehicles and public transportation (given the electricity needs of the subways). In this paper, we evaluate the effects of recent investments in subway systems in China on air quality.

There are mixed findings in the literature on rail transit congestion and commute modes. Most studies suggest that public transit has a small impact on total vehicle miles travelled and traffic congestion (e.g., Stopher 2003, Duranton and Turner 2011). Anderson (2014) found that public transit in Los Angeles has a minimal impact on total vehicle miles travelled, but that it reduces traffic congestion significantly. Nelson et al. (2007) found that rail transit's congestion-reducing benefit exceeds rail subsidies. As for Beijing, Yang et al. (2015) found that the opening of new subway lines in Beijing increases subway ridership and reduces bus ridership but has little impact on traffic congestion. Xie (2012) found that rail transit expansion in Beijing reduced automobile usage by commuters who reside in the traffic zones where the distance to the nearest station decreases, but the reduction of auto use is not enough to improve traffic congestion in Beijing overall.

There is little literature on the effects of China's recent subway investments on air quality. Some cost-benefit analyses on urban rail transit systems in China documented the substitution of travel modes, but they did not include potential air quality improvement as benefits of rail expansion (Miao 1996, Fan 2012, Lu 2012). Chen and Whalley (2012) found that the opening of Taipei's subway system reduced carbon

[2] Data source: http://mic-ro.com/metro/metrostats.html

[3] Data source: http://www.urbanrail.net/news.htm#nowopen

monoxide by 5 to 15 percent and had little effect on ground level ozone. Gendron-Carrier et al. (2015) examined 147 subway openings and expansions between 2000 and 2014 around the world, including 11 openings in China. However, their data do not include the dozens of openings in China since 2013, when subways in China grew rapidly.

In this paper, we study the 15 subway line openings in 2013 and 2014 in China, using hourly data of air quality measured in several indices. Compared to Gendron-Carrier et al. (2018)'s monthly data, the high-frequency data used in this paper allows us to control for potentially nonlinear relationships among air pollution and other factors that change frequently, such as weather. It also enables us to compare the effects across periods of a day, motivated by the fact that daytime pollution is more harmful than nighttime because more people are staying outside during the daytime. We examine five air pollutants that are associated with vehicle use (CO, $PM_{2.5}$, $PM_{10}$, $NO_x$, and $SO_2$)[4], as well as one summary measure of air quality (AQI), whereas Chen and Whalley (2013) examine CO and NOx, and Gendron-Carrier et al. (2018) use a proxy based on satellite observations.

Similar to Chen and Whalley (2013) and Gendron-Carrier et al. (2018), we take a Discontinuity Based OLS approach (DB-OLS) and test for a structural break in air pollution levels after the opening or expansion. We find that opening subways alleviated air pollution, especially during non-rush hours in the daytime; the effects are smaller in the cities with higher income and more subway lines; and the effects are larger in the cities with higher population density. Furthermore, the effect of the first subway line opening is stronger, compared to expansion of an existing subway system.

The remainder of the paper proceeds as follows. Section 2 introduces the data set combined from different data sources. Section 3 presents the methodology adopted for analysis. Section 4 presents the results from regressions. Section 5 investigates the heterogeneity in the effect. Section 6 concludes.

## 2. Data

We collect and merge four major data sets, including timing of subway openings in China since 2013, hourly air pollution data at the level of air quality monitors all over China, daily weather data at the level of meteorological stations, and data on city characteristics. The data sources and the details of the variables are discussed below.

[4]Although $SO_2$ was taken as a pollutant which is not caused by traffic in Chen and Whalley (2012), the case in mainland China is different. During 2013 and 2014, the period studied in this paper, the gasoline in China followed the third and the fourth standards, which require the sulfur content to be under 150mg/kg or 50mg/kg, respectively. This indicates that vehicle exhaust pipes can be a source of $SO_2$.

*Timing of subway openings.* We collect timing of subway openings for all the cities which had new subway lines open in 2013 and 2014 and summarize them in Table 1. As shown by Table 1, 15 subway lines in 11 cities opened new subway lines in 2013 and 2014. In total, 237 stations were added.

The cities which had new subway lines are scattered in the east and middle parts of China, as shown on Figure 1. We notice that Dalian, Changsha and the second line of Kunming's opening date is around May 1st, the Labor Day holiday, which is followed by a long vacation. The commuting behavior during the long vacation is different from ordinary days. We therefore exclude the three subway lines from the following analysis.

One concern about the choice of opening date is that the opening of a line might be correlated with unobserved determinants of air pollution. For example, if openings are typically on a particular weekday or weekend day, and the vehicle use pattern on that day is different from other days, it could confound the estimation. From Table 1, we see that almost all the opening dates are at the beginning or the end of a month, not a fixed weekday or weekend. So, the concern above is alleviated.

*Air pollution data.* We collect hourly air quality data from the website of the Ministry of Environmental Protection (MEP). The data start from January 2013 and cover all of the hundreds of air quality monitors in large cities in China. The concentrations of $PM_{2.5}$, $O_3$, $PM_{10}$, $SO_2$, $NO_2$, and CO are recorded by the monitors and reported by a real-time reporting system of MEP[5]. The Air Quality Index (AQI) is calculated based on the concentration data of these recorded pollutants. Summary statistics of air pollution variables and other variables discussed below are presented in Table 2.

As shown in Table 2 and Figure 2, there is large variation in air quality, and the concentrations of all monitored air-borne pollutants are significantly lower after subway opening than before the opening. The significance of the difference in $O_3$ (ozone) concentration, which is not emitted from automobile exhaust pipes, indicates the importance of controlling for confounding factors.

*Weather data.* To detangle the effect of weather on air quality, we collect daily data on wind, precipitation, temperature, and pressure data from the China Meteorological Administration.

*City characteristics.* To investigate the heterogeneity in the effects of subway opening on air pollution, we collect information from yearbooks on city characteristics, such as automobile ownership rate and per person income.

---

[5]Zhang and Mu (2017) found the quality of air pollution data has improved significantly since 2010, and the real-time reporting system of MEP contributes to the improvement.

## 3. Regression Analysis

We design our research based on Chen and Whalley (2012). They used a discontinuity-based approach in that they have high frequency (daily) measurements of air pollution. They estimated whether there is a discrete improvement in air quality around the time of the subway opening. They controlled for air pollution trends unrelated to the subway opening by fitting a polynomial to the daily air pollution and allowed for a discrete jump in the trend at the time of the opening.

We push forward their research in this paper. Rather than focusing on a subway opening in one city, we pool events across multiple cities[6]. We use a discontinuity based approach as in Chen and Whalley (2012), but after the estimation of the average effect, we investigate the heterogeneity in the effect and look into the factors that drive the heterogeneity.

We identify the effect of subway opening by estimating the following equation:

$$y_{it} = \beta_0 + \beta open_{it} + \boldsymbol{P}(\boldsymbol{t})'\boldsymbol{\delta} + \boldsymbol{P}(\boldsymbol{t})'\boldsymbol{\gamma}\, open_{it} + \boldsymbol{X}'_{\boldsymbol{it}}\boldsymbol{\theta} + \alpha_i + \epsilon_{it},$$
$$where - h < t < h \qquad (1)$$

where $y_{it}$ is the log of air quality in city $i$ and time $t$; $open_{it}$ is a dummy variable which equals one if $t$ is after the subway opening in city $i$; $\boldsymbol{P}(\boldsymbol{t})$ is a polynomial, controlling for time trend; $\boldsymbol{X}_{\boldsymbol{it}}$ is a set of variables that affect the concentration level of air-borne pollutants, which include weather (wind speed, precipitation, pressure, temperature, etc.) and whether the day is a holiday, a weekend, or during the winter, when district-wide heating is provided in northern cities; $\alpha_i$ is a city fixed effect; $\epsilon_{it}$ is the error term; and $h$ is the bandwidth.

We estimate Equation (1) separately for each of the air pollutant variables. According to previous literature, automobile exhaust mainly contributes to CO, $PM_{2.5}$, $PM_{10}$, $NO_x$ and $SO_2$ concentration levels. We therefore expect negative effects on these five pollutants as well as AQI, and little effect on ozone.

## 4. Findings

### 4.1 Main Result

We estimate Equation (1) and summarize the results in Table 3. For each pollutant, we first run a regression without control variables, then add the weather variables and dummies of cities, the day of the week, and whether the day is a holiday. Furthermore,

[6] Cities in which the subway opening date is around May 1st are excluded because May 1st is the Labor Day holiday, and auto use behavior is different on holidays than on weekdays. These excluded subway lines are Changsha Line2, Dalian Line12, and Kunming Line2.

we add the interaction of city dummy and weather. For each regression, only the estimated coefficient of the opening dummy, which shows the effect of subway on air quality, is reported.

The results are stable across specifications for most of the pollutants. We take the specification with the most controls as the main regression. It shows that the opening of subways reduced $PM_{2.5}$ by 17%, $PM_{10}$ by 16%, , $NO_2$ by 12% , CO by 0.4%, and $SO_2$ by 21% on average. In total, the opening of a subway decreases the AQI by 15%.

To provide a visual analysis, we plot the pollutants against days before or after the opening date, after taking out the weather effects dummy variables. As shown by Figures 3 and 4, there are drops on the opening day for the pollutants that are caused by traffic, while there is no drop for ozone. Moreover, we also see a change in the trend of air pollution. Therefore, we report the estimators of trends before and after subway openings in Table 4. It indicates that subway openings contribute to the decrease in the time trend of pollution levels. On average, it decreases the AQI trend by 0.15% per day.

### *4.2 Identifying Assumption Validity*

We examine the validity of the identifying assumption in this part. Figures 5 and 6 show the daily pollution level in the cities with subway openings in 2013 and 2014. Each plot in Figure 5 is for one air-borne pollutant for which concentration can be affected by auto exhaust, including $PM_{2.5}$, $NO_2$, $PM_{10}$, CO, and $SO_2$, as well as the comprehensive air quality index (AQI). The plot in Figure 6 is for ozone, for which concentration is not likely to be affected by auto exhaust. The row dots are fitted by kernel regressions separately for those before and after the opening (the red vertical line). The results are in line with intuition. There are obvious drops in pollution levels after the opening for all the pollutants, except for ozone.

We further test the changes in air quality by regressions, specified in Equation (2).

$$y_{it} = \beta_0 + \beta_j D_{jt} + \boldsymbol{X'_{it}\theta} + \alpha_i + \epsilon_{it},, where - h < t < h \qquad (2)$$

where $D_{jt}$ equals 1 if the day is more than j days after the subway opening, and j is between -14 and 14.

We plot all the statistics of the Wald test $\beta_j$ = 0 and the day relative to the opening. The result is shown in Panel A of Figure 7. The statistics increase before the subway opening and get to a peak right after the subway opening, then decrease. This suggests there is a structural break on the day of the subway opening.

As the usual event study diagram, we regress air quality on lags and leads for opening, and including all the other controls, as specified in Equation (3).

$$y_{it} = \beta_0 + \sum_{j\neq 0,-h<j<h} \beta_j D_{jt} + \boldsymbol{X}'_{\boldsymbol{it}}\boldsymbol{\theta} + \alpha_i + \epsilon_{it},, where - h < t < h \qquad (3)$$

We then plot all the coefficients on leads and lags against time, with confidence intervals, as shown in Panel B of Figure 7. Because we drop the dummy of the opening day, the coefficients of dummies indicate the difference in air quality between the opening day and days before or after subway opening. This shows that the coefficients of lags are generally positive; the coefficients of leads are negative; and there is a drop on the opening day. This indicates that the subway opening decreased air pollution.

Our identification strategy is valid only when there is no discontinuity in the covariates on the day of the subway opening. In our case, the covariates are weather variables. They are exogenous to the openings of subway lines. As shown in Figure 8, the weather variables distribute smoothly around the opening date. There is no obvious change in the weather on the days of the subway opening.

### *4.3 Robustness Checks*

To test for the robustness of the results, we do several robustness checks and summarize the results in Table 5. The first test is on bandwidth. The opening of a new subway line could be an important event that attracts public attention. Therefore, on the opening day, and in the week or two following the opening, there would be a large increase in ridership as people try out the new subway line. We therefore double the optimal bandwidth to test whether the pollution reducing effect is stable in the long term. We also take half of the optimal bandwidth as a robustness check. The results are listed in Panel A of Table 5. All signs of the estimators remain negative and the magnitudes remain stable. This suggests that subway openings' alleviation of air pollution is not just because of the rush to a new transportation option; the effect of subways on pollution alleviation can last for the long run. In fact, the new subways are not the first subways of the cities in the sample, therefore they can only attract limited attention.

Second, as the method of local regression may affect the result, we change the order of polynomial and the method of local regression. The result, reported in Panel B of Table 5, remains similar to the main regression. When the order of the polynomial is quadratic, the estimators become smaller, giving a 7% decrease in AQI. The estimators of traffic-associated pollutants are still negative and statistically significant.

Third, we control for lags of the dependent variable. As noted by Henderson (1992), the pollution may last for hours. Therefore, endogeneity may exist due to the missing variables. So, we add the pollution of the last four hours in the equation. The results are listed in Panel C of Table 5. After the time lag terms were added, the sign of subway opening is still negative, but the coefficients become smaller. The reason is the existence of lagged pollution, so that the estimator measures a short time effect, as Chen and Whalley (2012) noted. We also calculate the total effect after adding the lag term; the magnitude is almost half of that in the main regression, which is near the result of using a quadratic polynomial. The significance of subway opening is slightly weakened: the estimates of AQI, $PM_{2.5}$, $PM_{10}$ and CO are significant at the 10% level, while $NO_2$ and $SO_2$ are significant at the 5% level. The reason may be that, after controlling for the lag of pollutants, the magnitude of the point estimator becomes smaller, which influences the significance. As the dependent variable turns to the traffic-unrelated pollutant ozone, there is no significant change, as we expected.

At last, we carry out a matching difference in difference method and present the results in Table 6. For each city with opening subways, we match on income per capita, GDP, and population density. We then choose the most similar cities from each province as a control group. For Xian, Kunming and Wuhan, for which we do not have data for other cities in their province around the opening date, we pick the most similar city in their neighboring province. For Shanghai, a megacity, we choose the second largest city, Beijing, as its control. The city characteristics of the treatment group and the control group are presented in Panel A of Table 6. There is no significant difference in income and GDP; however, population is a little denser in the treatment group. Then, we use a difference in difference to identify the treatment effect of subway opening, controlling for these city characteristics. As shown by Panel B of Table 6, the estimated effects are similar to those from the main regression.

## 5. Heterogeneous Effects

In this section, we investigate possible heterogeneity in the effects of subway opening on air quality.

Given that auto use behavior and public transit experience are different during rush hours and non-rush hours, we expect that the effect of subways on air quality varies across time periods in a day. At night, when the subways do not operate, a subway opening is expected to have no effect on air quality. During rush hours, travel time is usually one of the factors when people make travel decisions, and public transit is crowded during rush hours. Then subways are not a good alternative to autos. Substitution of autos to subways is likely to be stronger during non-rush hours. To explore the heterogeneity of the effect in time periods, we regress the average level of pollutants at night（0am to 5am), rush hour (7am to 9am) and non-rush hour (11am to

3pm) on the same independent variables. As shown by Table 7, the results are the same as expected: the subway has the largest effect on non-rush hours, and the air quality at night is not affected.

Next, we explore heterogeneous effects across cities with different income and population density. Results are presented in the first two panels of Table 8. They show that the effect is smaller in the cities with higher income, and larger in the cities with higher population density. The reasons could be that automobile users with higher income are less likely to be diverted to subways, and that subways have a stronger diversion effect in the cities that are more congested.

Finally, we measure the marginal effect of subway opening. The magnitude of the effect is determined by the extent to which the subway diverts automobile users. Given that the first subway in a city is likely to attract more passengers than an extension line to an existing system, the effect of the first subway is expected to be larger. As the number of existing stations increases, new stations are more likely to be located in places with lower population density, so the marginal effect of adding a subway station decreases. Estimation results presented in Panels C and D of Table 8 confirm these two hypotheses, showing that the first subway line has a stronger effect, and, as the number of existing stations increases, the influence decreases.

## 6. Conclusion

Exposure to air pollution has substantial negative effects on health. In this paper, we investigate whether urban rail transit expansion improves air quality. By exploiting the regression discontinuity based OLS approach, we find that opening subways significantly reduced traffic-related air pollution and had no effect on air pollution not related to traffic. This indicates that investment in public transportation infrastructure improved air quality.

Furthermore, we explore the heterogeneous effect among cities. We found that subway opening alleviated more air pollution in the daytime than at night, especially during non-rush hours. We also find that the alleviating effects are lower in the cities with higher income per capita, and higher in the cities with higher population density; the more existing subway stations, the weaker the effect of opening a new subway line; and the first subway line has the strongest effect on air pollution. The heterogeneous effects can provide more information to policy makers on the timing of building new subway lines.

## References

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

Beijing Traffic Analysis Report. (2013).

Beijing Transportation Commission. (2014).

Blackman, A., Li, Z., and Liu, A. A. (2018). Efficacy of command-and-control and market-based environmental regulation in developing countries. *Annual Review of Resource Economic*s, *10*(1), annurev-resource-100517–023144.

Beijing Municipal Environmental Protection Bureau. (2014).

Cao, J., Wang, X., and Zhong, X. H. (2014). Did driving restrictions improve air quality in Beijing? *China Economic Quarterly, 13*, 1091–1126.

Chay, K. Y., and Greenstone, M. (2003). The impact of air pollution on infant mortality: Evidence from geographic variation in pollution shocks induced by a recession*. The Quarterly Journal of Economics, 118*, 1121–1167.

Chen, S., Yang, J., Qin, P., and Xu, J. (2015). How traffic jams affect air quality: Analysis of the nonlinear relationship between traffic congestion and air pollution in Beijing. Environment for Development (EfD). (Working Paper).

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

Cohen, A. J., Anderson, R. H., Ostro, B., Pandey, K. D., Krzyzanowski, M., Kunzli, N., Gutschmidt, K., et al. (2004). Urban air pollution. *Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors*, (1) 1353–1434. ed. Majid Ezzati, Alan D. Lopez, Anthony Rodgers, and Christopher J.L. Geneva: World Health Organization.

Currie, J., and Neidell, M. (2005). Air pollution and infant health: What can we learn from California's recent experience? *The Quarterly Journal of Economics, 120*, 1003–1030.

Duranton, G., and Turner, M. (2011). The fundamental law of road congestion: Evidence from US cities. *American Economic Review, 101*(6), 2616–2652.

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

EPA. (2004). Air quality criteria for particulate matter. EPA/600/P-99/002aF (October).

Fan, Z. (2012). The economic effect analyssi of the uran rail transit aystem. (Doctral Dissertation). Southwest Jiaotong University (in Chinese).

Gendron-Carrier, N., Gonzaleznavarro, M., Polloni, S., and Turner, M. (2018). Subways and urban air pollution. (Nber Working Papers).

Gibson, M., and Carnovale, M. (2015). The effects of road pricing on driver behavior and air pollution. *Journal of Urban Economics, 89*, 62–73.

Henderson, J. V. (1996). Effects of air quality regulation. *American Economic Review*, *86*(4), 789–813.

Hester, R. E., and Harrison, R. M. (2009). Air quality in urban environments. *Royal Society of Chemistry, 28*.

Kain, J. F. (1968). Housing segregation, negro employment, and metropolitan decentralization. *The Quarterly Journal of Economics*, *82*(2), 175–197.

Lu, M. (2012). Research on comprehensive benefits of rrban rail transit system. (Doctral Dissertation). Beijing Jiaotong University (in Chinese).

Miao, Y., Had, P., and Wang, S. (1996). The cost-benefit analysis of the urban rail transit system. *Journal of Dalian Railway Institute, 01,* 17–23 (in Chinese).

Neidell, M. J. (2004). Air pollution, health, and socio-economic status: The effect of outdoor air quality on childhood asthma. *Journal of Health Econonomics, 23,* 1209–1236.

Nelson, P., et al. (2007). Transit in Washington, DC: Current benefits and optimal level of provision. *Journal of Urban Economics, 62*(2), 231–251.

Stopher, P. 2004. Reducing road congestion: A reality check. *Transport Policy, 11*(2), 117–131.

Sun, C., Zheng, S. Q., and Wang, R. (2014). Restricting driving for better traffic and clearer skies: Did it work in Beijing? *Transport Policy, 32,* 34–41.

Viard ,V. B., and Fu, S. (2015). The effect of Beijing's driving restrictions on pollution and economic activity. *Journal of Public Economics, 124,* 98–115.

Vickrey, W. S. (1969). Congestion theory and transportation investment. *American Economic Review, 59(*2), 251–260.

Xie, L. (2016). Automobile usage and urban rail transit expansion: Evidence from a natural experiment in Beijing, China. *Environment and Development Economics*, 1–24.

Yang, J., Liu, A. A., Qin, P., and Linn, J. (2017). The effect of vehicle ownership restrictions on travel behavior: Evidence from the Beijing License Plate Lottery.

Yang, J., Chen, S., Qin, P., and Lu, F. (2015). Will subway expansion alleviate traffic congestion? Evidence from Beijing. (Working paper).

Zhang, J., and Mu, Q. (2018). Air pollution and defensive expenditures: Evidence from particulate-filtering facemasks. *Journal of Environmental Economics and Management, 92*, 517-536.

## Tables and Figures

**Table 1. New Subway Lines Opened in 2013 and 2014 in China**

| City | Line | Opening Date | Existing Stations | New Stations | Income (yuan) | Population Density (people/km$^2$) |
|---|---|---|---|---|---|---|
| Kunming | 1 | 5/20/2013 | 3 | 12 | 28354 | 594.97 |
| Xi'an | 1 | 9/15/2013 | 17 | 19 | 33100 | 1621.33 |
| Guangzhou | 6 | 12/28/2013 | 129 | 21 | 42049 | 1786.73 |
| Wuhan | 4 | 12/28/2013 | 47 | 15 | 29821 | 1885.84 |
| Zhengzhou | 1 | 12/28/2013 | 0 | 20 | 26615 | 5119.8 |
| Shanghai | 12 | 12/29/2013 | 338 | 15 | 43851 | 3809 |
| Shanghai | 16 | 12/29/2013 | 338 | 11 | 43851 | 3809 |
| Changsha | 2 | 4/29/2014 | 0 | 19 | 36826 | 1589.06 |
| Kunming | 2 | 4/30/2014 | 15 | 14 | 31295 | 599.8 |
| Dalian | 12 | 5/1/2014 | 12 | 8 | 33591 | 1185.31 |
| Ningbo | 1 | 5/30/2014 | 0 | 20 | 41731 | 932.74 |
| Wuxi | 1 | 7/1/2014 | 0 | 24 | 41731 | 1494.77 |
| Nanjing | 10 | 7/1/2014 | 53 | 14 | 42568 | 984.85 |
| Nanjing | S1 | 7/1/2014 | 53 | 8 | 42568 | 984.85 |
| Nanjing | S8 | 8/1/2014 | 75 | 17 | 42568 | 984.85 |

*Notes:* The table lists cities opening new subways during 2013 to 2014 in China. Column 2 to Column 4 report the name of the new subway line, the opening date, number of subway stations existing in the city and the number of stations of the new subway line. Column 5 and Column 6 report average income and population density of each city in the year of opening new subways.

**Table 2. Summary Statistics**

| | | Before Subway Opening (72818 city hours) | | After Subway Opening (2875 city days) | | Difference |
|---|---|---|---|---|---|---|
| | Variable | Mean | Std. Dev. | Mean | Std. Dev. | |
| Pollutants Caused by Traffic | AQI | 100.8084 | 66.90311 | 86.33328 | 58.96354 | 14.47507 *** |
| | $PM_{2.5}$ | 70.23 | 58.11 | 57.77 | 51.40 | 12.45699 *** |
| | $PM_{10}$ | 114.1469 | 85.96324 | 97.76359 | 76.09637 | 16.38333 *** |
| | $NO_2$ | 48.13263 | 26.33313 | 42.72271 | 22.69197 | 5.409921 *** |
| | CO | 1.559721 | 15.06225 | 1.194638 | 0. 6991027 | 0. 3650838 *** |
| | $SO_2$ | 32.52876 | 33.14853 | 25.18093 | 30.82956 | 7.347832 *** |
| Pollutants Not Caused by Traffic | $O_3$ | 66.18237 | 49.53516 | 57.87509 | 49.33356 | 8.307283 *** |
| Weather Condition | wind speed | 23.15825 | 9.020403 | 21.19821 | 8.299235 | 1.96 *** |
| | precipitation | 32.18027 | 103.4444 | 37.15488 | 93.44784 | -4.974603 *** |
| | pressure | 9990.742 | 392.9302 | 9595.228 | 762.1698 | 395.5138 *** |
| | temperature | 179.2843 | 89.35125 | 189.878 | 78.84315 | -10.59363*** |

*Notes:* The range of the data is from January 2013 to October 2014. Columns 1 and 2 report mean value and standard deviation of air pollution and weather before subway opening in cities listed in Table 1. Columns 3 and 4 report mean value and standard deviation after subway opening. The last column reports the difference and the t test result before and after the opening of the subway. *** p<0.01, ** p<0.05, * p<0.

**Table 3. Effects of Subway Opening on Air Quality**

| | Pollutants Caused by Traffic | | | | | | Pollutants Not Caused by Traffic |
|---|---|---|---|---|---|---|---|
| | AQI | $PM_{2.5}$ | $PM_{10}$ | $NO_2$ | CO | $SO_2$ | $O_3$ |
| Without Control Variables | -0.235*** | -0.288*** | -0.245*** | -0.070** | -0.021 | -0.207*** | -0.055 |
| | (0.032) | (0.041) | (0.035) | (0.032) | (0.028) | (0.045) | (0.051) |
| With Control Variables | -0.135*** | -0.150*** | -0.154*** | -0.098*** | -0.039* | -0.140*** | 0.123*** |
| | (0.026) | (0.032) | (0.029) | (0.024) | (0.020) | (0.031) | (0.037) |
| Further Control Interaction of City Dummy and Weather | -0.147*** | -0.171*** | -0.160*** | -0.119*** | -0.004 | -0.214*** | 0.001 |
| | (0.025) | (0.031) | (0.028) | (0.025) | (0.021) | (0.032) | (0.041) |
| Bandwidth | 15 | 15 | 15 | 15 | 15 | 15 | 15 |
| Observations | 5,869 | 5,869 | 5,869 | 5,869 | 5,869 | 5,869 | 5,869 |

*Notes:* Discontinuity based OLS. The control variables of the second row include wind speed, precipitation, pressure, and dummies of holiday, dummy of hours in a day, heating period, weekday and city, and the interaction of weather and hour dummies; we further control the interaction of city dummy and weather in the third row. They enter the regressions in a linear form. Bandwidth is 15 days, giving a total 30 days window. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

**Table 4. Effects of Subway Opening on Air Quality Trend**

| | Pollutants Caused by Traffic | | | | | | Pollutants Not Caused by Traffic |
|---|---|---|---|---|---|---|---|
| | AQI | $PM_{2.5}$ | $PM_{10}$ | $NO_2$ | CO | $SO_2$ | $O_3$ |
| Before opening | 0.00053*** | 0.00052*** | 0.00066*** | 0.00054*** | 0.00025*** | 0.00121*** | -0.00019 |
| | (0.00009) | (0.00011) | (0.00010) | (0.00008) | (0.00007) | (0.00010) | (0.00012) |
| After opening | -0.00096*** | -0.00098*** | -0.00114*** | -0.00081*** | -0.00015 | -0.00201*** | -0.00002 |
| | (0.00012) | (0.00014) | (0.00013) | (0.00011) | (0.00009) | (0.00014) | (0.00017) |
| Difference | -0.00149*** | -0.0015*** | -0.0018*** | -0.00135*** | -0.0004*** | -0.00322*** | 0.00017 |

*Notes:* Discontinuity-based OLS. The specification is same as the second row of Table 3. The first row reports the slope before subway opening and the second reports the slope after subway opening. The last row reports the difference. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

**Table 5. Robustness Checks**

Panel A: Different Bandwidth

| | Pollutants Caused by Traffic | | | | | | Pollutants Not Caused by Traffic |
|---|---|---|---|---|---|---|---|
| | AQI | $PM_{2.5}$ | $PM_{10}$ | $NO_2$ | CO | $SO_2$ | $O_3$ |
| A Week Bandwidth | -0.107*** | -0.151*** | -0.187*** | -0.203*** | -0.079** | -0.393*** | -0.004 |
| | (0.040) | (0.049) | (0.044) | (0.037) | (0.031) | (0.044) | (0.051) |
| A Month Bandwidth | -0.184*** | -0.230*** | -0.152*** | -0.046*** | 0.049*** | -0.075*** | 0.101*** |
| | (0.018) | (0.023) | (0.020) | (0.015) | (0.014) | (0.020) | (0.024) |

Panel B: Different Order Polynomials and Local Regression Method

| | Pollutants Caused by Traffic | | | | | | Pollutants Not Caused by Traffic |
|---|---|---|---|---|---|---|---|
| | AQI | $PM_{2.5}$ | $PM_{10}$ | $NO_2$ | CO | $SO_2$ | $O_3$ |
| Quadratic Polynomials | -0.077** | -0.086** | -0.110*** | -0.081** | -0.073*** | -0.168*** | 0.018 |
| | (0.036) | (0.044) | (0.039) | (0.033) | (0.028) | (0.042) | (0.051) |
| Kernel Regression | -0.114*** | -0.120*** | -0.143*** | -0.116*** | -0.055*** | -0.202*** | 0.089*** |
| | (0.028) | (0.034) | (0.032) | (0.024) | (0.020) | (0.031) | (0.034) |
| Observations | 5,869 | 5,869 | 5,869 | 5,869 | 5,869 | 5,869 | 5,869 |

Panel C: Controlling Lag Pollutants

| | Pollutants Caused by Traffic | | | | | | Pollutants Not Caused by Traffic |
|---|---|---|---|---|---|---|---|
| | AQI | $PM_{2.5}$ | $PM_{10}$ | $NO_2$ | CO | $SO_2$ | $O_3$ |
| Opening Subways | -0.013* | -0.015* | -0.017* | -0.030** | -0.020* | -0.054*** | 0.035 |
| | (0.007) | (0.009) | (0.009) | (0.015) | (0.011) | (0.020) | (0.024) |
| First lag dependent variable | 1.090*** | 1.162*** | 1.051*** | 0.729*** | 0.656*** | 0.665*** | 0.617*** |
| | (0.014) | (0.014) | (0.014) | (0.015) | (0.015) | (0.015) | (0.014) |
| Second lag dependent variable | -0.084*** | -0.168*** | -0.030 | 0.118*** | 0.204*** | 0.143*** | 0.181*** |
| | (0.021) | (0.022) | (0.021) | (0.020) | (0.018) | (0.018) | (0.017) |
| Third lag dependent variable | -0.028 | -0.030 | -0.108*** | 0.013 | 0.032* | 0.033* | 0.075*** |
| | (0.021) | (0.022) | (0.020) | (0.020) | (0.018) | (0.018) | (0.017) |
| Forth lag dependent variable | -0.023 | -0.005 | 0.033** | -0.009 | -0.008 | 0.018 | -0.036** |
| | (0.014) | (0.014) | (0.014) | (0.015) | (0.015) | (0.015) | (0.015) |
| Total effect | -0.06907 | -0.08383 | -0.08573 | -0.10336 | -0.06793 | -0.1747 | 0.110351 |
| Observations | 4,969 | 4,969 | 4,969 | 4,969 | 4,969 | 4,969 | 4,969 |

*Notes:* Discontinuity-based OLS. In Panel A, the specification is the same as the second row of Table 3. The difference is the bandwidth. In Panel B, we control a quadratic polynomial time trend in the first row, instead of the linear polynomial in Table 3. Further, in the second row, we use the kernel regression to control the time trend. Other control variables and the bandwidth are the same as the second row of Table 3. In Panel C, we control the pollutants one to four hours earlier as well as other control variables in the main regression. We report estimators and the total effect considering the lag pollutant. The calculation method is based on Henderson (1996). Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

**Table 6. Results from a Matching Difference in Difference Approach**

Panel A: Cities Characteristics

| | Treatment | | Control Group | | Difference |
|---|---|---|---|---|---|
| | Mean | Std. Dev. | Mean | Std. Dev. | |
| Income Per Capita (yuan) | 37445.24 | 6648.23 | 34757.94 | 9574.30 | -2687.29 |
| GDP (thousand million yuan) | 971.00 | 640.00 | 831.00 | 580.00 | -140.00 |
| Population Density (thousand people per $km^2$) | 1.03 | 0.53 | 0.66 | 0.41 | -0.37** |

Panel B: Difference-in-Difference Result

| | Pollutants Caused by Traffic | | | | | | Pollutants Not Caused by Traffic |
|---|---|---|---|---|---|---|---|
| | AQI | $PM_{2.5}$ | $PM_{10}$ | $NO_2$ | CO | $SO_2$ | $O_3$ |
| | -0.168*** | -0.221*** | -0.193*** | -0.024 | -0.007 | -0.068*** | 0.009 |
| Open*Treatment | (0.018) | (0.022) | (0.020) | (0.016) | (0.014) | (0.021) | (0.025) |
| Open | 0.020 | 0.044*** | 0.037** | -0.040*** | 0.016 | -0.012 | 0.031* |
| | (0.013) | (0.017) | (0.015) | (0.012) | (0.010) | (0.016) | (0.019) |
| Treatment | -1.906*** | -2.134*** | -3.230*** | -3.849*** | -2.424*** | -1.557*** | 4.107*** |
| | (0.249) | (0.313) | (0.277) | (0.229) | (0.193) | (0.294) | (0.346) |
| Observations | 9,655 | 9,655 | 9,655 | 9,655 | 9,655 | 9,655 | 9,655 |

*Notes:* Difference in difference. The control variables in the main regression are also controlled. The sample contains 2 weeks before and after subway opening. We report standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

**Table 7. Effects of Subway Opening on Air Quality of Various Time Periods**

| | Pollutants Caused by Traffic | | | | | | Pollutants Not Caused by Traffic |
|---|---|---|---|---|---|---|---|
| | AQI | $PM_{2.5}$ | $PM_{10}$ | $NO_2$ | CO | $SO_2$ | $O_3$ |
| Night Time | -0.049 | -0.073 | -0.012 | -0.062 | -0.021 | 0.128 | 0.061 |
| | (0.126) | (0.151) | (0.139) | (0.108) | (0.082) | (0.124) | (0.165) |
| Rush Hour | -0.186 | -0.228 | -0.187 | -0.118 | -0.124 | -0.037 | 0.287** |
| | (0.124) | (0.146) | (0.140) | (0.079) | (0.082) | (0.136) | (0.144) |
| Non-Rush Hour | -0.239** | -0.275* | -0.235* | -0.168* | -0.143* | -0.173 | 0.311*** |
| | (0.118) | (0.144) | (0.126) | (0.092) | (0.082) | (0.120) | (0.107) |
| Observations | 261 | 261 | 259 | 261 | 261 | 261 | 261 |

*Notes:* Discontinuity-based OLS. The specification and the bandwidth are the same as in Table 3. The dependent variables are pollution of different periods of a day on daily level. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

## Table 8. Heterogeneous Effects

| | AQI | $PM_{2.5}$ | $PM_{10}$ | $NO_2$ | CO | $SO_2$ |
|---|---|---|---|---|---|---|
| Panel A: Income | | | | | | |
| Open*Income | 0.636*** | 0.812*** | 0.638*** | 0.425*** | 0.681*** | 0.381*** |
| | (0.068) | (0.084) | (0.062) | (0.053) | (0.074) | (0.080) |
| Open | -6.797*** | -8.660*** | -6.789*** | -4.494*** | -7.291*** | -4.123*** |
| | (0.711) | (0.875) | (0.647) | (0.558) | (0.780) | (0.841) |
| Income | -0.508*** | -0.620*** | -0.671*** | -0.385*** | -0.453*** | -0.050 |
| | (0.070) | (0.086) | (0.063) | (0.055) | (0.076) | (0.082) |
| Panel B: Density | | | | | | |
| Open*Density | -0.097*** | -0.172*** | 0.067*** | -0.107*** | -0.035 | -0.051** |
| | (0.019) | (0.024) | (0.018) | (0.015) | (0.021) | (0.023) |
| Open | 0.577*** | 1.107*** | -0.580*** | 0.738*** | 0.111 | 0.232 |
| | (0.142) | (0.174) | (0.130) | (0.111) | (0.156) | (0.167) |
| Pop. Density | 0.331*** | 0.526*** | -0.131 | 0.146** | 0.305*** | -0.010 |
| | (0.093) | (0.115) | (0.086) | (0.073) | (0.103) | (0.110) |
| Panel C: First line | | | | | | |
| Open*First Line | -0.237*** | -0.250*** | -0.327*** | -0.100*** | -0.272*** | -0.074*** |
| | (0.023) | (0.029) | (0.021) | (0.018) | (0.025) | (0.028) |
| Open | -0.021 | -0.029 | 0.058** | 0.009 | -0.021 | -0.101*** |
| | (0.028) | (0.035) | (0.025) | (0.022) | (0.031) | (0.033) |
| First Line | 0.598 | 0.791 | -4.884*** | -1.543*** | -1.305** | -2.955*** |
| | (0.492) | (0.607) | (0.444) | (0.387) | (0.539) | (0.583) |
| Panel D: Number of Existing Stations | | | | | | |
| Open*Stations | 0.001*** | 0.001*** | 0.002*** | 0.000*** | 0.001*** | -0.000*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Open | -0.172*** | -0.186*** | -0.191*** | -0.064*** | -0.195*** | -0.114*** |
| | (0.027) | (0.033) | (0.024) | (0.021) | (0.029) | (0.032) |
| Number of | 0.002*** | 0.002*** | 0.001*** | -0.000 | 0.002*** | 0.003*** |
| Existing Stations | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |

*Notes:* In this table, we add income of citizen, population density, dummy of the first line and number of existing stations and their interaction with the dummy of opening subway respectively into Equation (1). Control variables and bandwidth are the same as in Table 3. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

**Figure 1. Cities with Subway Openings and Extensions in 2013 and 2014**

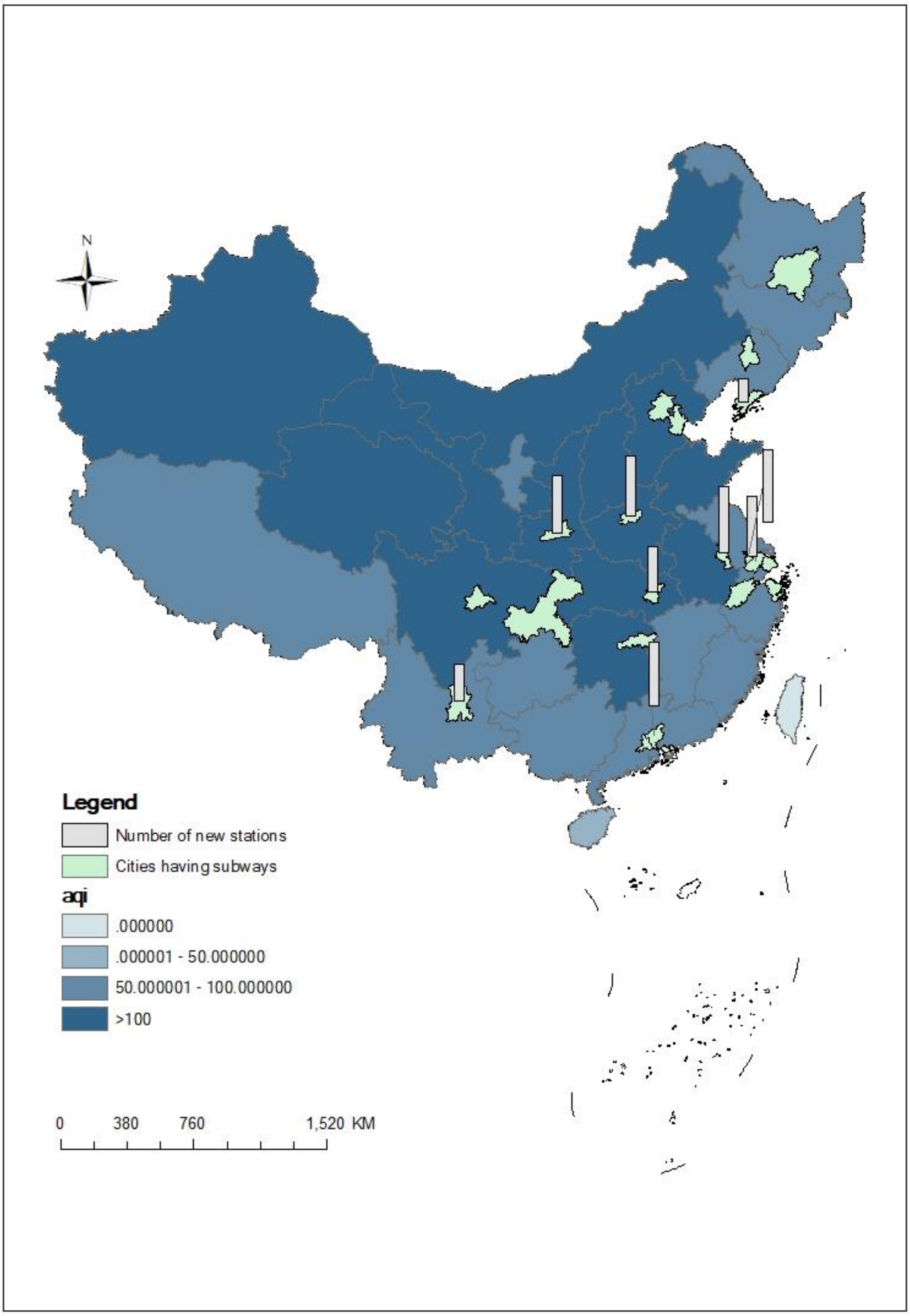
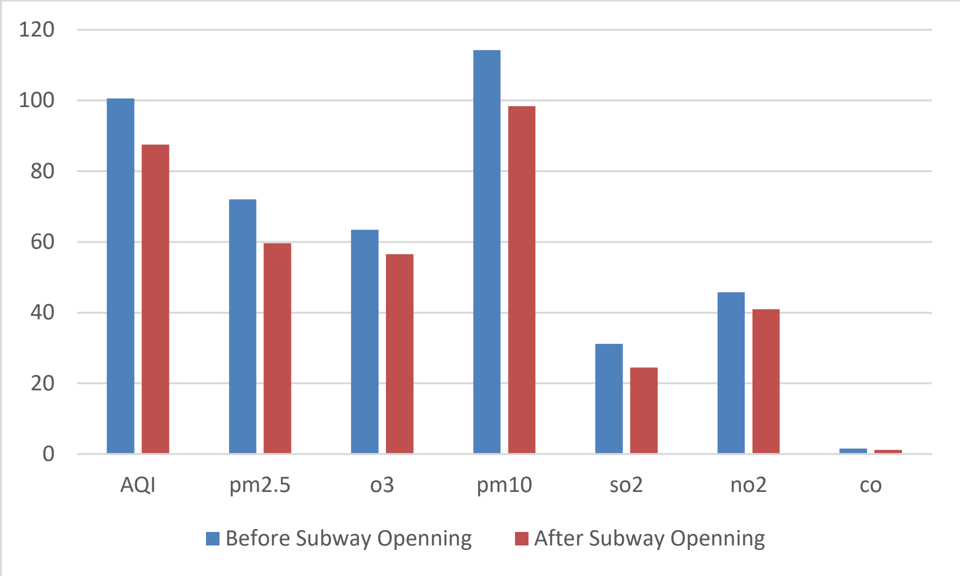
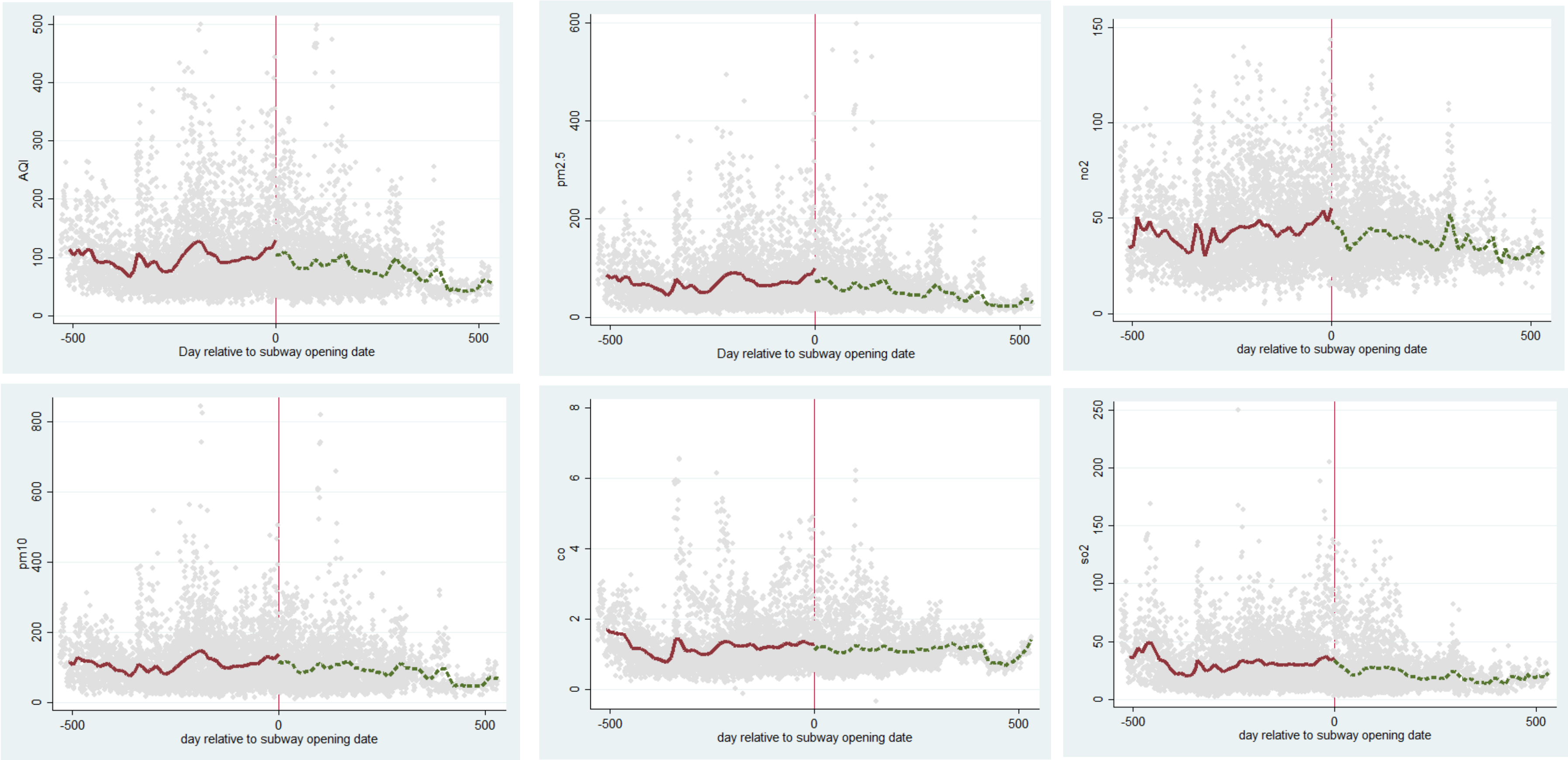


**Figure 2. Average Pollutant Level Before and After Subway Opening**



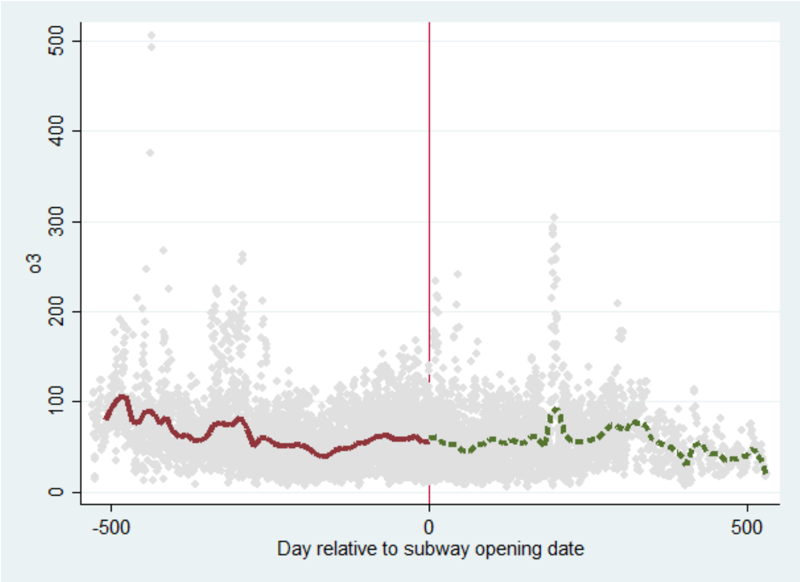

*Notes*: This figure shows the air quality before and after subway opening in cities listed in Table 1. The time range is between January 2013 and October 2014.

**Figure 3. Daily Pollution Level in the Cities with Subway Opening in 2013 and 2014**




*Notes:* Each plot is for one air-borne pollutant, including $PM_{2.5}$, $NO_2$, $PM_{10}$, CO, and $SO_2$, as well as the comprehensive air quality index (AQI). All cities listed in Table 1 are included. The studied period is from January 2013 to October 2014. The red vertical line denotes the opening of a subway line. The row dots are fitted by kernel regressions separately for those before and after the opening.

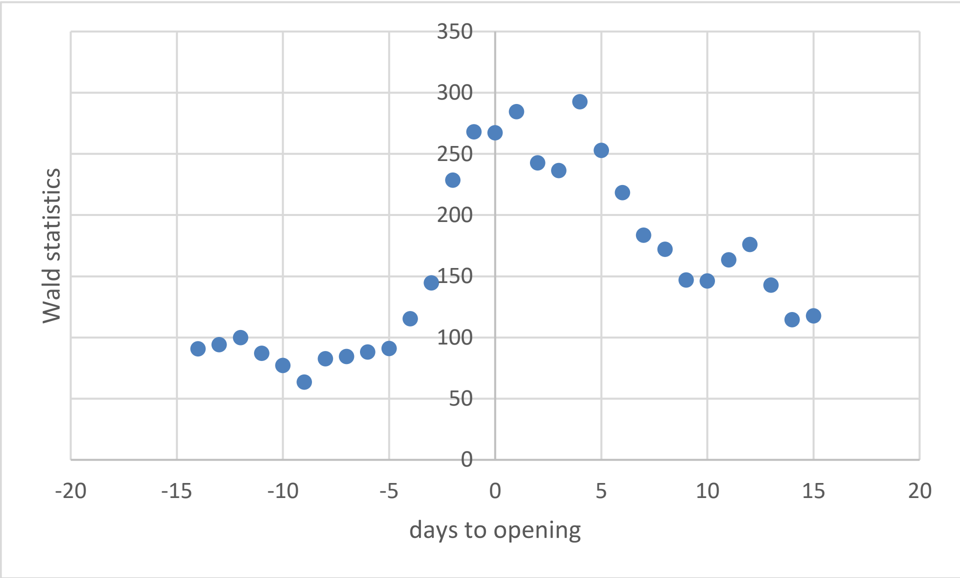
**Figure 4. Daily Pollution (Ozone only) Level in the Cities with Subway Opening in 2013 and 2014**



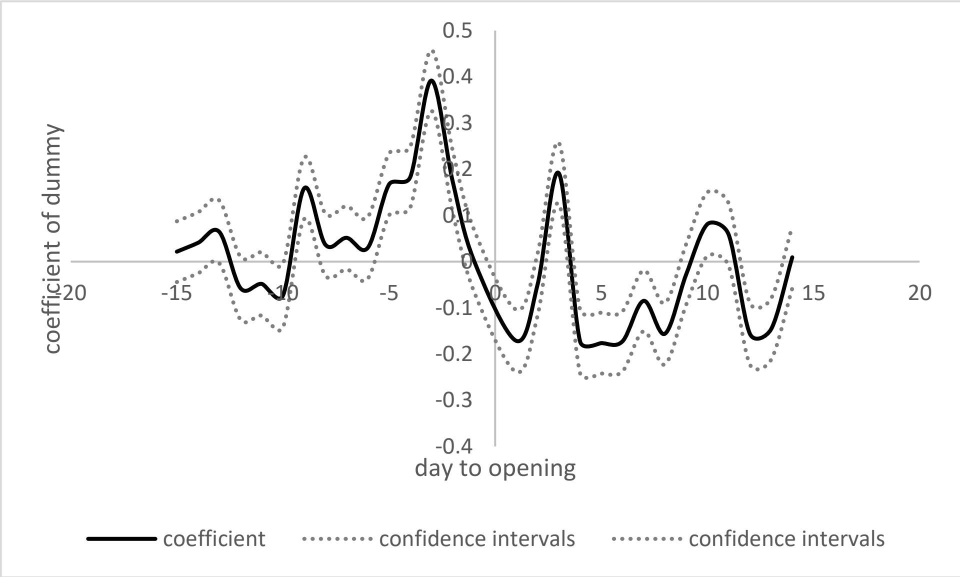

*Notes:* The plot is for ozone. All cities listed in Table 1 are included. The studied period is from January 2013 to October 2014. The red vertical line denotes the opening of a subway line. The row dots are fitted by kernel regressions separately for those before and after the opening.

## Figure 5. Structural Break

Panel A: Structural Break

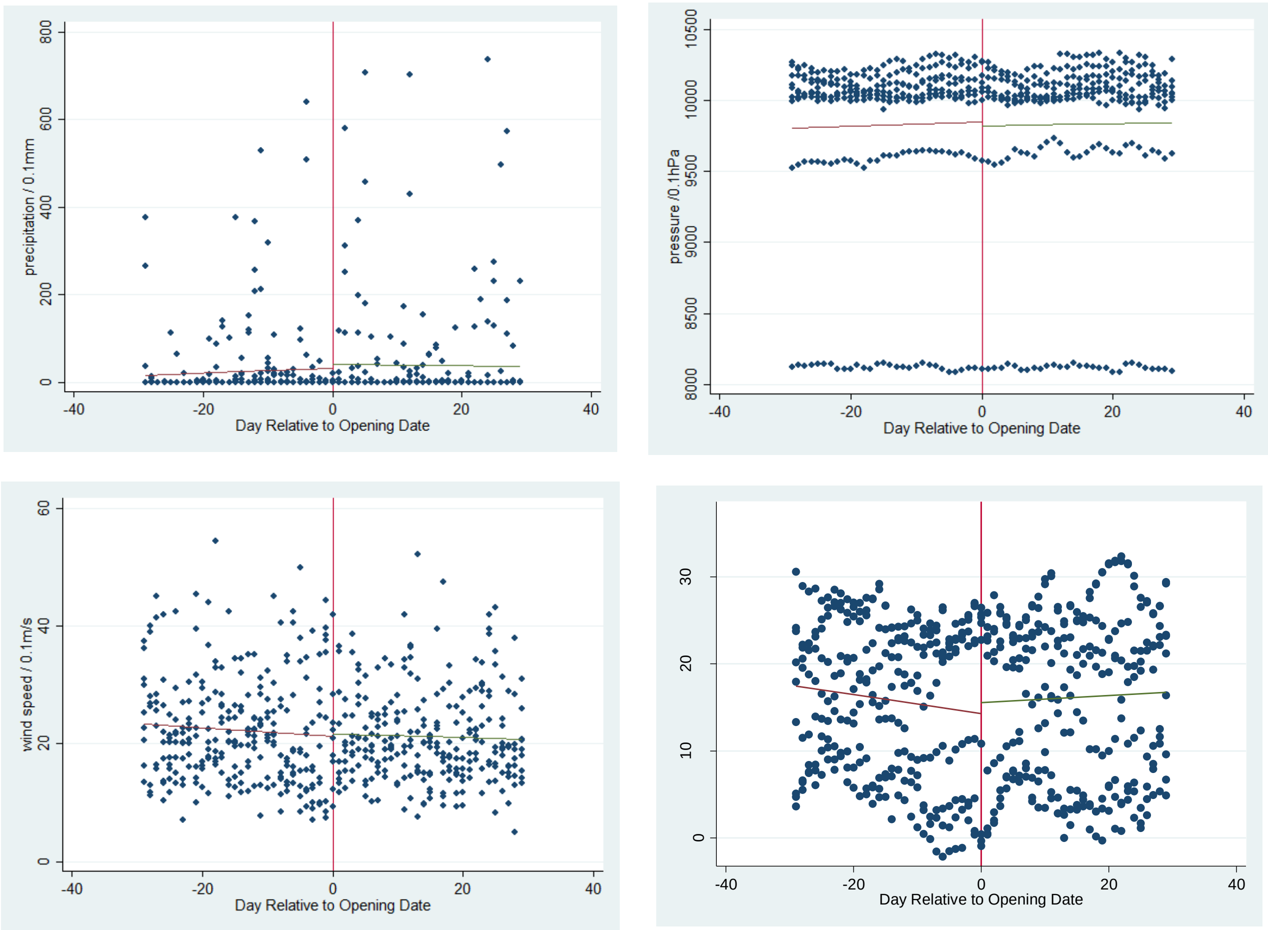



Panel B: Event Study



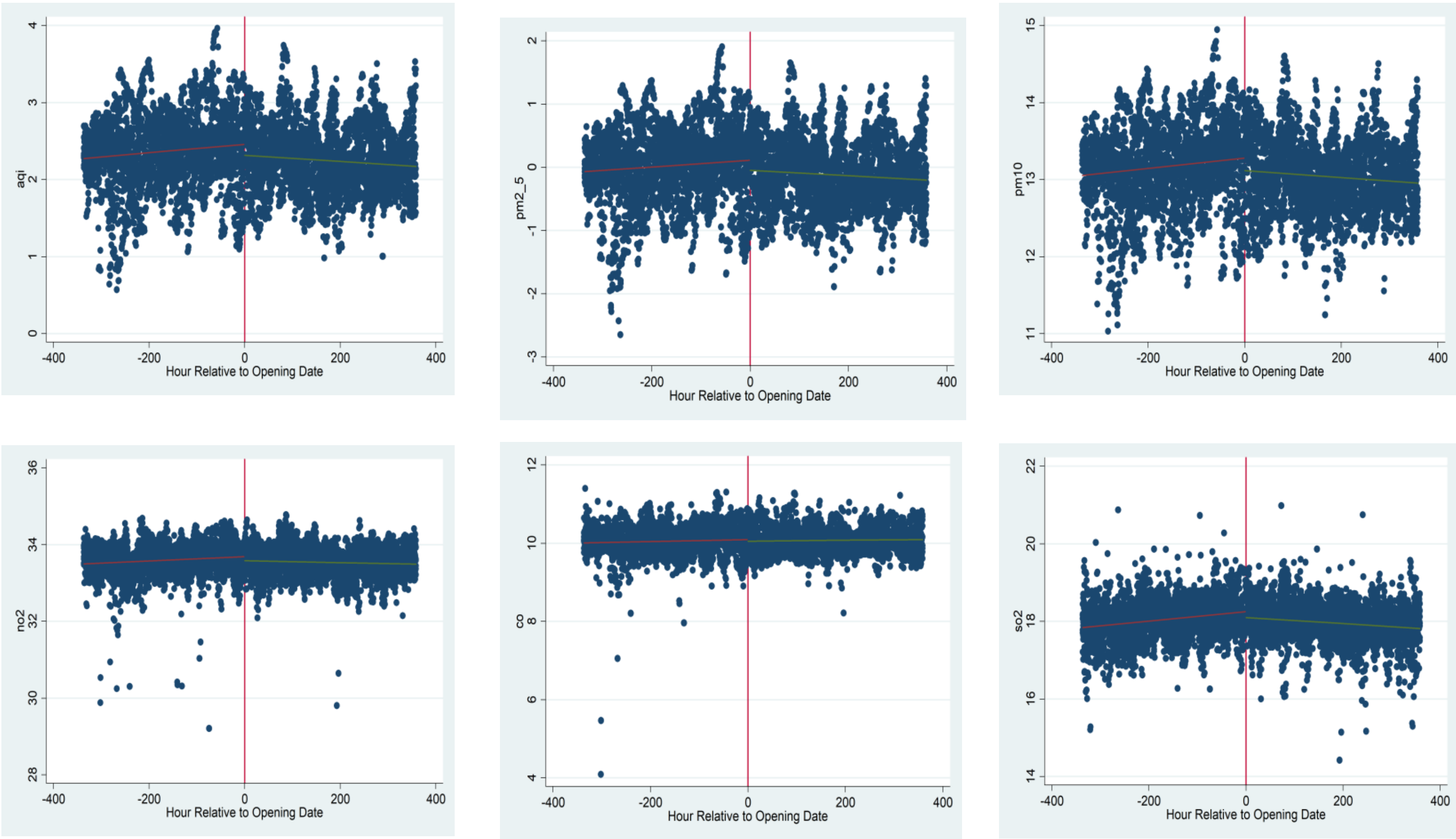

*Notes*: Panel A plots the Wald test statistics of $\beta_j = 0$ in equation (2), control variables including wind speed, precipitation, pressure, and dummies of holiday, dummy of hours in a day, heating period, weekday and city, and the interaction of weather and hour dummies. Panel B plots the coefficients of lags and leads of the opening day. Control variables are the same as Panel A.

## Figure 6. Weather Conditions Before and After Subway Opening Days



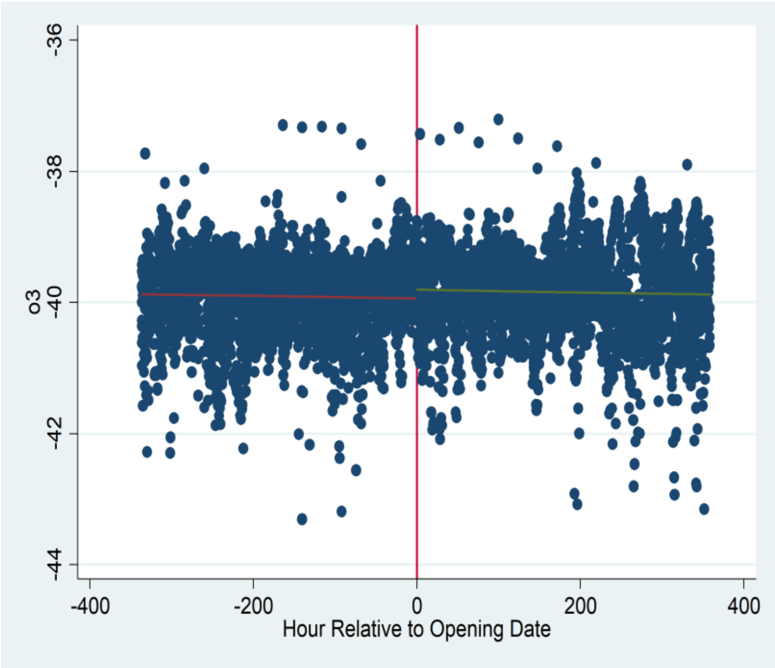

*Notes:* The four figures plot the distributions of wind speed, pressure, precipitation and temperature of the cities, respectively. The range is a month before and after subway opening. The dots are fitted by linear regressions separately for the days before and days after the openings.

## Figure 7. Distribution of Adjusted Pollutants Caused by Traffic on Time




*Notes:* The plots are the pollution value after taking out the effect of control variables and dummies. The time range is within the bandwidth used in the main regression. The red and green lines are the fit line by linear regression.

Figure 8. Distribution of Adjusted Pollutants Not Caused by Traffic on Time




*Notes*: The plots are the pollution value after taking out the effect of control variables and dummies. The time range is within the bandwidth used in the main regression. The red and green lines are the fit line by linear regression.