Tenure Stability and Environmental Performance: A Study of Chinese Cities

Chun Song, Purdue University, song404@purdue.edu
Juan Sesmero, Purdue University, jsesmero@purdue.edu

Selected Paper prepared for presentation at the 2017 Agricultural & Applied Economics Association Annual Meeting, Chicago, Illinois, July 30-August 1

Abstract

We empirically examine the link between Chinese local leader's tenure stability and their environmental performances. Capturing this relationship is difficult because tenure are not randomly assigned (selection bias), and because tenure length may be affected by performance evaluation that includes measures of air quality (endogeneity). We address these issues by combining generalized propensity score with a control function estimator. Results show that optimal tenure length, in terms of air quality, should be at least five years, while very long tenure is associated with less effective air pollution reduction. These findings reconcile theoretical predictions with empirical evidence. They also shed light on the reasons behind apparently inconsistent findings in previous empirical work.

1 Introduction

China is facing severe air pollution. The annual average $PM_{2.5}$[1] in the country was 58 $\mu g/m^3$ in 2016, almost six time higher than WHO standard (WHO, 2006). In 2016, 254 out of 338 cities failed to meet the national ambient air quality standard (NAAQS). Major sources of air pollution are power plants, cement and steel industries, as well as vehicles (Wang & Hao, 2012). The pervasive air pollution is associated with huge treatment cost which is around $140-160 billion annually (Crane & Mao, 2015), while posing increasing concern to Chinese public health.

Since the beginning of 21st century, the Central Government has been addressing air pollution problem in its 10th, 11th and 12th Five Year Plans (which play fundamental role in establishing national development strategies). In 2013, Central Government issued Air Pollution Prevention and Control Action Plan and later revised its Air Law. The Plan, commonly known as Ten National Guidelines, anticipates an overall improvement of air quality nationwide within five years, as well as a significant drop of pollution level in regions with high population density. These national directives has established a comprehensive framework for local government to implement their own policies, with emphasis on approaches that are expected to substantially improve air quality, such industrial restructuring, more efficient coal use, the promotion of clean vehicles, prevention of transboundary pollution, etc.

However, when these central policies get implemented at local level, local agents seems to have preferences of those short term approaches rather than mid-long term ones. Compared to fast measures such as vehicle control (Huang, 2014), temporarily shutting down factories and restaurants (Wang X. , 2016), and spaying water on street more often, long term approaches that have been strongly promoted by central government are much less adopted. Meanwhile, most city

[1] $PM_{2.5}$ is fine particulate matter less than 2.5 microns in in aerodynamic diameter.

leaders tend to stress that the air quality in their areas will be improved within a short time period (Beijing News, 2013; NewsChina, 2014, Linfen News, 2016)., ignoring the fact that air pollution treatment has been proved to be highly time consuming in many developed countries that used to face similar issue. There seems to be a mismatch between central initiatives and local reactions in terms of air pollution mitigation. If such preference continues to shape local policies, it will be very difficult to essentially improve air quality in the country.

Studies have found various institutional factors that might contribute to this implementation gap (Ghanem, 2014; Qin, 2015; Zhou, 2009), one factor that has gained increasing attention recently is the short tenure, or frequent switch, of local leader. This claim has been strongly suggested previously by Eaton & Kostka (2014) in the case of central government's green initiatives. We believe that this is likely to be an important reason that hinders the implementation of central directives regarding air pollution treatment at local level. The next section offers a brief analysis of the institutional setting at local China and the policy impact of short tenure.

2 Chinese local tenure system and its impact on environmental policy

After Chinese central government releases the overall guidelines, these directives will be conveyed to provincial level and then to the municipal. There are two chief executives on the municipal level, the mayor (Chinese : 市长) and the Secretary of City Party Committee[2]（Chinese:市委书记). The Secretary is the highest-ranking official in the city and has more control power in making important decisions and local policies (Zhang, 2006), therefore our study focus on secretaries and their policy making behaviors.

[2] Under some circumstances, the title of party secretary and mayor are held by the same person, but it is very rare.

In China, the legal term length for both leader is five years and no more than two consecutive terms. Yet in practice, tenure is not fixed: leaders can be removed (including being promoted, demoted or assigned to another position with equal administrative rank) any time and does not necessarily follow the national or provincial political cycle. The average tenure of secretary is 4.05 years from 2000 to 2014, with 65% of secretaries leaving the position before serving the first term in full. Such short tenure is believed to benefit both central government and local leader, making it difficult to alter in reality despite various official documents that attempts to stabilize local leader's tenure (Eaton & Kostka, 2014). Studies have argued that short time horizon the tenure of individual leader is likely to cast impact on their policy priorities. We summarize several possible channels through which leader's tenure might affect their policy priority.

First, under the merit-based promotion system (Wu, 2013), local leaders compete for limited promotion opportunities which rely highly one their actual work achievements. To win the "promotion tournament" (Li & Zhou, 2005), the most efficient way is to allocate energy and resource to those fields that can quickly produce tangible results, rather than to fields that tend to take long time to generate effects. Second, frequent leader switch can lead to discontinuous policy. In China, new leaders prefer to cease the old projects while creating new policies in order to show "competency and creativity". It also creates a highly uncertain policy making environment, or as the Chinese idiom describes: "issue order in the morning and change it at night" (Chinese: 朝令夕改). Third, the negative impact of short tenure might be exacerbated by the current adopted floating tenure system, under which the time horizon for leader is highly uncertain (Xu X. , 2010).

However, very long tenure is no necessary optimal: If the leading leader has been in office for longer than average, his expectation of being promoted will decrease and is likely to display passive action in implementing policies. It is also easier for business-government collusion to be formed, and leaders might benefit from the rent-seeking behavior of polluting companies. In some cities, executive leaders even become protector of polluting firms by intervening on law enforcement or help those firms pass environmental assessments.

Based on these facts, we argue that unstable tenure creates an unstable anticipation for local leader, thus affects their policy priority and preference, and eventually effect on air pollution mitigation. Generally speaking, short-tenured leaders are less likely to put effort in long term policies because they expect that such input is very unlikely to generate tangible improvement by the time their tenure ends. Such impact might be further strengthened by the powerful role played by local leaders in affecting implementation processes and outcome.

3 Literature review

There has been an extensive literature analyzing the impact of leader's time horizon on their policy making behaviors. According to government failure theory, government officials, who assumed to be social planners who maximize total social welfare in classic economics (Mankiw, 1986), also use their personal utility maximization when making gubernatorial decision (LeGrand, 1991). When the social optimal choices conflict with leader's personal objectives, government failure occurs, as the private benefit/cost of official from making certain decision does not match the social benefit/cost (Wolf, 1979).

Political science and political economy studies have shown how the length of tenure affects the politician's behavior. Olson (1993) believes that "Whenever an autocrat expects a brief tenure, it pays him to confiscate those assets whose tax yield over his tenure is less than their total value", while "stationary bandit" will monopolizes and rationalizes theft in the form of tax. Nordhaus (1975) uses a conceptual model to show that politician's choice between different political objectives can change over their tenures. Politician might manipulate of those objectives in short run for the purpose of winning election, and this is referred to as the political business cycle. Following his line of reasoning, later studies have incorporated more realistic assumptions (Hibbs, 1977; Alesina & Roubini, 1997).

More recent literatures generally believe that the current political cycles in many countries are not necessarily the ideal solution to cope with problems which easily encompass whole political or administrative careers to successfully manage them (Sprinz D. , 2012). It is the relationship between the relatively "fast electoral clock" and the long-term policy environmental policy challenges which poses the question of institutional design to master asynchronous clocks (Sprinz D. F., 2009). The complexity lies in that the short-term political cycle and long-term policy goals should be simultaneously modeled, in order to solve for an optimal option (Hovi, 2009). Yet more recent studies that rely on field investigations show a rather complicated tenure impact (Underdal, 2010). The notion of discounting has been used to illustrate policy problems with very long time-lags between action and effect. For local politicians, that means higher uncertainty in conducting such policy within their term (Hepburn, 2010). Study based on game theory also find that longer the tenure, the smaller the critical discount factor of regional environmental cooperation is, and the greater chance to sustain regional environmental cooperation (Li & Wang, 2016). A field research in China takes a close investigation on the tenure impact by interviewing stakeholders especially government officials (Eaton & Kostka, 2014).

Based on existing literature, it is likely that tenure length will affect local leader's behavior and policy choice towards air pollution treatment, and thus affect air quality. However, existing literature fails to reach conclusion on the direction of influence, and there is little qualitative evidence regarding this issue. Therefore, simply summarizing those pros and cons does not show a clear influencing direction. In order to test the effect of tenure on air quality, more empirical examinations are necessary.

4 Data and variables

We are interested in quantitatively examining the effect of tenure length (treatment) on air quality (outcome). A natural measure would be average tenure length in the city; but then two cities having same frequency of switching leader in a given time period would end up with same average tenure. This also result in a cross sectional data set with only 86 observations. A city-year panel also faces the challenge of precisely capturing tenure. Therefore, we use the individual leader (i.e. the city secretary) as our unit of analysis, which allows a straightforward measuring of tenure as well as leader characteristics, and is also easy to incorporate city attributes.

We measure air quality using the average growth rate of the Air Quality Index (AQI). The AQI is constructed as a piecewise function of major air pollutants. Chinese Environmental Bureau started reporting daily air quality data of 86 cities since 2000[3], and it was later updated in 2013 by adding more pollutants and changing the frequency of daily measurements. To ensure that outcomes are comparable, we examines the period before 2013. This is an unbalanced panel data as some cities began reporting AQI only after 2004 or after 2005. We aggregate data from daily to annually: We first calculate the annual AQI growth rate, then calculate the average AQI growth rate during the leader's tenure. Positive average growth rates indicate an increase in air pollution, the larger the number, the faster the air pollution growth speed, and vice versa. For those secretaries who only stay in office for only one year, the growth rates are calculated relative to the basis of

[3] In some documents, this index is called API before 2013, or Air Pollution Index, yet the official website calls this index "AQI". To be consistent with the data source, we refer this index as AQI. See website of Chinese Environmental Bureau Data Center: http://datacenter.mep.gov.cn/index)

the value from their predecessor. This also reduces the potentially impact of possible data manipulation brought forward by previous authors (Ghanem & Zhang 2014).

The treatment variable is tenure length. The original tenure data comes from a dataset published by Fudan University Social Science Center in 2015. The center collects information on secretaries in all major cities with tenure length measured in years. As indicated by the Center, if a secretary started his or her tenure between January 1st and June 30st, that year is counted as the first year in office. If the secretary's tenure starts between July 1st and December 31st, then next year is counted as the first year in office. The timely coarseness of this definition has many disadvantages. Therefore, we disaggregate the original data by measuring tenure in months.

We drop those secretaries whose tenures were not completed during the time period covered by this study. We also drop abnormal personnel changes such as death and being arrested. That leaves us with observations on 272 secretaries. Four secretaries out of 272 appear twice, they are considered as different observations because they have worked in different cities. We exclude data on the four provincial-level cities, or "直辖市" in Chinese: Beijing, Shanghai, Tianjin, and Chongqing, as they are under the direct governance of central government and their leaders are of higher administrative hierarchy than those of regular cities.

We include important secretary-level and city-level variables. Secretary-level characteristics include education, age, gender, etc. Moreover, whether a secretary was born in the same region can potentially affect the leader's preferences and knowledge of the area and, ultimately, his/her decisions. China has adopted the Rule of Avoidance to avoid localism and only very few leaders are actually from the same city. Yet, many leaders were born in the same province. Therefore, we have included a dummy variable that equals one if the city leader was born in the province where the city is located, and takes a value of zero otherwise. Previous studies shows that leaders from minority ethnic are more likely to experience longer tenures. Therefore we include a race dummy. We also adopt a similar method known as the Shih's method (Shih, Adolph, & Liu, 2012) to control for personal connections: we include an indicator variable to denote whether the city leader used to shares a workplace with the incumbent provincial Party secretary or governor.

Following previous studies, we include a time trend to capture trends in tenure length. We include dummies indicating whether the city is capital of the province, and whether the city is coastal city. Other controls were obtained from the China City Statistical Yearbook and include GDP per capita, and GDP growth rate during the secretary's tenure relative to that of the last secretary in the same city. These variables are intended to capture potential tradeoffs between economic and environmental goals embedded in the incentive structure faced by secretaries. Finally, we control for the initial level of AQI.

5 Method

5.1 Identification issue

Isolating causal effect of leader's tenure on air quality improvement is difficult because of two reasons: first, tenure are not randomly assigned (selection bias), and second, tenure may be affected by performance evaluation that includes measures of air quality (endogeneity).

The first source of bias arises from the complex and dynamic interactions between institutional/administrative factors, the environment, and economic development. One example we notice is that minority autonomous areas may display longer tenures as well better air quality. In China, some cities are known as ethnic minority autonomous areas. These "regional autonomies" guarantee ethnic minorities the freedom to use and develop their own ethnic languages, and to maintain their cultural and social customs. In addition, the central government has provided preferential treatment and aid to areas where ethnic minorities live. These area tend to have leaders that are from the same ethnic group as they have better understanding of local situation, and these leaders tend to serve longer tenures. This is to ensure stable governance in those areas with complex cultural and historic backgrounds. However, most of these cities are also characterized by a lower degree of urbanization and lower reliance on manufacturing and, thus, lower pollution level.

Table 1 Comparison between ordinary cities and minority autonomous cities

	AQI	Average secretary tenure	GDP per capita	Obs.
Minority autonomous cities	67.43	4.59	30.41	53
Ordinary cities	75.18	3.92	42.10	219
Total sample	73.67	4.05	39.82	272

The second potential source of bias comes from the merit-based promotion system that take a leader's environmental achievements into the promotion criteria. The link between a secretary's environmental achievements and his/her chances of promotion depend upon the weight of such achievements in promotion criteria, and the actual enforcement of such criteria. Some scholars have even argued that those involved in secretary promotion decisions may even be inclined to maintain a secretary in the same city, as a result of good environmental achievements (Sturm, 2004). This is not to be interpreted as a punishment, but as a way to reward environmental stewardship. It is, then, unclear if a reverse causality problem even exists; and if one exists, what is the direction of the causal effect. Existing evidence on this is mixed.

Zheng et al. (2014) finds that reduction of air pollution can contribute to higher chances of promotion. While Wu (2013) finds that spending on environmental improvements is significantly negatively related to the secretary's chances of promotion in that year. Therefore, the mechanistic existence and direction of reverse causality between tenure length and air quality improvements is unclear. Despite the mixed results from empirical studies, it is reasonable to expect an impact of outcome on treatment through promotion mechanism that potentially cause endogeneity.

5.2 Econometric strategy

In order to simultaneously address both selection bias and endogeneity, we adopt a procedure that combines Generalized Propensity Score (GPS) with instrumental variables. Specifically, we use GPS to reduce the selection bias in covariates. This approach has been widely used to restore randomness of continuous treatment. It was first outlined by Imbens (2000), and later developed by Hirano and Imbens (2004). GPS is defined as the conditional density of treatment (Hirano & Imbens, 2004). Two identification assumptions are made: first, conditional on the covariates, the treatment assignment is independent of the potential outcomes; this is the weak ignorability assumption which is not untestable. As explained in the data section, we carefully choose observable covariates to minimize the possibility that this assumption is violated. Second, the GPS model relies on a balancing property which requires that, given the propensity function, the conditional distribution of the treatment is independent of covariates (Hirano & Imbens, 2004). We will test this assumption later.

The GPS does not address the reverse causality issue. Here we incorporate instrumental variables to correct for endogeneity. The key to adequately addressing the endogeneity of tenure is to identify instruments that are both relevant (correlated with tenure) and exogenous (uncorrelated with the growth in AQI). Previous studies (Zhang and Gao, 2008) have used the secretary's age to instrument for tenure. Current personnel policies require the leaders of many cities to retire by a certain age – generally less than 65 years old. Therefore, leaders who are older at the time they start their tenure tend to serve shorter periods.

Figure 1 below shows the relationship between leaders' ages at the time when they start their tenures and the average tenure length. The size of the bubbles denote number of observations in the age cohort. The trend revealed in Figure 1 is consistent with our expectation: for young leaders who are in the early stages of their political careers, tenure is relatively long. This is either because they lack important personal connections, or because they have not yet made achievements deserving of a promotion. For relatively older leaders (e.g., 50 years old and above), tenure is typically much shorter. This is either because they are forced to retire, or because they quickly are promoted to another position before they reach the retirement age for the local position. We use a quadratic specification of age as an instrument for tenure in our first stage regressions. In addition to being relevant, the validity of age can be established: beyond its correlation with tenure, it is difficult to imagine the age of the secretary having any impact on air quality, as shown in the right plot in figure 1.

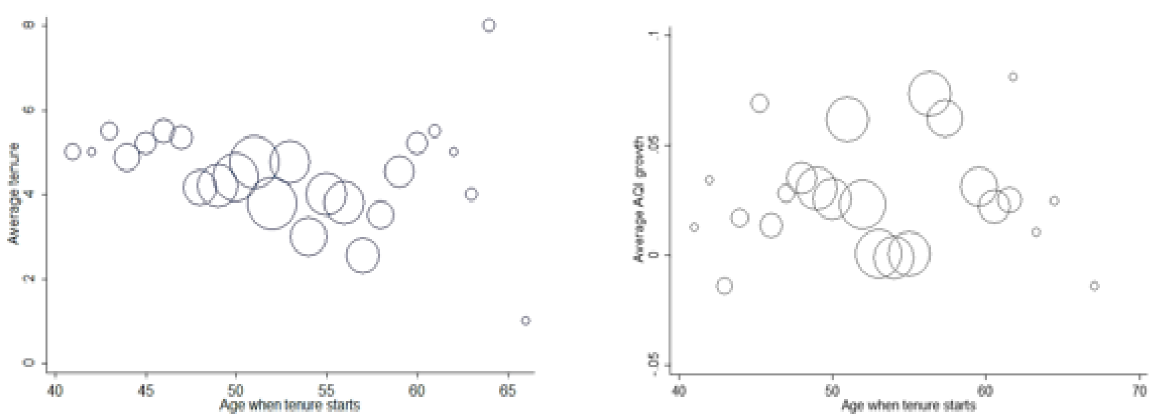

Figure 1 Scatter plots of age
(Left: age and tenure, right: age and AQI growth. Circle represents sample size)

With the GPS and instrumental variable, we combine these two approaches in our estimation. First, we estimate the conditional expected value of treatment given IV and covariates. This step is also known as selection function:

$$T_i = \beta_0 + IV_i\beta_1 + \boldsymbol{X}_i\beta_2 + u_i. \quad (1)$$

The predicted values, $\hat{T}_i$, from this regression are used to obtain the conditional probability of treatment (the GPS), under the assumption that the residual from equation (1) follows a normal distributional. Then the GPS is estimated as follows:

$$\widehat{R_i} = \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} exp\left[-\frac{1}{2\hat{\sigma}^2}\left(T_i - \hat{T}_i\right)^2\right]. \quad (2)$$

With the estimated GPS, we then estimate the outcome equation. We regress the outcome (the average AQI growth rate during a leader's tenure) on the tenure length, the estimated GPS from equation (2), their quadratic terms and interactions, the initial AQI. In particular, we include the residual from equation (1) in the regression to control for the endogeneity as in a standard control function approach:

$$A_i = \rho_0 + \rho_1 T_i + \rho_2 {T_i}^2 + \rho_3 \widehat{R_i} + \rho_4 \hat{R}_i^2 + \rho_5 T_i * \hat{R}_i + \rho_6 AQI_{i0} + \rho_7 \hat{u}_i + \eta_i. \quad (3)$$

The failure to reject the null hypothesis that $\rho_7 = 0$ would indicate the existence of endogeneity. The GPS is used to adjust for any imbalance in the covariates. Finally, we generate the dose response function by using equation (4) to calculate the average outcome over a particular treatment level:

$$E\left(\hat{A}_i|t\right) = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{\rho}_0 + \hat{\rho}_1 \mathrm{t} + \hat{\rho}_2 \mathrm{t}^2 + \hat{\rho}_3 \hat{R}(t, X_i) + \hat{\rho}_4 \hat{R}(t, X_i)^2 + \hat{\rho}_5 t\hat{R}(t, X_i) + \hat{\rho}_6 AQI_{i0} + \hat{\rho}_7 \hat{u}_{\mathrm{i}}\right)$$

(4)

That is, the estimate GPS is averaged over X_i but evaluated at each $t \in (\underline{t}, \overline{t})$ where $(\underline{t}, \overline{t})$ represents the range of support over which the response function is evaluated.

We also estimate the following three models as comparison to our main model: (1) OLS regression that does not address either source of bias; (2) estimation based on control function that addresses the potential endogeneity but not the selection bias; and (3) a standard generalized propensity score that address selection bias but not endogeneity. In the OLS model, we regress A_i on tenure variables and all the covariates as well as initial AQI; in the model with control function estimator, we first regress tenure on instrumental variables and all the covariates, then use the residual in estimating second stage equation; in the GPS model, we regress tenure on covariates but not instrumental variables, and use the residual in estimating GPS. Comparison across models reveal important information regarding this source of endogeneity.

6 Results

Table 2 reports the result from estimating equation (1). The dependent variable is tenure length of secretary as defined previously. The results suggest that tenure length is negatively correlated with starting year, indicating that tenures are generally getting shorter. The secretary's performance in terms of economic growth is negatively correlated with tenure length, perhaps suggesting that high economic growth increases the secretary's chances of promotion, as implied by previous literatures (Wu, 2013). The instruments turns out to be statistically significantly correlated with treatment level with Consistent with our prior expectations, tenure length is negatively correlated with age.

Table 2 Estimation of the selection function

Variable	Estimate	Std. Error	P (>|t|)
(Intercept)	293.948***	82.155	0.000
Age	-0.808***	0.237	0.001
Age2	0.008***	0.002	0.001
Experience=1	0.264	0.258	0.309
Year when tenure starts	-0.134***	0.041	0.001
Male=1	0.256	0.657	0.697
Han nationality=1	-0.161	0.453	0.722
Bachelor degree=1	0.336	0.531	0.528
Master degree=1	0.263	0.532	0.622
PhD degree=1	0.604	0.590	0.307
Northern city=1	-0.314	0.387	0.418
Eastern city=1	-0.045	0.325	0.891
Middle city=1	-0.632	0.453	0.164
Northwestern city=1	-0.661	0.543	0.225
Northeastern city=1	-0.083	0.495	0.866
Southern city=1	-0.255	0.292	0.383
Leader from local=1	0.060	0.252	0.813
Capital city=1	0.171	0.280	0.540
Coastal city=1	-0.415	0.360	0.250
Minority dwelling area=1	0.320	0.339	0.347
Relative GDP growth rate	-0.869**	0.506	0.087
GDP per capita	0.001	0.003	0.739
R^2	0.13		
Obs.	272		

Notes: (1) Robust standard error

(2) *** significant at 1% level, ** significant at 5% level, * significant at 10% level

We then use the residuals from the above estimation to estimate the GPS using the formula given by Equation (2). We check the balancing property using the algorithm called "blocking the score" (Hirano & Imbens, 2004). First, observations are divided into groups (e.g., three) based on the quantile of tenure length. For group one, we estimate the GPS with respect to the median length of tenure in the group, and stratify observations into blocks based on the quantile of GPS. For observations in the rest two groups, we calculate their GPS with respect to group one and assign them into corresponding blocks. This way we are able to define treated and control observations: by definition, GPS is the conditional probability of receiving a certain level of treatment, those observations that have similar GPS but actually belong to different tenure groups are therefore considered counterfactuals.

Then we calculate the mean difference of covariate values. This exercise is repeated for multiple GPS blocks within each treatment group, and for each of the covariates in the selection function. A weighted average of those differences for each covariate is used to calculate the t-statistics.

Figure 2 reports the difference in mean t-statistics for all 21 covariates before and after adjusting for GPS. Large differences in covariate means across groups, translate into high dispersion of points around zero. If differences fall between -2 and 2, mean across treatment groups are not statistically significantly different from zero, and the data is balanced. The fact that many points fall outside of this range before balancing, and within this range after balancing, suggests an effective balancing by the GPS approach.

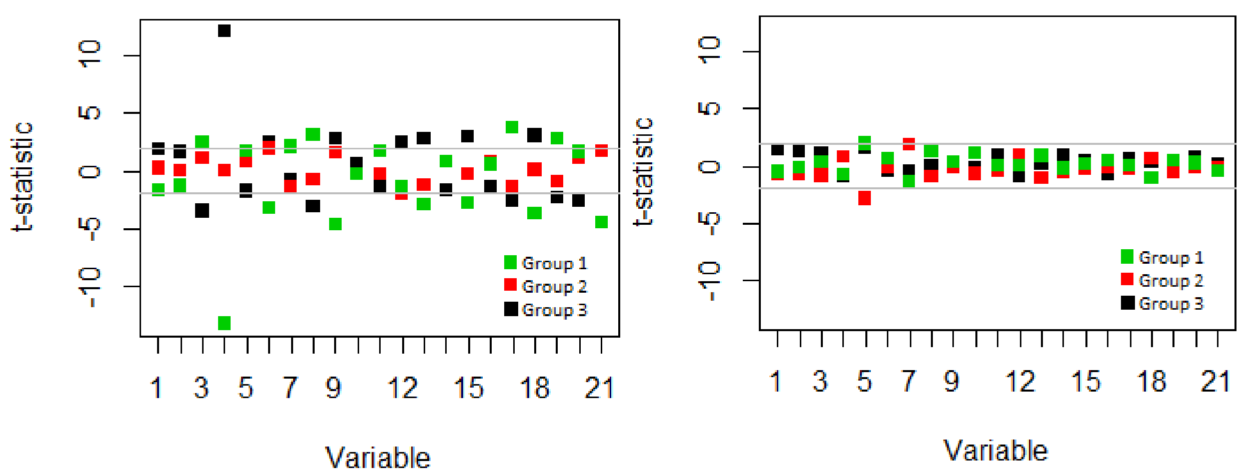

Figure 2 Balance of covariate before and after accounting for the GPS.
(Left: unadjusted t statistics; right: GPS adjusted t-statistics).

With the improvement in balancing property, we now proceed to estimate the outcome model using equation (3). Results are reported in Table 3. Both tenure variables have statistically significant (and quantitatively relevant) effects on air quality outcome. Both linear and quadratic terms of tenure are statistically significant, confirming our prior expectation of a nonlinear relationship between tenure length and pollution growth hypothesis. Moreover, the residual from first step is significant at 1% level, suggesting the existence of endogeneity. Finally, results show that the initial AQI level still picks up the non-linear nature of the equation of motion for air pollution.

Table 3 Result from estimating the outcome function

Variable	Estimate	Std. Error	P (>|t|)
Intercept	0.169	0.087	0.053
$Tenure$	-0.076***	0.025	0.003
$Tenure^2$	0.006*	0.003	0.086
GPS	0.060	0.688	0.930
GPS^2	1.800	1.142	0.120
$Tenure$* GPS	-0.032	0.061	0.601
Residual	0.054***	0.011	0.000
Initial AQI	-0.001***	0.000	0.000
R^2	0.25		
Obs.	272		

Notes: (1) Dependent variable: average annual growth rate of AQI over tenure.

(2) *** significant at 1% level, ** significant at 5% level, * significant at 10% level

Figure 3 displays the average impact of tenure on outcome at each treatment level. Results reveal a non-linear U-shaped relationship between tenure length and average pollution growth rate. As tenure becomes longer, the pollution keeps increasing but in a slower speed. This is consistent with theoretical predictions, in that short tenure tends to shorten the leaders' time horizon, incentivizing ill-designed and short-sighted policies (e.g. Sprinz 2012). After tenure reaches five years, pollution growth becomes negative, and the vertex happens in eighth year which is associated with an effective pollution reduction rate during the leaders' administration. For tenure longer than 8 years, pollution growth is negative but increasing, meaning that the pollution mitigation is slowing down. This is also consistent with theoretical predictions, which postulate that long tenure periods may be conducive to 1) collusion between local governments and special interests, and 2) the "quiet life" effect by which lack of incentives hinder efforts to reduce pollution.

On the other hand, in line with the arguments above, medium-to-long-length tenure periods might weaken the forces underlying the lack of environmentally friendly policies. Results seem to confirm this hypothesis as medium-length tenure periods (8 years) seem to prompt the most effective improvement in air quality, as opposed to simply slowing down the rate of pollution growth.

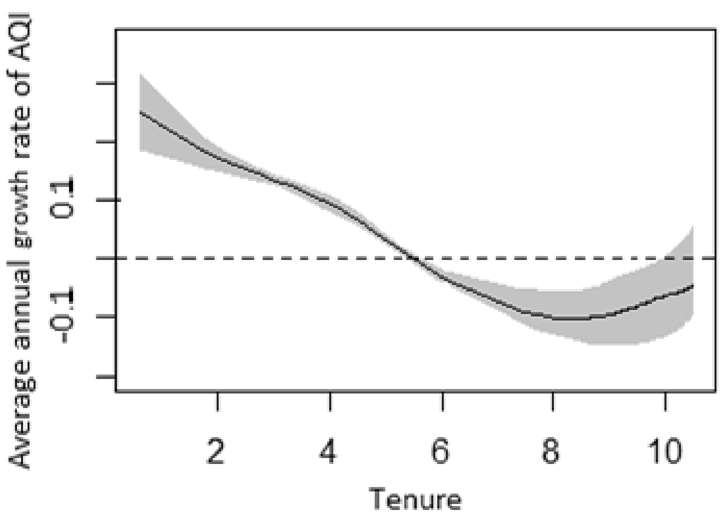

Figure 3 Dose response curve based on GPS-IV estimation

We also test the relationship with linearity restriction by excluding the quadratic term in the outcome function. The estimated coefficient on tenure is -0.05 and is significant at 5% level. This suggests that the overall trend associated with longer tenure is decreasing pollution growth, which can be seen from Figure 3 as well. Yet such linear setting veils the nonlinear relationship and is subject to b endogeneity.

In Table 4, we compare the results from the main model to the rest three models described in section 5. First, in terms of statistical significance, the GPS with instrumental variable is the only model that shows significant impact of tenure on air quality. This suggests that biases from non-random selection of tenure length and reverse causality can mask this important institutional link. Addressing at least one of these sources results in a U-shape link (albeit statistically weak) between tenure and pollution growth rate. Addressing both simultaneously also results in a U-shape link between treatment and outcome, but the link is also statistically significant.

Table 4 Comparison of results from different models

	OLS	Control Function	GPS	GPS with IV
Tenure	0.014	-0.014	-0.015	-0.076***
$Tenure^2$	0.001	0.001	0.001	0.006***
R^2	0.27	0.36	0.11	0.25
Obs.	272	272	272	272
Address endogeneity	No	Yes	No	Yes
Address selection bias	No	No	Yes	Yes

To further test such non-linear relationship supported by the results, we extract the subsample from the dataset whose tenures are at least eight years long, and test the trend associated with the pollution growth rate over their tenures. There are 17 leaders who have tenure at least eight years long. If the tenure is longer than eight years, we examine the first eight years. We calculate the mean as well as median AQI growth rate for each year over their tenures as plotted in figure 9. Overall we find consistent statistical evidence with our empirical results: for those long tenured leaders, both the mean and median shows that the pollution growth tend to be a little higher at the beginning of administrations, maintaining relative low or negative during the tenure until the tenure reaches a rather long length.

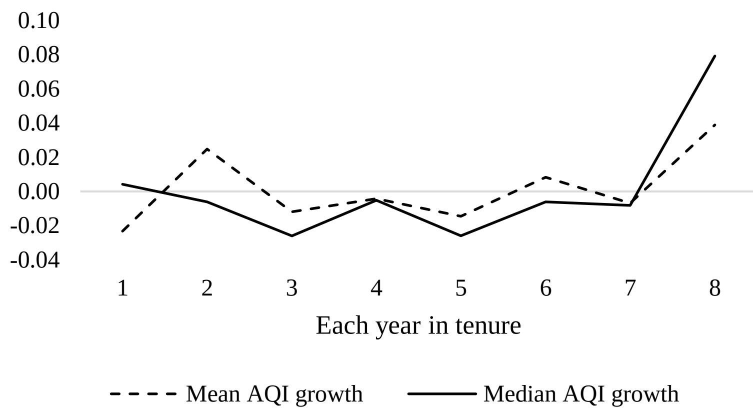

Figure 4 Average AQI growth rate for each year over tenure

7 Discussion

This paper examines the empirical relationship between tenure length and air pollution, measured as average growth rate of air pollution. Inspired by theoretical predictions, we hypothesize that the relationship between these variables may be best characterized by a U-shape function. While correlations between tenure and air quality are easily detected, identification of the causal effect of tenure on air pollution is hampered by multiple sources of bias, i.e. endogeneity and selection bias.

We combine a Generalized Propensity Score approach with a control function approach to identify treatment effect. Our results confirm our prior expectations and reveal that very short tenures (< 5 years) result in substantial increases in air pollution, while very long tenures (> 8 years) is associated with negative yet increasing growth of air pollution, meaning that after a critical point, the mitigation of air pollution slows down. The results advances our understanding of the role of institutional design on air quality in two key ways. First, it reconciles apparent inconsistencies between empirical studies. Second, it reconciles inconsistencies between theoretical predictions and empirical findings.

Other empirical studies have concluded that short tenure periods lead to increased air pollution and, by linear extrapolation, longer tenure always leads to pollution reduction (Hovi, 2009). This is, however, inconsistent with theoretical predictions from the political economy literature. Theoretical analyses anticipated that very long tenure periods can also lead to increased pollution due to collusion between political leaders and special interest, and/or due to a lack of incentives to implement abatement policies (Eaton and Kostka 2014). Our analysis reveals that the monotonic link between tenure length and air quality may have been an artifact of the linearity assumption imposed in those studies. By accommodating a non-linear treatment effect, our study reveals that empirical evidence aligns with theoretical predictions; very short tenure seems to lead to environmental deterioration, while very long tenure is not necessary the optimal choice either with respect to pollution mitigation.

Our findings suggest that the current floating tenure system of local Chinese governments might create perverse incentives for local leaders. This system adds an element of uncertainty or instability to tenure length, creating incentives for local leaders to allocate resources in fields that can generate observable achievements, which possibly translates into short-sighted behaviors on the part of local leaders. Therefore, local governments have, on average, limited success in implementing the central government's directives. This is perhaps the most likely explanation for what has come to known as the "implementation gap". While our analysis suggests that increased stability of tenure might help close the implementation gap, it also shows that tenure periods should not be too long. In fact, our analysis indicates that when tenure length exceeds eight years, pollution reduction is slower.

The central government in China has shown some awareness of the disadvantages associated with short, unstable tenures. In fact, certain directives by the central government echo the principle that "success does not have to be realized in my tenure" (Eaton & Kostka, 2014). This principle was embraced by President Xi during his administration at Jiangxi Province. However, the central government was slower to implement actual incentives operationalizing this principle. Appropriate incentives would have to, somehow, attribute improvements in certain indicators to past decisions by local leaders, and reward that leader, even after leaving office. But linking changes in outcome variables to specific decisions made in the past can be extremely challenging. In this sense, "success does not have to be realized in my tenure" can also be understood as "mistakes do not have to be punished in my tenure". Perhaps one way to incentivize long term input of local leader is reward nominally positive actions with longer payoff periods. Broadly speaking, central government can incorporate diverse measurements into the evaluation system instead of solely measuring the apparent and tangible achievements.

References

Alesina, A., & Roubini, N. (1997). *Political cycles and the macroeconomy.* Cambridge: MIT Press.

Chan, C. K. (2008). Air pollution in mega cities in China. *Atmospheric environment, 42*(1), 1-42.

Eaton, S., & Kostka, G. (2014). Authoritarian Environmentalism Undermined? Local Leaders' Time Horizons and Environmental Policy Implementation in China . *China's Quaterly*, 359-380.

Fleisher, B. L. (2010). Human capital, economic growth, and regional inequality in China. *Journal of development economics, 92*(2), 215-231.

Forman, S. E. (1993). *The life and selected writings of Thomas Jefferson.* New York: Random Hous.

Hepburn, C. (2010). Environmental policy, government, and the market. *2*(26), 117-136.

Hibbs, D. (1977). Political parties and macroeconomic policy. *American political science review*.

Hirano, K., & Imbens, G. W. (2004). The propensity score with continuous treatments. *Applied Bayesian modeling and causal inference from incomplete-data perspectives*(226164), 73-84.

Hovi, J. S. (2009). Implementing long-term climate policy: Time inconsistency, domestic politics, international anarchy. *3*(9), 20-39.

Kostka, G. (2014). *Barriers to the implementation of environmental policies at the local level in China.* Washington, DC: World Bank Group.

LeGrand, J. (1991). The theory of government failure. *British journal of political science, 4*(21), pp. 423-442.

Li, H., & Zhou, L. (2005). Political turnover and economic performance: the incentive role of personnel control in China. *Journal of public economics, 89*(9), 1743-1762.

Li, M., & Wang, Q. (2016). Study on the Impact of Government Term on Regional Environmental Cooperation Stability Based on Game Theory. *China population resource and environment (Chinese version), 3*(26), 83-88.

Mankiw, N. G. (1986). Free entry and social inefficiency. *The RAND Journal of Economics*, pp. 48-58.

Nordhaus, W. D. (1975). The Political Business Cycle. *The Review of Economic Studies*, pp. 169–190.

Olson, M. (1993). Dictatorship, Democracy, and Development. *American Political Science Review, 87*(03), 567-576.

Qin, L. (2015, 4 9). *China's new Air Pollution Law omits key measures in war on smog*. Retrieved from China Dialogue: https://www.chinadialogue.net/article/show/single/ch/8156-China-s-new-Air-Pollution-Law-omits-key-measures-in-war-on-smog

Saich, T. (2010). Diversity with Unity. In T. Saich, *Governance and politics of China* (pp. 13-16). New York: Palgrave Macmillan.

Shih, V., Adolph, C., & Liu, M. (2012). Getting ahead in the communist party: explaining the advancement of central committee members in China. *American Political Science Review, 106*(01), 166-187.

Sourthern China Post. (2011, 12 7). *东莞林业局局长胡炽海：领导一换届 树也跟着换 (Director of the Forestry Bureau of Dongguan: when leader rotate, trees also change)*. Retrieved from Nanfang City News: http://city.sina.com.cn/focus/t/2011-12-07/103725508.html

Sprinz, D. (2012). Long-Term environmental policy: challenge and research. *9*(3), 67-77.

Sprinz, D. F. (2009). Long-term environmental policy: definition, knowledge, future research. *3*(9), 1-8.

Underdal, A. (2010). Complexity and challenges of long-term environmental governance. *Global Environmental Change,, 3*(20), pp. 386-393.

Wang, X. (2009). Local Governor transition and economic growth. *8*(4), 1301-1328.

WHO. (2006). *Air quality guidelines: global update 2005: particulate matter, ozone, nitrogen dioxide, and sulfur dioxide.*

Wolf, C. (1979). Theory of nonmarket failure: framework for implementation analysis. *The Journal of Law and Economics*, pp. 107-139.

World Bank. (2007). *Cost of pollution in China: Economic estimates of physical damage* . Washington, D.C.: World Bank.

Wu, J. D. (2013). Incentives and outcomes: China's environmental policy. *NBER Working paper 18754*.

Xu, X. (2010). Behavior of Regional Governor in China: Appointment System. *China Economic Quarterly, 9*(4), pp. 1447-1466.

Zhang, X. (2006). Fiscal decentralization and political centralization in China: Implications for growth and inequality. *Journal of comparative economics, 34*(4), 713-726.

Zuo, C. V. (2015). Promoting city leaders: the structure of political incentives in China. *The China Quarterly, 224*, 955-984.