# Pollution Regulations, Air Quality, and the Local Economy

Chen, Ying

School of Economics The Wang Yanan Institute for Studies in Economics, Xiamen University

February 2020

# Pollution Regulations, Air Quality, and the Local Economy

Ying Chen*

February 2019
(latest version here)

## Abstract

Air quality is an important amenity that affects the labor supply to a local economy while regulations aiming at improving it can be costly and consequently reduce the local labor demand. This article studies how air quality and its regulation respectively and jointly affect the local economy through these two channels by exploiting China's first national air pollution regulation and migration reform as natural experiments. I propose an instrumental variable for local pollution levels by applying rich remote-sensing data to the engineering considerations of power plant construction. The estimation results suggest that heavy air pollution has driven out high-skilled workers when migration costs fall, while the regulation to curb pollution has led to a reduction in manufacturing employment in targeted locations and sectors. Additional results show relatively slower firm and wage growth in more regulated prefectures and sectors, and a modest local employment reallocation from heavy-to light-polluting industries.

---

[*] School of Economics & The Wang Yanan Institute for Studies in Economics, Xiamen University, Xiamen 361000, China. ychen@xmu.edu.cn. Financial support from the Fundamental Research Funds for the Central Universities (grant number 20720201045) is acknowledged. I am extremely grateful for the generous guidance from Vernon Henderson. This paper also benefited greatly from my discussions with Olmo Silva, Henry Overman, Maisy Wong, Gilles Duranton, and Steve Gibbons. I am solely responsible for any errors.

# 1 Introduction

Environmental regulations are becoming more prevalent in developing countries as they begin to tackle pollution. How does such attempt in achieving better air quality affect local economic growth? Evidence based on developed country experience suggests that better air quality is a desirable amenity that attracts human capital, which benefits the local economy (Glaeser et al., 2001; Chay and Greenstone, 2003; Heblich et al., 2016; Lin, 2018). In order to achieve this, however, the corresponding regulations often impose costs on the industries and workers in regulated locations, which can negatively affect the local economy through reduced labor demand (Becker and Henderson, 2000; Greenstone, 2002; Walker, 2013; Roback, 1982).

The trade-offs between the potential benefits and costs of tightened environmental regulations in consideration of local economic growth is particularly ambiguous for developing countries. The recent rapid structural change and industrialization in many developing countries make it difficult to assess the efficiency of environmental regulations and its resulting impact on the local labor demand. On the other hand, the body of empirical literature on how people's preferences for clean air in developing countries affect the local labor supply differentials across locations is limited and inconsistent. Moreover, the current literature has so far paid little attention to the joint local labor demand and supply implications of pollution regulation and environmental conditions, which is of keen interest to policy makers.

This paper studies the impacts of China's first national pollution regulation on local labor demand while taking into account the contemporaneous air quality effect on local labor supply across space, in order to analyze their respective and joint impacts on the local economy and contribute to our understanding of this question. Following the Ninth Five-Year Plan, China started implementing the Air Pollution Prevention and Control Law (APPCL) in 1998 by setting a national air quality standard for all prefectures to achieve. As a result, prefectures with higher $SO_2$ pollution levels in 1998 were subjected to more stringent regulation in the following years. Concurrent to this regulation was the reform in China's labor mobility restrictions (*hukou* system), which relaxed the previously curtailed internal migration and allowed workers to relocate across space (Chan, 1994; Au and Henderson, 2006a). All else being equal, the mobility reform would allow workers to move away from strictly regulated prefectures due to the reduced labor demand, their preference for clean air, or a combination of the two.

The analysis of this paper exploits these two concurrent policy changes to empirically estimate the resulting local employment and city growth outcomes, following the framework of the Rosen-Roback model of spatial equilibrium. While the existing $SO_2$ pollution level serves as a reasonable proxy for both the local regulation intensity (Shi and Xu, 2018) as well as air quality, its non-random distribution across space poses identification challenges. Prefectures with higher $SO_2$ levels tend to have endogenously distributed industrial compositions or productivity levels (Chay and Greenstone, 2003). Additionally, development process such as urbanization can further exacerbate levels of air air pollution (Zheng and Kahn, 2013), raising concerns of reverse causality. To address such challenges, I propose a novel instrumental variable (IV) strategy using the power plant suitability index widely applied in Civil Engineering practices for its site selection[1] (e.g., Barda et al., 1990; Pohekar and Ramachandran, 2004; Choudhary and Shankar, 2012). The suitability index is a nonlinear transformation of a set of locational characteristics to measure cost, accessibility, and safety considerations in building power plants, which I construct using detailed remote-sensing data. As thermal power plants consistently contribute to about 50% of China's $SO_2$ emissions (Lu et al., 2010), the suitability index captures a location's likelihood to have higher ambient $SO_2$ levels because of its potential, rather than actual status, of hosting thermal power plants[2].

The baseline identification assumption is that a prefecture's potential to have hosted thermal power plants at the time of the two policy changes will affect its future economic outcomes only through its increased likelihood to face stringent regulations due to worse local $SO_2$ pollution, conditional on regional fixed effects, individual components of the suitability index, climatic controls, and covariates of pre-existing political and economic characteristics. Controlling for each of the comprising factors of the suitability index is crucial in ensuring that the identifying variation comes from the non-linear transformation functions and weights determined by civil engineers for largely non-economic considerations. Another key institutional setup to support the validity of the exclusion restriction is that electricity distribution and pricing are centralized across regional offices and supervised by the State Council, so firms and households located near power plants do not enjoy preferential electricity rates. I conduct balance tests given available data to show that prefectures more suited to host power plants do not have differential pre-existing growth trends in terms of city size, human capital skill composition, manufacturing and total employment. In addition, the baseline results are highly robust to a wide range of different specifications, restrictive samples, and alternative computations of the instrumental variable.

The estimation results find that China's environmental regulation has caused a substantial reduction in local industrial employment growth in the regulated prefectures relative to non-regulated ones. As the secondary employment in an average Chinese prefecture by 2010 only accounts for 22% of its total employment, the negative regulatory effect appears to have no impact on the overall employment growth. The contemporaneous *hukou* reform, on the other hand, have reduced migration costs and led to greater outflow of high-skilled workers (i.e., workers with college education and above) in high pollution prefectures relative to low pollution ones. However, changes in the population of low-skilled workers – whose employment prospects were more adversely affected by the APPCL – do not show any significant differences.

In order to better interpret these overall reduced form effects, I proceed to investigate the potential channels at work. The first question that I address is to confirm that the observed negative employment impact is indeed driven by environmental regulation (i.e., the productivity-based channel), rather than the alternative cognitive- or health-based channels that pollution level itself adversely affects worker productivity or work duration (Chang et

---

[1] The state of Wisconsin, for example, has its own *Common Power Plant Siting Criteria* set by its Public Service Commission.

[2] The actual power plant locations may still be endogenous for reasons such as political nepotism.

al., 2019; He et al., 2019; Currie et al., 2009; Hanna and Oliva, 2015; Fu et al., 2017; Kahn and Li, 2019). As the APPCL has stipulated additional costs on heavy coal-users, unaffected employment growth patterns in the primary and tertiary sectors serve as the first piece of evidence supporting the negative regulatory impact. Further heterogeneity analysis using detailed industrial firm survey data also shows that this effect is driven by the heavy coal-using (i.e., $SO_2$-emitting) and state-owned firms. Such heterogeneity by polluting intensity and ownership type would be unlikely if the adverse employment effect was driven by prefecture-level heavy pollution. Therefore, the APPCL lowers the local labor demand growth in the $SO_2$-emitting industries via the productivity-based channel by imposing additional costs to their production.

The additional results are in support of a productivity-based channel and shed light on the local labor market adjustments to environmental regulations. The main findings are that: (i) there is evidence of some employment reallocation from the regulated heavy polluting industries to the non-polluting ones locally, but not enough to negate the adverse regulatory impact; (ii) the negative regulatory impact occurred in the first half of the sample period while the employment reallocation to non-polluting industries happened in the later half; and (iii) patterns in the changes of firm counts and wage suggest that more stringent environmental regulation has adverse impacts on the local economy along both the extensive and intensive margins: it leads to firm exits as well as reduction in worker productivity.

Next, I proceed to interpret the asymmetric $SO_2$-induced migration patterns of high- and low-skilled workers in response to the *hukou* reform that reduced migration cost. Migration restriction prior to the *hukou* reform indicates that the spatial distribution of population could adjust to workers' locational preferences and labor market prospects as migration cost declines after the reform. The asymmetric migration response in this context relates to (i) the migration cost reduction is skill-biased and hence more salient for high-skilled workers than low-skilled workers[3]; and (ii) the negative environmental regulatory impact disproportionately affects low-skilled workers' employment prospects because of the skill composition of the regulated industries[4]. The outflow of high-skilled workers in high pollution prefectures relative to low pollution ones, therefore, reflects high-skilled workers' preferences for clean air as they "vote with their feet" once migration cost is reduced.

Meanwhile, the lack of a statistically significant estimate for the population change in low-skilled workers implies that their post-reform migration cost remains relatively high. A similar asymmetric migration pattern by urban-rural *hukou* status corroborates the importance of considering different migration costs by skill[5]. Additionally, it is also possible that low-skilled workers' preference for clean air is weaker than high-skilled workers. Although the scope of this paper is limited in disentangling these two channels, it is clear that the low-skilled workers bear the regulatory cost of environmental protection policies.

---

[3] Section 2 provides institutional details on the reform and its skill-biased nature.

[4] According to the Chinese Household Income Project, one of the most comprehensive and representative household survey series for China, more than 95% of the secondary sector employment is comprised of low-skilled workers during my sample period.

[5] Holders of rural *hukou* face additional barriers when moving across prefectures (more in Section 2).

The findings in this paper, that the environmental protection policies impose a negative regulatory impact on the local labor market, reconciles some conflicting patterns seen in developed versus developing country experience. While sizable evidence in developed countries substantiates considerable costs of environmental regulation on industry profits and local employment (e.g., Becker and Henderson, 2000; Greenstone, 2002; Walker, 2013), the existing evidence based on China mostly finds non-negative relationships (Tanaka et al., 2014; Yang et al., 2014; Zhao et al., 2018; Li et al., 2019)[6]. The common difference-in-differences (DID) design and the lack of pre-period data raise concerns over such findings, particularly in China's context of rapid industrialization. This paper is the first to document an adverse regulatory impact of China's first wave of environmental protection policies on the local employment, which is contrary to existing work but consistent with findings from developed country experience.

The $SO_2$-induced out-migration of high-skilled workers following the *hukou* reform, on the other hand, adds to a growing literature emphasizing the skill heterogeneity in which air quality affects migration (Chen et al., 2017a; Lin, 2018; Freeman et al., 2019; Khanna et al., 2019), and the effect of migration costs on China's local economies (Au and Henderson, 2006b; Tombe and Zhu, 2019). Relative to the existing work, this article draws attention to a different question of policy interest. Specifically, I study China's environmental regulation and migration reform as large-scale natural experiments to document how regulations affect the local employment growth and adjustment, and to study the role of skill-biased migration cost in changing the local skill composition with respect to air quality.

This paper employs an instrumental variable strategy to exploit exogenous variations in the spatial distribution of air pollution. The two widely applied IV's in the existing literature are the phenomenon of thermal inversion (Arceo et al., 2016; Sager, 2019; Chen et al., 2017a; Fu et al., 2017) and the distribution of pollution from its sources by wind direction and speed (Schlenker and Walker, 2016; Freeman et al., 2019; Lin, 2018). However, thermal inversion can be defined somewhat arbitrarily across space in a country like China with a wide range of elevation[7]. And for using wind features, careful identification also requires exogenous variation in the original source of pollution (e.g., Schlenker and Walker, 2016). The instrumental variable used in this paper instead exploits the non-linear transformation of local geographic variations, stemming from the widely applied civil engineering practice of power plant site selection based on construction safety, accessibility, and cost considerations. This suitability index can serve as a complementary IV in pollution-related research as a cross-sectional empirical tool on its own, or a time-varying one when interacted with wind features or changes in the utility network of electricity production and transportation.

The rest of the paper is organized as follows. The next section describes the details of

---

[6] The small strand of exceptions are Hering and Poncet (2014); Cai et al. (2016); Shi and Xu (2018), who find negative regulatory impacts but only when focusing on exporting firms or local in-bound FDI.

[7] For example, the elevations and corresponding pressure levels of which a thermal inversion episode occurs in Shanghai (near sea level) would not have measurable ground temperatures in Yunnan (roughly 2km in elevation).

the two policies, the APPCL and the *hukou* reform, as well as other relevant institutional background and data sources. Section 3 presents the empirical strategy. Section 4 discusses the estimation results and explores potential mechanisms. Section 5 concludes.

# 2 Background and Data

## 2.1 $SO_2$ and the APPCL revision

This paper focuses on the regulation of sulfur dioxide ($SO_2$) emission, the most severe environmental challenge China faces at the time: its annual national emission of 20 million tons since the 1990s has contributed to one quarter of the global emission and more than 90% of East Asia emission (Lu et al., 2010). It is the major by-product of burning fossil fuels such as coal by power plants and related industrial facilities. High levels of $SO_2$ in the air is harmful to the human respiratory system. It is also a primary precursor of acid rain that damages agricultural land, buildings, and overall ecosystem. When interacted with other compounds in the atmosphere, it contributes to the formation of haze and particulate matter (US Environmental Protection Agency, 2017b,a).

The objectives of the first APPCL revision (China's Ninth Five-Year Plan, 1996-2000) were to address the severity of $SO_2$ pollution and acid rain in China. The Two-Control Zone (TCZ) policy was added as a new chapter to the general provision for implementation starting in 1998, marking China's first serious attempt in addressing environmental concerns (Gao et al., 2009). The TCZ policy was named after the strategy of targeting two types of zones: the high sulfur dioxide zone and the high acid rain pollution zone, covering 175 prefectures in China. Enforcement of the State Council approved practices included 1) limiting the usage of high-sulfur coal; 2) installing desulfurization facilities, upgrading boilers and kilns and treating effluent gas in high coal-usage plants; 3) shutting down low-efficiency coal users and restricting new construction of thermal power plants; and 4) levying emission charges on extremely heavy polluters (Gao et al., 2009; Goulder, 2005).

Based on the above policy instruments, I conceptualize the TCZ policy implementation starting in 1998 as an adverse productivity shock to the regulated firms (heavy coal users) and prefectures. World Bank (2003) documents that low-sulfur coal is twice as expensive as high-sulfur coal. Evidence from the US Clean Air Act also estimates that upgrading equipment and treating pollutant increases the firm's average cost by 17% (Becker and Henderson, 2000).

## 2.2 $SO_2$ pollution trends

The Chinese context is different from well-studied developed-country experience both in terms of the extent of $SO_2$ pollution and the implementation of environmental protection policies. Similar to other developing countries going through rapid industrialization, the pollution levels in this discussion are magnitudes higher (Greenstone and Hanna, 2014). Moreover, the $SO_2$ pollution levels has been consistently high in China even with the implementation of environmental protection policy. Figure 1 shows the $SO_2$ time trends for China (consistently high), the United States (steady decline), and the European Union 33 countries (steady and the fastest decline) since 1990.

In addition to the stark contrast in time trends across countries, China's $SO_2$ levels during the sample period of this paper can be divided into two sub-periods correspondingly to two important policy years. The first turning point is around 1998, the year of APPCL implementation. Between then and 2002, China's overall $SO_2$ pollution levels showed no upward sign. Starting from 2003, however, it went up substantially as China's joining of the World Trade Organization spurred further industrialization. This sharp change in trend is closely related to China's overall economic cycle and structural change. Therefore, potential heterogeneity of the two sub-periods are examined closely in the later analysis.

Figure 2 further shows the important role of China's structural change in considering the APPCL implementation and $SO_2$ time trend. The sub-figures (a) and (b) compare the $SO_2$ level versus $SO_2$ per GDP levels of 1998-2010 change against the 1998 level. Without scaling by GDP levels, part (a) shows no convergence relative to initial levels. After scaling by GDP in part (b), the underlying correlation suggests that initially more polluted and hence more stringently regulated prefectures grow slower in terms of their $SO_2$ per GDP levels. This pattern is suggestive of the efficacy of the APPCL in suppressing $SO_2$ pollution growth once the GDP growth is also considered.

## 2.3 The *Hukou* reform and its skill heterogeneity

Another concurrent policy change crucial in understanding the overall local economic impact of air pollution and its regulation is related to China's internal migration restrictions imposed by the *hukou* system. The *hukou* works similarly to an internal passport. When a Chinese citizen is born, he or she inherits a local citizenship tied to his or her mother's *hukou* place of residence (as specific as to a census tract level). In addition to the specific location, this citizenship also classifies individuals as entitled to either a rural or urban status. This dichotomy of legal status along with location differences determine one's rights to the varying quality of local housing, schooling, job opportunities, health care, and even "grain ratios" prior to 1990s (Chan, 1994).

Because one's legal rights are so strictly tied with his or her *hukou* location and status, China's internal migration has been strongly curtailed through the *hukou* system. Prior to the reform in late 1990's, changing one's *hukou* status or moving across locations legally has been extremely difficult. The rare channels are through job assignments after university education, or job relocation within the government and large state owned enterprises, or marriage, etc. Each means comes with high friction and monetary cost[8]. Therefore, such

---

[8] People could otherwise move illegally as "unregistered" migrant or legally but temporarily, and in both cases they are entitled to little local public provision at the receiving (Au and Henderson, 2006b).

inflexibility disincentives anyone from moving and any spatial adjustment of labor supply is largely inelastic.

The gradual *hukou* reform in the late 1990's substantially relaxes (but does not eliminate) the migration restrictions. As a result, China's internal migration grows in scale. This reform allows people to relocate with legal rights to local employment and housing. However, access to other social benefits varies by skill. Those with university degree and above can relocate their *hukou* at much lower cost with full access to the local social benefits. For low-skilled workers, migration cost is much higher in terms of limited access to health care, schooling for their children, etc. (Chan, 2013).

Hence in this paper, *hukou* reform is treated as a skill biased reduction in migration cost from 1998 to 2010. The amenity channel through which the variation in air quality is then more salient for high skilled workers because of their lower mobility cost and higher valuation of amenity. Whereas for low-skilled workers whose valuation of amenity could be similar or lower, *hukou* reform makes it possible to identify whether they sort across places when their employment outlook is negatively affected by APPCL.

## 2.4 Data

The analysis in this paper uses three main data sets merged to the unified prefecture level: the Chinese Population Censuses (1982, 1990, 2000, 2010), the Medium and Large Industrial Enterprise Surveys from 1998 to 2007, the spatial and geographic data based on satellite-derived imageries, remote sensing data, and geological surveys. The documentation and reports from China's State Council and Bureau of Statistics complement the main data with additional institutional information.

The Chinese Population Census is the most reliable and comprehensive source of data that reflects local characteristics. It allows me to compute the following prefecture-level variables in the four census years: population, population of high- and low-skilled workers (i.e., with and above college education versus below), total employment, and employment by sector (primary, secondary[9], and tertiary).

The firm survey data has been widely used in the literature as its sample covers over 90% of China's total industrial output (Brandt et al., 2012). Although its sample emits small non-state owned firms whose annual sales are less than 5 million RMB (Brandt et al., 2014), but its rich details allow me to aggregate information such as employment, capital, value-added, and wage variables by firm ownership and sector to the prefecture level. In this paper I focus on the heterogeneity between state and non-state owned firms. The state ownership is defined as having 50% or above paid-in capital from the state. Combining its 3-digit industrial code and the 1997 Input-Output Table of China from the Bureau of Statistics, I also define heavy and light $SO_2$-emitting firms by their coal usage intensity. I categorize the top 12 industries (out of 106 total industries in the input-output matrix) in coal usage (in

---

[9] Secondary sector is industry plus raw material extraction.

monetary value) as heavy coal users. Each of the defined heavy coal-using industries account for 2% and above of the national coal consumption, and they jointly consume more than 70% of coal in China.

This paper also makes use of and processes a large amount of spatial data. I collect and compute prefecture-level annual average $SO_2$ density, temperature, wind speed, and precipitation using satellite derived spatial data from various National Aeronautics and Space Administration (NASA) projects. These satellite derived measures provide more granular and consistent coverage than ground measures, and are the only source of climate variables for this sample period. For more recent years where both the satellite and ground measures of climatic data are available, Chen et al. (2017a) show that they share consistent trends. In order to compute the instrumental variable, I process remote sensing and geo-survey records from the United States Geological Survey (USGS) and the Harvard WorldMap library. Appendix Table A3 documents the type of data and their corresponding sources, time span spatial resolution. Another key component of the research design is the location of coal-fired power plants, which I collect from the non-profit coal monitoring website SourceWatch Coal Issues[10]. The power plant coordinates are geo-located based on official addresses and confirmed using Google Map Satellite imageries. Lastly, this paper benefits from the generosity of Baum-Snow et al. (2017), whom kindly shared their road and rail maps for China and the unified prefecture boundaries having accounted for the frequent boundary changes throughout the sample period.

The rest of the information in use of this paper, such as the list of TCZ cities, has been collected from official documentation. To supplement the main data, I also utilize additional variables from the Chinese Urban Yearbooks and the 1995 firm survey data. Table 1 presents the summary statistics of key variables by TCZ status and for the full sample.

# 3 Research design

The relationships between prefecture-level ambient $SO_2$ pollution and local growth in the context of China's environmental protection policies and labor mobility reform are formalized as a Roback-style model (Roback, 1982) in the Appendix (section A2), based on the works of Moretti (2011) and Allcott and Keniston (2017). Empirically, I estimate firstly the regulatory impact of the APPCL on the local labor market and secondly the effect of air quality on city size adjustment following the *hukou* reform. The effects are estimated respectively based on the general form of a growth equation:

$$\ln(y_{i,2010}) - \ln(y_{i,2000}) = \xi_1 + \rho_1 Regulate_{i,1998} + \phi_1 Z_i^{-1} + \epsilon_i, \tag{1}$$

$$\ln(y_{i,2010}) - \ln(y_{i,2000}) = \xi_2 + \rho_2 AirQuality_{i,1998} + \phi_2 Z_i^{-1} + u_i, \tag{2}$$

where $i$ denotes the level of analysis of Chinese prefectures. The change in $y$ between 2000 and 2010 represent the set of outcomes of interest, capturing adjustments in the local labor market or city sizes. Equations 1 and 2 differ in the choice of treatment variables to

---

[10] Accessed in June 2017.

separately examine the regulatory effect of APPCL and the amenity effect of air quality on the local economy. To measure environmental regulation, I define $Regulate_{i,1998}$ as the log of prefecture average $SO_2$ density in 1998. It directly determines how stringently a prefecture is regulated under the APPCL revision and thus serves as a valid proxy for regulating intensity[11]. As an alternative, I run a specification where pollution regulation is measured by the binary Two-Control Zone (TCZ) status of the prefectures in 1998. Given that *hukou* reform started in the late 90s and the relative ranking of prefecture pollution levels remains consistent between 1998 and 2010, $AirQuality_{i,1998}$ represents air quality variation across prefectures that affects household migration choices. Therefore, the log of prefecture average $SO_2$ density in 1998 (or alternatively the PM2.5 level), can also serve as the treatment variable of air quality. In the baseline results, I use the log of prefecture average $SO_2$ density in 1998 as the proxy for both $Regulate_{i,1998}$ and $AirQuality_{i,1998}$ to make the later discussion of point estimates from both equations consistent and comparable in the context of the local economy. Results using the alternative measures are reported in the robustness checks.

To control for observable differences across prefectures that may be associated with the outcome and treatment variables, I include a vector of pre-existing prefecture political, economic and climatic characteristics, $Z_i$. The set of political characteristics includes regional dummies, provincial-level city and provincial capital dummies to address the concern that higher administrative status might be correlated with higher pollution levels, industrial productivity, and more stringent regulation. Concerning pre-existing economic conditions, I control for the completeness of the transportation network (log number of rail and road rays), log distance to the coast line, 1995 share of employment in heavy coal-using industries, and the log of college educated population in 1990. These controls aim to address the concern that prefectures with worse air quality and more stringently regulation differ fundamentally in their transportation infrastructure and connectivity, industrial as well as human capital compositions. $Z$ also includes climatic characteristics that may affect local $SO_2$ pollution formation: average temperature, wind speed, and precipitation. Both of the error terms, $\epsilon_i$ and $u_i$, could be correlated across prefectures facing similar regulation intensity for reasons such as pollution spillovers. Therefore, I cluster the standard errors at the level of provinces.

## 3.1 Power plant suitability as the instrumental variable

While $Z_i$'s allow prefectures of varying pollution levels to be as comparable as possible in terms of observed differences, unobserved characteristics could still pose challenges for identification. Variation in ambient $SO_2$ across prefectures is far from randomly distributed. For example, places that are more economically active tend to pollute more in aggregate (Chay and Greenstone, 2003). As a result, $\rho_1$ may be biased as more productive places likely face stricter environmental regulation under policy implementation. Similarly, $\rho_2$ could be biased if the local skill composition correlates to the industrial composition and influences the city sizes.

---

[11] Refer to section 2 for more details.

### 3.1.1 Coal-fired power plants in the Chinese context

To address these endogeneity concerns, my research design uses power plant suitability to instrument for variations in prefecture level $SO_2$ pollution. Several attributes of the coal-fired power plants in the Chinese context make this a valid identification strategy. The first is the paramount role of coal-fired power plants in China's energy generation and coal consumption. Currently, over 70% of China's electricity production is still from coal sources, according to the International Energy Agency[12]. Secondly, coal-fired power plants consistently account for around 50% of China's coal consumption throughout the 2000s[13]. As the main contributor to $SO_2$ emission, coal-fired power plants are hence the key target for China's $SO_2$ regulatory policies since the late 1990's. Based on these two attributes, prefectures hosting coal-fired power plants are likely to face more stringent environmental regulation under APPCL.

Additionally, the power generation and supply networks in China are separated. Electricity is produced locally, but distributed centrally by regional offices and supervised by the State Council (China State Council, 1993). Therefore, firms do not have the incentive to locate near thermal power plants for accessing cheap electricity within a given region (hence the importance of controlling for regional fixed effects). Firms may still locate near thermal power plants for indirect reasons like access to coal source or transportation networks. I return to this particular concern in section 3.2.

Despite these nice contextual features of thermal power plants, there could still be unobservable factors correlated to actual plant locations and economic prospects of their hosting cities, such as political favoritism. It is possible that a coal-fired power plant is strategically placed to favor the local politician for his or her future career or nepotism motivations. These places and their incumbent firms can both be affected by such unobserved factors. Therefore in this paper, I take a further step to utilize the exogeneity in the selection process of power plant hosting sites by borrowing the power plant suitability index from the engineering literature.

### 3.1.2 The local suitability in power plant hosting

The power plant suitability index measures the probability of a given location in hosting coal-fired thermal power plants subject to construction cost, safety, and feasibility considerations from a civil engineering perspective. Both the engineering literature (Barda et al., 1990; Pohekar and Ramachandran, 2004; Choudhary and Shankar, 2012) and practical guidelines (e.g., the *Common Power Plant Siting Criteria* set by the State of Wisconsin Public Service Commission) generally agree on the factors to be considered: topography, land use, water bodies, fuel supply, and so on. The transformation of factors to compute the suitability

---

[12] Data sources: IEA statistics: China, People's Republic of: Electricity and Heat. Data accessed on August 16, 2017. The percentage of electricity generated by coal in 1990, 2000, 2010, and 2014 are: 71%, 78%, 77%, and 73% respectively.

[13] Industrial production accounts for another 40% and the rest is in domestic use or transportation (Lu et al., 2010).

index follows the general form:

$$S_j = \sum w_k f_k(k_j) \tag{3}$$

where suitability for hosting thermal power plant at location $j$ is calculated as the weighted sum of rescaled value of factor $k$. Both of the non-linear scaling function $f$ and weights $w$ are factor-specific.

The exact formula I employ comes from an Iranian government commissioned site selection project lead by engineers Zoej et al. (2005)[14]. Table 2 lists the detailed weights for each factor and their non-linear rescaling functions. For instance, the exact value of a location's distance to small rivers is reclassified to four scores, the higher the more suitable. When the location is either too far (>20km) or too close (<0.5km) to the river, it receive a score of 0. The optimal distance is between 0.5 and 10km with a score of 10, followed by the 10-20km range of a score of 5. After this factor-specific non-linear conversion of all the relevant factors, the suitability index is then computed as their weighted sum. In practice, the spatial resolution of the "location" is that of the coarsest layer of remote sensing data – $0.08° \times 0.08°$ in this case. In other words, the suitability index is computed for every $9 \times 9$ km$^2$ of China.

Figure 3a maps the spatial distribution of the suitability index across China prefectures in greyscale, with the darker color indicating higher suitability for hosting coal-fired power plants. Next, Figure 3b maps the distribution of ambient $SO_2$ density across China, with darker pixel indicating higher $SO_2$ pollution. Both part of Figure 3 also show the location of coal-fired power plants built by 1998. Visually, there is clear spatial overlapping patterns between power plant suitability and their actual locations, as well as $SO_2$ pollution and power plant locations.

## 3.2 Identifying assumptions and supporting evidence

The link between this suitability index and the local ambient $SO_2$ density is established through that it predicts coal-fired power plant locations, which are the biggest contributor to $SO_2$ emission in China throughout the study period. Because the analysis is carried out at the prefecture level, I use the prefecture-level mean of suitability as the instrument in the main specifications. The rationale is that while the suitability of an individual pixel might be difficult to detect for engineers and planners back in the 90s, prefecture-level comparison in terms of their respective average suitability would have been more readily available. As discussed later in the results section, the baseline findings are robust with alternative measures.

As the suitability index captures a location's probability rather than actuality of hosting coal-fired power plants, it further mitigates the endogeneity concerns than actual power plant locations. Furthermore, to address the concern over constituting factors of suitability

---

[14] While this approach is common in engineering (Barda et al., 1990; Pohekar and Ramachandran, 2004; Choudhary and Shankar, 2012), Zoej et al. (2005) is the only published paper with specific weights and non-linear transformations.

having direct impact on local economy, I include them linearly in the regression. The channel through which this instrumental variable works is effectively the set of weights and rescaling function from the engineering literature. It is a valid instrument under the assumption that the non-linear functional form of this index is orthogonal to treatment.

The first stage is formulated as follows:

$$\ln \mathrm{SO}_{2i,1998} = \alpha + \beta \mathrm{Suitability}_i + \gamma \mathrm{X}_i^{-1} + \varepsilon_i, \tag{4}$$

where the prefecture-level mean of suitability in prefecture $i$ is used to predict its 1998 ambient $SO_2$ density level that serves as a proxy for regulation intensity or air quality in 1998. In addition to the setup of equations 1 and 2, the linear forms of constituting components of suitability index are included in $X_i$. They are: log average elevation, log number of raw energy fields (gas, oil, and coal), dummy if near volcanoes, air fields, or earthquake spots, and urban area gas supply coverage. The geographic covariates and prefecture characteristics are computed using the same remote sensing data for the suitability computation. The only exception is the measure of urban area gas supply coverage, which comes from the 1999 China Urban Yearbook as it provides better accuracy at the prefecture level.

Table 3 summarizes the various linkages between suitability, power plant hosting, ambient $SO_2$ levels, and regulation intensity at the prefecture level. Firstly, panel A shows that prefectures with higher average thermal power plant suitability values are more likely to host power plants by 1998, both with and without controlling for prefecture economic characteristics. Panel B then establishes the positive correlation between thermal power plant hosting status and local ambient $SO_2$ pollution levels. Lastly, panel C provides support for using 1998 ambient $SO_2$ levels as valid proxy for regulation stringency. Prefectures with higher ambient $SO_2$ levels in 1998 are more likely to be listed as TCZ prefectures, which received stricter regulation under APPCL amendments.

Table 4 presents the first-stage results based on equation 4 to summarize the above relationship. Both the mean and top quartile of city-level suitability index are strongly predictive of ambient $SO_2$ in 1998, as well as for $NO_2$ levels in 2005 as a robustness check. Next, predicted local $SO_2$ at the prefecture level is used in the specifications of equations 1 and 2, further including the linear forms of constituting components of the suitability index, to obtain the two-stage least square estimates.

Before turning to the results, it is important to show that variations in the instrument do not correlate to pre-existing prefecture growth trends that may themselves affect the outcomes of interest. Table 5 presents results for this balance check: the prefecture level power plan suitability measure is not correlated to the pre-treatment growth trends in city size, high-skilled worker population, total employment and manufacturing employment. These results further substantiates the likelihood of this instrumental variable in satisfying the the exclusion restriction, as prefectures of varying degrees of suitability to host power plants do not differ statistically in terms of their population, skill composition, and local labor market characteristics predating pollution regulations and *hukou* reform.

# 4 Estimation results and channels at work

## 4.1 Effects on the local economy

Table 6 presents both the 2SLS and OLS estimates of specifications 1 (in panel A) and 2 (in panel B) to jointly show the local economic impacts. Based on equation 1, panel A focuses on the growth of employment by regulated and non-regulated sectors to test for regulatory effect of the APPCL revision in 1998 at the prefecture level. The results show that employment growth in the regulated sectors is relatively slower in prefectures under more stringent environmental regulations (columns 1 and 2) and it has no impact on the employment in non-regulated sectors (columns 3 and 4)[15] .

In column 1, the total prefecture-level industrial employment (resource extraction, manufacturing and utilities) is aggregated from the firm-level survey data, where small firms (less than 5 million RMB in sales) are excluded from the sample. Column 2 instead uses the employment variables from the population census, which comes from a comprehensive and representative sample. Although their estimates are of the same sign, the sample discrepancy between the industrial survey data (column 1) and population census (column 2) results in substantial differences in magnitude and statistical significance. This is worth noting for interpreting later analysis that makes use of the rich firm information from the industrial survey data.

While China's industrial sector has been growing rapidly between 2000 and 2010, the baseline estimate in column 2 suggests that a 10% higher local $SO_2$ and hence stricter regulation leads to a 1.7% decline in relative employment growth in the secondary sector of an average prefecture. Equivalently, a one standard deviation increase in 1998 log $SO_2$ level, which is approximately a 114% increase in ambient $SO_2$ levels (see table 1), leads to a reduction in local secondary employment growth of around 17.6%. This estimate has the same sign but much smaller magnitude than findings based on the US Clean Air Act (e.g., Greenstone, 2002; Walker, 2011).

The next two columns of panel A examine the change of employment growth in the primary (agriculture and fishery) and tertiary (service) sectors. They do not differ across prefectures with varying pollution regulatory intensity, which provides evidence that the negative effect on the regulated sectors is a direct result of environmental regulation. Despite the fact that China is one of the fastest growing manufacturing countries during this sample period, its secondary industry accounts for an average of 22% of the national employment in 2010 (16% in 2000), while the primary and tertiary sectors account for 51% and 26% respectively. Partly due to this underlying employment composition, the 1998 APPCL revision does not have an impact on prefecture total employment as shown in column 5.

Given the APPCL-induced negative local labor demand shock in the secondary industry, panel B of Table 6 proceeds to explore whether it translates into any changes in city sizes

---

[15] Estimation based on the alternative specification using the TCZ dummy to measure regulation is presented in Table A2 and discussed later.

as the *hukou* reform reduces migration costs. Columns 1 and 2 splits the population into low- and high-skilled workers for two considerations. Firstly, the regulation-induced negative local labor demand shock largely affects low-skilled workers as less than 5% of workers in the secondary sector has education levels beyond high school[16]. Secondly, the *hukou* reform (details in section 2.3) has been skill-biased. The migration cost of high-skilled workers is lowered substantially and quickly but that of the low-skilled workers remains relatively high (Au and Henderson, 2006b; Tombe and Zhu, 2019).

This skill-biased nature of the *hukou* reform is reflected in the corresponding estimates of changes in population of low- versus high-skilled workers in columns 1 and 2. Despite their worsened employment prospects, the population growth of low-skilled workers in highly regulated prefectures are not statistically different from their less regulated counterparts. As the 1998 $SO_2$ levels is positively associated with both regulation intensity and local air quality, the column 1 estimate implies that low-skilled workers are still facing high migration costs between 2000 and 2010. Based on the modified Roback-style urban model (Moretti, 2011), an alternative possibility is that low-skilled workers have negative valuations of air quality, and this negative valuation is large enough to offset the adverse local labor demand shock. Although this paper is limited in its scope to test for this possibility, it is an unlikely case.

The estimate on the population change of high-skilled workers, on the other hand, is negative and statistically significant. Following the assumption that high-skilled workers are almost not affected by the environmental regulation, this estimate is suggestive of their high valuation of air quality. In other words, following the *hukou* reform, a 10% increase of $SO_2$ pollution at the prefecture level causes its relative population growth of high-skilled workers to decline by 1%. This effect is consistent with findings in existing literature that high-skilled workers in developing countries respond sensitively to amenity differences such as air quality (e.g., Chen et al., 2017a; Khanna et al., 2019). The high-skilled workers "vote with their feet" to reveal their strong preference for amenities such as air quality once their migration cost is lowered. Comparing with the non-effects in the balance test (Table 5) for decades prior to 2000, this amenity-based sorting of the high-skilled population is made possible because of the *hukou* reform.

Next, columns 3 and 4 of panel B show the equation 2 specification for changes in the population of different *hukou* types. As discussed in section 2.3, holders of urban *hukou* face lower migration cost than rural *hukou* holders. The discrepancy in point estimates, that the urban *hukou* holders are moving away from highly polluted prefectures while their rural counterparts are not, further confirms the role of migration cost reduction in making the spatial adjustment of population possible. Lastly for panel B, column 5 tests for change in the overall population growth trends. Since the high- versus low-skilled composition in an average Chinese prefecture has a ratio of roughly 1:9, it is therefore not surprisingly to find no effect statistically on the overall population change.

Table 6 also presents the corresponding OLS estimates in the bottom of each panel.

---

[16] Based on my calculation using the Chinese Household Income Project data.

These estimates show the necessity of using the IV strategy, as OLS estimates would lead to different conclusions in both signing and magnitude. For example, the OLS estimates would suggest no regulatory effect on the secondary employment or no spatial adjustment in the population growth of high-skilled workers. The potential bias discussed in section 3 is related to that high pollution prefectures may be more industrialized prior to the APPCL implementation.

In addition to the results reported in this section, the Appendix includes a series of robustness and sensitivity checks concerning the IV validity and estimated baseline effects on the local economy. Table A1 presents supporting evidence on the IV validity with several variations: (i) excluding the linear measure of individual suitability comprising factors in the specification, (ii) replacing the current IV with one computed using a limited, less economic set of factors (highlighted in Table 1), (iii) replacing the current IV with the prefecture top quartile of suitability.

The Appendix Table A2 also provides a comprehensive set of sensitivity tests. The main results are robust to: (i) excluding prefectures in the bottom 10% of the 1998 $SO_2$ distribution, (ii) controlling for the initial level of the outcome variable as an additional covariate, (iii) excluding the 1995 share of employment in heavy coal-using industries as covariate, (iv) controlling for prefectures that provide coal-fired winter heating, i.e., the Huai River policy dummy as in Ebenstein et al. (2017), (v) using the Two Control Zone dummy as the treatment variable, (vi) replacing the treatment variable as 1998 $SO_2$ per capita, (vii) using the 1998 $PM_{2.5}$ level as treatment variable, (viii) using the 2010 $SO_2$ level as treatment variable. In summary, the results discussed in this section are robust in their signing, magnitude, and statistical significance.

## 4.2 Local labor market adjustment and channels at work

The previous section presents evidence that the APPCL causes decline in the secondary sector employment growth in more relative to less regulated prefectures. As high-skilled workers avoid bad air quality by migrating out of more polluted prefectures with *hukou* reform, there has been limited population relocation across space for low-skilled workers, whose employment prospects have been adversely affected by the environmental regulation. This section makes use of the rich firm information in the industrial survey data to explore heterogeneity, local labor market adjustments and potential channels at work.

The first task is to pinpoint the type of firms most affected by the APPCL. Panel A in Table 7 presents equation 1 estimation by coal usage intensity and firm ownership types. The first two columns show that the adverse pollution regulation effect occurred only in heavy coal-using firms, the ones targeted by APPCL. This provides further reassurance that the environmental regulation is the main cause of the observed adverse employment effect, rather than heavy pollution itself.

The next four columns further break down the sample by ownership of state owned versus non-state owned firms. Two curious patterns emerge. The first is that the adverse APPCL effect is driven solely by heavy coal-using state owned firms, and private firms do not appear to be affected. Given vast existing evidence on the governmental favoritism toward state owned firms that allowed them to be less efficient in the usage of capital and labor inputs (e.g., Hsieh and Klenow, 2009; Chen et al., 2017b), one may expect the APPCL to be larger for private rather than state owned firms. In fact, existing research on environmental regulations shows no convincing evidence for any ownership advantages by SOE. For example, Wang and Jin (2007) show that state- and privately owned polluting firms behave similarly under regulation.

Beyond differences in firm behaviors by ownership, three plausible explanations arise from the research and institutional contexts. The first straightforward explanation would be the previously highlighted data limitation of the industrial survey data. The non-state owned firms with annual sales less than five million RMB excluded from this data could well be the group of firms most affected by the APPCL. In particular, the implementation of APPCL directly imposes cost on production, which may have a disproportionately greater average cost impact on small firms.

The second reason relates to resource limitation in the enforcement of environmental regulations. Becker and Henderson (2000) document a non-uniform enforcement of regulation across firm size in the US under the Clean Air Act. As a result of limited regulatory resources, heavy polluters that are usually larger in size turn out to be the targets of regulation, whereas the enforcement agencies do not have enough resources to monitor smaller polluters as closely. This offers one good basis for the pattern observed here. Figures 4 and 5 show kernel density of firm size for state versus privately owned firms in heavy and light $SO_2$ polluting industries respectively. The state owned firms in heavy coal-using and hence $SO_2$-polluting industries are much larger in size compared to privately owned firms, and this pattern is consistent over time in both 1999 and 2007. For non-polluting industries, however, firm size distributions are very similar across ownership types. Therefore, it is likely that state owned firms in heavy polluting industries attract greater regulatory attention simply because they are large and conspicuous in polluting activities.

The third explanation is embedded in the Chinese contexts. There is evidence that the state owned enterprises (SOEs) have been the main targets of the APPCL implementation during this time as China undergoes a large scale privatization of the SOEs between 1997 and 2001 (Fang, 2002; Jefferson and Su, 2006). Heavy polluting SOEs are deemed unproductive by the State and need to be eliminated. For example, a State Council 1999 announcement declares, "SOEs that waste resources, pollute heavily, employ outdated technologies need to be forced into bankruptcy and shut down" (China State Council, 1999). Therefore, it is likely that environmental regulations accelerated the downward employment shock on polluting SOEs during the national privatization.

On the other hand, the opposite effects of columns 3 and 4 in panel A imply that there is substantial labor reallocation from regulated to non-regulated state owned firms. But the magnitude of point estimates suggests that the reallocation is not enough to make up for the employment decline brought by the APPCL. As the rapid SOE privatization stabilized in 2003, I break down the sample further into periods before and after this year in Table 7 panels B and C. Doing so shows that the observed labor reallocation in stringently regulated prefectures has a time lag. The employment decline in heavy coal using SOEs occurs within the first five years after the APPCL implementation, and the relatively faster employment growth in light coal-using SOEs in more regulated prefectures occurs after 2003. This local labor reallocation may suggest that the APPCL-induced adverse labor demand shock is short-term as it disappears after the first five years. However, China's national $SO_2$ emission rises sharply again starting in 2003 (Figure 1) before declining again in 2007. This is suggestive of a temporary cease in the regulation enforcement or a reflection of the surge in industrial growth and continued structural change since China's joining of WTO in late 2001.

Next, table 8 provides additional analysis to explore potential channels through which the APPCL affects the local economy. Based on the previous findings, I focus on the heterogeneity within state owned firms by their coal usage for the periods before and after 2003. Panels A - C examine along the dimensions of total local value-added, capital and wage bills aggregated from the industrial survey data. They present a consistent pattern of adverse regulatory impact on the targeted sector before 2003 and a small degree of reallocation onto the non-targeted sector after 2003. Panel D estimates show that both the adverse regulatory impact and the reallocation are driven by the number of firms in the relevant sector. In other words, more stringent environmental regulation caused faster firm exits in the regulated sector in the short term. Over time, however, some of the resources reallocated to the non-regulated sector and led to their faster relative growth.

For completeness, I addresses the question of whether the additional costs imposed by APPCL affect productivity in the regulated sectors. Under the assumption of competitive market, workers are paid their marginal productivity. Therefore, panel E of table 8 shows results on prefecture growth in average wage across sectors to shed light on worker productivity changes. The estimates suggest that prefectures under more stringent APPCL regulation have experienced substantially slower average wage growth relative to less regulated prefectures in the heavy coal using sector within the first five years. Interestingly, more regulated prefectures have caught up on it over time by exhibiting relatively faster average wage growth in the same sector. The patterns shown in table 8 suggests that the APPCL has an initial adverse effect on the local economy through eliminating firms, lowering growth and productivity in the regulated sector. However, worker productivity appears to have bounced back fully over time while the within-prefecture reallocation from regulated to non-regulated sectors makes up only some of the adverse impact.

# 5 Conclusion

In the recent decades, China went through rapid economic growth while its air quality remained high. As per capital income rises, more people may demand for cleaner air (Zheng and Kahn, 2013). In this case the government may be under greater pressure to tighten up environmental regulations. This is a common pattern seen throughout many emerging economies. Therefore, it is paramount for policy makers to understand how the environmental protection policies and air quality jointly affect the local economic growth.

The findings in this paper show that China's environmental regulation caused local manufacturing employment in heavily polluted prefectures to decline. This impact is concentrated on firms in the polluting industries, and particularly state owned firms that are of larger firm size and monitored closely during the national privatization reform. Within the local labor market over time, some unemployment is reallocated from regulated industries to non-polluting industries. However, the overall net effect is still negative over a ten-year period. Moreover, low-skilled workers whose job prospects adversely affected by APPCL do not appear to be reallocating to other prefectures. This implies persistently high migration costs for this group of workers. On the other hand, high-skilled workers whose job prospects are unlikely to be affected by APPCL migrate out of highly polluted prefectures as migration costs fall for them, showing their preference for cleaner air. This loss of human capital has long term implications for highly polluted prefectures.

"Produce and pollution first, clean up later" is a frequent choice made by policy makers. Low efficiency fuels are the cheaper way to generate energy and increase economic output, but the later cost of cleaning up and losing human capital can be substantial. Notably, these costs are additional to the well-documented health costs of heavy pollution.

Findings in this paper also shed light on the distributional impacts of environmental regulations and migration reform by skill. Low-skilled workers are the group bearing the regulatory costs. The high migration costs they face further worsens the adverse shock. Migration reform is biased toward those with college degree and above, giving them easier access to places where they are better off.

Lastly, this environmental regulation is not uniformly enforced across firm size and ownership. As a result of limited regulatory resources, larger firms become the main target. There is also evidence to suggest that the State used environmental regulation as a mean to eliminate firms with outdated technology, which is largely SOEs in the case of China during this sample period.

# References

**Allcott, Hunt and Daniel Keniston**, "Dutch Disease or Agglomeration? The Local Economic Effects of Natural Resource Booms in Modern America," *Review of Economic Studies*, 2017.

**Arceo, Eva, Rema Hanna, and Paulina Oliva**, "Does the Effect of Pollution on Infant Mortality Differ Between Developing and Developed Countries? Evidence from Mexico City," *The Economic Journal*, mar 2016, *126* (591), 257–280.

**Au and J V Henderson**, "Are Chinese Cities Too Small?," *Review of Economic Studies*, 2006, *73*, 549–576.

**Au, Chun-Chung and J Vernon Henderson**, "How migration restrictions limit agglomeration and productivity in China," *Journal of Development Economics*, 2006, *80* (2), 350–388.

**Barda, O.Haluk, Joseph Dupuis, and Pierre Lencioni**, "Multicriteria location of thermal power plants," *European Journal of Operational Research*, apr 1990, *45* (2-3), 332–346.

**Baum-Snow, Nathaniel, Loren Brandt, J. Vernon Henderson, Matthew A. Turner, and Qinghua Zhang**, "Roads, Railroads and Decentralization of Chinese Cities," *Review of Economics and Statistics*, 2017, *99(3)* (July), 435–448.

**Becker, Randy and Vernon Henderson**, "Effects of Air Quality Regulations on Polluting Industries," *Journal of Political Economy*, apr 2000, *108* (2), 379–421.

**Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang**, "Creative Accounting or Creative Destruction? Firm-Level Productivity Growth in Chinese Manufacturing," *Journal of Development Economics*, 2012, *97* (2), 339–351.

**_ , _ , and _** , "Challenges of working with the Chinese NBS firm-level data," *China Economic Review*, 2014, *30*, 339–352.

**Cai, Xiqian, Yi Lu, Mingqin Wu, and Linhui Yu**, "Does environmental regulation drive away inbound foreign direct investment? Evidence from a quasi-natural experiment in China," *Journal of Development Economics*, 2016, *123*, 73–85.

**Chan, Kam Wing**, *Cities with Invisible Walls: Reinterpreting Urbanization in Post-1949 China*, Oxford University Press, 1994.

**_** , "China, Internal Migration," in Immanuel Ness and Peter Bellwood, eds., *The Encyclopedia of Global Migration*, Wiley-Blackwell, 2013.

**Chang, Tom Y., Joshua Graff Zivin, Tal Gross, and Matthew Neidell**, "The effect of pollution on worker productivity: Evidence from call center workers in China," *American Economic Journal: Applied Economics*, 2019, *11* (1), 151–172.

**Chay, Kenneth Y and Michael Greenstone**, "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession," *Quarterly Journal of Economics*, 2003, *118*, 1121–1167.

**Chen, Shuai, Paulina Oliva, and Peng Zhang**, "The Effect of Air Pollution on Migration: Evidence from China," Technical Report, National Bureau of Economic Research, Cambridge, MA nov 2017.

**Chen, Ying, J Vernon Henderson, and Wei Cai**, "Political favoritism in China's capital markets and its effect on city sizes," *Journal of Urban Economics*, 2017, *98* (March), 69–87.

**Chen, Yuyu, Avraham Ebenstein, Michael Greenstone, and Hongbin Li**, "Evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy," *Proceedings of the National Academy of Sciences*, 2013, *110 (32)*, 12936–12941.

**China State Council**, "State Council Document 1993 No. 115," 1993.

_ , "Zhonggong zhongyang guanyu guoyou qiye gaige he fazhan ruogan zhongda wenti de jueding (Decisions on major issues in the privatization and development of state owned enterprises)," 1999.

**Choudhary, Devendra and Ravi Shankar**, "An STEEP-fuzzy AHP-TOPSIS framework for evaluation and selection of thermal power plant location: A case study from India," *Energy*, jun 2012, *42* (1), 510–521.

**Currie, Janet, Eric A. Hanushek, E. Megan Kahn, Matthew Neidell, and Steven G. Rivkin**, "Does pollution increase school absences?," *Review of Economics and Statistics*, nov 2009, *91* (4), 682–694.

**Ebenstein, Avraham, Maoyong Fan, Michael Greenstone, Guojun He, and Maigeng Zhou**, "New Evidence On The Impact of Sustained Exposure to Air Pollution on Life Expectancy From China's Huai River Policy," *Proceedings of the National Academy of Sciences*, 2017, *114* (39), 10384–10389.

**Fang, Gang**, "Progress in Ownership Changes and Hidden Risks in China's Transition," *Transition Newsletter*, 2002, *13* (1-5).

**Freeman, Richard, Wenquan Liang, Ran Song, and Christopher Timmins**, "Willingness to Pay for Clean Air in China," *Journal of Environmental Economics and Management*, dec 2019, *94*, 188–216.

**Fu, Shihe, Brian Viard, Peng Zhang, V Brian Viard, and Cheung Kong**, "Air Quality and Manufacturing Firm Productivity: Comprehensive Evidence from China," Technical Report 2017.

**Gao, Cailing, Huaqiang Yin, Nanshan Ai, and Zhengwen Huang**, "Historical Analysis of $SO_2$ Pollution Control Policies in China," *Environmental Management*, 2009, *43*, 447–457.

**Glaeser, Edward L., Jed Kolko, and Albert Saiz**, "Consumer city," *Journal of Economic Geography*, jan 2001, *1* (1), 27–50.

**Goulder, Lawrence H.**, "China's Pollution Levy System: Theoretical Capabilities and Practical Challenges," *Journal of Comparative Studies*, 2005, *21.*

**Greenstone, Michael**, "The Impacts of Environmental Regulations on Industrial Activity: Evidence from the 1970 and 1977 Clean Air Act Amendments and the Census of Manufactures The Environmental Protection Agency district office in Philadelphia complied with my Freedom of Inform," *Journal of Political EconomyJournal of Political Economy*, 2002, *110* (6), 1175–1219.

_ **and Rema Hanna**, "Environmental Regulations, Air and Water Pollution, and Infant Mortality in India," *American Economic Review*, 2014, *104* (10), 3038–3072.

**Hanna, Rema and Paulina Oliva**, "The effect of pollution on labor supply: Evidence from a natural experiment in Mexico City," *Journal of Public Economics*, feb 2015, *122*, 68–79.

**He, Jiaxiu, Haoming Liu, and Alberto Salvo**, "Severe air pollution and labor productivity: Evidence from industrial towns in China," *American Economic Journal: Applied Economics*, 2019, *11* (1), 173–201.

**Heblich, Stephan, Alex Trew, and Yanos Zylbergerg**, "East Side Story: Historical Pollution and Persistent Neighborhood Sorting," 2016.

**Hering, Laura and Sandra Poncet**, "Environmental policy and exports: Evidence from Chinese cities," *Journal of Environmental Economics and Management*, 2014, *68* (2), 296–318.

**Hsieh, Chang-Tai and Peter J. Klenow**, "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, nov 2009, *124* (4), 1403–1448.

**Jefferson, Gary H. and Jian Su**, "Privatization and restructuring in China: Evidence from shareholding ownership, 1995–2001," *Journal of Comparative Economics*, mar 2006, *34* (1), 146–166.

**Kahn, Matthew and Pei Li**, "The Effect of Pollution and Heat on High Skill Public Sector Worker Productivity in China," Technical Report, National Bureau of Economic Research, Cambridge, MA feb 2019.

**Khanna, Gaurav, Wenquan Liang, A. Mushfiq Mobarak, and Ran Song**, "The Productivity Consequences of Pollution-Induced Migration in China," *Working paper*, 2019.

**Li, Juan, Min Zhao, and Yien Yang**, "Environmental regulation and firms' performance: a quasi-natural experiment from China," *Chinese Journal of Population Resources and Environment*, jul 2019, *17* (3), 278–294.

**Lin, Yatang**, "The Long Shadow of Industrial Pollution: Environmental Amenities and the Distribution of Skills," 2018.

**Lu, Z, D G Streets, Q Zhang, S Wang, G R Carmichael, Y F Cheng, C Wei, M Chin, T Diehl, and Q Tan**, "Sulfur dioxide emissions in China and sulfur trends in East Asia since 2000," *Atmospheric Chemistry and Physics*, 2010, *10* (13), 6311–6331.

**Moretti, Enrico**, "Chapter 14 - Local Labor Markets," in David Card and Orley Ashenfelter, eds., *Handbook of Labor Economics*, Vol. 4, Part B of *Handbook of Labor Economics*, Elsevier, 2011, pp. 1237–1313.

**Pohekar, S.D. and M. Ramachandran**, "Application of multi-criteria decision making to sustainable energy planning—A review," *Renewable and Sustainable Energy Reviews*, aug 2004, *8* (4), 365–381.

**Roback, Jennifer**, "Wages, Rents, and the Quality of Life," *Journal of Political Economy*, 1982, *90* (6), 1257–1278.

**Sager, Lutz**, "Estimating the effect of air pollution on road safety using atmospheric temperature inversions," *Journal of Environmental Economics and Management*, nov 2019, *98.*

**Schlenker, Wolfram and Reed Walker**, "Airports, Air Pollution, and Contemporaneous Health," *The Review of Economic Studies*, 2016, *83 (2)*, 768–809.

**Shi, Xinzheng and Zhufeng Xu**, "Environmental regulation and firm exports: Evidence from the eleventh Five-Year Plan in China," *Journal of Environmental Economics and Management*, may 2018, *89*, 187–200.

**Tanaka, Shinsuke, Wesley Yin, and Gary H. Jefferson**, "Environmental Regulation and Industrial Performance: Evidence from China," 2014.

**Tombe, Trevor and Xiaodong Zhu**, "Trade, Migration and Productivity: A Quantitative Analysis of China," *American Economic Review*, 2019, *5*, 1843–72.

**US Environmental Protection Agency**, "Acid Rain Program Overview," 2017.

_ , "Sulfur Dioxide ($SO_2$) Pollution," https://www.epa.gov/so2-pollution/sulfur-dioxide-basics 2017.

**Walker, Reed**, "Environmental Regulation and Labor Reallocation: Evidence from the Clean Air Act," *American Economic Review: Papers & Proceedings*, 2011, *101* (3), 442–447.

_ , "The Transitional Costs of Sectoral Reallocation: Evidence From the Clean Air Act and the Workforce," *Quarterly Journal of Economics*, 2013, *128 (4)*, 1787–1835.

**Wang, Hua and Yanhong Jin**, "Industrial Ownership and Environmental Performance: Evidence from China," *Environmental & Resource Economics*, 2007, *36*, 255–273.

**World Bank**, “China: Air pollution and Acid rain control,” Joint UNDP/World Bank Energy Sector Management Assistance Programme 2003.

**Yang, R. R., R. Y. Long, T. Yue, and Y. Y. Zhuang**, “Environmental regulation and industrial performance: An empirical analysis of manufacturing industry in China,” *International Journal of Sustainable Development and Planning*, 2014, *9* (6), 769–777.

**Zhao, Xiaomeng, Chuanjiang Liu, and Mian Yang**, “The effects of environmental regulation on China’s total factor productivity: An empirical study of carbon-intensive industries,” *Journal of Cleaner Production*, 2018, *179* (1), 325–334.

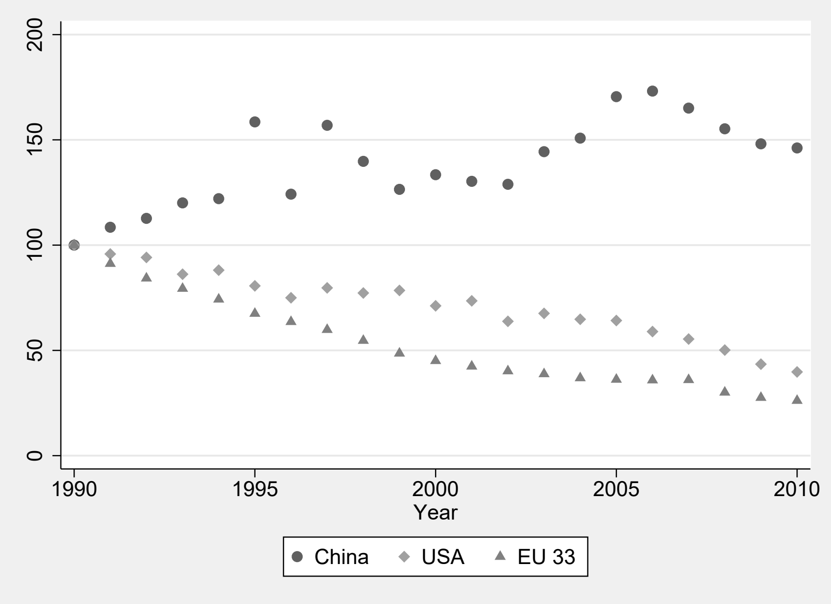
**Zheng, Siqi and Matthew E Kahn**, “Understanding China’s Urban Pollution Dynamics,” *Journal of Economic Literature*, sep 2013, *51* (3), 731–772.

**Zoej, M. J. Valadan, M. S. Mesgari, S. Beheshtifar, M. Karimi, R. Samadi, H. Yousefi, and T. Sohrab**, “Using GIS for Site Selection of Combined-Cycle Power Plants,” in “26th Asian Conference on Remote Sensing, ACRS 2005 and 2nd Asian Space Conference” Hanoi, Vietnam 2005, pp. 509–515.

# 6 Figures and Tables

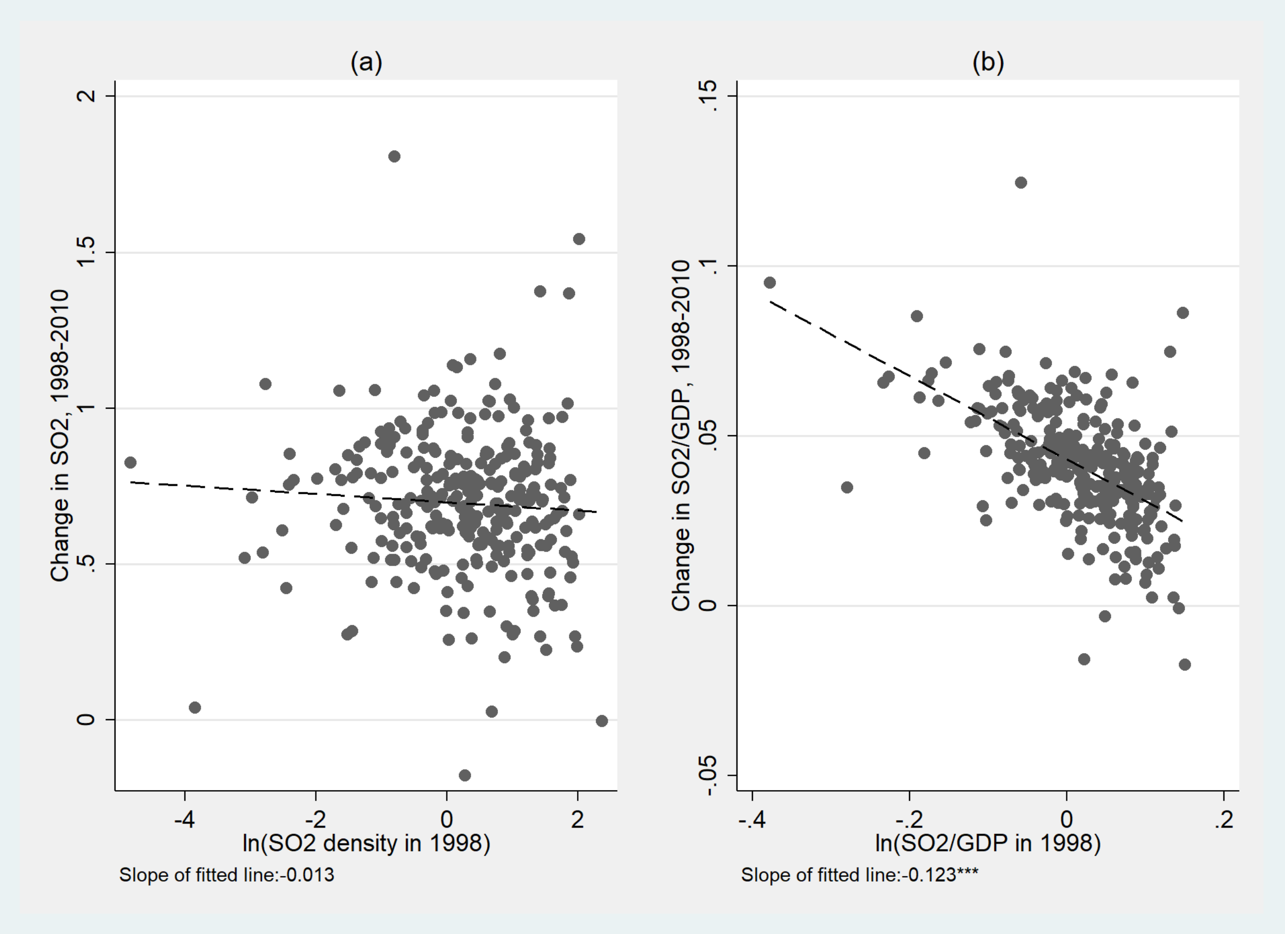
## Figures

Figure 1: The severity of $SO_2$ emission in China, 1990-2010



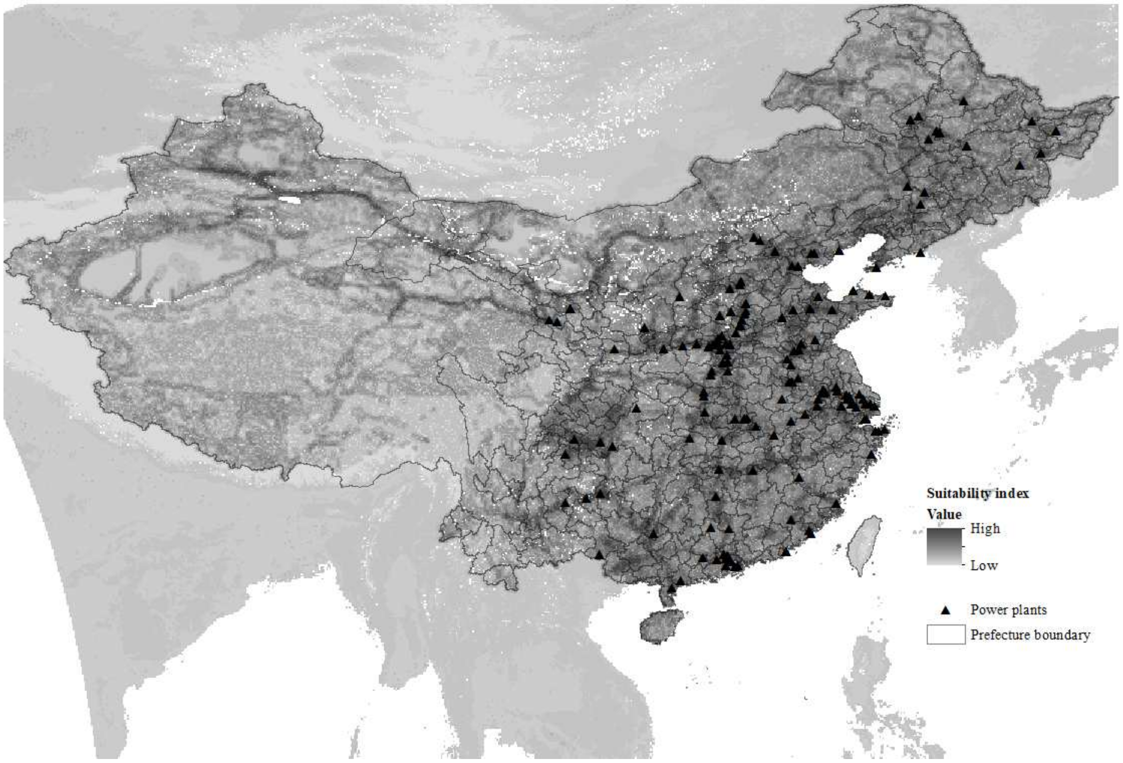
*Notes:* Data is converted to relative terms with respect to 1990 (=100).
*Data sources:* China Ministry of Environmental Protection; U.S. Environmental Protection Agency; European Environment Agency. Accessed on August 9, 2017.

Figure 2: Convergence in $SO_2$ levels after the APPCL implementation

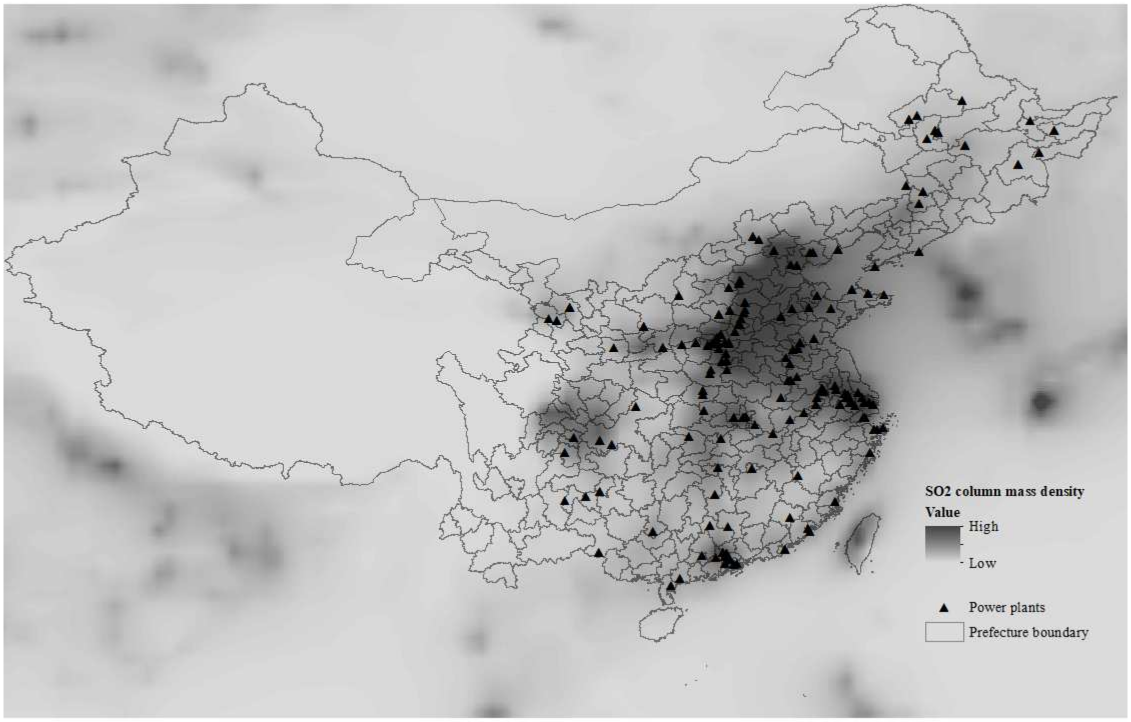


*Notes:* These two figures show the potential "convergence" since the APPCL implementation by plotting the change between 1998 and 2010 as a function of the initial 1998 levels. Part (a) uses $SO_2$ pollution levels and part (b) uses $SO_2$ per GDP levels. Without accounting for GDP, part (a) shows non-convergence in the sense that $SO_2$ pollution levels changed little since the APPCL implementation. Part (b), however, shows that initially more polluted prefectures have slower growth in terms of $SO_2$ per GDP (i.e., some convergence over time).

Figure 3: Spatial distributions of coal-fired power plants by 1998, suitability index, and ambient $SO_2$ density



(a) Coal-fired power plant suitability and actual locations for those built by 1998



(b) Ambient $SO_2$ density and coal-fired power plant locations for those built by 1998

Figure 4: Kernel density plot of firm size by ownership types, **heavy** $SO_2$ polluters

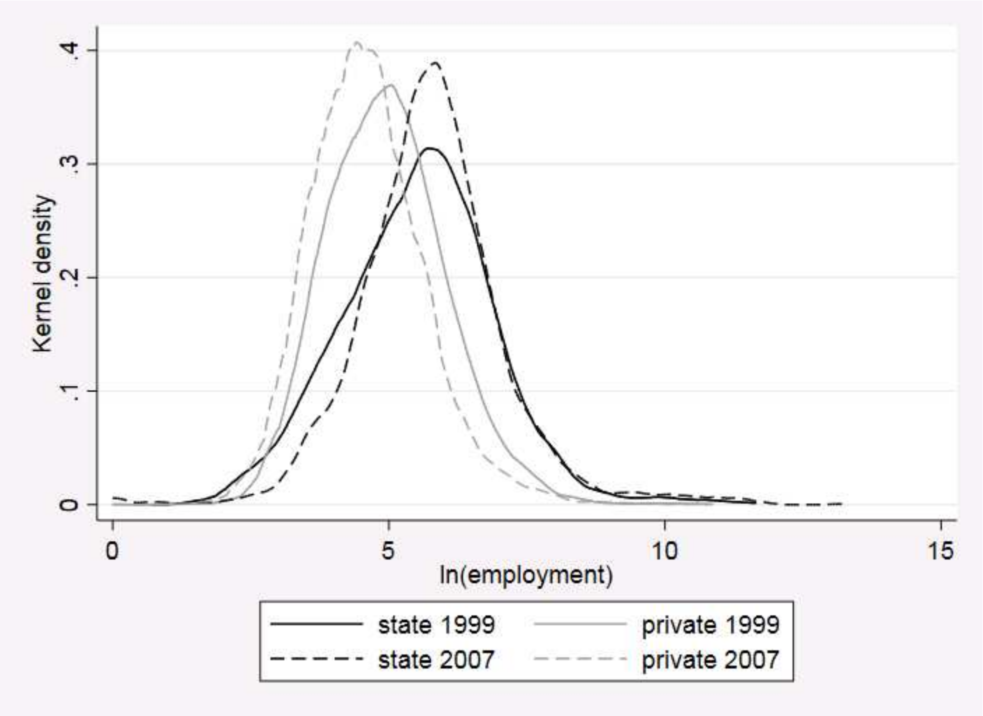
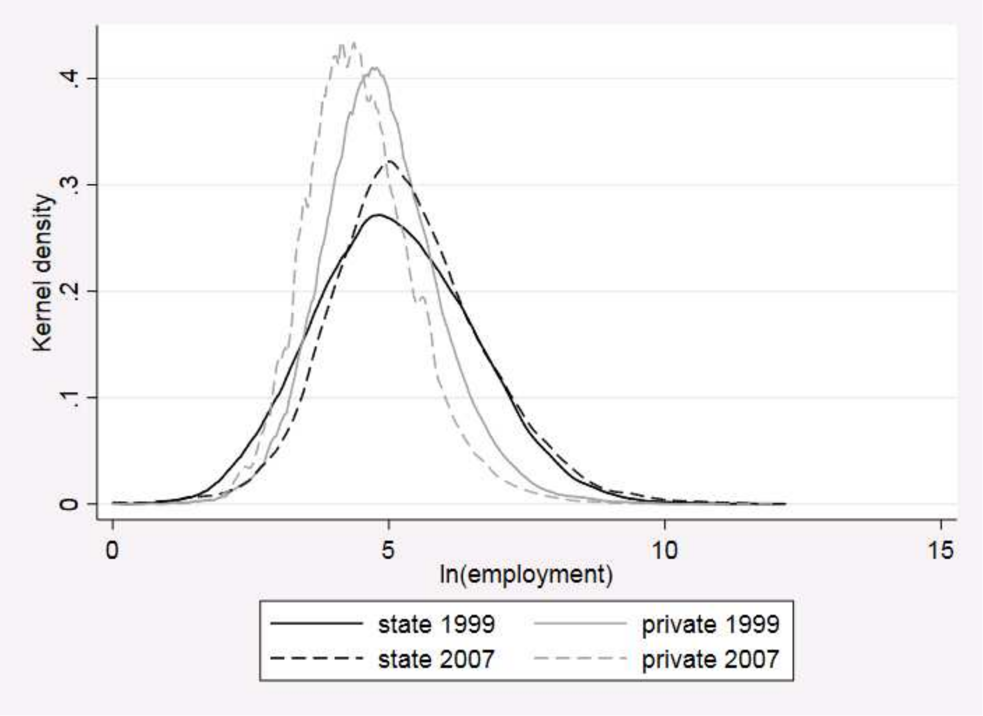


Figure 5: Kernel density plot of firm size by ownership types, **light** $SO_2$ polluters



# Tables

Table 1: Summary statistics of key variables

| Variables | Non-TCZ | TCZ | Total |
|---|---|---|---|
| ln($SO_2$ density in 1998) | -0.056 | 0.494 | 0.252 |
| | (1.307) | (0.874) | (1.118) |
| 1 if hosting thermal power plant by 1998 | 0.286 | 0.406 | 0.353 |
| | (0.454) | (0.493) | (0.479) |
| ln(thermal power generating capacity) | 1.721 | 2.656 | 2.244 |
| | (2.780) | (3.268) | (3.093) |
| Power plant suitability index: mean | 4.320 | 4.663 | 4.512 |
| | (0.663) | (0.524) | (0.613) |
| Power plant suitability index: top quartile | 4.947 | 5.359 | 5.178 |
| | (0.722) | (0.571) | (0.673) |
| Prefectures in eastern China | 0.278 | 0.487 | 0.395 |
| | (0.450) | (0.501) | (0.490) |
| Prefectures in western China | 0.278 | 0.194 | 0.231 |
| | (0.450) | (0.396) | (0.422) |
| Provincial capital or provincial-level city | 0.024 | 0.138 | 0.087 |
| | (0.153) | (0.345) | (0.283) |
| Number of road and rail rays | 3.968 | 4.588 | 4.315 |
| | (2.043) | (2.084) | (2.086) |
| ln(distance to nearest coast in km) | 5.549 | 5.193 | 5.350 |
| | (1.963) | (1.724) | (1.838) |
| ln(distance to nearest river in km) | 4.414 | 3.767 | 4.052 |
| | (1.673) | (1.980) | (1.876) |
| Annual average temperature in 2000 | 284.774 | 287.581 | 286.344 |
| | (5.914) | (4.765) | (5.474) |
| Annual average wind speed in 2000 | 3.044 | 3.063 | 3.054 |
| | (0.749) | (0.736) | (0.740) |
| Annual average precipitation in 2005 | 2.870 | 3.288 | 3.104 |
| | (1.256) | (1.229) | (1.256) |
| ln(average elevation) | 5.539 | 5.381 | 5.450 |
| | (1.632) | (1.434) | (1.524) |
| ln(min distance to energy sources) | 0.702 | 0.685 | 0.692 |
| | (0.925) | (0.761) | (0.836) |
| 1 if near volcanoes, airfields, or earthquake spots | 0.627 | 0.681 | 0.657 |
| | (0.486) | (0.467) | (0.475) |
| Prefecture gas access (%) | 0.608 | 0.801 | 0.716 |
| | (0.337) | (0.236) | (0.300) |

*Notes:* Standard deviations are printed in parentheses.

Table 2: Factor transformation and weights for computing the suitability index

| Factor | Sub-factor value | Reclassified values |
|---|---|---|
| **Elevation** | 0-1000 m | 10 |
| 0.06 | 1000-1400 m | 8 |
|  | 1400- 1800 m | 4 |
|  | >1800 m | 0 |
| **Slope** | 0-6 % | 10 |
| 0.05 | 6-10 % | 7 |
|  | >10 % | 0 |
| Road | 0-500 m | 0 |
| 0.08 | 0.5 - 10 Km | 10 |
|  | 10-20 Km | 7 |
|  | 20-40 Km | 3 |
|  | >40 Km | 0 |
| Rail | 0 - 500 m | 0 |
| 0.14 | 0.5 - 10 Km | 10 |
|  | 10-20 Km | 7 |
|  | 20-40 Km | 3 |
|  | >40 Km | 0 |
| Distance to urban area | 0-10km | 0 |
| 0.05 | 10-20km | 10 |
|  | 20-50km | 7 |
|  | 50-100km | 4 |
|  | >100km | 0 |
| **Distance to coal sources** | 0-5km | 10 |
| 0.05 | 5-50km | 5 |
|  | > 50km | 0 |
| **Cultivation** | Yes | 5 |
| 0.04 | No | 10 |
| **Gas pipe line** | 0-500 m | 0 |
| 0.08 | 0.5- 5 Km | 10 |
|  | 5-10 Km | 8 |
|  | 10-20 Km | 6 |
|  | 20-40 Km | 3 |
|  | >40 Km | 0 |
| Large river | 0- 500 m | 0 |
| 0.08 | 0.5 - 10 Km | 10 |
|  | 10 - 20 Km | 5 |
|  | >20 Km | 0 |
| **Small river** | 0- 500 m | 0 |
| 0.07 | 0.5 - 10 Km | 10 |
|  | 10 - 20 Km | 5 |
|  | >20 Km | 0 |

*Continued on next page*

Table 2 – *Continued from previous page*

| Factor | Sub-factor value | Reclassify values |
|---|---|---|
| **Distance to earthquake spots** | 0-1km | 0 |
| 0.05 | >1km | 10 |
| **Distance to volcanoes** | 0-1km | 0 |
| 0.02 | >1km | 10 |
| **Distance to airfields** | 0-5km | 0 |
| 0.05 | >5km | 10 |
| Distance to coal mines | 0-5km | 10 |
| 0.1 | 5-50km | 5 |
| | > 50km | 0 |
| Distance to oil and gas fields | 0-5km | 10 |
| 0.08 | 5-50km | 5 |
| | > 50km | 0 |

Table 3: The relationships among suitability, power plant hosting, $SO_2$ pollution levels, and regulation intensity

| | (1) Without any covariates | (2) With economic covariates |
|---|---|---|
| **Panel A: The relationship between suitability and power plant hosting status (probit)** | | |
| Dependent variable: power plant hosting dummy | | |
| Prefecture suitability: mean | 0.660*** | 0.540*** |
| | (0.146) | (0.171) |
| N | 286 | 286 |
| **Panel B: The relationship between power plant hosting and $SO_2$ pollution (linear)** | | |
| Dependent variable: $SO_2$ pollution levels | | |
| Power plant hosting dummy | 0.603*** | 0.310*** |
| | (0.127) | (0.119) |
| N | 286 | 286 |
| $R^2$ | 0.07 | 0.26 |
| **The relationship between $SO_2$ pollution levels and Two Control Zone status (probit)** | | |
| Dependent variable: TCZ dummy | | |
| ln(SO2 density in 1998) | 0.292*** | 0.280*** |
| | (0.068) | (0.079) |
| N | 286 | 286 |

*Notes:* Panel A shows the predictability of thermal power plant suitability (prefecture mean) in prefecture power plant hosting status. Panel B presents the positive association between power plant hosting and local $SO_2$ pollution levels. Panel C suggests that more polluted prefectures in 1998 are more likely to be listed as a Two Control Zone by the State Council. Column (1) presents the simple correlation between the dependent and independent variables and column (2) adds pre-existing economic conditions to the specification.

* $p<0.1$, ** $p<0.05$, *** $p<0.01$. Robust standard errors in parentheses.

Table 4: Predictive power of suitability index for main thermal power plant pollutants (IV first stage)

| Dependent variable: | $\ln(SO_2 \text{ in } 1998)$ | | $\ln(NO_2 \text{ in } 2005)$ | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Prefecture suitability: mean | 0.896*** | | 0.809*** | |
| | (0.083) | | (0.072) | |
| Prefecture suitability: top quartile | | 0.767*** | | 0.708*** |
| | | (0.072) | | (0.061) |
| N | 286 | 286 | 286 | 286 |
| $R^2$ | 0.75 | 0.74 | 0.72 | 0.72 |

*Notes:* All regressions followed the specification of equation 4. The first two columns show results on the main treatment variable, $SO_2$, using prefecture mean and top quartile suitability respectively. The last two columns instead uses another main pollutant from coal-fired power plants, $NO_2$. Its 2005 values are used due to data limitations.

* $p<0.1$, ** $p<0.05$, *** $p<0.01$. Standard errors (in parentheses) are clustered at the province level.

Table 5: Balance in pre-period outcomes

| Dependent variable:<br>Change in logarithm between: | Population<br>1982-1990<br>(1) | Pop. high school<br>1982-2000<br>(2) | Pop. college<br>1990-2000<br>(3) | Employment<br>1982-1990<br>(4) | Manu. emp.<br>1982-1990<br>(5) |
|---|---|---|---|---|---|
| Prefecture suitability: mean | 0.021 | 0.037 | -0.005 | 0.025 | 0.009 |
| | (0.013) | (0.066) | (0.042) | (0.017) | (0.054) |
| N | 286 | 286 | 286 | 286 | 286 |
| $R^2$ | 0.07 | 0.21 | 0.30 | 0.04 | 0.21 |
| Mean of dependent variable | 0.13 | 0.71 | 1.14 | 0.24 | 0.48 |

*Notes:* This table reports the OLS estimates of the instrumental variable on pre-2000 outcomes as falsification tests. All regressions controlled for the same set of covariates as the baseline regression. The dependent variable in columns 2 and 3 are the changes in population with high school degree and above and college degree and above. The pre-period years are inconsistent across columns due to data limitations.

* p<0.1, ** p<0.05, *** p<0.01. Standard errors (in parentheses) are clustered at the province level.

Table 6: Effects of APPCL and *hukou* reform on the local economy

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| **Panel A: change in ln(employment)** | Regulated sectors | | Non-regulated sectors | | Total |
| Sectors: | Industrial | Secondary | Primary | Tertiary | All |
| Sample period | 1998-2007 | 2000-2010 | 2000-2010 | 2000-2010 | 2000-2010 |
| **IV:** | | | | | |
| ln(SO2 density in 1998) | -0.095* | -0.175*** | 0.005 | -0.002 | 0.023 |
| | (0.049) | (0.046) | (0.033) | (0.026) | (0.025) |
| N | 286 | 286 | 286 | 286 | 286 |
| First-stage F stat | 117.32 | 117.32 | 117.32 | 117.32 | 117.32 |
| Mean of dependent variable | 0.19 | 0.46 | -0.25 | 0.40 | 0.05 |
| **OLS:** | | | | | |
| ln(SO2 density in 1998) | -0.014 | -0.019 | -0.047* | 0.032** | -0.000 |
| | (0.039) | (0.029) | (0.023) | (0.014) | (0.013) |
| N | 286 | 286 | 286 | 286 | 286 |
| $R^2$ | 0.35 | 0.42 | 0.47 | 0.23 | 0.27 |
| **Panel B: change in ln(population)** | Education levels | | *Hukou* status | | Total |
| | Low-skilled | High-skilled | Urban *hukou* | Rural *hukou* | All |
| **IV:** | | | | | |
| ln(SO2 density in 1998) | 0.003 | -0.102** | -0.085*** | -0.035 | 0.014 |
| | (0.022) | (0.040) | (0.032) | (0.039) | (0.025) |
| N | 286 | 286 | 286 | 285 | 286 |
| First-stage F stat | 117.32 | 117.32 | 117.32 | 115.20 | 117.32 |
| Mean of dependent variable | 0.01 | 0.99 | 0.39 | -0.16 | 0.05 |
| **OLS:** | | | | | |
| ln(SO2 density in 1998) | -0.009 | -0.028 | -0.005 | -0.041* | -0.008 |
| | (0.010) | (0.032) | (0.025) | (0.021) | (0.013) |
| N | 286 | 286 | 286 | 285 | 286 |
| $R^2$ | 0.23 | 0.38 | 0.24 | 0.16 | 0.37 |

*Notes:* Panel A follows the equation 1 specification and presents the IV and OLS estimates on the effect of the APPCL environmental regulation on local employment growth by sector. Panel B follows the equation 2 specification to show the IV and OLS estimates on the effect of air quality and environmental regulation on local population growth by education levels and *hukou* urban/rural status.

* $p<0.1$, ** $p<0.05$, *** $p<0.01$. Standard errors (in parentheses) are clustered at the province level.

Table 7: Local labor market adjustment following APPCL, IV estimates

| Dependent variable: change in ln(employment): | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Ownership type: | All firms | | State owned firms | | Non-state owned firms | |
| Coal usage intensity: | heavy | light | heavy | light | heavy | light |
| **Panel A: 1999-2007** | | | | | | |
| ln(SO2 density in 1998) | -0.202** | 0.081 | -0.515** | 0.335** | -0.020 | -0.082 |
| | (0.084) | (0.086) | (0.211) | (0.153) | (0.127) | (0.083) |
| N | 286 | 286 | 283 | 286 | 284 | 286 |
| First-stage F stat | 117.32 | 117.32 | 113.34 | 117.32 | 107.13 | 117.32 |
| Mean of dependent variable | -0.05 | 0.22 | -1.04 | -1.51 | 0.83 | 0.98 |
| **Panel B: 1999-2003** | | | | | | |
| ln(SO2 density in 1998) | -0.215** | 0.118 | -0.529*** | 0.133 | -0.004 | -0.029 |
| | (0.100) | (0.082) | (0.156) | (0.160) | (0.117) | (0.048) |
| N | 286 | 286 | 285 | 286 | 284 | 286 |
| First-stage F stat | 117.32 | 117.32 | 117.78 | 117.32 | 107.13 | 117.32 |
| Mean of dependent variable | -0.16 | -0.03 | -0.58 | -0.76 | 0.38 | 0.47 |
| **Panel C: 2003-2007** | | | | | | |
| ln(SO2 density in 1998) | 0.014 | -0.037 | -0.014 | 0.202** | -0.015 | -0.052 |
| | (0.058) | (0.030) | (0.106) | (0.102) | (0.082) | (0.054) |
| N | 286 | 286 | 283 | 286 | 285 | 286 |
| First-stage F stat | 117.32 | 117.32 | 113.34 | 117.32 | 109.41 | 117.32 |
| Mean of dependent variable | 0.11 | 0.25 | -0.47 | -0.74 | 0.45 | 0.51 |

*Notes:* This table presents the environmental regulation effect on local employment growth by firm ownership types and coal usage intensity (columns 1-6) as well as by sub-periods (panels A-C). All regressions followed the baseline IV specification, where ln($SO_2$ density in 1998) serves as a proxy for the APPCL enforcement intensity. The estimation results shed light on potential local labor market adjustments by sector and over time.

* $p<0.1$, ** $p<0.05$, *** $p<0.01$. Standard errors (in parentheses) are clustered at the province level.

Table 8: Regulatory mechanisms of APPCL

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| State owned firms by two periods: | 1999-2003 | | 2003-2007 | |
| Coal usage intensity: | heavy | light | heavy | light |
| **Panel A: change in ln(value-added)** | | | | |
| ln(SO2 density in 1998) | -0.432** | 0.206 | 0.296 | 0.350*** |
| | (0.191) | (0.152) | (0.222) | (0.102) |
| N | 276 | 284 | 278 | 284 |
| First-stage F stat | 120.09 | 118.78 | 111.57 | 115.52 |
| Mean of dependent variable | 0.10 | -0.01 | 0.51 | 0.11 |
| **Panel B: change in ln(capital)** | | | | |
| ln(SO2 density in 1998) | -0.730*** | 0.153 | 0.318 | 0.277** |
| | (0.239) | (0.148) | (0.212) | (0.120) |
| N | 285 | 286 | 283 | 286 |
| First-stage F stat | 117.78 | 117.32 | 113.34 | 117.32 |
| Mean of dependent variable | -0.08 | -0.30 | 0.03 | -0.29 |
| **Panel C: change in ln(wage bill)** | | | | |
| ln(SO2 density in 1998) | -0.660*** | 0.139 | 0.177 | 0.281** |
| | (0.184) | (0.172) | (0.152) | (0.111) |
| N | 285 | 286 | 283 | 286 |
| First-stage F stat | 117.78 | 117.32 | 113.34 | 117.32 |
| Mean of dependent variable | -0.04 | -0.24 | 0.23 | -0.13 |
| **Panel D: change in ln(number of firms)** | | | | |
| ln(SO2 density in 1998) | -0.180** | 0.037 | 0.060 | 0.213*** |
| | (0.086) | (0.073) | (0.082) | (0.079) |
| N | 285 | 286 | 283 | 286 |
| First-stage F stat | 117.78 | 117.32 | 113.34 | 117.32 |
| Mean of dependent variable | -0.54 | -0.77 | -0.54 | -0.88 |
| **Panel E: change in ln(wage per worker)** | | | | |
| ln(SO2 density in 1998) | -0.132** | 0.006 | 0.191** | 0.079 |
| | (0.061) | (0.043) | (0.078) | (0.076) |
| N | 285 | 286 | 283 | 286 |
| First-stage F stat | 117.78 | 117.32 | 113.34 | 117.32 |
| Mean of dependent variable | 0.54 | 0.52 | 0.70 | 0.61 |

*Notes:* This table studies the APPCL regulatory mechanism by examining the prefecture-level growth in terms of value added, capital, total wage, number of firms, and average wage in state owned firms, the most affected firms. All regressions followed the baseline IV specification. The estimates in panels A-C are driven by slower growth of firms (panel D). Panel E results indicate that firm productivity growth is slower in more regulated prefectures.

* $p<0.1$, ** $p<0.05$, *** $p<0.01$. Standard errors (in parentheses) are clustered at the province level.

# Appendix A1 Robustness and sensitivity checks

In Tables A1 and A2, I present comprehensive tests to exhibit the robustness of the identification strategy and results.

Table A1 checks the validity of the instrumental variable. With all key results laid out across the columns, panel A re-prints the IV baseline estimates for comparison. Panel B removes the set of covariates that are themselves factors in the suitability construction. Results only deviate slightly in the magnitude of point estimates. Their similarity suggests that this instrumental variable works through the weights assigned by engineers for each factor, rather than the factor values.

Panel C uses an alternative instrumental variable where the suitability measure is constructed with a limited subset of factors, printed in bold in Table 2. This alternative IV is hence based on factors that are less likely to have economic impacts themselves. For example, distance to small rivers is kept while that to large rivers is dropped. Arriving at close estimates in this panel shows that the most effective parts of the baseline IV are contributed by non-economic factors.

Panel D addresses the concern over whether the prefecture level mean of suitability measure is a sensible choice. Results presented here use the top quartile of suitability measure at the prefecture level as IV. Using the top quartile rather than the mean shows a slightly better first-stage F-stat, while the point estimates are very similar.

The following table, Table A2, explore the sensitivity of estimates across various alternative specifications and treatment variables, with baseline results again printed in panel A. Panel B excludes 25% of the sample that are in the bottom quartile of 1998 $SO_2$ pollution level. This address the concern that some prefectures are too clean and undeveloped for comparison. Using the top 75% of the sample arrives at the same results, although point estimates are slightly larger in magnitude as expected.

Panel C includes the initial level of the outcome variable as covariate. While including the initial level deals with concern over regression toward the mean, it does not change the conclusion. Panel D answers the question of whether or not initial industrial composition affect baseline results. Having excluded 1995 share of employment in heavy coal-usage industries, results are consistent while point estimates vary slightly. Panel E includes a heating dummy following the research design of Huai River Policy used in Chen et al. (2013). As Panel E shows similar results to the baseline, there is little concern over potential bias from systematic sources of energy use such as heating.

In Panel F, the TCZ dummy is used as an alternative treatment variable to directly reflect a prefecture's status of being regulated stringently. It is set to 1 if the prefecture has been listed as the key target prefecture by the State Council in 1998, and 0 otherwise. This is a more direct treatment variable proxying for regulatory cost. However its binary nature is unable to capture for as much variation as a continuous variable such as the 1998 $SO_2$

level. Moreover, the TCZ status is not measuring regulation versus no regulation. Rather, it switches on for being on the State Council's list, potentially facing stricter enforcement and closer monitoring than the rest. Lastly, using 2010 $SO_2$ levels in panel F reiterates the fact that the spatial variation of ambient $SO_2$ levels had not changed much between 1998 and 2010, a point made earlier in section 3.

Panels G through I present estimates based on alternative pollution-based treatment variables. Panel G uses per capital 1998 $SO_2$ pollution levels. Panel H uses the 1998 $PM_{2.5}$ levels. And Panel I uses the 2010 $SO_2$ levels from satellite images of much better resolution (the NASA OMI mission). All three alternative treatment variables address different concerns but they all show results consistent to the baseline.

Table A1: IV validity checks

| | college edu 00-10 (1) | total pop 00-10 (2) | secondary emp 00-10 (3) | employment 00-10 (4) | coal emp 99-07 (5) | coal emp state 99-07 (6) |
|---|---|---|---|---|---|---|
| **Panel A: IV baseline** | | | | | | |
| ln(SO2 density in 1998) | -0.102** | 0.014 | -0.175*** | 0.023 | -0.202** | -0.515** |
| | (0.040) | (0.025) | (0.046) | (0.025) | (0.084) | (0.211) |
| N | 286 | 286 | 286 | 286 | 286 | 283 |
| First-stage F stat | 117.32 | 117.32 | 117.32 | 117.32 | 117.32 | 113.34 |
| **Panel B: excluding IV components as covariates** | | | | | | |
| ln(SO2 density in 1998) | -0.138*** | 0.010 | -0.158*** | 0.012 | -0.188*** | -0.290** |
| | (0.031) | (0.017) | (0.036) | (0.021) | (0.052) | (0.137) |
| N | 286 | 286 | 286 | 286 | 286 | 283 |
| First-stage F stat | 130.15 | 130.15 | 130.15 | 130.15 | 130.15 | 128.87 |
| **Panel C: IV constructed with limited geographic factors** | | | | | | |
| ln(SO2 density in 1998) | -0.102** | 0.014 | -0.174*** | 0.025 | -0.197** | -0.504** |
| | (0.040) | (0.025) | (0.045) | (0.026) | (0.081) | (0.206) |
| N | 286 | 286 | 286 | 286 | 286 | 283 |
| First-stage F stat | 118.16 | 118.16 | 118.16 | 118.16 | 118.16 | 114.06 |
| **Panel D: using top quartile suitability as instrument** | | | | | | |
| ln(SO2 density in 1998) | -0.114*** | 0.016 | -0.190*** | 0.024 | -0.194** | -0.492** |
| | (0.040) | (0.023) | (0.044) | (0.027) | (0.080) | (0.197) |
| N | 286 | 286 | 286 | 286 | 286 | 283 |
| First-stage F stat | 112.57 | 112.57 | 112.57 | 112.57 | 112.57 | 100.93 |

* $p<0.1$, ** $p<0.05$, *** $p<0.01$. Standard errors (in parentheses) are clustered at the province level. Each panel is a different specification showing estimates on the main results.

Table A2: Robustness and sensitivity tests

| | college edu<br>00-10<br>(1) | total pop<br>00-10<br>(2) | secondary emp<br>00-10<br>(3) | employment<br>00-10<br>(4) | coal emp<br>99-07<br>(5) | coal emp state<br>99-07<br>(6) |
|---|---|---|---|---|---|---|
| **Panel A: IV baseline** | | | | | | |
| ln(SO2 density in 1998) | -0.102** | 0.014 | -0.175*** | 0.023 | -0.202** | -0.515** |
| | (0.040) | (0.025) | (0.046) | (0.025) | (0.084) | (0.211) |
| N | 286 | 286 | 286 | 286 | 286 | 283 |
| First-stage F stat | 117.32 | 117.32 | 117.32 | 117.32 | 117.32 | 113.34 |
| **Panel B: excluding bottom 25% in 1998 $SO_2$** | | | | | | |
| ln(SO2 density in 1998) | -0.104** | 0.034 | -0.263*** | 0.050 | -0.276** | -0.630** |
| | (0.052) | (0.032) | (0.064) | (0.032) | (0.112) | (0.311) |
| N | 257 | 257 | 257 | 257 | 257 | 254 |
| First-stage F stat | 75.75 | 75.75 | 75.75 | 75.75 | 75.75 | 72.96 |
| **Panel C: conditional on initial level** | | | | | | |
| ln(SO2 density in 1998) | -0.104*** | -0.006 | -0.173*** | 0.004 | -0.068 | -0.490** |
| | (0.033) | (0.021) | (0.047) | (0.022) | (0.075) | (0.233) |
| N | 286 | 286 | 286 | 286 | 286 | 283 |
| First-stage F stat | 105.68 | 104.30 | 93.58 | 109.36 | 85.15 | 85.21 |
| **Panel D: exclude share of heavy polluting employment in 1995** | | | | | | |
| ln(SO2 density in 1998) | -0.096*** | 0.012 | -0.174*** | 0.010 | -0.158** | -0.410** |
| | (0.037) | (0.023) | (0.048) | (0.025) | (0.068) | (0.182) |
| N | 286 | 286 | 286 | 286 | 286 | 283 |
| First-stage F stat | 114.43 | 114.43 | 114.43 | 114.43 | 114.43 | 111.73 |
| **Panel E: controlling for Huai River policy dummy** | | | | | | |
| ln(SO2 density in 1998) | -0.104** | 0.014 | -0.179*** | 0.028 | -0.213** | -0.548*** |
| | (0.042) | (0.026) | (0.048) | (0.027) | (0.087) | (0.209) |
| N | 286 | 286 | 286 | 286 | 286 | 283 |
| First-stage F stat | 102.12 | 102.12 | 102.12 | 102.12 | 102.12 | 94.99 |
| **Panel F: treatment variable is Two-Control Zone status** | | | | | | |

*Continued on next page*

Table A2 – *Continued from previous page*

| | college edu<br>00-10<br>(1) | total pop<br>00-10<br>(2) | secondary emp<br>00-10<br>(3) | employment<br>00-10<br>(4) | coal emp<br>99-07<br>(5) | coal emp state<br>99-07<br>(6) |
|---|---|---|---|---|---|---|
| Two-control zone | -0.475** | 0.065 | -0.817*** | 0.108 | -0.940*** | -2.446*** |
| | (0.204) | (0.118) | (0.255) | (0.124) | (0.338) | (0.915) |
| N | 286 | 286 | 286 | 286 | 286 | 283 |
| First-stage F stat | 18.08 | 18.08 | 18.08 | 18.08 | 18.08 | 18.60 |
| **Panel G: treatment variable is per capita $SO_2$ levels** | | | | | | |
| ln(SO2 density/person in 1998) | -0.077*** | 0.011 | -0.132*** | 0.018 | -0.152** | -0.389** |
| | (0.029) | (0.018) | (0.031) | (0.018) | (0.062) | (0.156) |
| N | 286 | 286 | 286 | 286 | 286 | 283 |
| First-stage F stat | 132.05 | 132.05 | 132.05 | 132.05 | 132.05 | 121.72 |
| **Panel H: treatment variable is 1998 $PM_{2.5}$** | | | | | | |
| ln(PM2.5 in 1998) | -0.383** | 0.052 | -0.659*** | 0.087 | -0.758** | -1.970** |
| | (0.168) | (0.094) | (0.227) | (0.098) | (0.329) | (0.914) |
| N | 286 | 286 | 286 | 286 | 286 | 283 |
| First-stage F stat | 33.63 | 33.63 | 33.63 | 33.63 | 33.63 | 31.34 |
| **Panel I: treatment variable is 2010 $SO_2$ levels** | | | | | | |
| ln(SO2 in 2010) | -0.103** | 0.014 | -0.178*** | 0.024 | -0.205** | -0.535** |
| | (0.041) | (0.025) | (0.048) | (0.026) | (0.085) | (0.221) |
| N | 286 | 286 | 286 | 286 | 286 | 283 |
| First-stage F stat | 91.85 | 91.85 | 91.85 | 91.85 | 91.85 | 86.96 |

# Appendix A2 Theoretical framework

The simple theoretical model presented here aims to formalize the relationship between prefecture-level ambient $SO_2$ pollution, regulation intensity, and local growth prospects in the context of China's policy changes on environmental regulation and labor mobility. It follows largely the works of Moretti (2011) and Allcott and Keniston (2017). There are two prefectures, $c \in \{a, b\}$ with prefecture specific air quality, $A_c$. Workers have skill types either high or low $k \in \{h, l\}$. Each prefecture has two labor-intensive production sectors, $j \in \{m, s\}$. $m$ represents the non-polluting sector and $s$ for the polluting sector. Both sectors employ low-skilled workers (at skill level $l$). Additionally, there is a skill-intensive sector $f$, employing workers of skill level $h$. All three goods are internationally traded and have exogenously determined prices $P_m$, $P_s$, and $P_f = 1$.

## A2.1 Production

### A2.1.1 Demand for low-skilled workers

Each sector $j \in \{m, s\}$ has a composite firm employing $N^l_{jc}$ low-skilled workers and earning revenue $R^l_{jc} = X^l_{jc}(N^l_{jc})^{1-\gamma}$, where $\gamma \in (0, 1)$. $X^l_{jc}$ is revenue productivity[17], and change in $X^l_{sc}$ will be used to capture cost of APPCL regulation in the polluting sector. Low-skilled workers have homogeneous productivity and hence wage level is equalized across the two sectors $m, s$ within a prefecture. Firms in a given labor-intensive sector have profit equation $\Pi^l_{jc} = R^l_{jc} - W^l_c N^l_{jc}$. Assuming that workers are paid their marginal productivity, firm profit maximization gives the wage equation: $W^l_c = (1-\gamma)X^l_{jc}(N^l_{jc})^{-\gamma}$. Rearranging it obtains demand of low-skilled workers by a typical firm in $j$ and prefecture $c$: $N^l_{jc} = (\frac{(1-\gamma)X^l_{jc}}{W^l_c})^{\frac{1}{\gamma}}$. Aggregating across $m, s$ to obtain prefecture-level labor demand $N^l_c = N^l_{mc} + N^l_{sc}$, which is in detail:

$$N^l_c = (\frac{(1-\gamma)}{W^l_c})^{\frac{1}{\gamma}}({X^l_{mc}}^{\frac{1}{\gamma}} + {X^l_{sc}}^{\frac{1}{\gamma}})$$

Taking logs (denoted with corresponding lower cases):

$$n^l_c = \frac{1}{\gamma}\ln(1-\gamma) - \frac{1}{\gamma}w^l_c + \widetilde{x}_c, \text{ where } \widetilde{x}_c \equiv \ln({X^l_{mc}}^{\frac{1}{\gamma}} + {X^l_{sc}}^{\frac{1}{\gamma}})$$

Demand for low-skilled workers in prefecture $a$ relative to prefecture $b$ is then:

$$n^l_a - n^l_b = -\frac{1}{\gamma}(w^l_a - w^l_b) + (\widetilde{x}_a - \widetilde{x}_b) \tag{5}$$

Relative demand of low-skilled workers in prefecture $a$ will be higher if its relative wage for low-skilled workers is lower and their relative productivity is higher.

---

[17] Revenue productivity by definition is $X^l_{jc} = P_j\Omega^l_{jc}$, price of the product times and physical productivity.

### A2.1.2 Demand for high-skilled workers

For simplicity, the skill-intensive sector $f$ is similar to $m, s$ except for the skill type. The composite firm employs $N_{fc}^h$ high-skilled workers and earning revenue $R_{fc}^h = X_{fc}^h (N_{fc}^h)^{1-\gamma}$. Therefore, the relative labor demand for high-skilled workers in prefecture $a$ is:

$$n_a^h - n_b^h = -\frac{1}{\gamma}(w_a^h - w_b^h) + \frac{1}{\gamma}(x_{fa}^h - x_{fb}^h) \tag{6}$$

## A2.2 Housing market

Following Moretti (2011), I assume that the number of housing units in prefecture $c$ equals to the number of workers with housing supply elasticity of $g$: $p_{hc}^k = g n_c^k + g_0$. $1/g$ is thus the price elasticity of housing supply that does not vary by prefecture or skill. Similar to the local labor market, the housing market is also segregated by skill. The housing price difference between $a$ and $b$ is:

$$p_{ha}^k - p_{hb}^k = g(n_a^k - n_b^k) \tag{7}$$

The implication is straightforward that the relative housing price of a prefecture for a given skill type will be higher as more workers of that skill type move in.

## A2.3 Consumers and workers

Individuals decide where to live and how much to consume by supplying 1 unit of labor. A representative individual consumes $C_{im}, C_{is}, C_{if}, C_{ih}$ units of labor-intensive polluting, non-polluting, and skill-intensive goods, as well as housing. Their utility also depends on amenity, $A_c$, for which in this paper I focus on variation of local $SO_2$ level. I also incorporate the idiosyncratic taste term $E_{ic}$ from Moretti (2011) to have imperfect labor mobility throughout. Both $A_c$ and $E_{ic}$ are skill-specific. Specifically, high-skilled workers value air quality more than low-skilled workers.

In order to incorporate migration costs that realistically describe the effect of *hukou*, I introduce a wage multiplier term $D_c^k$ to consumer budget constraint. $D_c^k \geq 1$ represents the difference in local social benefits one receives at their *hukou* prefecture $c$ relative to the non-*hukou* city. This is similar to that in Tombe and Zhu (2019) but instead of a flow measure I model it as sunk cost. This term is again skill-biased to capture the fact that migration costs imposed by *hukou* have been lower for high-skilled workers (i.e., $D_c^l > D_c^h$).

Individuals in each skill type maximize their Cobb-Douglas utility:

$$U_{ic}^k = C_{ih}^{\alpha} C_{is}^{\beta} C_{im}^{\lambda} C_{if}^{1-\alpha-\beta-\lambda} A_c^k E_{ic}^k$$

subject to $W_c^k D_c^k - P_m C_{im} - P_s C_{is} - P_f C_{if} - P_{hc} C_{ih} \geq 0$, with $\alpha, \beta, \lambda, 1-\alpha-\beta \in (0,1)$.

The indirect utility is then $u_{ic}^k = w_c^k + d_c^k - \alpha p_{hc} - \beta p_s - \lambda p_m + a_c^k + e_{ic}^k$ plus some constant.

Individual with *hukou* in prefecture $a$ is indifferent between prefectures $a$ and $b$ if $u^k_{ia} = u^k_{ib}$, which is equivalent to:

$$w^k_a - w^k_b = \alpha g(n^k_a - n^k_b) - (a^k_a - a^k_b) - (e^k_{ia} - e^k_{ib}) - d^k_a$$

By assuming that $e^k$ is distributed type I extreme value with scale parameter $(s^k)^2$ where $s^k \in (0, \infty)$, we have $e^k_{ia} - e^k_{ib} = -s^k(n^k_a - n^k_b)$. As defined in Moretti (2011), $s$ characterizes the degree of labor mobility based on preference. Larger $s$ means that workers have stronger attachment to their current location and are therefore less likely to migrate. Further combining with equation 7, we obtain the relative labor supply in prefecture $a$:

$$n^k_a - n^k_b = \frac{1}{\alpha g + s^k}[(w^k_a - w^k_b) + (a^k_a - a^k_b) + d^k_a] \tag{8}$$

For both skill types, the relative labor supply in prefecture $a$ will be higher if its relative wage is higher and air quality is better. Additionally, better social benefits that workers receive in their *hukou* prefecture $d^k_c$ is the opportunity cost to migrate. Lastly, these effects are scaled by skill-specific locational preference $s^k$ and housing price elasticity $\alpha g$.

## A2.4 Equilibrium conditions

In the labor market equilibrium, there is equalization within skill type between the prefecture-level labor demand and supply. Combining equations 5, 6 and 8, we obtain population differences between $a$ and $b$ by skill type:

$$\text{low-skilled: } n^l_a - n^l_b = \frac{1}{\alpha g + \gamma + s^l}[\gamma(\widetilde{x}_a - \widetilde{x}_b) + (a^l_a - a^l_b) + d^l_a] \tag{9}$$

$$\text{high-skilled: } n^h_a - n^h_b = \frac{1}{\alpha g + \gamma + s^h}[(x^h_{fa} - x^h_{fb}) + (a^h_a - a^h_b) + d^h_a] \tag{10}$$

As standard Roback models would suggest, the population difference for both skill types is determined by differences in city-level productivity (positively) and amenity/air quality (positively). However, the effects are scaled differently by skill as high- and low-skilled workers are assumed to have different distributions of taste for their hometown ($s^k$) and valuation of amenities such as air quality ($a^k$).

This model has been adapted for the context of this paper by adding the term of migration cost, $d^k_c$. The standard Roback effects driving population differences across prefectures are muted under the *hukou* system due to the large discrepancy in social benefits between *hukou* and non-*hukou* locations that an individual is entitled to. The scale of this migration cost keeps the actual relative population sizes of both skills in prefecture $a$ artificially higher than they otherwise would be if $d^k_c = 0$.

## A2.5 The impact of APPCL on the local economy

This section analyzes the effect of APPCL implementation in light of *hukou* reform. Prior to these two concurrent policy changes, $d_c^k$ is assumed to be infinitely large so that there is no migration between the two prefectures. The predicted effects in discussion are therefore relative to the initial arbitrary population distribution across prefectures.

### A2.5.1 Effects on low-skilled workers

By design, the implementation of APPCL only affects the polluting, low-skilled labor-intensive sector. This regulation intensity is correlated to the local air quality, $a_c$, so that the more polluted prefecture will be more stringently regulated. Suppose that both prefectures are equally productive in their labor-intensive sectors, but $SO_2$ pollution is exogenously more severe in prefecture $a$, $A_a < A_b$. Then the implementation of APPCL will be more stringent in $a$. As a result, the negative productivity shock is stronger in prefecture $a$, i.e., $\widetilde{x}_a - \widetilde{x}_b < 0$. This will then reduce demand for low-skilled workers in the polluting sector, and in prefecture $a$ relative to prefecture $b$ depending on the degree of reallocation from polluting to non-polluting sectors locally (equation 5).

How does this adverse environmental regulation effect on the labor demand for low-skilled workers translate into local population changes is further determined by the preference for air quality and migration costs of low-skilled workers (equation 9). Since pollution is assumed to be more severe in prefecture $a$ and its relative ranking does not change over time (as shown in Figure A1), $(a_a^l - a_b^l)$ is also negative. Under *hukou* reform, the migration cost $d_c^l$ becomes smaller but still remains substantially large for low-skilled workers as discussed in 2.3, i.e., $d_a^l$ is large and positive.

In other words, the APPCL reduces the polluting sector employment in prefecture $a$ relative to prefecture $b$. When changes in the local low-skilled population is considered, the worse air quality in prefecture $a$ could add to the outflow of low-skilled workers. However, both of the negative effects can be moderated by *hukou* imposed high migration cost on low-skilled workers.

### A2.5.2 Effects on high-skilled workers

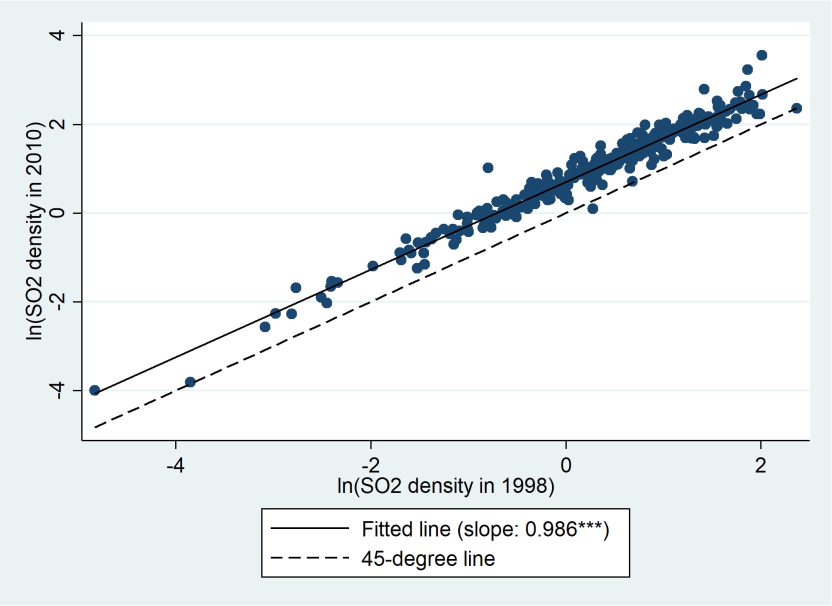
The local employment prospects for high-skilled workers are not affected by APPCL by assumption. In other words, $x_{fa}^h - x_{fb}^h$ is orthogonal to $a_a^h - a_b^h$. But concurrent to the APPCL implementation, *hukou* reform substantially lowered the migration cost $(d_c^h)$ for high-skilled workers, allowing for their preference for air quality to manifest. Since prefecture $a$ has worse air quality, equation 10 suggest that high-skilled population in prefecture $a$ will decrease relative to prefecture $b$.

## Appendix A3 Data source

Table A3: Sources of data for the computation of suitability index

| Name: | Source: | Details: |
|---|---|---|
| $SO_2$ column density | NASA Global Modeling and Assimilation Office | Time span: 1998-2010, annual means<br>Scale: $0.5° \times 0.625°$; unit: kg $m^{-2}$ |
| Infrastructure & land use | United States Geological Survey: Coal Geology, Land Use, and Human Health in China Compiled by Alex W. Karlsen et al. | List of shapefiles compiled by 2001<br>Key variables: airfields, small/large rivers, earthquake spots, volcanoes, coal bearing fields |
| Coal mines, oil & gas fields | USGS Compilation of GIS Data Representing Coal Mines and Coal-Bearing Areas in China | Surveyed by 2001, published in Jan 2015<br>Key info: coal mines, oil and gas fields |
| Gas pipelines | China Natural Gas Pipelines Dataset, Harvard ChinaMap | Gas pipelines and nodes |
| Elevation | Harvard China Historical GIS V5 DEM | 1km pixel resolution based on GTOPO-30 data from USGS |
| Land cover | Global Land Cover Facility, MODIS Land Cover | $5' \times 5' \approx 0.083°$ resolution, yearly 2001-2012 |

Figure A1: Relative ranking of ambient $SO_2$ levels between 1998 and 2010



The downtrend of $SO_2$ change between 1998 and 2010 in prefectures (as shown in figure 2) is however not enough to alter prefectures' rankings of pollution levels.