The London School of Economics and Political Science

Essays on Environmental and Urban Economics

Yatang Lin

A thesis submitted to the Department of Economics of the London School of Economics for the degree of Doctor of Philosophy, London, July 2017

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of about 64000 words.

Statement of inclusion of previous work

I can confirm that Chapter 2 is extended from a previous study for the Master of Research (MRes) degree in Economics at the LSE, awarded in July 2014.

Statement of use of third party for editorial help

I can confirm that chapters 1 and 2 of my thesis were copy edited for conventions of language, spelling and grammar by Eve Richards.

Acknowledgement

This document is the product of the time and thoughtfulness of many people.I am deeply indebted to my supervisors, Daniel Sturm, Guy Michaels and Robin Burgess, who not only guide my development as a researcher and provide moral support whenever needed, but also inspire me with their enduring enthusiasm for research. I am also extremely grateful to Vernon Henderson and Alwyn Young, who provided very helpful advice and continued support throughout my PhD.

I had the opportunity to present my work at numerous LSE seminars, where I received valuable comments and equally valuable criticism. In particular, I wish to thank Jeremiah Dittmar, Gianmarco Ottaviano, Steve Pischke and Thomas Sampson for very helpful comments that definitely improved my research.

I would also like to thank my PhD colleagues and friends in London, including Clare Balboni, Marcus Biermann, Shiyu Bo, Hao Dong, Vicky Fan, Jason Garred, Jiajia Gu, Chao He, Yan Liang, Yu-Hsiang Lei, Tsogsag Nyamdavaa, Frank Pisch, Xuezhu Shi, Eddy Tam, Sophie Wang, Bixuan Xu, Junichi Yamasaki, Tianle Zhang, and many others, for making my LSE years special.

Finally, I wish to express my deepest gratitude to my parents and grandparents, whose love, support and guidance are behind every successful step of my life. I dedicate this thesis to the memory of my late grandmother, for bringing me up and being my first teacher of life.

Abstract

The thesis consists of three independent chapters on environmental and urban economics. A central theme explored in this thesis is what determines the distribution of economic activities across space. My exploration in this direction begins with the roles of industrial pollution and transportation infrastructure in shaping the spatial distribution of skills, and extends to evaluate the spatial allocation efficiency of renewable energy projects.

The first chapter,***"The Long Shadow of Industrial Pollution: Environmental Amenities and the Distribution of Skills"***, investigates the role of industrial pollution in determining the competitiveness of post-industrial cities, with a focus on their ability to attract skilled workers and shift to a modern service economy. I assemble a rich database at a fine spatial resolution, which allows me to track pollution from the 1970s to the present and to examine its impacts on a whole range of outcomes related to productivity and amenity, including house prices, employment, wages, and crime. I find that census tracts downwind of highly polluted 1970s industrial sites are associated with lower housing prices and a smaller share of skilled employment three decades later, a pattern which became evermore prominent between 1980 and 2000. These findings indicate that pollution in the 1970s affected the ability of parts of cities to attract skills, which in turn drove the process of agglomeration based on modern services. To quantify the contribution of different mechanisms, I build and estimate a multi-sector spatial equilibrium framework that introduces heterogeneity in local productivity and workers' valuation of local amenities across sectors and allows the initial sorting to be magnified by production and residential externalities. Structural estimation suggests that historical pollution is associated with lower current productivity and amenity; the magnitudes are higher for productivity, more skilled sectors and central tracts. I then use the framework to evaluate the impact of counterfactual pollution cuts in different parts of cities on nationwide welfare and cross-city skill distribution.

The second chapter, ***"Travel Costs and Urban Specialization: Evidence from China's High Speed Railway"*** examines how improvements in passenger

transportation affect the spatial distribution of skills, exploiting the expansion of high speed railway (HSR) project in China. This natural experiment is unique because as a passenger-dedicated transportation device that aims at improving the speed and convenience of intercity travel, HSR mostly affects urban specialization through encouraging more frequent intercity trips and face-to-face interactions. I find that an HSR connection increases city-wide passenger flows by 10% and employment by 7%. To further deal with the issues of endogenous railway placement and simultaneous public investments accompanying HSR connections, I examine the impact of a city's market access changes purely driven by the HSR connection of other cities. The estimates suggest that HSR-induced expansion in market access increases urban employment with an elasticity between 2 and 2.5. The differential impacts of HSR on employment across sectors suggest that industries benefiting more from enhanced market access are the ones intensive in nonroutine cognitive skills, such as finance, IT and business services. These findings highlight the role of improved passenger travel infrastructure in promoting the delivery of services, facilitating labour sourcing and knowledge exchange across cities, and ultimately shifting the specialization pattern of connected cities towards skilled and communication intensive sectors.

In the last chapter, ***"Where does the Wind Blow? Green Preferences and Spatial Misallocation in the Renewable Energy Sector"***, I focus on the spatial allocation efficiency of renewable energy projects. How efficiently are renewable energy projects distributed across the US? Are "greener" investors worse at picking sites? Using extensive information on wind resources, transmission, electricity prices and other restrictions that are relevant to the siting choices of wind farms, I calculate the predicted profitability of wind power projects for all possible locations across the contiguous US, use this distribution of this profitability as a counterfactual for profit-maximizing wind power investments and compare it to the actual placement of wind farms. The average predicted profit of wind projects would have risen by 47.1% had the 1770 current projects in the continental US been moved to the best 1770 sites. I also show that 80% and 42% respectively of this observed deviation can be accounted for by within-state and within-county distortions. I provide further evidence that a large proportion of the observed within-state spatial misallocation is related to green investors' tendency of invest locally and sub-optimally. Wind farms in more environmentally-friendly counties are more likely to be financed by local and non-profit investors, are closer to cities, are much less responsive to local fundamentals and have worse performance ex-post. The implementation of state policies such as Renewable Portfolio Standard (RPS) and price-based subsidies are related to better within-state locational choices through attracting more for-profit investments to the “brown" counties, while lump-sum subsidies have the opposite or no effects. My findings have salient implications for environmental and energy policy. Policy makers should take account of the non-monetary incentives of renewable investors when determining the allocative efficiency of policies.

Contents

List of Figures

List of Tables

Chapter 1

The Long Shadow of Industrial Pollution: Environmental Amenities and the Distribution of Skills

This paper presents theory and evidence on the role of environmental amenities in shaping the competitiveness of post-industrial cities. I assemble a rich database at a fine spatial resolution to examine the impact of historical pollution on the distribution of skilled workers and residents within cities today. I find that census tracts downwind of highly polluted 1970s industrial sites were associated with higher pollution levels in the 1970s but not after 2000. However, they were less skilled and had lower wage and housing values in 2000, a pattern which became more prominent between 1980 and 2000. These findings suggest the presence of skill sorting on pollution and strong subsequent agglomeration effects. To quantify the contribution of different mechanisms, I build and estimate a multi-sector spatial equilibrium framework that introduces heterogeneity in local productivity and workers' valuation for local amenities across sectors, and allows initial sorting to be magnified by production and residential externalities. Estimation of the model suggests that historical pollution is associated with lower current productivity and amenity levels. The effects are more pronounced for productivity, more skilled sectors and central tracts. I use the framework to evaluate the impact of counterfactual pollution cuts in different parts of cities on nationwide welfare and the cross-city distribution of skills.

1.1 Introduction

The comparative advantages of cities has slowly evolved over: during the industrial era, cities served as production centers where firms benefited from proximity to natural resources, shared infrastructure and pooled labour markets. More recently, cities have increasingly reshaped themselves into innovation hubs and consumption centres. During this transition, while some cities and towns have managed to adapt by shifting the focus of their production towards services, others have struggled, witnessing declining populations and deteriorating economic conditions. Urban decline associated with structural transformation is commonplace in many regions: from Northeast England in the first half of the twentieth century to the US Rust Belt in the past few decades, and North China's heavily industrialised cities today. The experiences of early industrialised cities in the developed world may provide valuable lessons for policymakers in recently industrialized countries that are beginning to feel the pressure of urban industrial pollution.

A central puzzle in examining the decline of industrial cities is why they fail to attract new service industries despite their convenient location, developed infrastructure, existing agglomeration benefits and declining land values. There have been many proposed explanations for this failed transition, which include the structural mismatch of jobs and workers, a lack of entrepreneurship due to the dominance of large corporations and misplaced public policies that attempted to subsidize the failing manufacturing sector. In this paper, I focus on a new explanation-how urban industrial pollution may cast a long shadow over the pattern of economic development within cities. In particular, I look at whether past pollution undermines the attractiveness of these cities today. In the post-industrial era, when local amenities play an increasingly important role in the location choice of skilled workers and agglomeration forces are increasingly shaped by the interaction of skilled workers and residents (Glaeser and Saiz (2004a); Moretti (2012)), past pollution may limit the ability of parts of the cities to attract skilled workers.

To examine the causal link between historical industrial pollution and the current distribution of economic activity, I assemble a rich dataset of historical land use, pollution and census tract level outcomes in US metropolitan areas. To mitigate concerns that areas that were more exposed to historical pollution differ in economic outcomes because of the endogenous placement of polluting industries and closer economic links to the manufacturing industry, I consider the role of wind in distributing pollutants by comparing areas that are downwind of historically polluting plants to those that are equally close but in upwind positions. The identification assumption is that equidistant tracts established similar economic links to the nearby industrial areas and experienced similar labour market shocks when these areas deindustrialised. My empirical analysis yields three sets of results. First, areas downwind of historical industrial sites were more heavily polluted during the 1970s but this was no longer the case in the 2000s, partly because pollution has largely been cut in American cities due to the decline of manufacturing. Second, a wide range of economic outcomes continued to be influenced by historical pollution.Census tracts which were close to and downwind of heavily industrial areas in the 1970s were less skilled and had lower housing value in 2000. Third, these tracts also experienced declining housing prices, wages and shares of skilled employment from 1980 to 2000. Taken together, these findings highlight the role of historical pollution, rather than current pollution, in shaping the current distribution of skills, and are indicative of the strong production and residential agglomeration effects that set cities on diverging paths of development.

What makes the areas that were more exposed to historical pollution underperform today? Were these areas simply becoming undesirable places to live because historical pollution changed the composition of surrounding communities, which in turn led to deteriorating amenities; or were they also evolving to be less productive as high-quality firms in emerging skilled service industries avoided them? I examine the mechanisms at work in two steps. First, I look for qualitative evidence on the relationship between historical pollution and endogenous amenities. If historical pollution is associated with higher crime rates today, it most likely operates through channels of residential sorting and externalities. Second, I estimate a structural spatial equilibrium model that allows me to recover sector-specific productivity and amenity parameters from observable data. With these estimates, we can tell how far IT firms value a particular location as a productive place relative to construction firms, and how far finance versus manufacturing workers prefer to live in certain location in order to enjoy a better quality of life. I then decompose the aggregate productivity and amenity estimates into exogenous parts driven by fundamentals and endogenous ones determined by agglomeration forces and check (1) the effects of historical pollution on current sector-specific productivity and perceived amenity; and (2) the contribution of production and residential agglomeration effects in accounting for these effects.

The spatial general equilibrium model, which builds on Ahlfeldt et al. (2015) and extends their analysis by embedding the internal structure of individual cities into a system of cities, introducing sectoral heterogeneity in productivity and people's valuation of local amenities and allowing the initial sorting of skills to be magnified by production and residential externalities. These features allow me to separately estimate sector-specific productivity and amenity parameters using information on sectoral wages and employment at the census tract level observed at both place of work and place of residence, tract-to-tract bilateral commuting costs and housing values in 2000 and 2010. By checking the correlation between early pollution exposure and estimated productivity and amenity parameters, I show that, controlling for current pollution, tracts that were more exposed to historical pollution due to their position downwind of industrial areas had both lower productivity and amenity in 2000. I also find that these negative effects are stronger for productivity, skill-intensive sectors and central tracts.

To account for the contribution of production and residential externalities to the estimated negative impacts of historical pollution on local productivity and amenities, I follow the literature by defining these externalities as increasing functions of the workplace and residence employment density of surrounding tracts, respectively. To capture more flexible cross-industry interactions of workers, I allow the productivity of a particular sector to depend on the employment density of different sectors according to the industrial linkages across sectors specified by Glaeser et al. (2015). On the residential side, I assume that residents' valuation of local quality of life depends on both residential density and the skill-mix of the neighbourhoods. I then decompose both estimated local productivity and amenity into exogenous components driven by fundamentals and endogenous ones determined by agglomeration effects. The endogenous agglomeration terms highly correlate with historical pollution explain a large proportion of the estimated effects of historical pollution on both current productivity and amenity

To evaluate the differences in the costs and benefits of pollution abatement across time and space, I consider policy experiments that cut pollution levels in central and non-central tracts and evaluate the overall welfare and distributional impacts. These exercises require estimates of the elasticities of local production and residential fundamentals to pollution changes. To identify the contemporaneous productivity and amenity benefits from cutting pollution, I exploit quasi-experimental reductions in Particulate Matter (PM10) concentrations from 2000 to 2010 induced by differential local regulator responses to the Clean Air Act Amendments (CAAA), studied in Auffhammer et al. (2009) and Bento et al. (2015). In an instrumental variable specification, I show that changes in productivity and perceived amenity by workers from different sectors respond to the evolution in the PM10 concentration level in a way that mirrors the responses from the cross-sectional estimation. The counterfactual policy analyses reveal that cutting pollution in central cities leads to both larger cross-city flows of skilled workers and higher welfare gains, especially in the case with endogenous local productivity and amenity levels.

1.2 Related Literature

The paper studies the long-run impacts of historical industrial pollution on the current distribution of skills. It contributes to a number of different strands of literature, such as urban pollution, the geographical sorting of households and workers and the persistence of economic activities within cities. While the existing literature in environmental economics has mainly focused on the contemporary effects of pollution on health, residents' quality of life and local labour productivity and supply (Currie et al. (2015), Sullivan (2016), Hanna and Oliva (2015), Zivin and Neidell (2012)), this paper reveals an oft-ignored hidden cost of industrial pollution: its effects on the attractiveness of industrial cities in the wake of industrial decline. This puts these cities at a disadvantage during the structural transformation process.

In examining the mechanisms at work, my findings highlight the importance of skill sorting at both workplace and residence, which complements a long history of literature examining the geographic sorting of different types of agents (Tiebout (1956); Epple and Sieg (1999); Kuminoff et al. (2013)). A small literature deal specifically with the sorting of households on environmental amenities, such as Wu (2006),Banzhaf and Walsh (2008) and Bayer et al. (2009).

These findings also build on previous research on the persistence of economic activities within cities. For instance, Hornbeck and Keniston (2014) and Siodla (2015) examine the long-lasting impacts of great fires in cities. Ambrus et al. (2015) find that historical cholera outbreaks are associated with poorer neighbourhoods decades later. Redding and Sturm (2016) and Dericks and Koster (2016) examine how the negative impacts of Second World War bombings in London persist today. Heblich et al. (2016) document, in particular, residential sorting responses to pollution in the 19th century that persist today in British cities. My contribution to these two strands of literature is to embed the sorting mechanism into a general equilibrium setting, which helps me to identify the impacts on local productivity and amenity separately, and quantify the contribution of local agglomeration effects. I find a greater impact of historical pollution on local productivity, which is relatively new to this literature, since previous empirical studies have focused on residential sorting and neighbourhoods effects.

On the theory side, this paper extends the single city spatial equilibrium framework of Ahlfeldt et al. (2015) by integrating their model of the internal structure of cities into a system of cities. This allows me to examine how the distribution of pollution within a city shapes its overall competitiveness and the distribution of population and skilled employment across cities. My theoretical framework also introduces sectoral heterogeneity in productivity and amenity to generate the sorting of jobs and residents in response to pollution, while still remaining tractable. My framework contributes to a burgeoning literature on modelling and structurally estimating the internal structure of cities, such as Ahlfeldt et al. (2015), Davis and Dingel (2014), and Allen et al. (2015).

Methodologically, this paper also contributes to the literature on measuring the quality of life in cities Blomquist et al. (1988); Albouy (2008)). To my knowledge, this paper is one of the first attempts to measure the quality of life at a highly disaggregated level. It is also one of the first to directly measure locational quality-of-life as perceived by workers with different characteristics. I find that the variation in measured quality of life within cities is much greater than between cities. Residents who work for different sectors differ in their valuation of the same observable amenities; in particular, workers from skilled sectors display greater aversion to urban disamenities such as pollution and crime.

There is a small literature documenting the recent gentrification of inner cities (Fee et al. (2012), Couture and Handbury (2015), Baum-Snow and Hartley (2016)). My paper contributes to this literature by proposing central city industrial pollution as a particular contributor to the excessive decentralization of post-war industrial cities. I also examine to what extent the clean-up of inner city pollution could revive city centres. My preliminary results suggest that central city amenities perceived by high-skilled workers improved disproportionately from 2000 to 2010, which is consistent with the results presented in Couture and Handbury (2015).

Another strand of the literature studies endogenous productivity and amenity changes in response to the density and composition of an area's residents. Moretti (2004) and Ciccone and Peri (2006) look at the response of a city's productivity of a city to its skill-mix, while Bayer et al. (2009), Guerrieri et al. (2013) and Diamond (2016) study residential sorting on the basis of neighborhood characteristics and the endogenous supply of amenities to the city level skill-mix. I incorporate both channels into my analytical framework and find both to be important in accounting for the negative impact of historical pollution on current economic outcomes.

My findings also relate to the literature on the structural transformation of cities from manufacturing agglomerates to innovation clusters and consumption centres, with the agglomeration of high-skilled labour playing an increasingly important role (Glaeser and Saiz (2004a), Moretti (2012), Diamond (2016)). In my paper, I take this structural shift as given and consider how it magnifies the importance of historical industrial pollution in determining a city's growing potential.

Finally, this paper also contributes to the literature on the causes and consequences of the decline of the Rust Belt. Most importantly, technological change and economic globalization had profound impacts on regions oriented towards goods-production, especially in the Rust Belt (Feyrer et al. (2007)). Glaeser and Ponzetto (2007) argue that the Rust Belt's location-specific advantage, which stemmed from easier access to waterways and railroads, declined over time. Alder et al. (2014) cite the lack of competition in both output and factor markets as a key element in Rust Belt decline from a macroeconomics perspective. Glaeser et al. (2015) further suggest that proximity to old mines leads to specialization in heavy industries, the dominance of big firms and subsequently dampened entrepreneurial human capital across several generations.

1.3 Data

To examine the relationship between historical industrial pollution and the contemporary distribution of skilled labour within and across cities, I assemble a database of tract level outcomes from 1940 to 2010. I draw my data from three main sources: census outcomes from 1940 to 2000, matched across years according to the Longitudinal Tract Data Base (LTDB); U.S. Geological Survey (USGS) land use data in the 1970s, and EPA pollutants ambient concentration data collected at each monitoring site from 1957.

1.3.1 Workplace and Residence Location Choices

The main outcome variables are obtained from the Census Transportation Planning Package (CTPP) of the Bureau of Transportation (BTS). The CTPP includes three parts: tabulations by place of residence, tabulations by workplace, and flows from residence to workplace, all at the census tract level. It provides detailed information on the counts of employed population by gender (2), industry (15), occupation (25) and race (4) who live or work in certain census tracts, the median earnings by industry of people who live or work in these tracts, the number of people who commute between any census tract pairs, and the average travel time by four modes of transportation[1] between any census tract pairs. The CTPP data are available at the census tract level for only a few cities in 1990 and all metropolitan areas in 2000 and 2010. To calculate bilateral travel cost across tracts, I obtain the estimated

[1] Automobiles, public transportation, walking or cycling, other

driving time between the centroids of census tracts from OpenStreetMap.

I complement the CTPP data with decennial tract level information on employment, skill composition housing value and quality at the place of residence from 1940 to 2010. To normalize the NHGIS and CTPP data to 2010 census tract boundaries, I use the Longitudinal Tract Database (LTDB) Logan et al. (2014). I end up with an unbalanced panel of census tracts from 1970 to 2010, with 52,210 observations in 1970 and 66,438 observations in 2010. I use constant 2010 CBSA boundaries. The CBD is defined according to the CTPP 1990 Urban Geographic Data. [2]

sector-specific employment and earnings data by both place of work and place of residence are essential for me to separately identify the amenity and productivity parameters in my structural model. More specifically, earning by industry, partialling out educational attainment, gender, race, and occupation, serves as a measure of labour productivity of each industry at the tract level. Tract level quality-of-life perceived by people from different industries can be backed out using information on the counts of employed population by industry, local housing value and their access to job opportunities within this industry, measured as industrial productivity of nearby tracts inversely weighted by travel cost.

1.3.2 Historical Land Use

I obtain information on the location of historical industrial areas from the Enhanced Historical Land-Use and Land-Cover Datasets of the U.S. Geological Survey (USGS). This dataset depicts land use and land cover from the 1970s and has previously been published by the U.S. Geological Survey (USGS) in other file formats.[3] The basic sources of land use compilation data are NASA high-altitude aerial photographs, and National High-Altitude Photography (NHAP) program photographs. Urban or built-up land use is classified as residential, commercial, industrial, transportation, communications and utilities, industrial and commercial complexes, or mixed urban land use. In my paper, I define industrial areas as the land exclusively allocated as industrial land use.

[2] http://www.transtats.bts.gov/TableInfo.asp?Table_ID=1279&DB_Short_Name=CTPP%201990&Info_Only=1

[3] The original digital data sets were created by the USGS in the late 1970s and early 1980s and were later converted by USGS and the U.S. Environmental Protection Agency (USEPA) to a geographic information system (GIS) format in the early 1990s. (Price et al. (2007))

1.3.3 Pollution and Wind Data

The contemporary and historical air pollution data are drawn from the Environmental Protection Agency (EPA) airdata. The EPA monitors different air pollutants over time and their earliest data dates back to 1957. The data is available at monitor level, and I match the monitors to census tracts according to their coordinates. In this paper, I focus my attention on Total Suspended Particles (TSP) or Particulate Matter smaller or equal to 10micron in diameter (PM10) because there are more TSP/PM10 monitors than those of other pollutants and they are closely related to industrial pollution. Before 1990, the EPA mostly monitored TSP, during the 1990s the EPA mostly monitored PM10[4].

In my empirical analysis, I leverage the quasi-experimental variation in wind direction differences in distributing pollution from industrial sources. I draw the information on local wind conditions from NOAA's Quality Controlled Local Climatological Data (QCLCD). It consists of monthly summaries of wind direction and wind speed for approximately 1,600 U.S. locations, from January 1, 2005 and continues to the present.

1.3.4 Amenity Data

I collect a diverse set of data on tract level local amenities, such as public schools and crime rates, for two purposes. First, by directly examining the relationship between historical pollution and endogenous amenities, I can test for the relevance of residential agglomeration effects in driving the main results. Second, the observable amenities can be compared to the estimated amenity parameters as an additional validation of the structural estimation.

Data on public schools are drawn from the National Center for Educational Statistics' Common Core of Data (CCD)[5]. The CCD is NCES' primary census database that includes annual information for the universe of all public elementary and secondary schools, school districts, and other educational administrative and operating units across the U.S. The CCD contains three types of data: descriptive information on school location and type; demographic data on students and staff, and fiscal data on revenues and expenditure. I match schools to census tracts using their coordinates, which allows me to infer the number of public schools, average

[4] In 1987, EPA replaced the earlier Total Suspended Particulate (TSP) air quality standard with a PM10 standard. TSP standard counts for particles that are smaller or equal to 50micron in diameter. While the new PM10 standard focuses on smaller particles that are likely to be responsible for adverse health effects because of their ability to reach the lower regions of the respiratory tract.

[5] https://nces.ed.gov/ccd/pubschuniv.asp

student teacher ratio, total revenues and expenditures at the census tract level.

I obtain crime data from the National Neighborhood Crime Study (NNCS), 2000[6]. It reports tract-level crime data pertaining to seven of the FBI's crime index offenses for 10,851 census tracts, as well as tract-level information on social disorganization, structural disadvantage, socioeconomic inequality, mortgage lending, and other control variables garnered from the 2000 United States Census of Population and Housing Summary File 3 (SF3) and other publicly available sources.

I also collect information on a variety of natural amenities, following Lee and Lin (2015), including distances to a water body, average slope, flood hazard risk, average 1971–2000 annual precipitation, July maximum temperature and January minimum temperature.

1.4 Reduced-form Evidence

In this section, I estimate the effects of early industrial pollution exposure on subsequent growth and specialization patterns within cities. To exploit variation in pollution independent from local economic conditions, I consider the role of wind in disseminating pollutants. I first investigate whether areas that are closer to and downwind of 1970 industrial areas are more polluted in the 1970s and today, before proceeding to explore the "reduced form" relationship between 1970s industrial activities and contemporary economic outcomes, exploiting variation in wind direction.

1.4.1 Proximity to Early Industrial Areas and Local Pollution

Empirical Settings

To examine the relationship between pollution and industrial activities in the 1970s, I adopt the following specification using monitor-level pollution data.

$$p_{i,1971-1979} = \sum_{k=1}^{4} \mathrm{I}_{\mathrm{ikm}}\beta_k + X_i'\gamma + \alpha_c + \epsilon_{ic} \tag{1.1}$$

where $p_{i,1971-1979}$ denotes the average total suspended particle (TSP) readings from 1971 to 1979 recorded by TSP monitor i, $\mathrm{I}_{\mathrm{ikm}}$ is an indicator variable for whether or not monitor i lies within a distance buffer k from the closest industrial area m in the 1970s; In practice I examine the results over four distance buffers: within 1

[6] http://www.icpsr.umich.edu/icpsrweb/RCMD/studies/27501

km, 1-2 km, 2-3 km and 3-4 km. α_c are Core Based Statistical Area (CBSA)[7] fixed effects; X_i is a vector of controls that includes the distances to natural amenities, CBD and transportation lines.

To exploit the variation in ambient air pollution driven by factors not directly related to industrial activities, I also estimate a model similar to Equation (1), where I interact the same set of distance buffers to the closest industrial areas with wind directions. Holding constant the distances to industrial areas, monitors that are downwind should be more exposed to industrial pollution, thus, they should capture relatively higher TSP readings. For monitor i in CBSA c, I estimate following specification:

$$p_{ic} = \sum_{k=1}^{5} \mathrm{I_{ikm}} * Downwind_{im} \beta_{k_1} + \sum_{k=1}^{4} \mathrm{I_{ikm}} * \beta_{k_2} + X_i' \delta + \alpha_c + \epsilon_{ic} \tag{1.2}$$

where $Downwind_{im}$ is an indicator of whether or not monitor i is downwind of its closest industrial area m. Apart from the interacted terms between the four distance buffers and the downwind dummy, I also include that between the excluded group (4 km or more away from the closest industrial area) and the downwind indicator.

The definition of downwind status is central to my empirical specification. The data I use for this purpose keeps track of monthly wind speeds and directions at over 1600 weather stations in the continental US from 2005 to 2014. As the purpose of this paper is to examine the long-run impacts of early industrial pollution on location choices of residents and firms, wind directions should be a concern only if they are stable enough throughout the year. Therefore, instead of pooling monthly observations to get an annual average of wind direction, I define the seasonal wind coverage ranges in spring-summer (April-September) and autumn-winter (October-March) and consider a monitor to be downwind of the closest industrial area only if it is exposed to this area through wind in both winter and summer. The lower part of Figure 1.1 illustrates the way I define these ranges. The winter/summer ranges are defined by the 10th and 90th percentile of all monthly observations on wind directions from 2005 to 2014. I drop observations with monthly average wind speed lower than 0.5 m/s and force the downwind dummy to be zero if any of the calculated wind ranges exceed 180 degrees (the 90th and 10th percentiles differ by over 180 degrees). In the end, this definition suggests that about 9% of monitors are downwind of its closest industrial area.

As a further attempt to link observed pollution to industrial activities, I exploit

[7] A CBSA is a US geographic area defined by the Office of Management and Budget (OMB) that consists of one or more counties around by an urban center of at least 10,000 people and adjacent counties that are socioeconomically tied to the urban center by commuting.

another layer of variation in the pollution intensity of each industrial area. It is highly likely that the exposure to industrial areas through wind may only matter when these areas are sufficiently polluted. Thus, we can go one step further to check if the additional effects of being downwind of industrial estates on pollution also appear to be larger around more heavily polluted plants, by dividing the full sample of monitors into two according to whether or not the pollution intensity of the industrial areas these monitors are closest to is above the median, and run specification (1) and (2) on both subsamples separately.

To obtain a pollution intensity measure of each industrial area, I match them to their nearest TSP monitors. The upper graph of Figure 1.2 maps the location of industrial areas and TSP monitors with at least one year of readings from 1971 to 1979 in central Detroit. It is apparent that most of the industrial areas in my sample have a TSP monitor close-by, largely because these monitors are intended to oversee the most polluted parts of the city. I define the pollution intensity of industrial area m to be the average TSP readings from 1971 to 1979 collected at its closest monitor. to mitigate concerns over measurement errors, I only keep industrial areas with a TSP monitor less than 2 km away. The middle figure of Figure 1.2 shows the assignment, with dark areas being industrial zones kept in the sample. The division of the full sample of monitors according to whether or not the industrial areas they are closest to are above of below the median pollution level is illustrates at the bottom of Figure 1.2. It is apparent that industrial areas of different pollution intensity are distributed quite evenly across central Detroit.

I run specifications (1) and (2) on the two subsamples and check if the magnitudes are larger for monitors that are closer to more heavily polluted industrial areas. Another possible approach to exploit both the variation in pollution intensity and wind direction is a triple difference design, where the variables of interest are the triple interaction terms of the distance-to-industrial-areas indicators with the level of historical pollution of these areas and a downwind dummy. I adopt the triple-difference estimation method as a robustness check and report results in Appendix B.

The main purpose of this paper is to explore the long run negative impacts of early industrial pollution. However, it is likely that more polluted areas in the 1970s also tend to be more polluted now, making it hard to tell the long-run effects of historical pollution from those of current pollution. Figure A1 plots the pollution level in 2000 (measured by PM10) against that in the 1970s (measured by TSP)[8].

[8]PM10 levels are adjusted to be comparable to TSP levels, according to the average PM10/TSP ratio among all the readings collected at monitors that record PM10 and TSP at the same time.

It is clear that although tracts that are heavily polluted in the 1970s do indeed appear to be more polluted three decades later, the average pollution level has fallen significantly from 1970 to 2000. There also exists a significant amount of churning in the pollution distribution during this period. In particular, some of the most polluted tracts (TSP$> 200\mu g/m^3$) in the 1970s appear near the bottom of pollution distribution in the 2000s. To check if the areas downwind of historical industrial sites still remain dirty today, I re-estimate Equations (1) and (2), replacing the dependent variable with monitor-level pollution measures of PM10 from 2000 to 2010.

Results

Table 1.1 reports estimation results of specifications (1) and (2), where the dependent variable is the average TSP level at monitors from 1971 to 1979. Each column represents a different regression, where columns (1) and (2) report estimates on a full sample of TSP monitors with at least one reading from 1971 to 1979 and not located within any industrial areas. Columns (3) and (4) report results on a subset of TSP monitors that are closest to industrial areas with above-median pollution intensity, while columns (5) and (6) report regression estimates on the subset of monitors adjacent to below-median polluted areas. Standard errors are clustered at CBSA level. The unit of TSP is $\mu g/m^3$ and the average TSP from 1971 to 1979 is 69 $\mu g/m^3$.

Columns (1) and (2) report the results on the full sample. Monitors that are close to industrial areas record higher measures of TSP and the effects drop as they move away from the sources of pollution. Monitors within one kilometre of the closest industrial area capture 23.7 more units of TSP, while those within 2 to 3 kilometres capture 6.5 more. Wind appears to be important in disseminating pollutants from the sources, as is apparent in Column (2). We observe that, conditional on the distance to the nearest industrial area, the downwind monitors record higher levels of TSP. However, the effects do not appear to fade away linearly with distance. In terms of the point estimate, downwind matters most for tracts that are within 1 kilometre and from 2 to 3 kilometres of the closest industrial area, leading to 8.9 and 11.9 extra units of TSP measured respectively. Wind directions do not appear to matter much for areas that are more than 4 kilometres away from industrial areas. In general, the observed sphere of influence is consistent with the nature of airborne particles' movements. Wilson and Suh (1997) suggests that the travel distance of coarse particles, such as TSP and PM10, is up to 10 kilometres. Hence, it is natural that the effect of wind largely drops when we move four kilometres away from the pollution source.

The estimation results on the subsample of monitors nearest to above-median polluted industrial areas are shown in columns (3) and (4). It is clear that the magnitudes of estimated coefficients for the interaction terms are higher, which is consistent with the intuition that wind matters more around more heavily polluted sources. Conversely, as shown in Columns (5) and (6), the estimated coefficients of both distance buffers and their interaction terms with the downwind dummy are smaller in the subsample of monitors that are nearest to below-median polluted industrial areas. Most of the coefficients on the interaction terms are not significantly greater than zero. It suggests that wind matters in pollution diffusion only around sufficiently polluted sources. Therefore, in our next step in exploring the relationship between historical pollution and tract-level outcomes, we focus our attention on the tracts that are close to heavily-polluted industrial areas only.

Table 1.2 reports the estimation results with monitor-level average PM10 levels from 2000 to 2010 as the dependent variable. Similarly, Columns (1)-(2) show estimates for the full sample, Columns (3)-(4) for a subsample of the PM10 monitors closest to the above-median polluted industrial areas and Columns (5)-(6) on a subsample of the monitors closest to the below-median polluted areas. The average PM10 from 2000 to 2010 is 21 $\mu g/m^3$.

Column (1) shows that the monitors close to historical industrial areas also appear to be more polluted in the 2000s, but the magnitudes are much lower. The PM10 monitors within 1 kilometre of the closest industrial area record an extra 3.2 units of PM10, which is about 15% above the median level, compared to about 33% more TSP recorded in the 1970s. It suggests that areas near historical industrial sites remain more polluted in 2000, but the pollution level has fallen significantly since 1970. It is clear from Column (2) that all the coefficients on the interaction terms are either not statistically significantly different from zero or negative, which suggests that areas downwind to 1970s industrial sites are not more polluted today, if not less polluted. I believe it is mostly because the transportation of pollutants by wind appears to be important only around sufficiently polluted areas.

The estimation results for the subsample with the monitors nearest to the above-median and below-median polluted industrial areas are shown in Columns (3)-(4) and Columns (5)-(6), respectively. Quite surprisingly, being closer to historically more polluted industrial areas induces less current pollution than being closer to less polluted areas. This suggests that, although industrial zones in the 1970s are more likely to be industrial zones today, the pollution intensity of each industrial area changed greatly over recent decades. One explanation is that the most polluted industrial areas in the 1970s were under stricter regulatory oversight after the Clean

Air Act, which led to significant industrial relocation. Meanwhile, the coefficients on the interaction terms between the distance buffers and the downwind dummy remain small and insignificant in both subsamples. Putting these pieces of evidence together, the long-term impact of historical pollution on current economic outcomes, identified through variations in both the pollution intensity of industrial areas and wind conditions, are not likely to confound the effects of current pollution.

1.4.2 Early Industrial Pollution Exposure and Current Economic Outcomes

Empirical Setting

In the previous section, I presented evidence on the impact of proximity to industrial areas on local pollution and the additional role of wind in disseminating pollutants. In this section, I further to examine the relationship between historical industrial pollution and current economic outcomes. Do tracts that were dirtier in the 1970s underperform in the post-industrial era? How much of their failure to attract skilled workers and residents can be attributed to historical air pollution? To test this, I look at the “reduced form" relationship between economic outcomes in 2000 and exposure to 1970s industrial pollution in the following specification:

$$y_{ic} = \sum_{k=1}^{5} \mathrm{I_{ikm}} * Downwind_{im}\beta_{k_1} + \sum_{k=1}^{4} \mathrm{I_{ikm}} * \beta_{k_2} + X_i'\delta + \alpha_c + \epsilon_{ic} \tag{1.3}$$

where y_{ic} denotes the economic outcomes of interest observed in 2000, which include housing prices, shares of high-skilled workers and residents, median earning of workers and residents and the share college graduates; α_c are CBSA fixed effects; $\mathrm{I_{ikm}}$ is an indicator that switches to one if tract i lies within a distance buffer k from the closest industrial area m in 1970s; $Downwind_{im}$ is a dummy variable that takes value one if tract i is located downwind of industrial area m; X_i are tract-level characteristics, which include the distances to natural amenities and transportation lines, indicators for different distance buffers to the CBD and route distance buffers to the same industrial area m and the predicted manufacture job growth from 1970 to 2000 based on the industrial composition in 1970.

The coefficients of interest here are β_{k_1}, which account for the additional effects of being downwind of an industrial area for census tracts that are within a distance buffer k of the industrial area. Controlling for CBSA fixed effects limits our attention to within-city variation. The identification of our main results relies on the assumption that, in the absence of pollution, tracts that are downwind and upwind of the same industrial areas are similar in economic prosperity and skill composition. However, estimates of β_{k_1} may be inconsistent if being downwind of industrial tracts correlates with other geographical features of tracts that are relevant for economic development. For instance, for coastal cities, the wind could mostly come from water, which makes coastal tracts more likely to be located downwind of industrial tracts. To deal with this, I control not only for the distance to a body of water in all my regressions, but also re-examine my main results in a sample that excludes coastal cities.

Another concern is that that the location choice of early industrial sites may take account of its pollution impact on nearby neighbourhoods, driven by wind patterns. More specifically, a potentially heavily polluted plant might avoid locations upwind of wealthy neighbourhoods if the latter could exert enough influence on industrial location. To confirm that the placement of early industrial sites does not weigh in differently the socioeconomic characteristics of downwind and upwind neighbourhoods, I run a set of falsification tests and replace the outcomes from specification (3) to those observed in 1940 or 1950. A problem with running the falsification tests using 1970s industrial areas directly is that some of these industrial sites might have been set up before 1970, or maybe even before 1940 and as a result we may capture partial early treatment effects in this specification. To circumvent this issue, I try to get an idea of the location of industrial areas emerging from 1950 to 1970 using information on pollution changes during this period. An industrial area is considered to be newly added if the TSP reading from monitors within 5 kilometres increased by 30% from 1950 to 1970[9]. Similarly, the full sample is split into two according to the pollution intensity of the industrial areas.

Finally, to make use of the variation in the pollution intensity of different industrial areas, I split the full sample of census tracts into two, according to whether or not the pollution intensity of the closest industrial area from each tract is above-median. Similarly, pollution intensity is defined as the average measure of TSP ambient concentration at the TSP monitor closest to each industrial area from 1971 to 1979. Only industrial areas that are within 2 kilometres to the closest industrial area are kept in the sample.

[9] I do not require the readings from 1950 and 1970 to come from the same monitor. In other words, an industrial area's pollution levels in the 1950s and 1970s can come from two different monitors as long as both are located within 5 km to the industrial area. I make this compromise because there are less TSP monitors with valid readings in the 1950s. On average, the TSP levels decreased by 20% during this period. So if we observe an opposite 30% increase it is a strong indication that a new industrial areas was added nearby.

Results

Table 1.4 presents regression results on 2000 economic outcomes. The upper panel reports estimates on a sample of census tracts whose closest industrial areas are above-median polluted. The outcomes reported are employment density, share of high-skilled[10] employment and median wage at the place of work, as well as density, share of high-skilled employment, median wage, housing value and share of college graduates at the place of residence. The key coefficients of interest are reported in the first four rows, which report the additional effects of being downwind of a 1970s industrial areas within 0-1, 1-2, 2-3, 3-4 and over 4 kilometres. It is clear that census tracts that are downwind of and close to heavily polluted industrial areas have lower housing prices, are occupied by less educated residents from less skilled sectors and earn less. Meanwhile, the workers who work in these tracts are also less likely to be employed in high-skilled sectors and are earning less, which is suggestive of a negative impact of historical pollution on current labour productivity.

Estimates from a sample of census tracts closest to the below-median industrial areas are reported in the lower panel. It is clear that most of the estimated coefficients on the interaction terms between distance buffers and downwind are much smaller and statistically insignificant, which is consistent with the results on historical pollution: a downwind position appears to be detrimental to current economic outcomes only when the nearby industrial areas are sufficiently polluted because the wind direction only significantly affects the pollution concentration around industrial areas that are sufficiently polluted, as reported in Table 1.1.

We can compare the magnitudes of the coefficients from the “reduced form" estimation to those from the first stage. In the upper row of Figure 1.3, I plot the coefficients on the interaction terms between several 500 metre distance buffers and the downwind dummy in regressions with the monitor-level TSP reading as the dependent variable. It appears that being downwind matters most when the monitor is within 3 km of the closest industrial area, and when this industrial area is heavily polluted. In Figure 1.4, I repeat the same exercises using 2000 outcomes as the dependent variables. The negative impact of downwind position on most economic outcomes also appears to matter most when the census tract is within 3 km of its closest industrial area.

As briefly discussed in Section 4.2.1, one concern over specification (3) is that the location choice of early industrial sites may take account of its pollution diffusion to nearby neighbourhoods driven by wind patterns. To address this concern, I run a

[10] "High-skilled" industries are defined as finance, insurance and real estates (FIRE), information and professional services.

set of falsification tests to check if the census tracts that are close to industrial areas added during 1950-1970 differ from other tracts in 1950 in terms of socioeconomic outcomes, and if the tracts that are both closer to and downwind of the same newly-added industrial areas appear to be poorer or less educated than the upwind ones. Table 1.5 presents the results. We can see that the tracts within 4 km of the industrial areas emerging from 1950 to 1970 do appear to be poorer, less educated and with a lower share of residents working as managers and in professional and technical occupations, especially when these industrial areas are more polluted (upper panel). This is a sign that the placement of industrial areas from 1950 to 1970 was not random but tended to be closer to poorer neighbourhoods. However, the interaction terms of the dummy of being within 4 km and downwind are close to zero in both samples. This suggests that, although the placement of early industrial areas avoided rich areas, it was not sophisticated enough to take wind directions into active consideration. In other words, downwind richer neighbourhoods do not seem to have deterred nearby industrial placement more than their upwind counterparts did.

I do not adopt an instrumental variable specification as my main specification, because only a small proportion of census tracts in my sample can be matched to TSP monitors with readings from 1971 to 1979. Nevertheless, in a subsample of tracts with TSP monitors nearby, I look into the relationship between economic outcomes and historical pollution instrumented by proximity to industrial areas and wind directions. The results are presented in Table A11. Even in such a small sample, we still observe a negative relationship between TSP readings in the 1970s and current distribution of skills, with historical pollution instrumented by distance buffers and a downwind dummy, although some of the coefficients are not significant due to the small sample size and relatively weak first stages.

1.4.3 Dynamic Effects from 1980-2000

We have established that census tracts that were more exposed to 1970s industrial pollution were poorer and less skilled in 2000. But it is unclear whether or not they were equally poor in the 1970s, or if they improved or declined with the subdued industrial activities during this period. Without any agglomeration forces in the trend of de-industrialization, these areas should have experienced an improvement in air quality and a loss of manufacturing jobs, both leading to a higher ratio of employment in the high-skilled service sector. However, if agglomeration forces in skill-intensive service sectors are strong enough, the failure to attract high-skilled workers in the wake of industrial decline put these areas at a disadvantage throughout the whole structural transformation process and they may have ended up with an even lower share of high-skilled service employment.

To explore the evolution of census tracts in the post-industrial era, I replace outcomes in specification (4) with the growth rates of key outcomes from 1980 to 2000, including housing prices, median income, share of college graduates and the employment of different sectors. As workplace outcomes are not available before 2000, I look only at the growth of outcomes counted at place of residence.

The estimates identified through variation in wind direction are presented in Table 1.6. It is apparent from the upper panel of Table 1.6 that the tracts downwind and close to the 1970 industrial areas are not only poorer and less skilled in 2000, but also experience slower growth in total employment, median income, housing price and the share of college graduates from 1980 to 2000. If we compare the results on manufacturing employment (Column (2)) to those on FIRE employment (Column (3)), it is clear that the tracts that were more exposed to historical pollution experience lower growth of the residents who work in FIRE sectors but not in manufacturing. It is clear that historical pollution is associated not only with lower housing prices and less skilled communities, but also with declining economic conditions and worsening skill compositions, which is the opposite to what we might have expected from cleaner air accompanying de-industrialization. A plausible story would be that tracts that were more severely affected by industrial pollution ended up with less-skilled neighbourhoods and labour pools by the end of 1970s, which became a huge disadvantage when they tried to attract newly-available service jobs, high-tech firms or college graduates as the country made the transition from a manufacturing economy to a service-oriented one.

1.4.4 Mechanisms

In the above sections I have shown that census tracts that were more exposed to industrial pollution in the 1970s were poorer and less skilled in 2000 and displayed lower growth rates in housing prices, income and employment from 1980 to 2000. These pieces of evidence strongly suggest the relevance of agglomeration forces in shaping the evolution of industrial cities during waves of de-industrialization. But we are still unclear about the nature of the agglomeration effects in operation here. In the rest of this paper, I approach this issue from two angles. First, in this section, I check if any observed endogenous amenities such as crime rates or the provision of public schools respond to early industrial pollution. Second, in the next two sections, I lay out a theoretical framework that helps me to recover local productivity and amenity estimates, so that I can check whether historically polluted areas fail to attract skilled workers and residents in the post-industrial era because they fail to offer high-quality jobs or a high quality of life. I then further decompose both estimated local productivity and amenity into exogenous components driven by fundamentals and endogenous ones determined by agglomeration effects under assumptions on production and residential externalities.

A straightforward way to test for check the existence of residential agglomeration forces is to examine the relationship between early industrial pollution exposure and observable endogenous amenities, such as local crime rates and public school shares. Higher crime rates or a lower provision of public schools in 2000 in historically more polluted tracts purely driven by being downwind of 1970 industrial areas could only be an outcome of the sorting of local residents and resultant changes in local community compositions or tax bases. Housing durability could also play a role here: if housing units constructed around more heavily polluted areas in the 1970s were of poorer quality, and housing stocks are persistent, they may still exist thirty years later and act as a particular kind of disamenity, especially to high-income residents.

Table 1.7 shows the results. The upper panel reports the estimation results on the sample of census tracts closest to above-median polluted industrial areas, and the lower panel on that to below-median polluted ones. For simplicity, I use only one distance dummy (within 4 kilometres) instead of four finer divisions. The interaction terms between this distance dummy and a downwind dummy are positive and significant for both violent crime rate and the number of public schools per capita in the heavily-polluted sample. This suggests that census tracts that are predicted to be more polluted in the 1970s due to wind direction end up being more dangerous and less accessible to public schools in 2000. The evidence on housing quality is less conclusive. I use the share of housing units without kitchen or plumbing devices as a proxy for low housing quality. Additional exposure to historical pollution through wind did not affect housing quality in 1980 or 2000. Therefore, housing durability does not appear to be an important mechanism here, which is different from the findings of Heblich et al. (2016).

1.5 Theoretical framework

The previous section presents the negative effects of historical pollution on the attractiveness of local areas to both skilled workers and residents. Naturally we would like to know if it is due to the fact that the areas that were more exposed to historical pollution simply offer lower quality of life, as a result of changing neighbourhood composition, or that they also offer worse job opportunities, due to diverging location choices of firms. Additionally, with qualitative evidence on endogenous amenities, it makes sense to quantitatively account for the contribution of agglomeration effects at both place of work and place of residence.

To achieve both ends, I develop a multisector model of internal city structure based on Ahlfeldt et al. (2015), which allows to me recover sector-specific local productivity and amenity parameters, and decompose them into exogenous and endogenous components. My model differs from Ahlfeldt et al. (2015) in two key dimensions. First, to take into consideration the sorting of skills around historical pollution, I extend their framework to allow for systematic sectoral heterogeneity in productivity and amenities at different locations. Second, to examine the cross-city implications of the historical distribution of pollution, I extend embed their single city framework into a system of cities, which allows me to examine the effects of a counterfactual pollution cut in a subset of cities on the cross-city distribution of skills.

In my model, ex-ante identical workers simultaneously sort across sectors, and choose locations (census tracts in data) to work and live based on the amenity and productivity at these locations. The same local fundamentals, such as clean air, the distance to natural resources or to the CBD, could be of different production and consumption value for workers from different sectors in a systematic way. Admitting sectoral heterogeneity in both production and amenity valuation enables me to account for the sorting of sectors around both local productivity and amenities.

We consider a set of discrete locations or tracts, indexed by $i = 1, ..., P$, exogenously distributed across C discrete cities. The whole economy is populated by H workers, who are perfectly mobile across all the locations, within or across different cities. The mass of H workers can also move costlessly across S sectors. Firms from different sectors produce a single costlessly-traded final good, which is chosen as the numeraire.

Locations differ in their final goods productivity, amenities, floor space supply and access to the transport network[11]. Commuting is allowed across different locations within a city but not across cities.

[11] Floor space is assumed to be fixed at each tract. It can be microfounded as the product of a fixed supply of land and capital. The density of development (floor space to land ratio) can vary across tracts but is assumed to be exogenous to this model.

1.5.1 Preferences

Worker o from sector s residing in tract i and commuting to tract j derives her utility from consumption of the single final good c_{ijso}, consumption of housing h_{ijso} and local amenities.

$$U_{ijso} = \frac{B_{is} z_{ijso}}{d_{ij}} \left(\frac{c_{ijso}}{\beta}\right)^{\beta} \left(\frac{h_{ijso}}{1-\beta}\right)^{1-\beta} \tag{1.4}$$

where B_{is} stands for common residential amenities that makes a particular location more or less attractive to live for workers from sector s; d_{ij} captures the disutility from commuting from tract i to j for work ($d_{ij} = e^{\kappa \tau_{ij}} > 1$), where τ_{ij} is the bilateral travel time between tract i and j, and κ regulates the response of bilateral commuting cost of travel time. Travel time is measured in minutes. In this model, commuting is only allowed within each city, so a worker cannot live and work in different cities.

Following Ahlfeldt et al. (2015), I assume heterogeneity in the utility that workers derive from living and working in different parts of the city as a employees of different sectors, and allow this idiosyncratic component of utility z_{ijso} to be drawn from an independent Frechet distribution. In my model, utility derived from living and working in different tracts is allowed to differ across workers' chosen sectors.[12] The heterogeneous utility that a worker o from sector s living in tract i and commuting to work in tract j, modelled as z_{ijso}, comes from the following Frechet distribution:

$$F(z_{ijso}) = e^{-T_{is} E_{js} z_{ijso}^{-\epsilon}} \tag{1.5}$$

where the scale parameter T_{is} determines the average utility derived from working for sector s and living in tract i; the scale parameter E_{is} determines the average utility derived from working for sector s in tract j; and the Frechet shape parameter ϵ governs the dispersion of idiosyncratic utility. After observing her realizations for idiosyncratic utility, each worker chooses a sector and a pair of locations to live and work in.

Solving the workers' utility maximization problem, taking local fundamentals, wages and prices, as well as other worker' sector and location choices as given, we are able to derive conditions on each worker' commuting probabilities. Using the

[12]Since the workers are ex-ante identical in preferences apart from the idiosyncratic component, the differences in amenities valuation across sectors are interpreted as characteristics specific to each sector, including the fixed skills or earning capabilities of workers from different sectors. One could think of it as a simplified version of a model where ex-ante heterogeneous workers sort across sectors first and choose their locations to live and work subsequently, where the systematic differences in utility realization across sectors are partially capturing the sorting of workers across sectors by their inherent characteristics.

feature that the maximum of a Frechet distribution is itself Frechet, the probability that a worker chooses to live in tract i, work in tract j and for sector s is:

$$\pi_{ijs} = \frac{T_{is}E_{js}(d_{ij}q_i^{1-\beta})^{-\epsilon}(B_{is}w_j)^{\epsilon}}{\sum_{o=1}^{O}\sum_{p=1}^{P}\sum_{s=1}^{S}T_{os}E_{ps}(d_{op}q_o^{1-\beta})^{-\epsilon}(B_{os}w_p)^{\epsilon}} \tag{1.6}$$

We can sum the probabilities across workplaces for a given residence i and sector s, which gives us the probability that a worker from sector s lives in tract i, π_{Rjs}; as well as across residences for a given workplace (j) and sector s, which gives us the probability that a worker from sector s works in tract j.

1.5.2 Production

The single final good in this model can be produced by different sectors. The productivity is allowed to differ across sectors in the same location and also allowed to be different across different locations for the same sector. The good is the costlessly-traded numeraire that takes common price $p = 1$ across all tracts. I follow assume the production technology to be Cobb-Douglas, and the final good production function of sector s in tract j to be:

$$y_{js} = A_{js}(H_{Mjs})^{\alpha}(L_{Mjs})^{1-\alpha} \tag{1.7}$$

where A_{js} is the location-sector specific productivity, H_{Mjs} and L_{Mjs} is the amount of labour and floor space hired by sector s at location j.

Firms choose a sector to specialize in and a location to produce. They take final goods productivity A_{js}, goods and factor prices, the utility distribution of workers, and the location decisions of other firms as given. Combining the first-order conditions of firms' profit maximization problem and a zero-profit condition, we have:

$$A_{js} = q_j^{1-\alpha}(w_{js})^{\alpha} \tag{1.8}$$

where q_j is the floor space price at location j, and w_{js} is the local wage of sector s. It is clear that conditional on local floor space prices, firms in tracts with higher productivity are able to pay their workers higher wages.

1.5.3 Housing market clearing

In characterizing housing market clearing, we assume each tract to be endowed with fixed floor space supply L_i, among which L_{Mis} is allocated to sector s for production purposes, and L_{Ris} is allocated to workers from sector s as residential floor space.

Solving for consumers' utility maximization problem yields:

$$(1-\beta)\frac{\mathbb{E}[w_{ps}|i]H_{Ris}}{q_i} = L_{Ris} \tag{1.9}$$

where $\mathbb{E}[w_{ps}|i]$ is defined as $\sum_{p=1}^{P}\frac{E_{ps}(w_{ps}/d_{ip})^{\epsilon}}{\sum_{r=1}^{P}E_{rs}(w_{rs}/d_{ir})^{\epsilon}}$, which stands for the expected wage a worker living in tract i could get.

Similarly, firms' maximization problem yields:

$$\Big(\frac{(1-\alpha)A_{js}}{q_j}\Big)^{\frac{1}{\alpha}}H_{Mjs} = L_{Mjs} \tag{1.10}$$

Housing market clearing has $\sum_{s=1}^{S}(L_{Ris}+L_{Mis}) = L_i$.

Floor space is also allowed to be allocated entirely to residential or commercial use of a particular sector. In a corner solution where all the floor space is allocated to residential use, $q_i = (1-\beta)\frac{\sum_{s=1}^{S}\mathbb{E}[w_{ps}|i]H_{Ris}}{L_i}$. When floor space is used for commercial purposes only, we have $q_i = \sum_{s=1}^{S}\left(\frac{(1-\alpha)A_{js}}{L_j}\right)^{\frac{1}{\alpha}}$.

1.5.4 Equilibrium

In a competitive general equilibrium, individuals maximize utility; final good producers maximize profits, and both labour and housing market clear. I follow Ahlfeldt et al. (2015) by starting with a benchmark of the model with exogenous location characteristics, before introducing agglomeration forces in section 6.1

Given the model's parameters $\alpha, \beta, \mu, \epsilon, \kappa$, exogenous location-sector specific characteristics $\mathbf{T}, \mathbf{E}, \mathbf{A}, \mathbf{B}, \mathbf{L}, \tau, H$, the general equilibrium of the model is referenced by vectors $\pi_{Rs}, \pi_{Ms}, L_{Ms}, L_{Rs}, \mathbf{q}, \mathbf{w}$

The equilibrium is characterized by the following equations:

$$\pi_{Ris} = \frac{\sum_{p=1}^{P}T_{is}E_{js}(d_{ij}q_i^{1-\beta})^{-\epsilon}(B_{is}w_{js})^{\epsilon}}{\sum_{o=1}^{O}\sum_{p=1}^{P}\sum_{s=1}^{S}T_{os}E_{ps}(d_{op}q_o^{1-\beta})^{-\epsilon}(B_{os}w_{ps})^{\epsilon}} \tag{1.11}$$

$$\pi_{Mis} = \frac{\sum_{o=1}^{O}T_{is}E_{js}(d_{ij}q_i^{1-\beta})^{-\epsilon}(B_{is}w_{js})^{\epsilon}}{\sum_{o=1}^{O}\sum_{p=1}^{P}\sum_{s=1}^{S}T_{os}E_{ps}(d_{op}q_o^{1-\beta})^{-\epsilon}(B_{os}w_{ps})^{\epsilon}} \tag{1.12}$$

$$A_{js} = q_j^{1-\alpha}(w_{js})^{\alpha} \tag{1.13}$$

$$(1-\beta)\frac{\sum_{p=1}^{P}\frac{E_{ps}(w_{ps}/d_{ip})^{\epsilon}}{\sum_{r=1}^{P}E_{rs}(w_{rs}/d_{ir})^{\epsilon}}H_{Ris}}{q_i} = L_{Ris} \tag{1.14}$$

$$\left(\frac{(1-\alpha)A_{js}}{q_j}\right)^{\frac{1}{\alpha}} H_{Mjs} = L_{Mjs} \tag{1.15}$$

$$\sum_{s=1}^{S}(L_{Ris} + L_{Mis}) = L_i; \tag{1.16}$$

1.6 Model Calibration and Estimation

In Section 4, I presented evidence of the long-run negative impacts of industrial pollution on the current distribution of skills. Additional results on growth trends and endogenous amenities suggest that the endogenous interactions between different types of agents play an important role in explaining the observed patterns.

The structural estimation of the theoretical framework presented in Section 6 further guides my empirical analysis in three ways. First, the model allows me to separately estimate the productivity and amenities perceived by workers from different sectors and test how they respond to historical pollution exposure. Second, I can quantify the contribution of production and residential externalities in magnifying the initial impact of pollution. Third, the model makes possible a quantitative evaluation of various counterfactual policies. For instance, I could look at the changes in the distribution of population and skilled labour across cities following counterfactual pollution cuts in particular parts of a subset of cities and check whether welfare gains can be achieved by moving industrial zones out of central cities, under different assumptions about production and residential agglomeration forces.

I calibrate the model's key parameters following Ahlfeldt et al. (2015), reported in the upper panel of Table 1.9. The data sources of the key variables used in recovering local productivity and amenity parameters are reported in the lower panel of Table 1.9. One major difference of my method from that of Ahlfeldt et al. (2015) is that with additional data on sectoral median earnings at workplace, I am able to back out the productivity parameters directly from observed earning data without relying on imputed commuting flows. In addition, my model introduces sectoral heterogeneity in both local productivity and amenities, which requires productivity and amenity parameters backed out by sectors. It is made possible by sector-specific wage and employment data observed both at workplaces and residence locations.

1.6.1 Productivity and Production Fundamentals

In the model, firm profit maximization and zero profit conditions imply

$$A_{js} = q_j^{1-\alpha}(w_{js})^{\alpha} \tag{1.17}$$

Therefore, we can directly back out local sector-specific productivity A_{js} from the observed sectoral earnings w_{js} and housing value q_j. In mapping the actual data to model variables, one empirical issue is that contrary to the model assumption of an ex-ante identical workforce, workers are heterogeneous in abilities in reality, and those of higher ability might sort into higher productive places (Combes et al. (2008), Behrens et al. (2014)). As a result, w_{js} may capture variations in both local productivity and local workers' average abilities. To deal with this concern, I follow Albouy (2008) and estimate a Mincerian-type wage equation:

$$ln(w_{ojs}) = X'_{ojs}\beta + \mu_{js} + \epsilon_{oj}$$

where o stands for individuals, j and s stand for census tracts and sectirs; $ln(w_{ojs})$ is the log earning of individual o who works for sector s in tract j, X_{ojs} are individual characteristics including gender, race, educational attainment and the occupations of workers who work for sector s in tract j; μ_{js} are tract level fixed effects that capture tract level sector-specific productivity. Using μ_{js} instead of w_{js} as a measure of local productivity minimizes the potential biases driven by the sorting of workers on observable characteristics. Although it remains a possibility that workers also sort on unobserved characteristics, such as the quality of education, I believe that this plays a less important role than sorting on observables. [13]

A practical issue in estimating μ_{js} using microdata is that the census microdata from the Integrated Public Use Microdata Series (IPUMS) sample only reports geographical locations of individuals at the Public Use Microdata Area (PUMA) levels. A PUMA is a geographical unit much larger than a census tract. To get around this difficulty, I first estimate the relationship between log earnings and workers' characteristics conditional within each sector, including gender (2), race (4), educational attainment (9), and occupation (25), from a Mincerian model using the 5% IPUMS microdata, and combine the estimated coefficients with tract-level averages of all the observable characteristics (share of male, college graduates, managers, etc. by

[13] De la Roca et al. (2014) discovers little sorting on unobservable ability within broad occupation or education groups across metropolitan areas, using NLSY97 data. In Appendix, I will compare results obtained using raw and residual median earnings data to check the sign and magnitude of potential biases driven by sorting on observables, which should be able to give us a basic idea on the extent of biases driven by sorting on unobservables.

each sector at each tract). For each sector s, I run the following specification

$$ln(w_{ojs}) = X'_{oijs}\beta + \alpha_{PUMA,s} + \epsilon_{ojs}$$

where $\alpha_{PUMA,s}$ are PUMA fixed effects. $\hat{\beta}_s$ captures how different individual characteristics account for the variations in earnings within sector and PUMAs. I then obtain $w_{js} = \mu_{js}$ as the residuals between the actual and predicted log sector-specific earnings at tract level:

$$\mu_{js} = ln(\bar{w_{js}}) - \bar{X_{js}}'\hat{\beta} \tag{1.18}$$

where $ln(\bar{w_{js}})$ is approximated as log median earning by industry at tract level, and $\bar{X_{js}}$ are individual characteristics averaged by industry at the tract level. I use μ_{js} in place of w_{js} in obtaining $A_{js} = q_j^{1-\alpha}(\mu_{js})^{\alpha}$.

I use the local gross rental rate collected from NHGIS as an estimate of q_j. In section 4.4, I do not find housing quality, measured by the share of housing stocks without kitchen or plumbing services, to differ systematically across tracts with different levels of historical pollution. Nevertheless, I still account for the possibility that housing quality differs across tracts of different amenity value in a systematic way by obtaining residuals of housing rental rate controlling for the number of rooms, kitchen and plumbing facilities, type and age of building and the number of residents per room, in a way similar to that described in obtaining μ_{js}.

The estimated sector-specific productivity can be further decomposed into an exogenous part that reflects local fundamentals and an endogenous part that reflects the spatial interactions of different agents.

I follow the literature in assuming that productivity externalities are dependent on the travel-time weighted sum of workplace employment density. I start with the simplest case by assuming production externalities to be the same for different sectors:

$$A_{js} = a_{js}\Upsilon_j^{\lambda}, \quad \Upsilon_j \equiv \sum_{p=1}^{P} e^{-\delta\tau_{jp}}\left(\frac{H_{Mp}}{L_p}\right) \tag{1.19}$$

where H_{Mp}/L_p is the workplace employment density; τ_{jp} is the travel time in minutes from tract j to p; δ governs the rate of spatial decay and λ controls the importance of agglomeration effects in determining local productivity.

The specification above implicitly assumes that the access to workers outside one's own industry is as important as that to workers within one's own industry. However, in reality, it is reasonable to argue that the interactions with workers from the same industry may be more important. Therefore, I consider a more general specification of production-side agglomeration effects. I assume the production externalities of a particular sector s to depend on the employment density of both its own density and the density of other industries, but its dependence on the employment density of other industries is smaller. The cross-sector production externalities are stronger if the sectors are more similar to each other in workers' characteristics and input-output relationship. In another word:

$$A_{js} = a_{js}\Upsilon_{js}^{\lambda}, \quad \Upsilon_{js} \equiv \sum_{p=1}^{P} e^{-\delta\tau_{jp}} \Big(\sum_{r=1}^{S} \frac{Sim_{rs} H_{Mpr}}{L_p} \Big) \tag{1.20}$$

where Sim_{rs} is the weighted[14] average of pairwise labour pool similarity and input-output similarity specified in Glaeser et al. (2015). This similarity index is standardized to range from 0 to 1 and $Sim_{ss} = 1$.

1.6.2 Amenities and Residential Fundamentals

With detailed information on sector-specific productivity at the tract level, the natural next step is to recover measures of amenity.

In my model, we start from the residential choice condition (Equation (11)), multiply both side of the equation with the population mass H[15], and assume adjusted sector-specific amenities $\tilde{B_{is}}$ and adjusted sector-specific wage $\tilde{w_{ps}}$ to be $\tilde{B_{is}} \equiv B_{is} * T_{is}$ and $\tilde{w_{ps}} \equiv w_{is} * E_{ps}$[16]

$$H_{Ris} = \sum_{p=1}^{P} \Big(d_{ip} q_i^{1-\beta} \Big)^{-\epsilon} \Big(\tilde{B_{is}} \tilde{w_{ps}} \Big)^{\epsilon} \tag{1.21}$$

where H_{Ris} stands for sector-specific employment by place of residence, $\tilde{w_{ps}}$ are sector-specific adjusted wages, which can be approximated by sector-specific median earnings collected at place of work. q_i is housing price of tract i, d_{ip} stands for commuting cost from residence tract i to workplace j. It is clear that the adjusted sector-specific amenity measures, $\tilde{B_{ps}}$, can be recovered as:

[14] Weight is determined by the contribution of these measures to coagglomeration patterns in Glaeser et al. (2015). I ignore technology similarity because the knowledge spillovers across most service industries that I am studying are not patented, and technology similarity plays a relatively small role in explaining coagglomeration patterns according to Glaeser et al. (2015).

[15] Normalization: $\sum_{o=1}^{O}\sum_{p=1}^{P}\sum_{s=1}^{S} T_{os}E_{ps}(d_{op}q_o^{1-\beta})^{-\epsilon}(B_{os}w_p)^{\epsilon} = \bar{H}$;

[16] As defined in the model, T_{is} and E_{ps} represent the Frechet scale parameters drawn at each residence and workplace for each industry, they enter the model isomorphically with amenity B_{is} and wage w_{ps} in determining the relative attractiveness of a location to residents and workers. Therefore, in my actual structural estimation, I treat w_{ps} and $\tilde{w_{ps}}$, B_{is} and $\tilde{B_{is}}$ as the same for simplicity.

$$\tilde{B_{is}} = H_{Ris}^{\frac{1}{\epsilon}} q_i^{1-\beta} / \Big(\sum_{p=1}^{P} \Big(\tilde{w_{ps}} / d_{ips} \Big)^{\epsilon} \Big)^{\frac{1}{\epsilon}} \tag{1.22}$$

Similarly, estimated amenities can also be decomposed into an exogenous part that captures local fundamentals and an endogenous part that captures the benefit from living closer to other people. I adopt similar specifications of residential externalities as for production externalities:

$$\tilde{B_{is}} = \tilde{b_{is}} \Omega_i^{\eta}, \quad \Omega_i \equiv \sum_{p=1}^{P} e^{-\rho \tau_{ip}} \Big(\frac{H_{Rp}}{L_p} \Big) \tag{1.23}$$

where H_{Rp}/L_p is the density of residents; residential externalities decay with travel time τ_{ip}; the importance of access to surrounding density in determining local amenities is governed by η. It is worth noting that the nature of residential externalities could be very different from agglomeration forces at workplace. Specifically, the distinction between own-industry and other industry access may not be as important in the residential case. Residents tend to care more about other aspects of their neighbours, such as income or education, other than sector. Therefore, apart from the basic specification, I consider an alternative case where the access to a higher-educated population is allowed to play an additional role in shaping local amenities apart from the access to total population:

$$\tilde{B_{is}} = \tilde{b_{is}} \Omega_i^{\eta} \Omega_{Hi}^{\nu_s}, \quad \Omega_i \equiv \sum_{p=1}^{P} e^{-\rho \tau_{jp}} \Big(\frac{H_{Rp}}{L_p} \Big), \quad \Omega_{Hi} \equiv \sum_{p=1}^{P} e^{-\rho \tau_{jp}} \Big(\frac{H_{RHp}}{L_p} \Big) \tag{1.24}$$

where H_{RHp}/L_p is the density of college graduates, here $Omega_i$ captures the access to all the neighbours from tract i, and Ω_{Hi} captures that only to the college-educated neighbours. When $\nu = 0$, it collapses to the standard one. This additional assumption echoes empirical findings on gentrification and neighborhood effects. For instance, Diamond (2015) shows that the endogenous supply of amenities such as safety and public school quality depends on the skill mix of cities. Such a specification requires estimation of the parameters ν_s, which capture the influence of access to highly-educated neighbours on the amenities perceived by workers from different sectors. For identification, I need an instrument that correlates not with local amenities but with the skill-mix of residents for each tract. A natural idea is to use productivity shocks that affect the local demand for workers with different skills. Here I base my identification on changes in amenities in response to changes in access to highly-educated neighbours from 2000 to 2010. I instrument changes in access to highly-educated neighbours with Bartik-style predicted shifts in the number of residents with college degrees depending on the industrial employment from 80 industries in each tract in 2000, the skill requirement of each industry and the national growth of each industry from 2000 to 2010. The predicted change in the number of college graduates living in tract i between 2000 and 2010 is:

$$\Delta \hat{H}_{RHi} = \sum \left(\frac{H_{Ris',2000}}{H_{Ri,2000}} \right) * CollegeShare_{s',2000} * \Delta H_{Rs'}$$

where $H_{Ris',2000}$ is the number of workers over age 25 from sector s' living in tract i in 2000; $H_{Ri,2000}$ is the total number of workers over age 25 who live in tract i in 2000; $CollegeShare_{s',2000}$ is the national share of college graduates for sector s'; and $\Delta H_{Rs'}$ is the national growth in employment in sector s'. Correspondingly, the change in access to college-educated neighbours $\Omega_{Hi} \equiv \sum_{p=1}^{P} e^{-\rho\tau_{jp}} (\frac{H_{RHp}}{L_p})$, can be instrumented by:

$$\hat{\Omega_{Hi}} = \sum_{p=1}^{P} e^{-\rho\tau_{jp}} \left(\frac{\Delta \hat{H}_{RHp}}{L_p} \right)$$

1.6.3 Productivity and Amenities Estimation Results

Table A5 shows the correlation between my estimated aggregate amenities $\tilde{B_{is}}$ by sector and some observable amenity measures at the tract level. My estimated amenities correlate as expected with these "real-world" measures of amenities, such as the number of public schools, crime rate, pollution[17] and access to beaches. Besides, the valuation of the same observable amenities varies across workers from different sectors. For example, living in a tract that encompasses an industrial area reduces the subjective utility of FIRE workers by 8.6% and that of manufacturing workers by only 3.4%. Access to a beach increases the subjective utility of FIRE workers by 7.6% and that of manufacturing workers by only 5.3%. Being closer to the CBD and railways appear to be disamenities.

As a more straightforward presentation of the estimated amenity measures, Figure 1.9, 1.10 and 1.11 map amenities perceived by FIRE and manufacturing workers to areas around New York, Detroit and San Francisco. It is clear that the perceived amenity varies considerably within cities. From Figure 1.9, we observe the stark contrasts in neighbourhood desirability between the Upper East Side and East Harlem of New York City. It is also apparent that in New York, FIRE workers, more

[17] Since data on crime rate and direct pollution measures at the census tract level are only available for a subset of tracts, I do not include them in my main specification. instead, the results are shown in Table A6 and Table A7

than manufacturing workers, prefer central Manhattan as a place to live and display weaker preferences over non-central locations. This pattern is reversed in Detroit, as shown in Figure 1.10, where FIRE workers exhibit stronger preferences over the suburbs than manufacturing workers. The San Francisco Bay Area is overall a much more desirable place to live and the preferences across different locations within the city do not vary as greatly as in New York or in Detroit. San Francisco downtown is a desirable place to live for both types of workers but more desirable for FIRE workers. From these maps, it can be seen that one difference distinguishing Detroit from the more successful cities such as New York or San Francisco is its considerably lower urban amenity in downtown, especially for FIRE workers.

It appears to be a piece of evidence consistent with the observed centralized poverty in former industrial towns. As these cities fail to create favourable living conditions for the skilled labour force around locations where the skilled sectors are more productive, they are becoming increasingly unattractive to these sectors.

Table A8 shows the correlation between my estimated aggregate productivity A_{is} by sectors with the same set of covariates. It is clear that being closer to industrial areas brings about productivity loss, especially in high-skilled sectors, but the magnitudes are small compared to those on amenity. In contrast to the findings on amenity, being closer to the CBD increases the productivity of all sectors, and the relationship is stronger for high-skilled sectors. Understandably the productivity of these sectors relies strongly on human interactions, which are more likely to be achieved in dense urban cores where commuting costs are relatively low. Being close to highways is also associated with higher productivity, but the effects do not vary systematically by the skill intensity of sectors.

1.6.4 How do Productivity and Amenity Respond to Early Industrial Pollution?

In the previous sections, I backed out sector-specific productivity and amenity parameters (A_{js} and $\tilde{B_{is}}$), and isolated parts of the productivity and amenities determined by local fundamentals, a_{js} and $\tilde{b_{is}}$, under several assumptions on the extent and forms of production and residential externalities. In this section, I will check how these four sets of parameters correlate with historical pollution exposure driven purely by wind patterns.

In Figure 1.12, I plot the estimated coefficients on the interaction term between an indicator of within 4 km from the closest 1970 industrial area and a downwind dummy, where the dependent variables are sector-specific productivity or amenity estimates in a sample with census tracts that are closest to the above-median polluted industrial areas. They presents the way in which historical industrial pollution affects estimated productivity and amenity today and how these impacts vary across industries. To make the point that early industrial pollution creates self-fulfilling skill sorting, I plot the estimated coefficients against the skill intensity of each industry. It is clear that more skilled industries suffer more in productivity as a result of being more exposed to industrial pollution in the 1970s. FIRE productivity in tracts that are close and downwind of 1970s industrial areas is 8% lower, compared to manufacturing productivity that is only 4% lower. To highlight the role of agglomeration effects, I present the estimated coefficients from a sample with only central tracts in the right figure[18]. The effects are three times as large, suggesting that agglomeration forces play an important role in this pattern, since the closer to CBD, the denser are the tracts and the stronger are the agglomeration benefits.

In Figure 1.13, I repeat the exercise with sector-specific amenity parameters as the outcomes. The patterns are quite similar except for relatively lower magnitudes. Employees from more skilled sectors find the historically dirtier tracts less pleasant now, more than do employees from other sectors. The estimated amenity perceived by FIRE employees in tracts that are close to and downwind of 1970s industrial areas is 2.5% lower and the amenity perceived by manufacturing employees does not appear to respond strongly to historical pollution.

In the next step, I decompose the estimated productivity and amenity into an endogenous agglomeration term and an exogenous local fundamental term. As discussed in sections 6.1 and 6.2, there are different assumptions on the functional forms of agglomeration effects. Here I consider the general case where sector-specific productivity depends on both own-industry employment and that from other industries according to the similarity between them, while sector-specific amenity depends on both residential density and the share of college graduates (Equations (20) and (24)).

If we believe the impact of early industrial pollution on contemporary variables is largely driven by agglomeration forces, we should observe much smaller estimated coefficients on the responsiveness of production and residential fundamentals (a_{js} and $\tilde{b_{is}}$) to pollution. The results are shown in Figure 1.14 and 1.15. In both figures, I plot the estimated elasticities of aggregate productivity/amenity to historical pollution exposure on the left panel and the those of fundamentals to pollution on the right. Comparing the magnitudes of the estimated impacts between the two panels in Figure 1.14, it is clear that accounting for the agglomeration effects explains about 50% of the variation in the impact of historical pollution on productivity.

[18]Central tracts are defined as tracts closest to the CBD holding up to 25% of total CBSA population.

Taking away the agglomeration effects also makes the slope of the estimated coefficients over sector-specific skill intensity much flatter, although most of them still remain negative. This is a clear sign that agglomeration effects play an important role in driving the observed relationship between historical pollution and sector-specific productivity, but are not sufficient to account for all the estimated impacts. There are two possible ways to explain the second observation. It might be that lingering pollution still exerted an instantaneous negative impact on productivity fundamentals in 2000, or that the actual production-side agglomeration forces that shaped the evolution of American neighbourhoods are stronger than those estimated in the literature. I believe the second one to be more relevant quantitatively, as Figure 1.3 shows that being downwind of a historical industrial site plays almost no role in explaining the variation in pollution after 2000. In addition, we are looking at changes during a period of rapid structural transformation and massive churning in the labour market. As a result, the initial distribution of skills across the country can be quite important in shaping local creativity and competitiveness at the beginning of a period of prolonged growth in skilled service sectors.

In Figure 1.15, we consider the responsiveness of sector-specific residential amenity parameters $\tilde{B_{is}}$ and fundamentals $\tilde{b_{is}}$ to pollution. The patterns are similar to those on productivity in the sense that accounting for the agglomeration effects explains a high proportion of the impacts of historical pollution, leaving the average of estimated pollution impacts on amenity measures across different industries close to zero. However, the slope of the estimated impacts against sectoral skill intensity does not change after deducting the contribution of access to residential density to amenity. Again, it confirms the role of residential agglomeration effects in accounting for the long-run impact of historical pollution, which is consistent with my findings on observable endogenous amenities, presented in Table 1.7. But we may need to be more flexible about the functional forms of agglomeration effects across different industries.

1.6.5 How do Productivity and Amenity Respond to Contemporary Industrial Pollution?

One of the major objectives of the theoretical framework is to lay a foundation for policy counterfactual analyses. One scenario is to examine the general equilibrium implications of cutting pollution in particular parts of cities. To achieve this, we need a set of reliable estimates on the elasticities of sectoral productivity and amenity to pollution.

To uncover the causal impacts of changes in air pollution on local productivity and amenity, I implement an identification strategy that exploits differential regulation intensity in response to the Clean Air Act Amendments (CAAA) after 1990, studies by Auffhammer et al. (2009) and Bento et al. (2015).

In 1990, the US Environmental Protection Agency (EPA) started to regulate PM10 seriously. The regulation under the 1990 CAAA assigns "nonattainment" status to counties with PM10 concentrations exceeding the federal ceiling. Industrial polluters in nonattainment counties would face stricter regulatory oversight. A county is designated to be out of attainment if at least one of the monitors within the county had daily or annual PM10 concentrations exceeding any of these standards. Following Bento et al. (2015), I capture the differential enforcement of CAAA within a nonattainment county by assigning nonattainment status to each monitor within the county and use monitor attainment status as an instrument for localized pollution reductions. The intuition is that local regulators have incentives to target the areas around nonattainment monitors as a strategy to comply with federal standards in the least costly way. As workplace employment and earnings data are available only for the years 2000 and 2010 in a full US sample, we focus our attention on changes in productivity and amenity during this period, and see how they respond to pollution cut driven by within-county differential implementation of CAAA.

The National Ambient Air Quality Standard (NAAQS) for PM10 before 2012 not only stipulated the maximum level of the PM10 annual average but also required the daily reading of PM10 concentration not to exceed a particular level more than once per year, which creates some randomness in regulatory stringency conditional on the initial level of pollution, in that a county could break the standard with only two bad days. Ideally, we could check how annual PM10 changes respond to monitor level nonattainment status in the following way:

$$\Delta PM10^j_{i,t} = \alpha_1 MonitorNonAtt^j_{i,t} + \theta_t + \Delta\epsilon_{i,t} \tag{1.25}$$

where $\Delta PM10^j_{i,t}$ is the change in PM10 readings from year $t-1$ to year t, at monitor i in county j. To minimize the concerns of mean reversion in pollution, I focus only on the nonattainment status that relies on the second highest reading throughout a year, defining $MonitorNonAtt^j_{i,t}$ to be a dummy that takes one if the second highest PM10 reading in year $t-1$ exceeds $150\mu g/m^3$, effectively breaking the second NAAQS standard. In other words, I capture the variation in pollution reduction induced by differential regulatory efforts as a result of two bad days in the past year, which is pretty random in nature. As my outcomes are observed only in decades, I need to aggregate the above specification to the whole period from 2000 to 2010 by adding up both sides of the equation. The adjusted first stage of my instrumental variable specification is accordingly:

$$\Delta PM10_{ic}^{2010-2000} = \sigma N_i + X_i'\mu + \alpha_c + \epsilon_i \tag{1.26}$$

where $\Delta PM10_i$ is the change in tract-level PM10 measures from 2000 to 2010; the instrument N_i is equal to the ratio of nonattainment years during 2001 to 2007.

In the second stage, I examine the impacts of regulation-induced PM10 cut on estimated productivity and amenity parameters, taking the form of:

$$\Delta ln(y_{is}^{2010-2000}) = \beta \Delta P\hat{M}10_i + X_i'\Gamma + \alpha_c + \epsilon_i \tag{1.27}$$

where y_{is} stands for sector-specific local productivity A_{is} or amenity $\tilde{B_{is}}$. $\Delta ln(y_{is})$ measures their log changes from 2000 to 2010. $\Delta P\hat{M}10_i$ is the change in PM10 from 2000 to 2010, instrumented by N_i, the ratio of nonattainment years during 2001 to 2007. X_i are time-invariant observable tract characteristics; μ and Γ capture changes over time in the premium to these characteristics; α_c are city fixed effects that control for city-wide shocks in productivity and amenities. I only keep the tracts within 2 km of the closest PM10 monitor with positive readings in year 2000,2002-2007 and 2010, and assign monitor-level pollution cut and nonattainment status to the closest tracts when I map monitor-level first stage results to the tract-level second stage.

The upper panel of Table 1.10 reports the responsiveness of estimated sector-specific amenity changes from 2000 to 2010 to the reductions of PM10 concentrations, instrumented by monitor level nonattainment. I also include changes in the number of public schools during the same period as a control variable to check if the measured amenity changes are able to pick up changes in other observable amenities.

It is clear that policy-induced pollution cuts lead to growth in perceived local amenity, especially among workers from skilled sectors. A unit decrease in PM10 leads to a 0.9% appreciation of local amenity perceived by FIRE workers, 1.48% by IT employees, but 0.08% by manufacturing workers. In this context, the regulated tracts (tracts with at least one year in nonattainment status) experienced 5 more units of PM10 decreases from 2000 to 2010, which translated into an average 5% increase in local amenity. In the meantime, amenity also responds positively to new public schools. One additional public school makes local FIRE employees 1.95% more appreciative of the tract as an ideal place to reside, compared to 1.16% for manufacturing workers.

The lower panel of Table 1.10 reports similar results on productivity changes. Although the coefficients are not significant for most of the outcomes, their signs are mostly negative, which suggests that a reduction in PM10 level induces a positive but insignificant growth in productivity. Figure 1.16 plot the coefficients from both productivity and amenity regressions on sectoral skill intensity. It is clear that the responsiveness of both sector-specific amenity and productivity to contemporary pollution cuts increases in the skill intensity of the industry.

1.6.6 Model Validation

In this section, I use my calibrated model of 1970 to predict how TSP cut in the 1970s cause housing prices to change during this period. I then compare the model predictions to reduced form estimates in the literature. The experiment I consider is cutting TSP by 10 $\mu g/m^3$ for a random sample of census tracts.

In Table 1.12, I report the simulated change in housing price and population. It appears that a 10 $\mu g/m^3$ cut in TSP leads to a significant 11.7% increase in housing price. Given the baseline TSP levels in the 1970s, the implied elasticity of housing prices to TSP reductions is -1.4, which suggests that a 1% cut in TSP leads to a 1.4% growth in housing value. I consider two pieces of reduced-form evidence from the literature. The first is Chay and Greenstone (2003)'s estimate of housing price responses to Clean-Air-Act-induced TSP cuts in the 1970s. The time frame and pollutant type match my experiment but their unit of analysis is county. It is clear that my estimates are 4-6 times as large as theirs. It is likely that the difference in the unit of observation is partially driving the observed discrepancy. Population mobility across census tracts should be larger than that across county, and one could argue that the lack of mobility limits the responses of housing prices. Another possibility is that pollution monitors are usually installed in the dirtier parts of a county, examining pollution changes captured by the weighted average of monitor readings might overestimate the extent of pollution cut, and eventually underestimate the response elasticity.

To address this, I bring in a second piece of evidence from Bento et al. (2015), where they consider the effects policy-induced PM10 cut from 1990 to 2000 on housing price changes in a sample of US census tracts. They find that 10 units cut in PM10 lead to about 10% increase in housing price, and the implied elasticity of house prices with respect to PM10 reductions is about -0.6, two times as large as Chay and Greenstone (2003). Their estimates are closer to the ones I get out of the model but still smaller. One possibility is that the effects I get are based on a model with both production and residential side agglomeration effects. So the housing price changes should be interpreted as the difference between initial housing price and the final equilibrium housing price. If the process of sorting and agglomeration lasts longer than a decade, the period of focus of both Chay and Greenstone (2003) and Bento et al. (2015), then their estimates are naturally smaller than mine since their changes are only halfway through the initial price and final equilibrium ones.

1.7 Counterfactual Analysis

In the previous sections, I presented empirical evidence linking historical pollution to the distribution of skills and long-run city development patterns. In this section, I will further quantify the aggregate impact of pollution reduction on welfare and the cross-city distribution of economic activities through the lens of the theoretical framework outlined in section 5.

1.7.1 Settings

The policy experiments I consider here are pollution reductions in different parts of cities. In practice, I divide the 310 metropolitan CBSAs in my sample into two halves and simulate the general equilibrium effects of partial pollution reduction in the "treated" half. To illustrate different implications of the same level of pollution cuts in different parts of a city, I consider the case of a pollution cut of 10 $\mu g/m^3$ in all central tracts, or in a subset of similar non-central tracts, within each "treated" city. Central tracts are defined as the top quartile of tracts in distance to the CBD. To mimic real-life tradeoffs between environmental externalities and economic benefits in environmental regulation decisions, I select the subset of non-central tracts based on their manufacturing productivity. I match each treated central tract with a non-central tract based on their estimated manufacturing productivity, and ensure the average manufacturing productivity of non-central treated tracts to be similar to that of central treated tracts. The idea is that the environmental benefits of relocating industries from the city centre to the suburbs might be compromised if this process is related to a large drop in manufacturing productivity and economic losses. So we want to make sure that we are comparing welfare gains from pollution reductions across tracts with similar costs from the contraction in manufacturing activities.

Throughout my framework, pollution cuts will be realized as changes in local production and amenity fundamentals (a_{is} and $\tilde{b_{is}}$). Section 6.5 discusses the procedure to obtain productivity and amenity elasticities to pollution cuts. Although the coefficients reported in Table 1.10 correspond to the elasticity of aggregate productivity and amenity (A_{is} and $\tilde{B_{is}}$) changes to pollution change, under the assumption that agglomeration forces are limited within ten years, we approximate these to be the estimates on the responsiveness of local production and residential fundamentals (a_{is} and $\tilde{b_{is}}$) to pollution cuts. As is apparent from Table 1.10, the coefficients on amenity are negative across the board, but positive on productivity for a few sectors, although the coefficients are not statistically significant in the second case. Since conceptually we believe pollution per se is detrimental to local productivity (Zivin and Neidell (2012)), I impose the elasticity to be zero in my counterfactual analysis if the coefficients are positive. For now I do not incorporate the cost of environmental regulation such as loss in manufacturing TFP and workers' welfare (Greenstone et al. (2012), Walker (2013)).

The model is calibrated using 2000 data, but we are more interested in the welfare implications of pollution reductions before the industrial decline. Therefore the parameters need to be adjusted in order to reflect realities in the 1970s. Here I make two key assumptions on the evolution of productivity and residential fundamentals from 1970 to 2000. First, the changes in residential fundamentals during this period are assumed to be induced by changes in air quality only. Second, the changes in productivity fundamentals from 1970 to 2000 correspond only to the national growth trends of different sectors[19]. I recalibrate the model with the assumed fundamental values to solve for the other variables. In the presence of agglomeration forces, there is the possibility of multiple equilibria. I impose the equilibrium selection rule of solving for the closest 1970 equilibrium to the observed equilibrium with original fundamental values.

To illustrate the quantitative relevance of agglomeration effects, I undertake two sets of counterfactual analyses, keeping and shutting down agglomeration mechanisms respectively. In a model without agglomeration effects, pollution reduction is assumed to affect aggregate productivity and amenity estimates a_{is} and $\tilde{b_{is}}$, directly. In a model with agglomeration effects, it exerts an initial impact on the fundamentals A_{is} and $\tilde{B_{is}}$, which in turn leads to further changes in aggregate productivity and amenity through changes in workplace and residence employment density. As discussed in sections 6.1 and 6.2, there are different assumptions on the functional forms of agglomeration effects. Here I consider the general case where sector-specific productivity depends on both own-industry employment and that from other industries according to the similarity between them, and sector-specific amenity depends on residential density only. (Equations (19) and (23))

[19]To be more specific, I assume the same percentage growths in productivity across all the census tracts in the US from 1980 to 2000. For each sector, the productivity growth is determined by growth in national sectoral employment from 1980 to 2000: $H_{Ms,2000}/H_{Ms,1980} = (A_{s,2000}/A_{s,1980})^{\epsilon}$

1.7.2 Results

Table 1.11 reports the simulated income and employment distributions after the proposed pollution cut experiments. Conceptually, the same levels of pollution reduction in different parts of a city may entail different aggregate welfare or distributional impacts for several reasons. First, similar initial impacts are exacerbated to a different extent in areas with different residential and workplace density through agglomeration effects. Second, commuting enables workers to access high productivity locations. As a result, productivity growth in well-connected tracts is more easily propagated to surrounding areas, and hence is expected to increase welfare more than similar growth in less well connected areas. Third, a fixed floor space supply means that uneven increases in productivity and amenity fundamentals will lead to a relocation of floor space across residential and business uses and a corresponding labour relocation. This channel may lead to negative overall welfare effects in the presence of agglomeration: an increase in amenity in highly productive areas with high workplace density will lead to a conversion of floor space use from business to residence and potentially a drop in workplace employment, which will then lead to lower productivity for the surrounding areas due to a drop in nearby workplace density. From a welfare perspective, this drop in production-side agglomeration benefits may outweigh the accompanying growth in local amenity and agglomeration benefits from higher residential density.

The upper panel reports the results obtained from an exogenous model when agglomeration effects are removed on the left and those from an endogenous model on the right. It is evident that pollution cuts always lead to overall welfare gains across the US, in the form of higher rent and labour incomes. In the exogenous model shutting down agglomeration effects, cutting pollution in 3500 central tracts from 155 treated cities leads to a 0.43% and 0.36% growth in rent and labour income respectively for the whole US. It also encourages a relocation of labour across industries, generating 2% growth in employment in the skilled sectors. Apart from the aggregate impacts, these empirically-realistic changes in pollution also result in substantial changes in the spatial distribution of employment across locations. Cities with pollution cuts in central tracts experience a 1.8% inflow of residents/workers and an even higher 6.3% growth in employment in skilled sectors, compared to a population loss of 1.4% in other cities. The aggregate welfare gains are only slightly higher under the endogenous model, with growth in rent and labour incomes of 0.46% and 0.43% respectively, compared to a much larger labour relocation effect. Pollution cutting in central tracts leads to 4.7% and 12.4% growth in employment and skilled employment in treated cities and a 1.9% and 2.9% drop in employment and skilled employment in other cities.

Moving to the bottom panel, we examine the welfare and distributional impacts of cutting pollution in non-central census tracts. Without agglomeration effects, a 10 $\mu g/m^3$ cut in TSP level leads to a 0.035% growth in rent income and 0.048% in labour income for the whole US, which are much smaller than the impact of cutting the same level of pollution in central tracts. The extent of both cross-city and cross-industry labour relocation is smaller than the previous policy experiment. In general, central tracts have higher productivity in the service sector and lower overall amenity than non-central tracts. As a consequence, cutting pollution in central tracts leads to a higher productivity change than the amenity change compared to non-central tracts. In the meantime, central tracts are more conveniently connected to other tracts, which means that extra productivity benefits are realized as a employment and income growth in a larger number of surrounding tracts.

A somewhat surprising result, presented in the lower-right panel of Table 1.11, is that the overall welfare gains are lower under the endogenous model than the exogenous model following pollution cuts in non-central tracts, although the same policy experiment leads to a greater labour relocation across cities and industries. A plausible explanation is that agglomeration effects are likely to drive larger inflows of employees and residents into suburban tracts from both central tracts and other cities. The agglomeration benefits from inflows in these tracts are outweighed by the loss in agglomeration benefits in tracts with population and employment outflows, as these suburban tracts are less dense and connected to surrounding tracts. A relevant lesson is that when we consider place-based policies, including targeted pollution abatement, we need to take full account of the possible agglomeration effects at both targeted and the surrounding areas in terms of population loss.

1.8 Conclusion

The prolonged decline of former industrial cities in recent decades has gathered considerable attention and interest within both academic and policy circles. In this paper, I evaluate the role of former industrial pollution in shaping the competitiveness of post-industrial US cities. Severe pollution incurred during an industrial city's manufacturing heyday affected the location choices of skilled workers and residents, which reduced their attractiveness to modern service jobs. This, in turn, affect their ability to transition into a new type of service economy, which relies on the presence of high skilled workers and residents. This is an important issue because apart from the contemporary negative impacts on health, productivity, and quality-of-life, pollution is associated with a more subtle long-run cost of shifting the development trajectory of cities. And this growth effect of pollution has largely been ignored in the existing literature. My findings from the US have profound contemporary implications in developing countries such as China and India, where high levels of pollution might cast a long shadow on the prospect of developing new industries within cities many decades later. The lessons from the US are particularly important and timely as early signs of deindustrialisation have already emerged in the Chinese economy.

To look at this issue, I assemble a rich database at a fine spatial resolution that allows me to link industrial pollution in the 1970s to current economic outcomes. To overcome the empirical challenge that the placement of polluting industries in the 1970s was not random, I use the fact that wind patterns generate quasi-exogenous variation in pollution exposure, holding constant local manufacturing activity. I first show that being downwind of an industrial site was associated with higher pollution levels in the 1970s but this was no longer the case in the 2000s, partly because pollution has largely been cut in American cities with manufacturing decline. However, a wide range of economic outcomes continues to be influenced by historical pollution. Census tracts downwind of highly polluted 1970s industrial sites have lower housing prices and a smaller share of skilled workers and residents three decades later, a pattern which became more prominent between 1980 and 2000. These findings suggest that early pollution limited the ability of cities to attract skilled workers, which in turn affect the subsequent agglomeration patterns.

To delve further into the mechanisms at work, I construct a structural model that allows me to back out sector-specific productivity and amenity parameters. As such, we can tell how far IT firms value a particular location relative to construction firms, and how far finance versus manufacturing workers prefer to live in certain areas in order to enjoy a better quality of life. I find the magnitude of effects of historical pollution on current productivity are higher than those on amenity, which suggests that the sorting of skilled workers and productive businesses away from historically-polluted areas plays a particularly important role in determining the post-industrial distribution of skills.

With the help of the model, I am also able to carry out a variety of policy counterfactual experiments. I look at the effects of cutting pollution in central or suburban tracts of a subset of cities on aggregate welfare and the distribution of skilled workers across industries and cities. I find the benefits to be much larger with pollution cut in central city census tracts. The key lesson is that the long-run social cost of pollution differs markedly across areas with different employment density and transportation connectivity. This should be taken into consideration by policy makers from countries at different stages of development. For industrialising countries in Africa, urban planners should consider the long-run negative externalities of industrial pollution in determining the location of polluting industries in cities, while in newly industrialised countries such as China and India, immediate actions should be taken to control pollution in central cities, which involves stricter standards on central city pollution or a relocation of manufacturing activities to suburban industrial zones.

There are several extensions worth exploring. First, one critical and to some extent surprising finding of this paper is the larger impact of historical pollution on current productivity than amenity, especially among skilled sectors. A meaningful follow-up project would be to delve further into the channels involved. I plan to examine the relationship between historical pollution exposure and the current location of firms and patenting activities to see if the documented productivity effects arise from the sorting of productive firms and the clustering of innovation activities. Second, there are a variety of model extensions and counterfactual analyses that could be considered, such as incorporating amenities derived at workplace into the model and simulating outcomes taking account of the economic losses from pollution abatement. Finally, as the findings of this paper offer particularly relevant lessons for newly industrialised countries such as China and India, it would be worthwhile to check for evidence of skill sorting on environmental amenities within Chinese or Indian cities, using a combination of online retailing and social network data to map out the distribution of young professionals within-city.

Figures and Tables

Table 1.1: Industrial Activities and Pollution in the 1970s

Variables	Monitor-level average TSP from 1971 to 1979					
Sample	Full Sample		Monitors closest to above-median polluted industrial area		Monitors closest to below-median polluted industrial area	
	(1)	(2)	(3)	(4)	(5)	(6)
$1(disind \in 0-1km)$	23.71***	23.48***	22.36***	22.83***	11.11***	9.137***
	(1.510)	(1.683)	(2.526)	(2.923)	(0.808)	(0.749)
$1(disind \in 1-2km)$	12.07***	12.14***	8.897***	11.20***	7.130***	5.935***
	(1.455)	(1.663)	(2.377)	(2.786)	(0.825)	(0.795)
$1(disind \in 2-3km)$	6.456***	5.005***	3.598	4.623	5.154***	4.631***
	(1.500)	(1.687)	(2.729)	(3.072)	(0.756)	(0.804)
$1(disind \in 3-4km)$	7.329***	6.996***	4.606	4.842**	4.813***	4.132***
	(1.752)	(1.940)	(3.330)	(2.342)	(0.875)	(0.889)
$1(disind \in 0-1km)$*Downwind		8.997**		11.75**		0.203
		(3.921)		(5.461)		(2.535)
$1(disind \in 1-2km)$*Downwind		4.403		9.872*		2.523
		(3.082)		(5.268)		(2.364)
$1(disind \in 2-3km)$*Downwind		11.98***		11.86***		-0.401
		(3.292)		(3.779)		(2.492)
$1(disind \in 3-4km)$*Downwind		5.224		3.300		1.226
		(5.674)		(9.281)		(3.978)
$1(disind > 4km)$*Downwind		1.112		-2.138		-2.624
		(2.930)		(6.520)		(3.040)
Observations	4,968	4,968	2,471	2,471	2,422	2,422
p-value (Sum of interactions=0)		0.0001		0.0028		0.5620

Notes: Dependent variables are the average measure of TSP ambient concentration from 1971 to 1979 collected at each TSP monitor with positive reading during this period. $1(disind \in a-bkm)$ is an indicator of whether or not the distance from a TSP monitor to the closest 1970s industrial area is within a and b km. Downwind is an indicator of whether or not the TSP monitor is located downwind of the industrial area, which is defined by Figure 1.1. Results from columns (1) and (2) are obtained from a full sample of monitors with positive TSP readings from 1971 to 1979, those from columns (3) and (4) are obtained from a sample of monitors that are closest to industrial areas with above-median pollution intensity in the 1970s, and those from columns (5) and (6) are from a sample of monitors closest to industrial areas with below-median pollution level. Controls include CBSA fixed effects, the distance from each monitor to transportation lines, natural amenities and the CBD. Robust standard errors are clustered at CBSA level.

Table 1.2: Industrial Activities in the 1970s and Current Pollution

Variables	Monitor-level average PM10 from 2000 to 2010					
Sample	Full Sample		Monitors closest to above-median polluted industrial area		Monitors closest to below-median polluted industrial area	
	(1)	(2)	(3)	(4)	(5)	(6)
$1(disind \in 0-1km)$	3.233***	3.252***	1.845	1.912	3.369***	3.321***
	(0.849)	(0.892)	(1.726)	(1.853)	(0.863)	(0.925)
$1(disind \in 1-2km)$	1.877**	1.832**	1.163	0.854	2.219**	2.336**
	(0.820)	(0.861)	(1.746)	(1.822)	(1.098)	(1.168)
$1(disind \in 2-3km)$	1.921**	1.887*	0.282	0.272	2.349*	2.277
	(0.902)	(0.978)	(1.710)	(1.862)	(1.250)	(1.412)
$1(disind \in 3-4km)$	2.328	1.982	4.129	3.795	1.191	1.278
	(1.799)	(1.931)	(3.859)	(4.321)	(1.275)	(1.330)
$1(disind \in 0-1km)$*Downwind		-2.566*		-3.710**		-0.142
		(1.362)		(1.876)		(2.093)
$1(disind \in 1-2km)$*Downwind		-1.235		1.067		-1.363
		(1.544)		(3.922)		(2.082)
$1(disind \in 2-3km)$*Downwind		-2.393		-5.743		-0.254
		(2.183)		(3.965)		(3.307)
$1(disind \in 3-4km)$*Downwind		2.015		-0.750		-3.225
		(3.647)		(5.336)		(6.775)
$1(disind > 4km)$*Downwind		-2.897*		-3.819		-1.343
		(1.475)		(2.755)		(2.439)
Observations	1,893	1,893	951	951	942	942
R-squared	0.299	0.301	0.358	0.360	0.377	0.378
p-value (Sum of interactions=0)		0.3975		0.3089		0.5625

Notes: Dependent variables are the average measure of PM10 ambient concentration from 2000 to 2010 collected at each PM10 monitor with positive reading during this period. $1(disind \in a-bkm)$ is an indicator of the distance of a PM10 monitor from the closest 1970 industrial area is within a and b km. Downwind is an indicator of whether or not the PM10 monitor is located downwind of the industrial area, which is defined by Figure 1.1. Results from columns (1) and (2) are obtained from a full sample of monitors with positive PM10 readings from 2000 to 2010, those from columns (3) and (4) are obtained from a sample of monitors that are closest to industrial areas with above-median pollution level in the 1970s, and those from columns (5) and (6) are from a sample of monitors closest to industrial areas with below-median pollution level. Controls include CBSA fixed effects, the distance from each monitor to transportation lines, natural amenities and the CBD. Robust standard errors are clustered at CBSA level.

Table 1.3: Industrial Activities in the 1970s and Current Pollution: Other Pollutants

	Monitor-level average pollutants from 2000 to 2010							
Variables	Ozone		CO		SO_2		NO_2	
$1(disind \in 0-1km)$	-0.00193***	-0.00160**	0.0790	0.0904*	5.892***	6.163***	10.21***	9.700***
	(0.000597)	(0.000630)	(0.0515)	(0.0545)	(1.431)	(1.474)	(1.853)	(1.801)
$1(disind \in 1-2km)$	-0.000373	-0.000448	0.0733*	0.0697	4.836***	5.786***	9.168***	8.671***
	(0.000631)	(0.000667)	(0.0434)	(0.0441)	(1.552)	(1.573)	(2.082)	(2.078)
$1(disind \in 2-3km)$	-0.00110	-0.00118	0.0157	0.0480	1.985	1.384	5.651**	4.981**
	(0.000674)	(0.000725)	(0.0533)	(0.0636)	(1.453)	(1.807)	(2.229)	(2.248)
$1(disind \in 3-4km)$	-0.00179**	-0.00207**	0.0315	0.0317	1.983	2.474	3.642	2.702
	(0.000905)	(0.000830)	(0.0645)	(0.0675)	(2.009)	(2.201)	(2.571)	(2.637)
$1(disind \in 0-1km)$*Downwind		-0.00289		-0.140		-1.596		2.249
		(0.00186)		(0.121)		(1.682)		(3.292)
$1(disind \in 1-2km)$*Downwind		0.000668		-0.0587		-6.899***		3.542
		(0.00162)		(0.149)		(2.535)		(3.463)
$1(disind \in 2-3km)$*Downwind		0.000794		-0.278**		3.296		2.785
		(0.00128)		(0.137)		(3.162)		(4.732)
$1(disind \in 3-4km)$*Downwind		0.00300		-0.0192		-3.502		9.853
		(0.00336)		(0.0950)		(5.277)		(6.886)
$1(disind > 4km)$*Downwind		-8.69e-05		0.125		2.151		-3.954*
		(0.000920)		(0.137)		(2.208)		(2.310)
Observations	1,330	1,330	594	594	664	664	538	538
R-squared	0.445	0.450	0.486	0.496	0.473	0.482	0.592	0.598

Notes: Dependent variables are the average measures of ambient concentration of four alternative pollutants (Ozone, CO, SO_2 and NO_2) from 2000 to 2010 collected at each monitor with positive reading during this period. $1(disind \in a-bkm)$ is an indicator of the distance of a PM10 monitor from the closest 1970 industrial area is within a and b km. Downwind is an indicator of whether or not the PM10 monitor is located downwind of the industrial area, which is defined by Figure 1.1. Controls include CBSA fixed effects, the distance from each monitor to transportation lines, natural amenities and the CBD. Robust standard errors are clustered at CBSA level.

Table 1.4: Economic outcomes in 2000: Downwind dummy

Outcomes in 2000: Tracts closest to above-median pollution level								
	By place of work			By place of residence				
VARIABLES	Employ density	Highskill ratio	Median earnings	Employ density	Highskill ratio	Median earnings	Housing value	College ratio
$1(disind \in 0-1km)$*Downwind	-0.0955	-0.0213*	-0.0346*	-0.0633	-0.0215***	-0.0124	-0.0957***	-0.0356**
	(0.102)	(0.0117)	(0.0194)	(0.0447)	(0.00709)	(0.0194)	(0.0291)	(0.0148)
$1(disind \in 1-2km)$*Downwind	-0.117	-0.0204***	-0.0280**	0.0114	-0.0320**	-0.0739***	-0.0814***	-0.0493**
	(0.0970)	(0.00629)	(0.0123)	(0.0389)	(0.0125)	(0.0260)	(0.0273)	(0.0155)
$1(disind \in 2-3km)$*Downwind	-0.115	-0.0100	-0.0190	-0.0954	-0.00510	-0.0192	-0.0124	-0.0190
	(0.116)	(0.00658)	(0.0176)	(0.0832)	(0.0125)	(0.0263)	(0.0308)	(0.0120)
$1(disind \in 3-4km)$*Downwind	-0.232	-0.00939	-0.0536**	-0.0196	-0.00364	-0.0236	0.0190	-0.0256
	(0.154)	(0.00985)	(0.0229)	(0.105)	(0.0166)	(0.0277)	(0.0537)	(0.0229)
$1(disind > 4km)$*Downwind	0.174	0.00614	0.0237	0.194	-0.00294	-0.0301	0.0225	-0.0206
	(0.185)	(0.0120)	(0.0217)	(0.135)	(0.0106)	(0.0407)	(0.0519)	(0.0193)
$1(disind \in 0-1km)$	0.0190	-0.0267	0.0428	-0.0136	-0.0344***	-0.163***	-0.188***	-0.0960**
	(0.238)	(0.0164)	(0.0362)	(0.252)	(0.00939)	(0.0460)	(0.0569)	(0.0219)
$1(disind \in 1-2km)$	0.0612	-0.0179	0.0299	0.0688	-0.00684	-0.0965**	-0.113**	-0.0580**
	(0.231)	(0.0140)	(0.0270)	(0.223)	(0.0107)	(0.0415)	(0.0470)	(0.0184)
$1(disind \in 2-3km)$	-0.0153	-0.0144	0.0131	0.113	-0.00966	-0.0788*	-0.0631	-0.0439**
	(0.145)	(0.0113)	(0.0300)	(0.172)	(0.00685)	(0.0466)	(0.0444)	(0.0213)
$1(disind \in 3-4km)$	0.0228	-0.00779	0.0279	0.0473	0.00579	-0.0251	-0.0108	-0.00287
	(0.106)	(0.00759)	(0.0215)	(0.0840)	(0.00660)	(0.0346)	(0.0452)	(0.0194)
Observations	8,173	8,173	8,023	8,203	8,203	8,020	7,933	8,176
R-squared	0.375	0.080	0.179	0.551	0.266	0.252	0.558	0.175
p-value (Sum of interactions=0)	0.1241	0.0159	0.0016	0.3498	0.2919	0.0870	0.0938	0.0027

Outcomes in 2000: Tracts closest to below-median pollution level								
	By place of work			By place of residence				
VARIABLES	Employ density	Highskill ratio	Median earnings	Employ density	Highskill ratio	Median earnings	Housing value	College ratio
$1(disind \in 0-1km)$*Downwind	-0.153*	-0.00808	-0.0204	-0.0228	-0.00320	0.0225	-0.00648	-0.0133
	(0.0781)	(0.00637)	(0.0133)	(0.0653)	(0.00627)	(0.0198)	(0.0268)	(0.0106)
$1(disind \in 1-2km)$*Downwind	0.00521	0.00123	-0.0179	0.0237	2.47e-05	-0.0124	-0.0233	-0.00582
	(0.104)	(0.00538)	(0.0140)	(0.0794)	(0.00614)	(0.0301)	(0.0367)	(0.0123)
$1(disind \in 2-3km)$*Downwind	-0.122	-0.00629	0.0191	-0.0525	0.00331	0.0238	-0.0192	-0.00780
	(0.126)	(0.00876)	(0.0127)	(0.112)	(0.00615)	(0.0322)	(0.0316)	(0.0135)
$1(disind \in 3-4km)$*Downwind	-0.177	0.00707	-0.0400*	-0.0489	0.00241	0.00793	0.00539	0.0144
	(0.136)	(0.0106)	(0.0238)	(0.110)	(0.0102)	(0.0253)	(0.0443)	(0.0158)
$1(disind > 4km)$*Downwind	-0.391**	-0.0157	-0.00880	-0.186	-0.00779	-0.0233	-0.0239	-0.0269
	(0.112)	(0.00903)	(0.0132)	(0.107)	(0.0122)	(0.0328)	(0.0551)	(0.0233)
$1(disind \in 0-1km)$	0.442***	-0.0330***	-0.0172	0.338**	-0.0403***	-0.179***	-0.262***	-0.109***
	(0.158)	(0.0101)	(0.0208)	(0.157)	(0.00817)	(0.0290)	(0.0355)	(0.0187)
$1(disind \in 1-2km)$	0.403***	-0.0287***	-0.0388*	0.406***	-0.0309***	-0.155***	-0.209***	-0.0875**
	(0.140)	(0.00917)	(0.0209)	(0.151)	(0.00676)	(0.0290)	(0.0332)	(0.0183)
$1(disind \in 2-3km)$	0.267**	-0.0170*	-0.0391*	0.296**	-0.0204***	-0.130***	-0.147***	-0.0632**
	(0.119)	(0.00883)	(0.0204)	(0.134)	(0.00707)	(0.0315)	(0.0350)	(0.0183)
$1(disind \in 3-4km)$	0.178*	-0.0119*	-0.0196	0.214**	-0.0103*	-0.0431*	-0.0741**	-0.0392**
	(0.103)	(0.00677)	(0.0128)	(0.0986)	(0.00530)	(0.0220)	(0.0300)	(0.0135)
Observations	8,139	8,139	7,740	8,139	8,139	7,734	8,038	8,139
R-squared	0.327	0.105	0.267	0.424	0.330	0.310	0.540	0.254
p-value (Sum of interactions=0)	0.0629	0.7770	0.1600	0.7035	0.8898	0.5932	0.6523	0.7199

Notes: Results from the upper panel are obtained from a sample of census tracts that are closest to 1970 industrial areas with above-median pollution intensity measure, and those from the lower panel from a sample of census tracts closest to industrial areas with below-median pollution level. Dependent variables of the first three columns are employment density, the ratio of FIRE, IT and professional services in total employment, and median earnings, counted at the place of work. The last five columns report results on employment density, the ratio of FIRE, IT and professional services in total employment, median earnings, median housing value and ratio of college graduates among the population over 25 years old, counted at their place of residence. $1(disind \in x-ykm)$ is an indicator of whether or not the distance from a tract to its closest industrial area is within x to y km. Downwind is an indicator of whether or not the TSP monitor is located downwind of the industrial area, defined by Figure 1.1. Controls include CBSA fixed effects, the distance from each tract to transportation lines, natural amenities and the CBD, as well as route distance buffers to the same industrial area, and predicted manufacture job growth from 1970 to 2000 based on the industrial composition in 1970. Robust standard errors are clustered at CBSA level.

Table 1.5: Placebo Checks: 1940 and 1950 outcomes

VARIABLES	%College graduates50	%College graduates40	log income 1950	Manager share50	Professional share50
	Tracts closest to above-median polluted industrial areas				
$1(disind \in 0-4km)$	-0.0402*	-0.0340***	-0.123***	-0.0185***	-0.00587
	(0.0232)	(0.0115)	(0.0401)	(0.00656)	(0.00848)
$1(disind \in 0-4km)*$ Downwind	0.0332	0.0196	0.0503	0.0174	0.00771
	(0.0274)	(0.0262)	(0.127)	(0.0136)	(0.0175)
Downwind	-0.0134	-0.0195**	0.00746	-0.00935	-0.00788
	(0.0100)	(0.00851)	(0.0547)	(0.00597)	(0.00842)
Observations	2,418	4,084	3,563	3,984	3,984
R-squared	0.248	0.218	0.321	0.219	0.132
	Tracts closest to below-median polluted industrial areas				
$1(disind \in 0-4km)$	0.00399	-0.0250**	-0.0985**	-0.0220***	-0.0109***
	(0.0139)	(0.0101)	(0.0466)	(0.00792)	(0.00403)
$1(disind \in 0-4km)*$ Downwind	-0.0114	-0.00426	0.00858	0.00750	-0.00484
	(0.0152)	(0.0120)	(0.0704)	(0.0120)	(0.00829)
Downwind	0.0105	0.0278*	0.0586*	0.0217***	0.0153
	(0.0174)	(0.0149)	(0.0328)	(0.00549)	(0.00969)
Observations	1,886	4,236	3,504	4,202	4,202
R-squared	0.261	0.231	0.251	0.151	0.131

Notes: Results from the upper panel are obtained from a sample of 1950 census tracts that are closest to industrial areas with above-median pollution level in the 1970s, and those from the lower panel from a subsample of tracts closest to industrial areas with below-median pollution level. Dependent variables are the share of college graduates in 1940 and 1950, log median income in 1950, the share of managers and professional/technical workers in total employment in 1950, $1(disind \in 0-4km)$ is an indicator of whether or not the distance from a tract to the closest industrial area is within 0 to 4 km. Downwind is an indicator of whether or not the TSP monitor is located downwind of the industrial area. Controls include CBSA fixed effects, the distance from each monitor to transportation lines, natural amenities and the CBD. Robust standard errors are clustered at CBSA level.

Table 1.6: Growth from 1980-2000: Downwind dummy

	Growth from 1980 to 2000: Tracts closest to above-median pollution level					
VARIABLES	Employment	Manufacture employment	FIRE employment	Median income	Housing value	College graduates
$1(disind \in 0-1km)$*Downwind	-0.0561*	-0.0372	-0.119**	-0.0315***	-0.0578***	-0.0822**
	(0.0285)	(0.0382)	(0.0478)	(0.00962)	(0.0201)	(0.0374)
$1(disind \in 1-2km)$*Downwind	-0.0641*	-0.00678	-0.146**	-0.0629***	-0.0530*	-0.143***
	(0.0351)	(0.0463)	(0.0683)	(0.0228)	(0.0274)	(0.0473)
$1(disind \in 2-3km)$*Downwind	-0.0162	0.00772	-0.0597	-0.0226	-0.0126	-0.0577
	(0.0351)	(0.0476)	(0.0424)	(0.0178)	(0.0206)	(0.0524)
$1(disind \in 3-4km)$*Downwind	-0.0454	-0.0768	-0.0381	-0.0209	-0.000550	-0.0258
	(0.0412)	(0.0799)	(0.0453)	(0.0208)	(0.0299)	(0.0497)
$1(disind > 4km)$Downwind	0.101	0.157	0.141	0.00478	0.00788	0.147
	(0.0971)	(0.138)	(0.0952)	(0.0282)	(0.0336)	(0.0930)
$1(disind \in 0-1km)$	0.165	0.364**	0.223	0.0175	0.139**	0.341
	(0.153)	(0.176)	(0.197)	(0.0324)	(0.0546)	(0.253)
$1(disind \in 1-2km)$	0.145	0.267*	0.243	0.0369	0.126**	0.300
	(0.125)	(0.153)	(0.167)	(0.0272)	(0.0524)	(0.204)
$1(disind \in 2-3km)$	0.0955	0.194	0.113	0.0140	0.0830***	0.208
	(0.100)	(0.129)	(0.125)	(0.0256)	(0.0314)	(0.153)
$1(disind \in 3-4km)$	0.0751	0.170	0.0791	0.0302	0.0646***	0.128
	(0.0711)	(0.120)	(0.0926)	(0.0195)	(0.0214)	(0.105)
Observations	7,983	5,895	7,584	8,035	7,683	7,921
R-squared	0.326	0.337	0.247	0.229	0.545	0.204
p-value (Sum of interactions=0)	0.0557	0.3579	0.0120	0.0036	0.0085	0.0027

	Growth from 1980 to 2000: Tracts closest to below-median pollution level					
VARIABLES	Employment	Manufacture employment	FIRE employment	Median income	Housing value	College graduates
$1(disind \in 0-1km)$*Downwind	0.0223	0.0697	0.0504	0.0124	-0.00289	0.0499
	(0.0439)	(0.0531)	(0.0542)	(0.0155)	(0.0196)	(0.0523)
$1(disind \in 1-2km)$*Downwind	0.00732	-0.00384	0.00881	0.0129	-0.00151	0.0558
	(0.0286)	(0.0391)	(0.0438)	(0.0167)	(0.0206)	(0.0379)
$1(disind \in 2-3km)$*Downwind	-0.0637	-0.0910	-0.0891*	0.0130	0.00270	-0.0146
	(0.0424)	(0.0604)	(0.0512)	(0.0139)	(0.0175)	(0.0439)
$1(disind \in 3-4km)$*Downwind	0.213**	0.211**	0.134*	0.00647	-0.0247	0.188**
	(0.0518)	(0.0688)	(0.0592)	(0.0216)	(0.0331)	(0.0807)
$1(disind > 4km)$Downwind	0.0585	0.0669	0.00883	0.0288	0.00547	0.0721
	(0.0971)	(0.138)	(0.0952)	(0.0282)	(0.0336)	(0.0930)
$1(disind \in 0-1km)$	-0.0883	-0.0118	-0.204***	-0.0394*	-0.0149	-0.213**
	(0.0629)	(0.0617)	(0.0771)	(0.0235)	(0.0358)	(0.0931)
$1(disind \in 1-2km)$	-0.0957*	-0.0366	-0.182***	-0.0446**	-0.0128	-0.214***
	(0.0525)	(0.0546)	(0.0661)	(0.0192)	(0.0277)	(0.0760)
$1(disind \in 2-3km)$	-0.0799	-0.0352	-0.137**	-0.0533***	-0.0145	-0.176***
	(0.0492)	(0.0541)	(0.0686)	(0.0181)	(0.0267)	(0.0673)
$1(disind \in 3-4km)$	-0.0559	-0.0106	-0.0875*	-0.0308*	0.000763	-0.113**
	(0.0391)	(0.0460)	(0.0512)	(0.0157)	(0.0193)	(0.0568)
Observations	7,786	6,284	7,533	7,827	7,627	7,761
R-squared	0.381	0.334	0.334	0.244	0.523	0.250
p-value (Sum of interactions=0)	0.0733	0.1950	0.3408	0.2258	0.6812	0.0141

Notes: Results from the upper panel are obtained from a sample of census tracts that are closest to industrial areas with above-median pollution level from 1971 to 1979, and those from the lower panel from a sample of census tracts closest to industrial areas with below-median pollution level during the same period. Dependent variables are growth in total, manufacture, FIRE employment, median household income, median housing value and the number of college graduates from 1980 to 2000, all counted at the place of residence. $1(disind \in x-ykm)$ is an indicator of whether or not the distance from a tract to its closest industrial area is within x to y km. Downwind is an indicator of whether or not the TSP monitor is located downwind of the industrial area. Controls include CBSA fixed effects, the distance from each tract to transportation lines, natural amenities and the CBD, as well as route distance buffers to the same industrial area, and predicted manufacture job growth from 1970 to 2000 based on the industrial composition in 1970. Robust standard errors are clustered at CBSA level.

Table 1.7: Observable amenities and housing quality

	Tracts closest to above-median polluted industrial areas			
VARIABLES	2000 Violent crime rate	2000 public school per capita	1980 share no kitchen/plumbing	2000 share no kitchen/plumbing
$1(disind \in 0-4km)*$ Downwind	0.000346*	-0.000227*	0.000964	-0.000063
	(0.000185)	(0.000144)	(0.00264)	(0.000541)
$1(disind \in 0-4km)$	0.00076	0.000603	-0.000074	-0.000173
	(0.000674)	(0.000377)	(0.00171)	(0.000933)
Downwind	0.00003	0.000029	0.00322	-0.00151
	(0.000299)	(0.000422)	(0.00452)	(0.00151)
Observations	2,315	4,841	2,206	7,939
R-squared	0.343	0.005	0.2	0.088

	Tracts closest to below-median polluted industrial areas			
VARIABLES	2000 Violent crime rate	2000 public school per capita	1980 share no kitchen/plumbing	2000 share no kitchen/plumbing
$1(disind \in 0-4km)*$ Downwind	-0.000163	0.000016	-0.000531	-0.00187
	(0.000198)	(0.000018)	(0.00161)	(0.00121)
$1(disind \in 0-4km)$	0.000350***	0.000061***	-0.000921	0.00144*
	(0.000115)	(0.000023)	(0.0023)	(0.000808)
Downwind	0.00125***	-0.000027	0.00085	0.00152
	(0.000307)	(0.000031)	(0.00374)	(0.00138)
Observations	1,209	3,822	2,224	5,925
R-squared	0.29	0.114	0.316	0.077

Notes: Results from the upper panel are obtained from a sample of census tracts that are closest to 1970 industrial areas with above-median pollution intensity measure, and those from the lower panel from a subsample of tracts closest to industrial areas with below-median pollution level. Dependent variables are tract level violent crime rate in 2000, the number of public schools per capita in 2000, share of housing units without kitchen or plumbing devices in 1980 and 2000. $1(disind \in 0-4km)$ is an indicator of whether or not the distance from a tract to the closest industrial area is within 0 to 4 km. Downwind is an indicator of whether or not the TSP monitor is located downwind of the industrial area. Controls include CBSA fixed effects, the distance from each monitor to transportation lines, natural amenities and the CBD. Robust standard errors are clustered at CBSA level.

Table 1.8: CBSA-level evidence

VARIABLES	Employment growth			Population growth		
Share of $1(disind \in 0-4km)$ tracts in total	-0.376***	-0.421***	-0.265***	-0.126	-0.244**	-0.0758
	(0.113)	(0.1000)	(0.101)	(0.119)	(0.114)	(0.123)
Share of $1(disind \in 0-4km)$*central tracts in total	-0.336		-0.458	-0.428		-0.606**
	(0.310)		(0.293)	(0.269)		(0.275)
Share of downwind tracts in total	-0.135	-0.117	-0.151	-0.400	-0.499	-0.542
	(0.387)	(0.454)	(0.419)	(0.433)	(0.489)	(0.467)
Share of $1(disind \in 0-4km)$*downwind tracts in total		-0.229	-1.079**		0.0868	-0.262
		(0.361)	(0.430)		(0.345)	(0.399)
Share of $1(disind \in 0-4km)$*central downwind tracts in total			6.407			10.33***
			(5.018)			(3.846)
Observations	198	198	198	198	198	198
R-squared	0.727	0.716	0.746	0.835	0.832	0.843

Notes: Dependent variables are CBSA level employment growth from 1983 to 2003 or population growth from 1980 to 2000. Share of $1(disind \in 0-4km)$ tracts in total is the ratio between the number of census tracts within 4 km to the nearest industrial area and the total number of tracts of a CBSA. Share of $1(disind \in 0-4km)$ tracts in total counts the share of tracts that are both within 4 km to the nearest industrial area and located in central city in the total number of tracts. Central tracts are defined as the top quartile of tracts in distance to the CBD. Share of $1(disind \in 0-4km)$*downwind tracts in total denotes the share of tracts in total that are both within 4 km to the nearest industrial area and downwind of it. Full control includes historical population level, census division dummies, physical geography and socioeconomic controls used in Duranton and Turner (2012). SE clustered at census division level.

Table 1.9: Parameter Calibration and Data Sources

Definition and Value of Key parameters		
Parameter	Definition	Value
$1-\beta$	Consumer expenditure share in land	0.25
$1-\alpha$	Firm expenditure share in land	0.2
ϵ	Frechet Shape parameter	6.83
κ	Semi-elasticity of commuting cost on travel time	0.01
λ	Production externalities elasticity	0.07
δ	Production externalities spatial decay	0.36
η	Residential externalities elasticity	0.15
ρ	Residential externalities spatial decay	0.75

Sources of Data Used in Structural Estimation		
Variables	Description	Source
H_{Mis}	Sectoral workplace employment	CTPP
H_{Ris}	Sectoral residents	CTPP
w_{is}	Sectoral wage at workplace	CTPP
	Educational attainment at workplace	LODES
	Gender,race and occupation by sector at workplace	CTPP
q_i	Gross rental rate	NHGIS
	Housing quality measures	NHGIS
τ_{ip}	Bilateral travel time between tracts	OpenStreetMap

Notes: The upper panel shows the calibrated value of key parameters of the model, all of which come from Ahlfeldt et al. (2015). The lower panel presents data sources of the observed variables used in my structural estimation.

Table 1.10: Change in amenity and productivity from 2000 to 2010

Industries	FIRE	IT	Edu Med	Professional	Public admin	Art Entertain	Manufacture	Wholesale	Retail	Utility
					Change in log aggregate amenity from 2000 to 2010					
$\Delta PM10$	-0.00923*	-0.0148*	-0.00667**	-0.00466*	-0.00488*	-0.00099	-0.0008	-0.0053	-0.0081	-0.00241
	(0.00556)	(0.00769)	(0.0029)	(0.000245)	(0.00289)	(0.00478)	(0.00653)	(0.0043)	(0.00619)	(0.00444)
Public Schools	0.0195***	0.00496	0.0140**	0.0146**	0.00983**	0.00719	0.0116**	0.0117*	0.00553	0.00762
	(0.00514)	(0.00580)	(0.00632)	(0.00651)	(0.00492)	(0.00490)	(0.00491)	(0.00618)	(0.00394)	(0.00525)
Observations	2,294	1,882	2,653	2,630	2,204	2,400	2,368	2,067	2,396	2,274
R-squared	0.792	0.844	0.734	0.730	0.826	0.936	0.817	0.744	0.923	0.811
					Change in log aggregate productivity from 2000 to 2010					
$\Delta PM10$	-0.00427	-0.0069	-0.0068	-0.0150	0.0170	0.0346	0.00139	-0.00544	-0.0031	-0.00206
	(0.0060)	(0.0063)	(0.0083)	(0.0109)	(0.0198)	(0.0269)	(0.0120)	(0.0268)	(0.0026)	(0.00191)
Public Schools	0.00657	0.0200	0.0537	0.000358	0.0164	0.0232	0.00370	-0.0185	-0.0104	-0.00402
	(0.0130)	(0.0219)	(0.0631)	(0.0231)	(0.0170)	(0.0258)	(0.0221)	(0.0244)	(0.0174)	(0.0203)
Observations	1,982	1,127	2,298	2,170	1,572	1,632	2,072	1,873	2,049	1,718
R-squared	0.088	0.085	0.154	0.171	0.114	0.119	0.151	0.109	0.078	0.104

Notes: Dependent variables in the upper panel are the log changes in aggregate amenity perceived by workers from different sectors from 2000 to 2010. Dependent variables in the upper panel are the log changes in aggregate productivity of different sectors from 2000 to 2010. The change in PM10 is instrumented by the share of monitor level nonattainment ratio in PM10 from 2002 to 2007. Only tracts within 2 km to a PM10 monitor station are kept in the sample. CBSA fixed effects are controlled for. Robust standard errors are clustered at CBSA level.

Table 1.11: Counterfactuals: Cutting pollution

	Experiment: Cutting TSP by 10 $\mu g/m^3$ in 3500 central tracts					
	Exogenouse case			Endogenous case		
%Change in	Whole U.S.	Treated cities	Other cities	Whole U.S.	Treated cities	Other cities
Rent income	0.430%	2.248%	-0.674%	0.464%	3.305%	-1.684%
labour income	0.364%	2.250%	-0.674%	0.426%	3.312%	-1.689%
Employment	0	1.807%	-1.386%	0	4.670%	-1.948%
Employment in skilled sector	1.966%	6.373%	-1.386%	3.741%	12.447%	-2.879%
	Experiment: Cutting TSP by 10 $\mu g/m^3$ in 3500 non-central tracts					
	Exogenous case			Endogenous case		
%Change in	Whole U.S.	Treated cities	Other cities	Whole U.S.	Treated cities	Other cities
Rent income	0.035%	0.424%	-0.250%	0.007%	0.745%	-0.535%
labour income	0.048%	0.437%	-0.237%	0.006%	0.746%	-0.537%
Employment	0	0.603%	-0.462%	0	1.135%	-0.870%
Employment in skilled sector	0.388%	1.506%	-0.462%	0.922%	3.283%	-0.873%

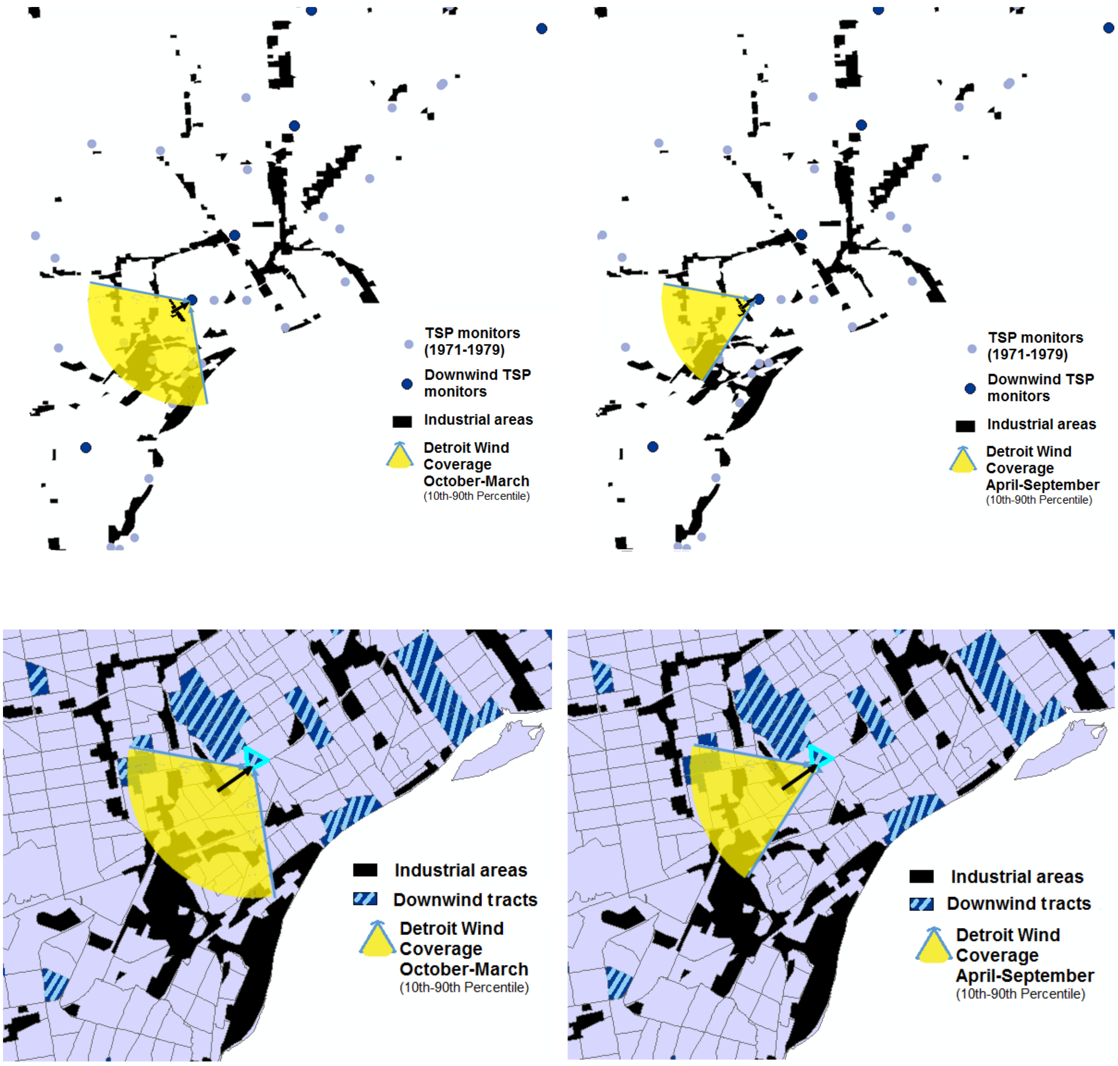
Notes: Policy experiment is cutting TSP levels by 10 $\mu g/m^3$ of either central city or non-central tracts randomly drawn from 155 out of 310 US CBSAs. Central tracts are defined as top quartile tracts in distance to the CBD. Counterfactual percentage changes in total rental income, labour income, total employment and employment in skilled sectors (FIRE, IT and professional services) obtained under a model with only exogenous productivity and amenity are presented in the left panel, while those under a model with agglomeration effects are presented on the right.

Table 1.12: Model Validation

Variable	Housing price		
Source	Model	Chay and Greenstone (2005)	Bento, Freeman and Lang (2014)
Treatment	10 units drop in TSP	10 units drop in TSP	10 units drop in PM10
Unit of obs.	Census tract	County	Census tract
Time period	From 1970	1970-1980	1990-2000
Estimated effects	0.117	0.022-0.036	0.09-0.133
Implied elasticity between pollution and housing value	-1.4	-0.2 to -0.35	-0.6

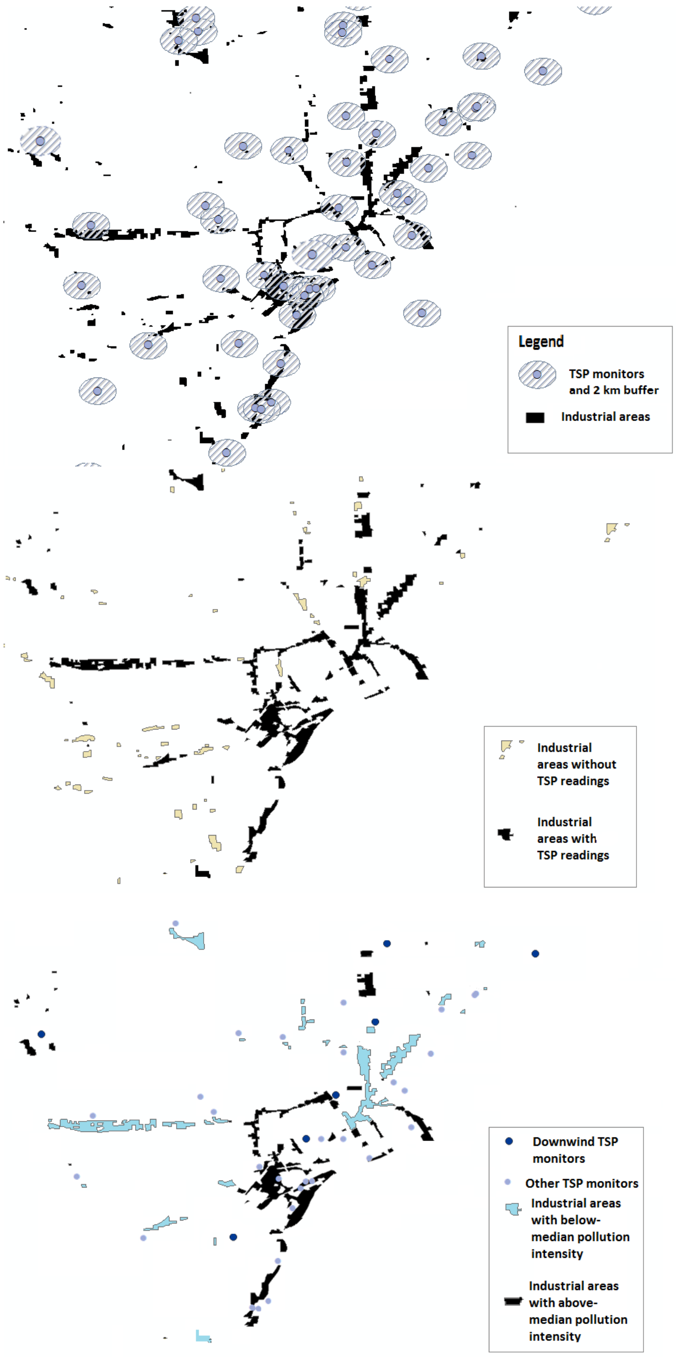
Notes: The first column reports the effects of a 10 units TSP cut on housing price predicted by my structural model. The second and third report estimated reduced form results from similar treatments in the literature. The last column reports the implied elasticity between housing price change and pollutants changes.

Figure 1.1: Definition of downwind



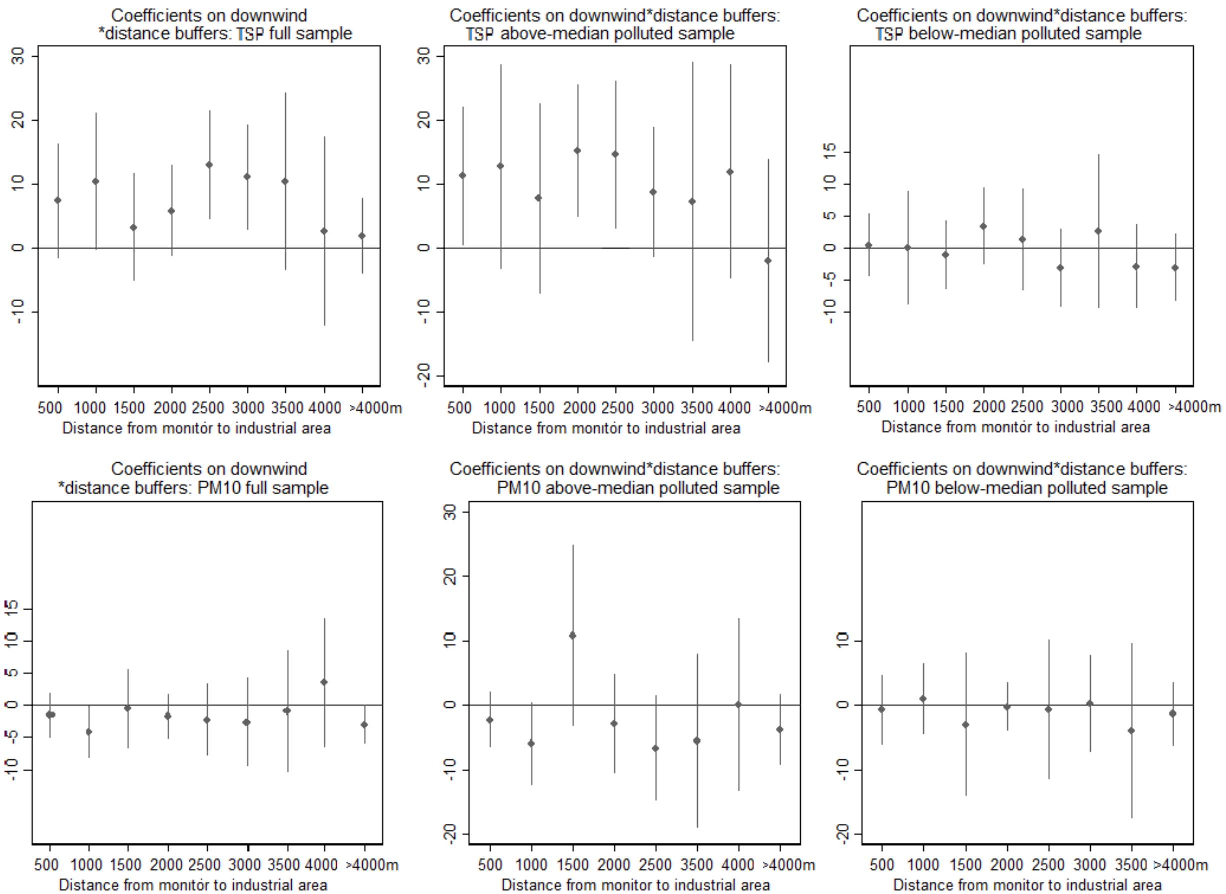
Notes: The upper graph maps the location of 1970 industrial areas (shaded areas) and TSP monitors with at least one year of readings from 1971-1979 around central Detroit. The lower graph shows a map of central Detroit census tracts and the same industrial areas. These graphs illustrate my definition of whether or not a TSP monitor or a census tract is in the downwind direction of its closest industrial area. The solid arrowed line depicts the direction from the closest industrial area to the highlighted census tract. The light shaded area in the left draws the range of summer wind direction around the highlighted monitor/tract. This range is defined by wind directions within the 10th and 90th percentiles of all the observations of April to September monthly wind direction from 2005 to 2014, after dropping months with wind speed lower than 0.5 m/s. The right light shaded area draws the range of winter wind direction, based on October to March monthly wind direction observations. A TSP monitor or a tract is only defined to be downwind of its closest industrial area if its direction (black arrowed line) to the area lies within BOTH wind ranges. In the upper graph, downwind TSP monitors are market as larger, darker dots and in the lower graph, downwind tracts are marked as shaded striped areas.

Figure 1.2: 1970 industrial areas and TSP monitors



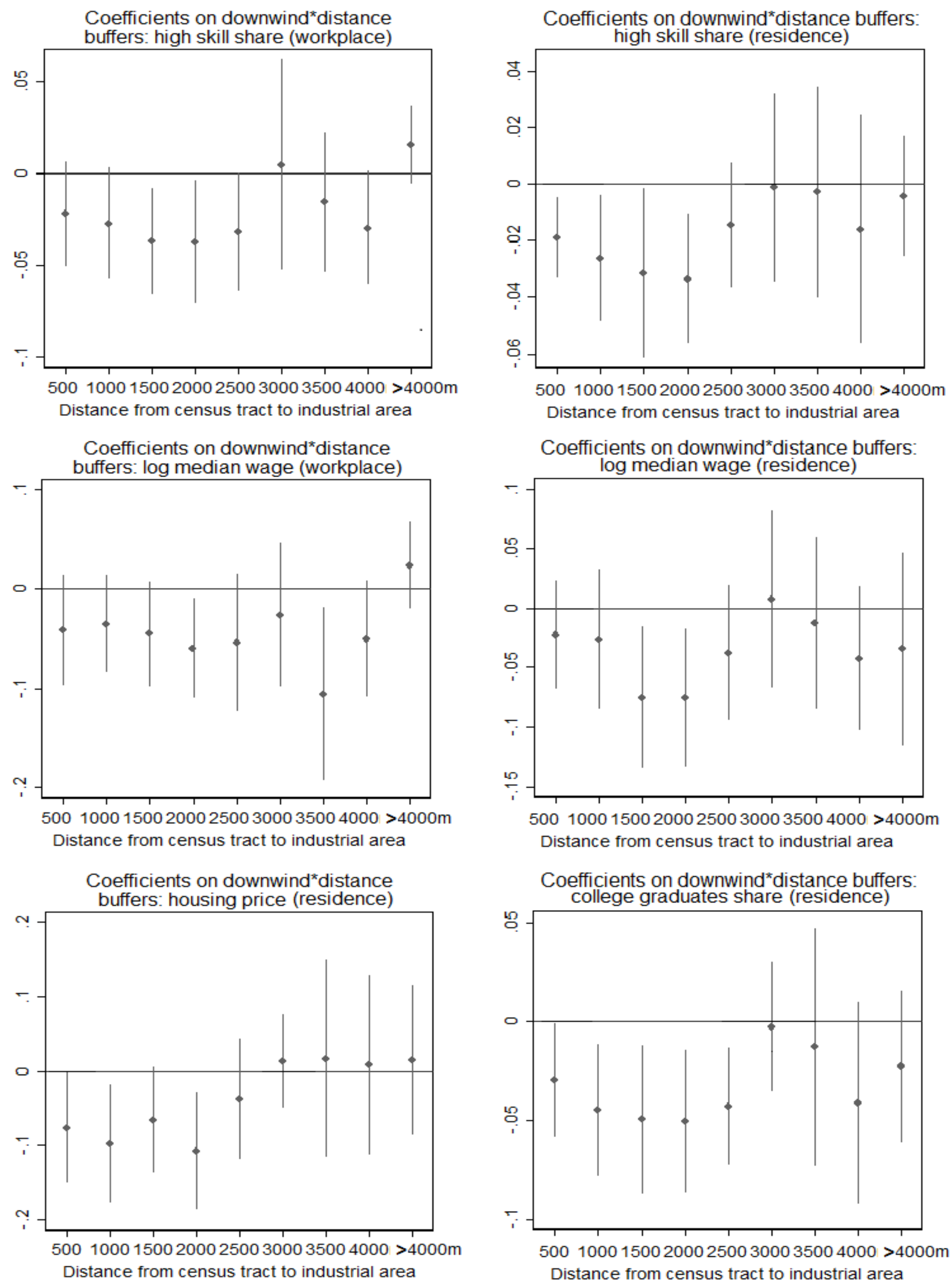
Notes: These graphs illustrate the way the pollution intensity of each industrial area is defined. The top map shows industrial areas and TSP monitors in the 1970s in central Detroit, along with a 2 km buffer around each TSP monitor. In practice, for each industrial area, I define its pollution intensity as the TSP reading of the closest TSP monitor to this area. I only keep industrial areas within 2 km to the closest TSP monitors. The middle map shows industrial areas with (darker) and without pollution intensity assignment. I also divide the industrial areas with pollution intensity assignment into two groups according to whether or not the pollution intensity is above the national median. The full samples of monitors or tracts are also divided into two subsamples according to whether or not they are nearest to above/below-median polluted industrial areas. The bottom map shows the distribution of above and below median polluted industrial areas around central Detroit, with the darker areas denoting above-median polluted ones.

Figure 1.3: Estimated coefficients of downwind on 1970 and 2000 pollution



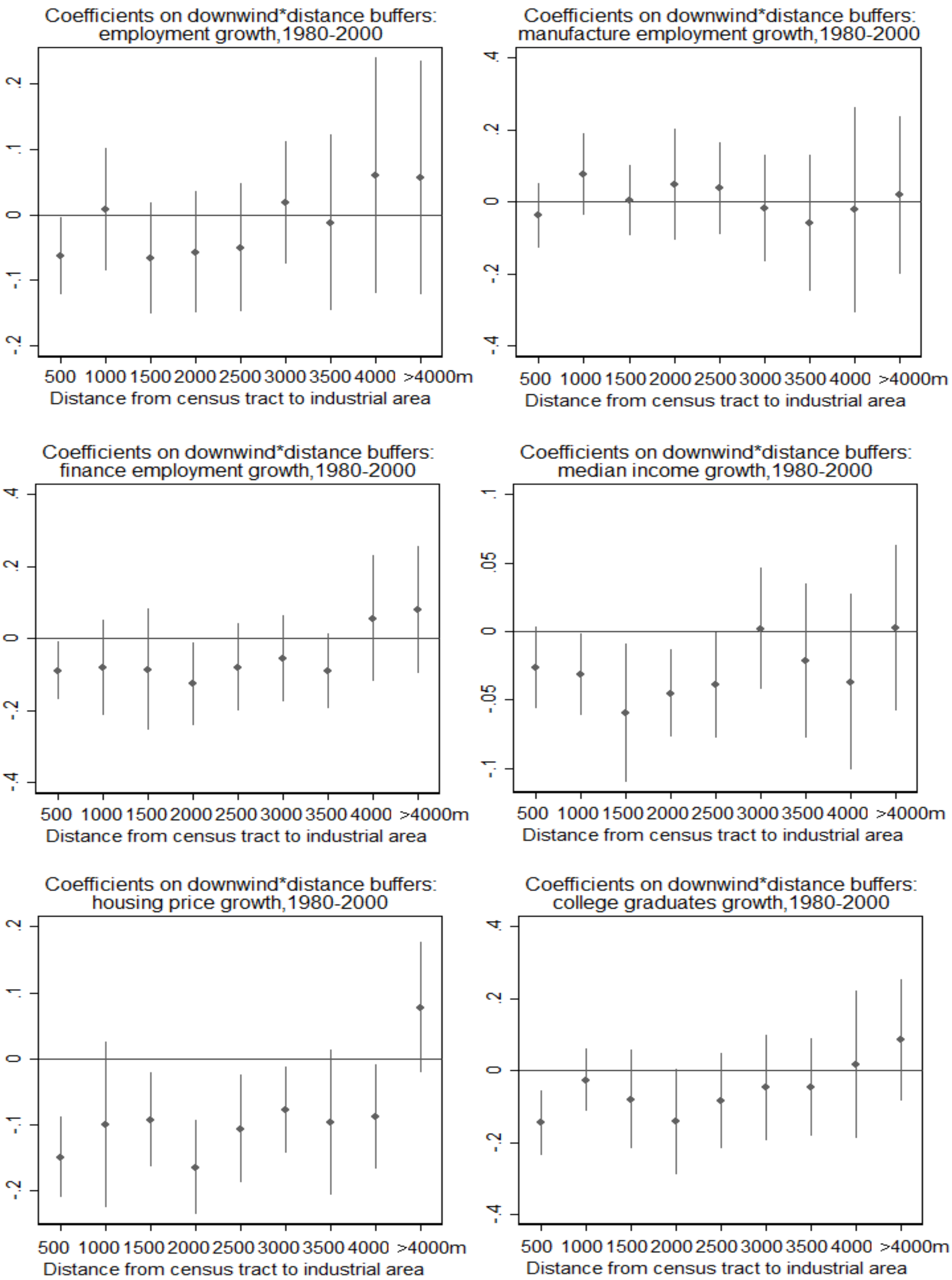
Notes: The three figures above display the estimated coefficients and 95% confidence intervals in regressions where the dependent variable is average TSP ambient concentration at monitor level from 1971 to 1979. The independent variables are dummies of whether or not the distance from each monitor to its closest industrial area falls into specific 500 metres buffers, interacted with a dummy of whether or not the TSP monitor is downwind of the industrial area. The three figures below display the estimated coefficients and 95% confidence intervals in regressions where the dependent variable is average PM10 ambient concentration at monitor level from 2000 to 2010. The independent variables are dummies of whether or not the distance from each monitor to its closest industrial area falls into specific 500 metres buffers, interacted with a dummy of whether or not the PM10 monitor is downwind of the industrial area. The graphs from the left to the right are coefficients obtained in the full sample, a sample with monitors closest to above-median polluted industrial areas and a sample with those closest to below-median polluted industrial areas.

Figure 1.4: Downwind interactions on above-polluted sample: Economic outcomes in 2000



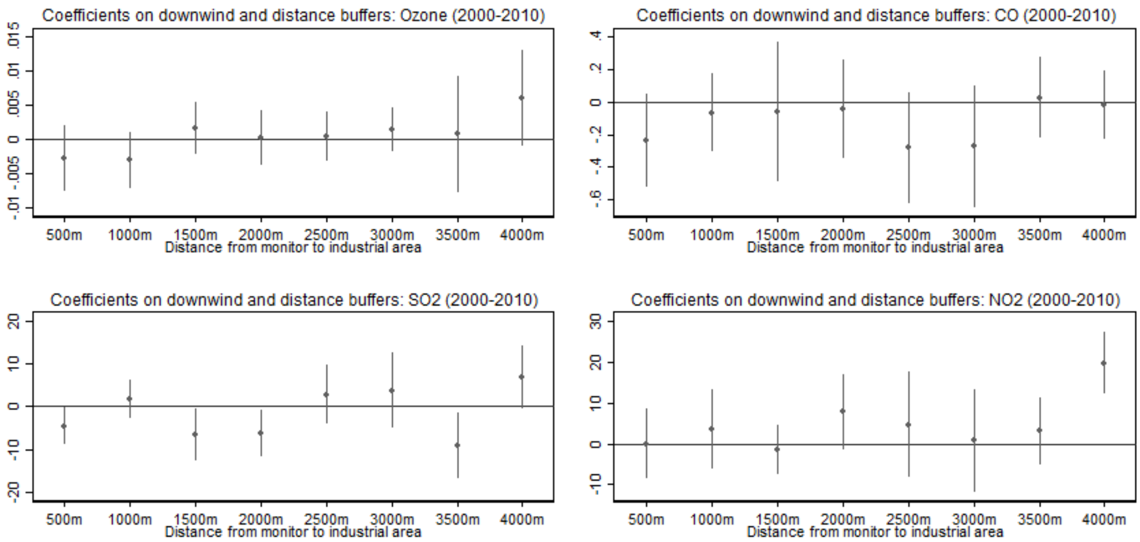
Notes: The figures display the estimated coefficients and 95% confidence intervals in regressions where the dependent variables from top-left to botton-right are high skilled employment share at workplace and residence, median wage at workplace and residence, housing price and college graduates share at residence. The independent variables are dummies of whether or not the distance from each tract to its closest industrial area falls into specific 500 metres buffers, interacted with a dummy of whether or not the tract is downwind of the industrial area. All the coefficients are obtained in regressions on a sample of census tracts that are closest to above-median polluted industrial areas. Those on the subsample of census tract nearest to below-median pollution industrial areas are reported in Figure 1.7.

Figure 1.5: Downwind interactions on above-polluted sample: Dynamic effects



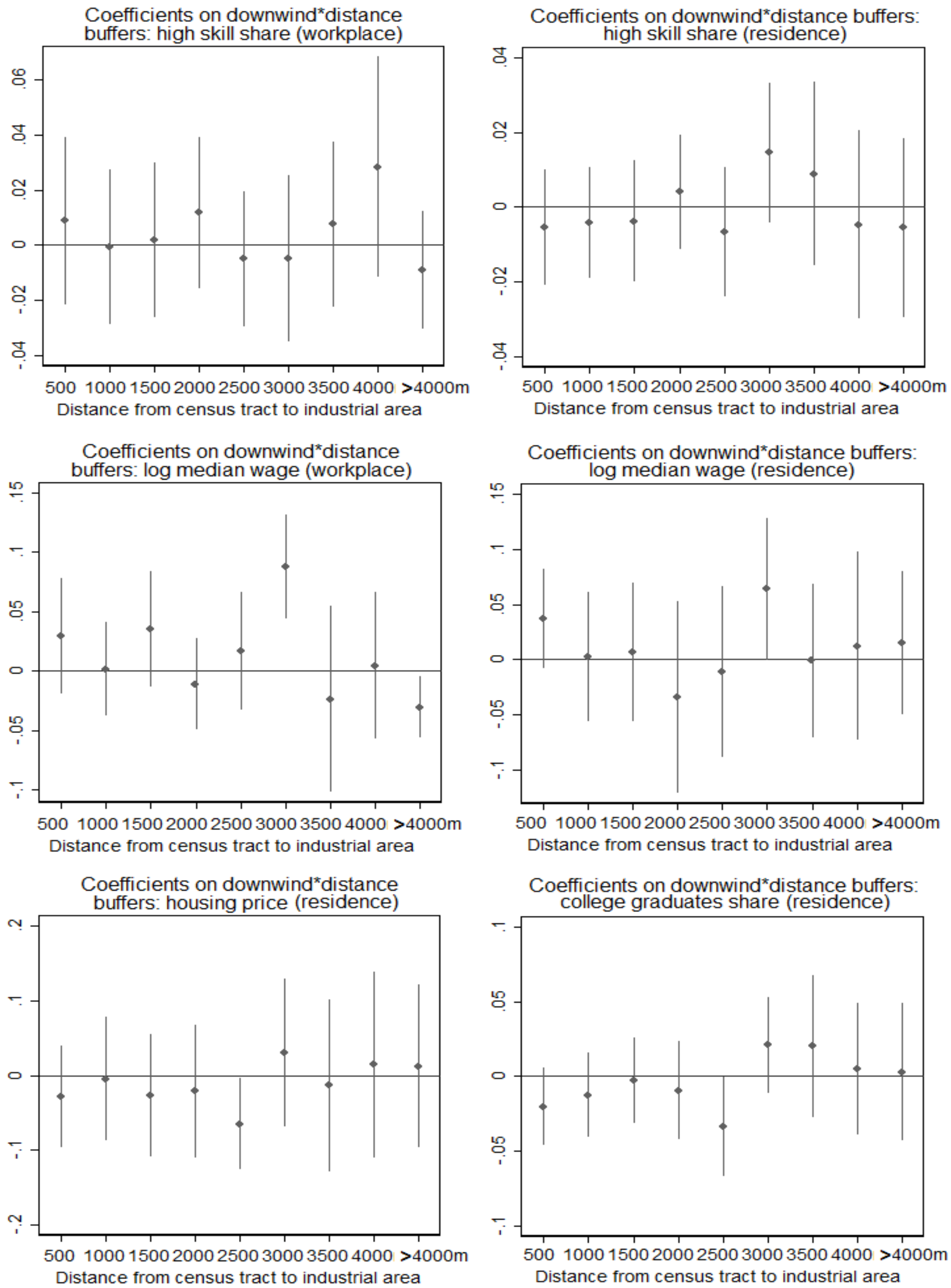
Notes: The figures display the estimated coefficients and 95% confidence intervals in regressions where the dependent variables from top-left to botton-right are growth rates from 1980 to 2000 in total employment, manufacture employment, FIRE employment, median income, housing price and the number of college graduates, all counted at place of residence. The independent variables are dummies of whether or not the distance from each tract to its closest industrial area falls into specific 500 metres buffers, interacted with a dummy of whether or not the tract is downwind of the industrial area. All the coefficients are obtained in regressions on a sample of census tracts that are closest to above-median polluted industrial areas. Those on the subsample of census tract nearest to below-median pollution industrial areas are reported in Figure 1.8.

Figure 1.6: Estimated coefficients of downwind on 2000 pollution: other pollutants



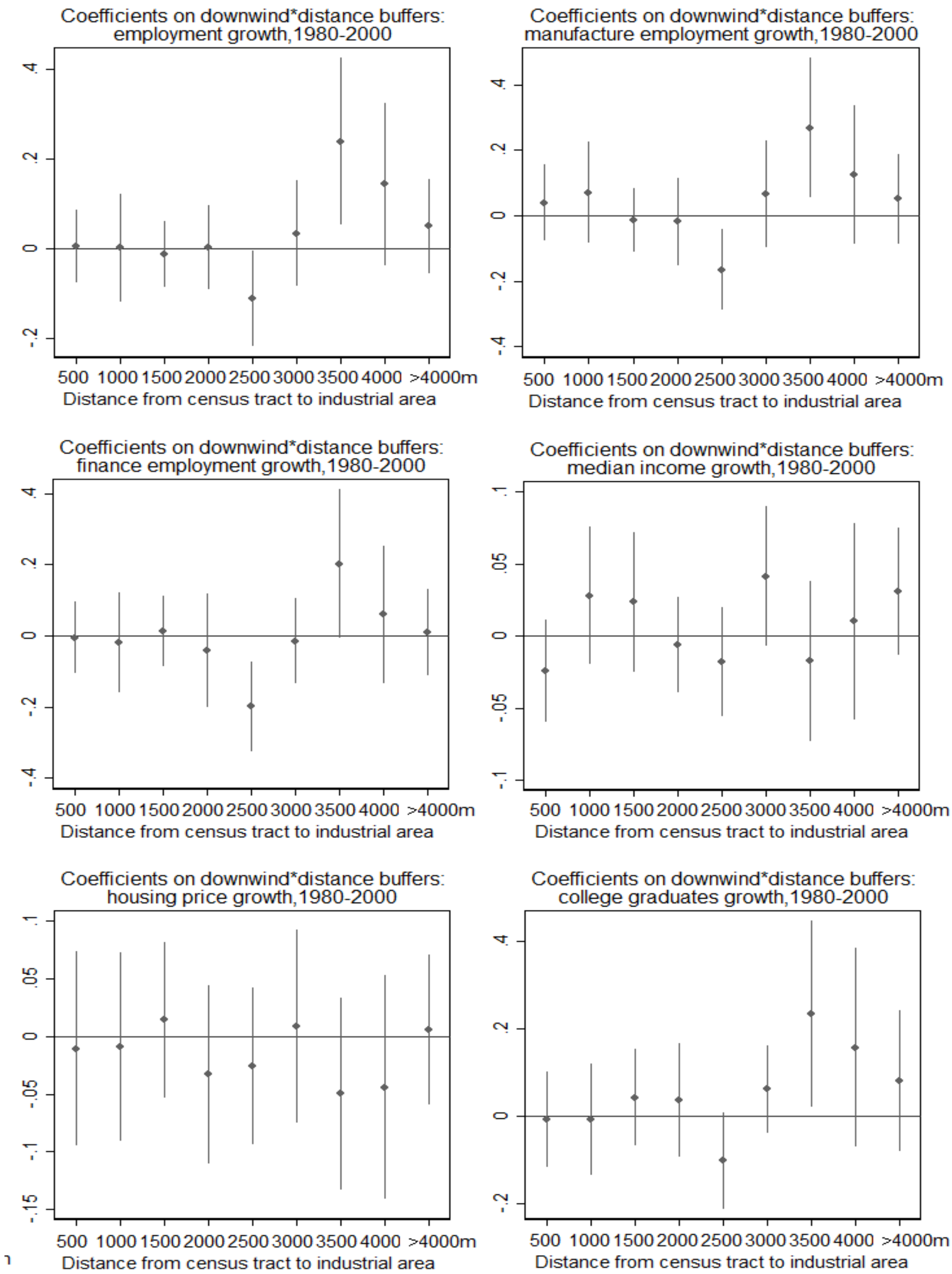
Notes: The figures display the estimated coefficients and 95% confidence intervals in regressions where the dependent variables from top-left to botton-right are high skilled employment share at workplace and residence, median wage at workplace and residence, housing price and college graduates share at residence. The independent variables are dummies of whether or not the distance from each tract to its closest industrial area falls into specific 500 metres buffers, interacted with a dummy of whether or not the tract is downwind of the industrial area. All the coefficients are obtained in regressions on a sample of census tracts that are closest to below-median polluted industrial areas. Those on the subsample of census tract nearest to above-median pollution industrial areas are reported in Figure 1.4.

Figure 1.7: Downwind interactions on below-polluted sample: Economic outcomes in 2000



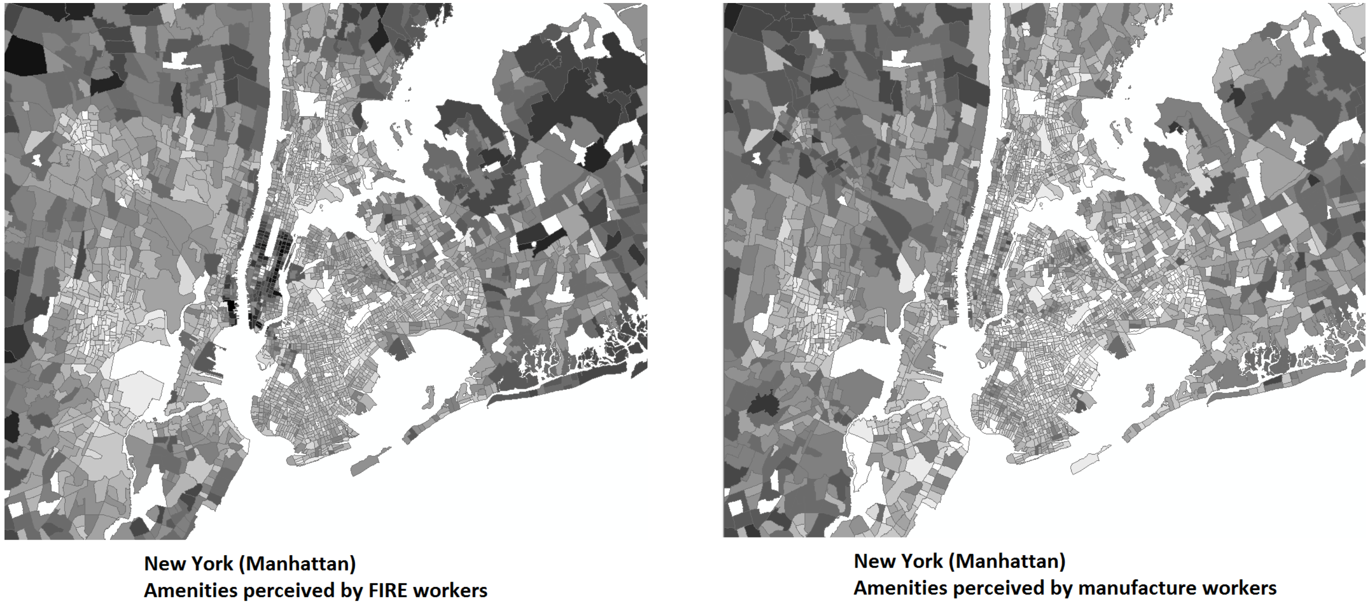
Notes: The figures display the estimated coefficients and 95% confidence intervals in regressions where the dependent variables from top-left to botton-right are average ambient concentrations of Ozone, CO, SO_2 and NO_2 at monitor level from 2000 to 2010. The independent variables are dummies of whether or not the distance from each tract to its closest industrial area falls into specific 500 metres buffers, interacted with a dummy of whether or not the tract is downwind of the industrial area. All the coefficients are obtained in regressions on a sample of census tracts that are closest to below-median polluted industrial areas.

Figure 1.8: Downwind interactions on below-polluted sample: Dynamic effects



Notes: The figures display the estimated coefficients and 95% confidence intervals in regressions where the dependent variables from top-left to botton-right are growth rates from 1980 to 2000 in total employment, manufacture employment, FIRE employment, median income, housing price and the number of college graduates. The independent variables are distance buffers of 500 metres interacted with downwind dummy. All the coefficients are obtained in regressions on a sample of census tracts that are closest to below-median polluted industrial areas. Those on the subsample of census tract nearest to above-median pollution industrial areas are reported in Figure 1.5

Figure 1.9: Perceived quality-of-life by FIRE and manufacture workers around Manhattan, 2010



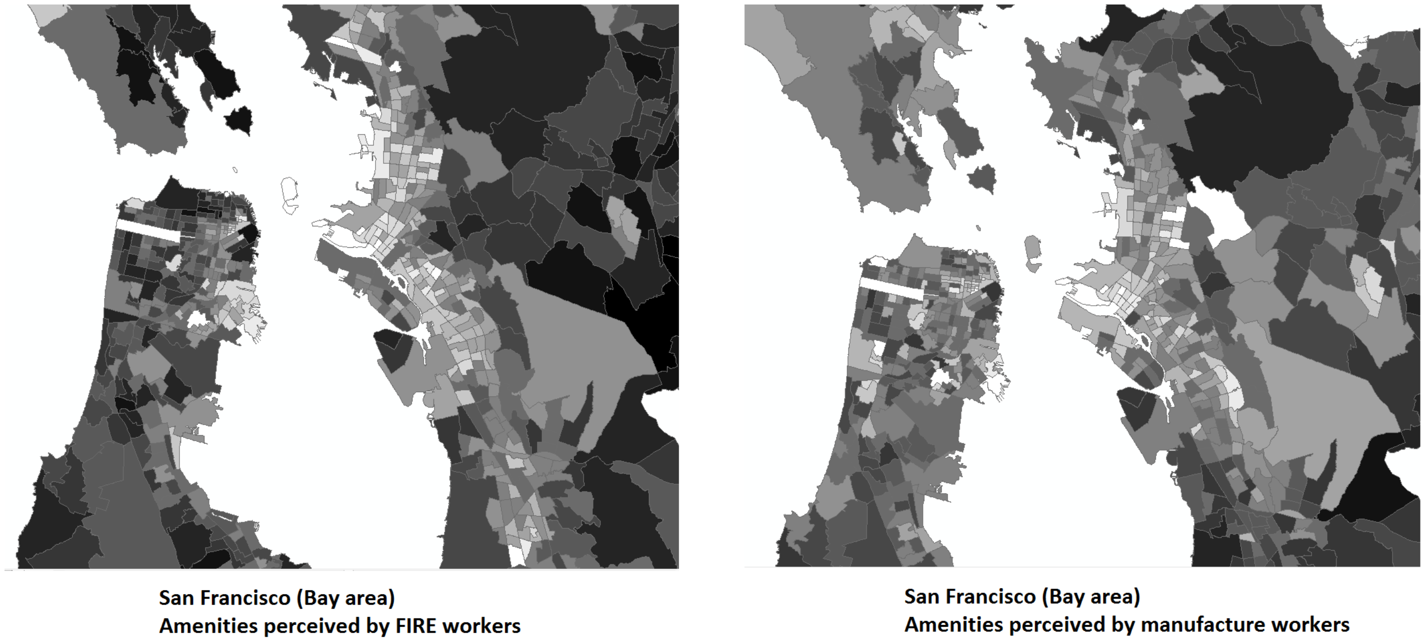
Notes: Figure 1.(A) maps the perceived amenity by FIRE workers in 2000 around Manhattan, New York. Figure 1.(B) maps the perceived amenity by manufacture workers in 2000 in the same area. Perceived amenity is defined in equation (22).

Figure 1.10: Perceived quality-of-life by FIRE and manufacture workers in Detroit, 2000

Detroit: Amenities perceived by FIRE workers **Detroit: Amenities perceived by manufacture workers**

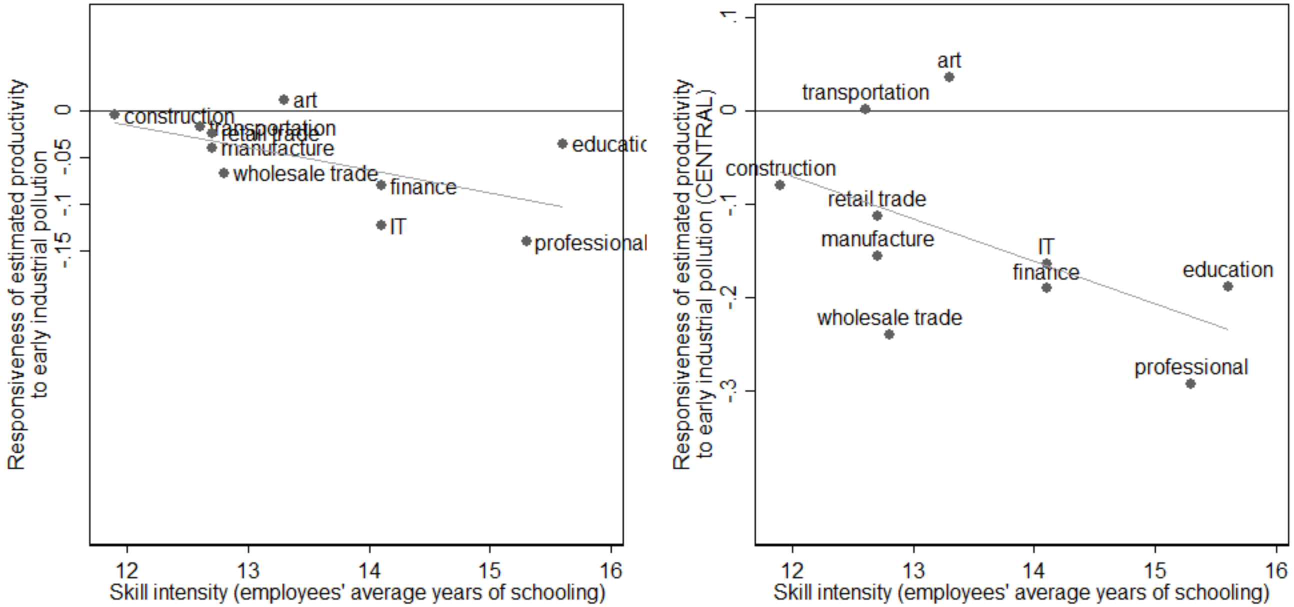
Notes: Figure 2.(A) maps the perceived amenity by FIRE workers in 2010 in Detroit. Figure 2.(B) maps the perceived amenity by manufacture workers in 2000 in Detroit. Perceived amenity is defined in equation (22).

Figure 1.11: Perceived quality-of-life by FIRE and manufacture workers in the San Fransisco Bay Area, 2000



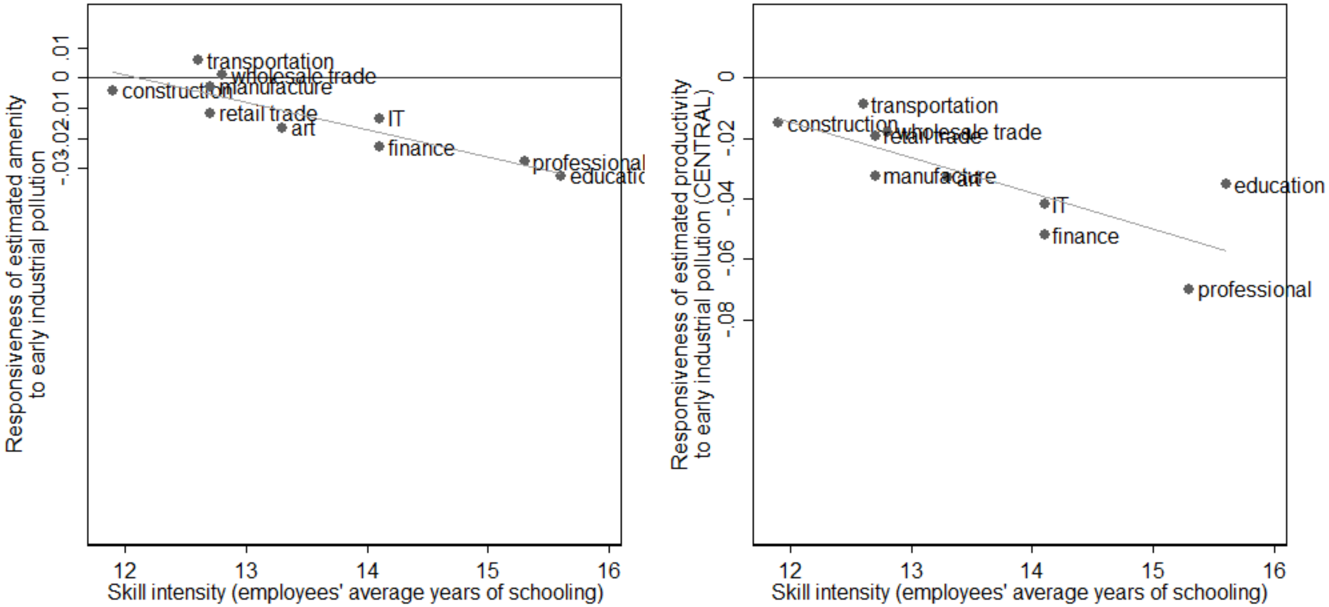
Notes: Figure 3.(A) maps the perceived amenity by FIRE workers in 2010 in the San Fransisco Bay Area. Figure 3.(B) maps the perceived amenity by manufacture workers in 2010 in the San Fransisco Bay Area. Perceived amenity is defined in equation (22).

Figure 1.12: Impacts of early pollution exposure on 2000 estimated sectoral productivity parameters



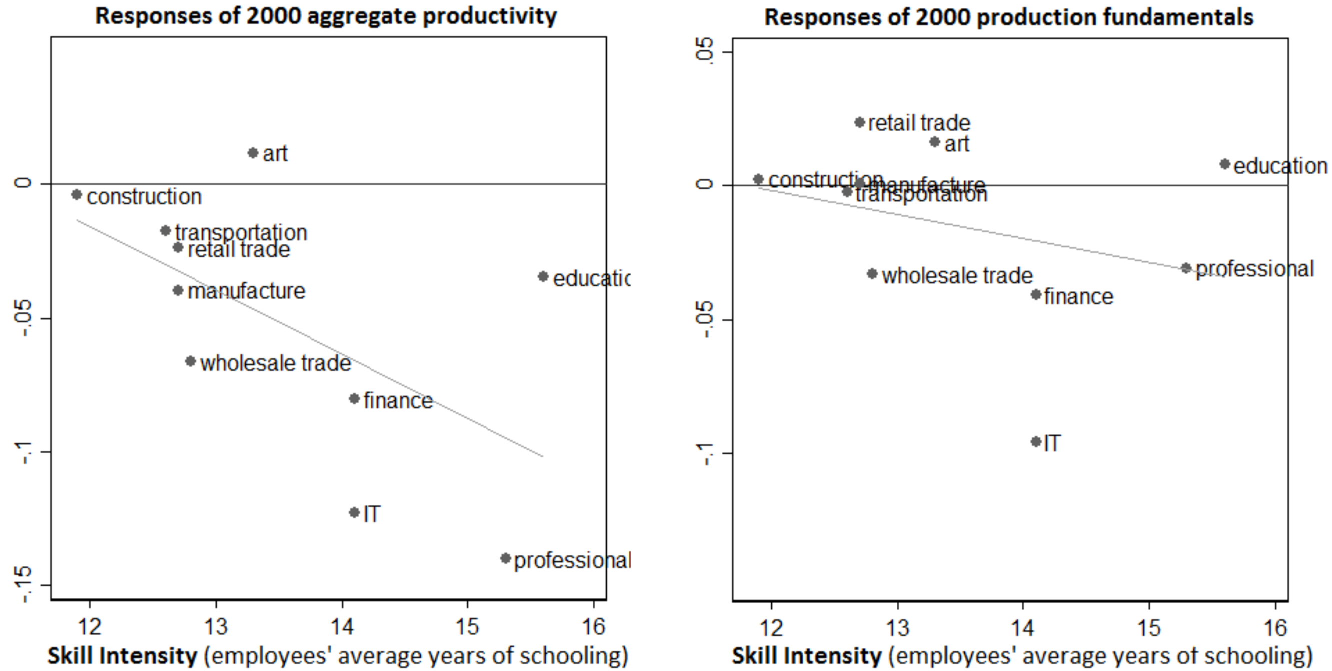
Notes: Figure 1.12(A) plots the impacts of 1970 pollution exposure on sector-specific aggregate productivity A_{js} in 2000 (y-axis) against the skill intensity of these industries (x-axis). The y-axis shows the estimated coefficients of the interaction term between the 0-4 km distance buffer from each tract to the 1970 industrial areas and a dummy of being downwind of these industrial areas, where the LHS variables are sector-specific productivity parameters estimated from the structural model. The sample is limited to census tracts that are closest to industrial areas with pollution intensity above the median. The x-axis plots the skill intensity of each industry proxied by the average years of schooling of their employees. Figure 1.12(B) is the same except that the estimates are obtained in a sample of tracts within the central city.

Figure 1.13: Impacts of early pollution exposure on 2000 estimated sectoral amenity parameters



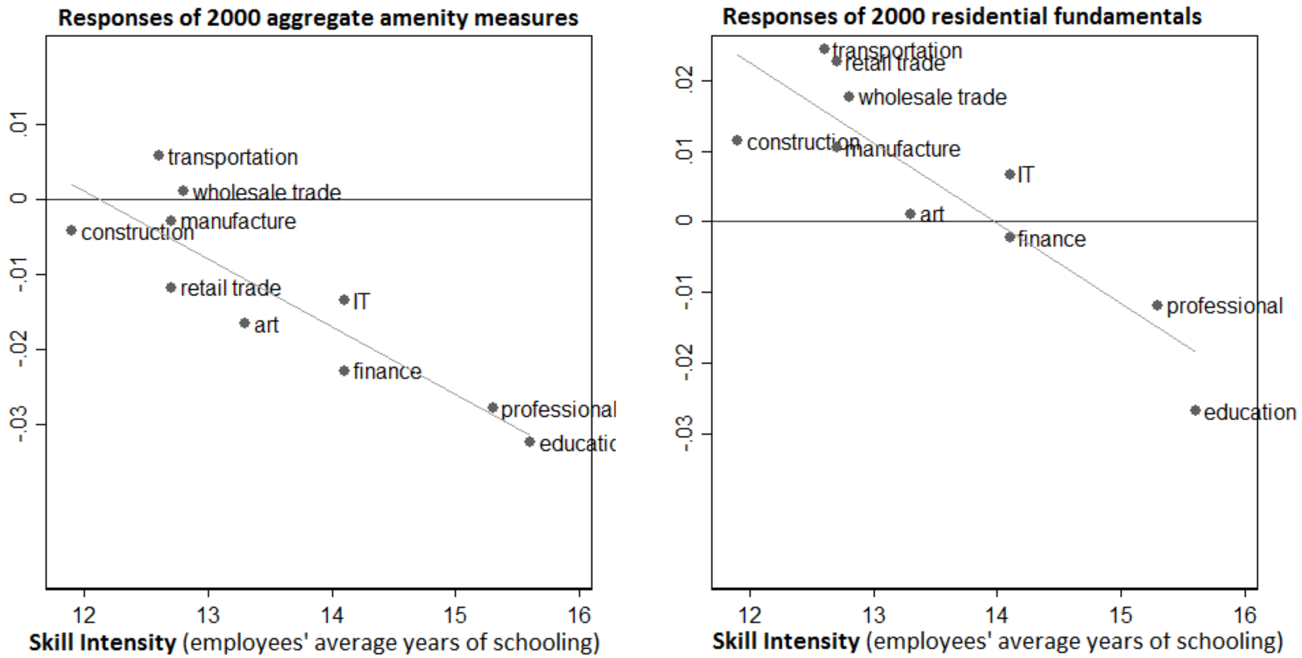
Notes: Figure 1.13(A) plots the impacts of 1970 pollution exposure on sector-specific aggregate amenity ($\tilde{B_{is}}$) in 2000 (y-axis) against the skill intensity of these industries (x-axis). The y-axis shows the estimated coefficients of the interaction term between the 0-4 km distance buffer from each tract to the 1970 industrial areas and a dummy of being downwind of these industrial areas, where the LHS variables are sector-specific amenity parameters estimated from the structural model. The sample is limited to census tracts that are closest to industrial areas with pollution intensity above the median. And the x-axis is the skill intensity of each industry proxied by the average years of schooling of their employees. Figure 1.13(B) is the same except that the estimates are obtained in a sample of tracts within the central city.

Figure 1.14: Impacts of early pollution exposure on 2000 estimated sectoral production fundamental parameters



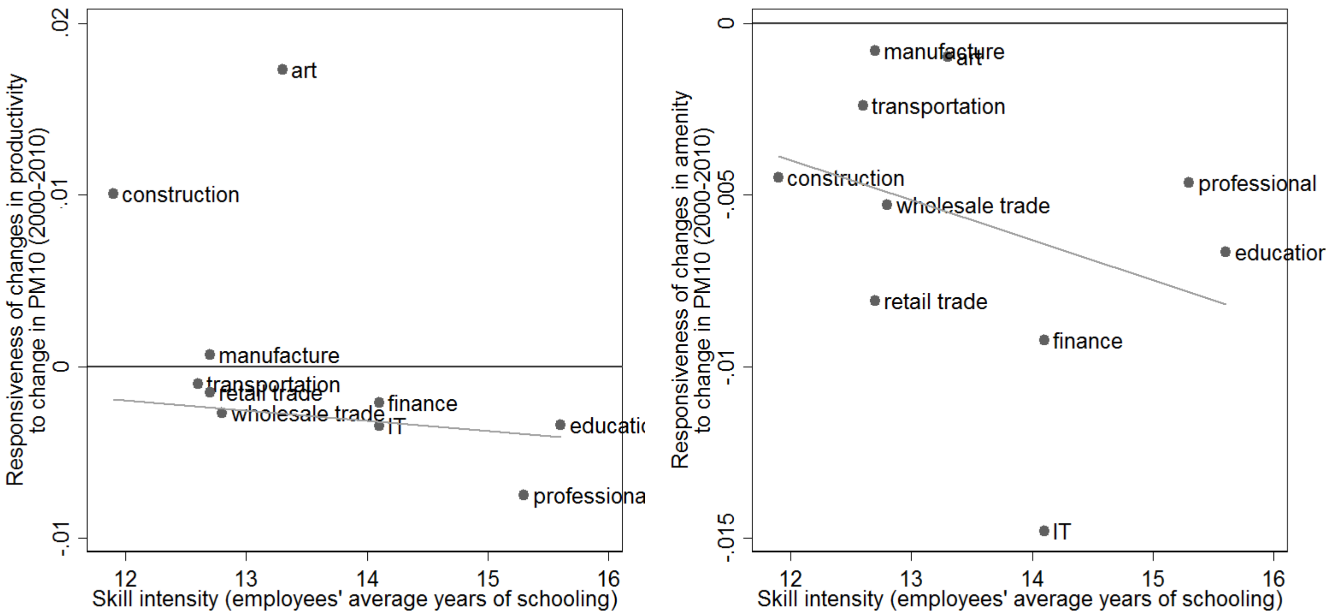
Notes: Figure 1.14(A) plots the impacts of 1970 pollution exposure on sector-specific production fundamentals (a_{js}) in 2000 (y-axis) against the skill intensity of these industries (x-axis). The y-axis shows the estimated coefficients of the interaction term between the 0-4 km distance buffer from each tract to the 1970 industrial areas and a dummy of being downwind of these industrial areas, where the LHS variables are sector-specific production fundamental parameters estimated from my structural model. Production fundamentals are defined as aggregate productivity divided by a workplace agglomeration intensity function. The sample is limited to census tracts that are closest to industrial areas with pollution intensity above the median. The x-axis plots the skill intensity of each industry proxied by the average years of schooling of their employees. Figure 1.14(B) presents similar coefficients with aggregate productivity estimates as LHS variables for comparison, the same as Figure 1.12(A).

Figure 1.15: Impacts of early pollution exposure on 2000 estimated sectoral residential fundamental parameters



Notes: Figure 1.15(A) maps the impacts of 1970 pollution exposure on sector-specific residential fundamentals ($\tilde{b_{is}}$) in 2000 to the skill intensity of these industries. The y-axis shows the estimated coefficients of the interaction term of 0-4 km distance buffer to the 1970 industrial areas and a dummy of being downwind of them where the LHS variables are sector-specific residential fundamental parameters estimated from my structural model. Residential fundamentals are defined as aggregate amenity divided by a residence agglomeration intensity function. The sample is limited to census tracts that are closest to industrial areas with pollution intensity above the median. The x-axis plots the skill intensity of each industry proxied by the average years of schooling of their employees. Figure 1.15(B) presents similar coefficients with aggregate amenity estimates as LHS variables for comparison, the same as Figure 1.13(A).

Figure 1.16: Responses of changes in productivity and amenity to PM10 changes (2000-2010)



Notes: Figure 1.16(A) maps the impacts of pollution reduction on sector-specific aggregate productivity changes from 2000 to 2010 to the skill intensity of these industries. The y-axis shows the estimated coefficients of the regression specified in equation (25), where the LHS variables are sector-specific productivity parameters changes and RHS variable is the instrumented PM10 change from 2000-2010. And the x-axis plots the skill intensity of each industry proxied by the average years of schooling of their employees. Figure 1.16(B) is the same except that the triple difference estimates are obtained in a sample of tracts within the central city.

Appendices

Appendix A

Appendix of "The Long Shadow of Industrial Pollution: Environmental Amenities and the Distribution of Skills"

A.1 Exploiting variation in topography

As briefly discussed in Section 4.1 and 4.2, apart from exploiting the variation in wind direction to identify the impact of historical pollution on current economic outcomes, we can also consider the role of topography in explaining pollution exposure. Significant elevation gaps between the locations of pollution sources and receptors will prevent pollutants from being transported. To test it, I estimate a model similar to Equation (3), taking the form

$$y_{ic} = \sum \mathrm{I_{ikm}} * SameElevation_{im}\beta_{k_1} + \sum \mathrm{I_{ikm}} * \beta_{k_2} X_i' \delta + \alpha_c + \epsilon_{ic} \tag{A.1}$$

where $SameElevation_{im}$ is an indicator of whether or not the elevations of monitor i and its closest industrial area m are the same. Elevation data come from the National Elevation Dataset (NED), available at one arc-second resolution (approximately 30 metres) for the continental US. It is noted that the emission of airborne industrial pollutants usually comes from tall smokestacks, so it is essential to account for the height of stacks. In practice, I define $SameElevation_{im}$ to be one if the monitor is at the same elevation or less than 100 metres higher than the industrial area it is exposed to, based on the idea that the height of a typical smokestack is lower than

100 metres. Moreover, apart from the elevation difference between monitor and industrial area, the elevation of areas between them also matters. The transportation of pollutants will be blocked if the tracts located in between the pollution source and receptor are higher in elevation than both. To account for this, I draw a straight line between a monitor and its closest industrial area and force $SameElevation_{im}$ to be zero if the maximum elevation of areas covered by this straight line exceeds the elevations at both ends. y_{ic} represents the outcomes of interest, which include pollution in the 1970s and after 2000, 2000 economic outcomes and growth rates from 1980 to 2000.

The results on historical and current pollution are reported in Table A1. It is clear that monitors that are not obstructed from the closest pollution source by terrain report higher TSP readings, and the effects are stronger at mid-to-long distance range. The TSP readings in unobstructed monitors are about 4-5 units higher when they are within 2-4 kilometres away from the closest industrial area than their obstructed counterparts. Similar to the results obtained using wind direction variation, coefficients on the interaction terms between distance buffers and the same elevation dummy are also larger in the sample of monitors closest to above-median industrial areas, as reported in Columns (2) and (3). The last three columns present results with PM10 from 2000 to 2010 as the outcome variables, and it is observed that the coefficients on the interaction terms remain small and insignificant in all three samples.

To examine the impacts of topography-driven variation in historical pollution on economic outcomes, the identification assumption here is that in the absence of industrial pollution, tracts that are at the same elevation as its nearby industrial area and those with significant elevation gaps from the same industrial area are similar in economic outcomes and growth prospects. A reasonable challenge here is that elevation gaps not only act as obstacles to the transmission of pollution but also weakens the economic linkages between industrial areas and nearby tracts. To check how big a concern is might be, I check the correlation between the route distance[1] to the closet industrial area and the same elevation dummy. It appears that elevation gaps do not make the nearby industrial areas significantly more difficult to access from each tract. Throughout my analysis, I control for route distance buffers to the same industrial area m, and a local ruggedness measure.

In Table A2, I report the results on 2000 economic outcomes. The patterns are largely consistent with those uncovered using the downwind variation. Census tracts that were more exposed to historical pollution due to a similar elevation to the

[1]Shortest distance to the closest industrial area by highway or railway.

closest industrial areas have lower housing value, the share of college graduates and median residents' income in 2000, and the effects are only statistically significant in a sample with tracts exposed to above-median polluted industrial areas. However, the patterns are less clear for workplace employment outcomes. One possible explanation is that the specification that exploits variation in elevation difference still suffers from biases driven by unobservable differential economic linkages to industrial areas. In this case, it is likely that stronger economic links to industrial areas lead to a higher wage, therefore the potential bias work against identifying the full effects of pollution on productivity.

Results on the dynamic effects are reported in Table A3. Similar to the results obtained using wind direction variation, census tracts that were more polluted in the 1970s due to topography characteristics witnessed slower growth in median income and college graduates from 1980 to 2000, although the effects are less apparent for other variables. Overall, this exercise of using elevation gaps in identifying the impacts of historical pollution on current economic outcomes are consistent with the main results using the wind variation.

A.2 Triple difference approach

A.2.1 Reduced form evidence

In the main text, I explore the variation in historical pollution driven by wind patterns by interacting a dummy of downwind with distance indicators. I also exploit the variation in the pollution intensity of historical industrial sites by splitting the sample into two according to the pollution intensity of industrial areas that each census tract is closest to, and find wind only matters for monitors/census tracts that are closest to industrial areas with above-median pollution intensity. Another way to exploit both the variation in pollution intensity and wind direction is a triple difference design. The main outcomes of interest are the triple-interaction terms across distance-to-industrial-areas indicators, the level of historical pollution of these areas and a dummy of downwind. The specification takes the form of:

$$y_{ic} = \sum \mathrm{I}_{\mathrm{ikm}} * TSP_m * Downwind_{im}\beta_{k_1} + \sum \mathrm{I}_{\mathrm{ikm}} * Downwind_{im}\beta_{k_2} + \sum \mathrm{I}_{\mathrm{ikm}} * \beta_{k_3} + \sum \mathrm{I}_{\mathrm{ikm}} * TSP_m\beta_{k_4} + \theta TSP_m + \gamma Downwind_{im} + \delta TSP_m * Downwind_{im} + X_i'\eta + \alpha_c + \epsilon_{ic} \tag{A.2}$$

where TSP_m is the average measure of total suspended particle reported by EPA from 1971-1979 around industrial area m, $Downwind_{im}$ is a dummy variable that takes value one if tract i is located in the downwind of industrial area m. Similarly, I assign TSP readings of the closest monitor averaged from 1971 to 1979 to industrial area m as the pollution intensity of it (TSP_m) I drop industrial areas that are not within 2 kilometres to the closest monitor, which takes up about 30% in my sample. I standardize the TSP measures to be of mean zero and standard deviation one for ease of interpretation.

From Table A9, It is clear that being more exposed to 1970s industrial pollution due to downwind position translates into higher 1970s pollution levels measured by average TSP reading from 1971-1979. The triple difference terms are negative across the board for most of my outcomes, apart from residential density. The patterns are quite comparable across variables counted at place of work and place of residence: early industrial pollution not only negatively affects the share of high-skilled workers who live in more exposed areas and their earnings, but also the share of skilled employees who work in more affected tracts. Median earnings of both workers who work and residents who live in dirtier tracts are lower. Housing prices and the share of college graduates are also lower in these tracts. It suggests that historically dirtier places are low in both productivity and amenity now. The coefficients on double difference terms of distance buffers to the closest industrial area and its historical pollution are negative and significant for most of the outcomes. But the signs and significance are mixed for coefficients on double difference terms of distance buffers and downwind dummy, which suggests that whatever the wind conditions are, being closer to an industrial area hurts, but being downwind of an industrial area only matters if this area is polluted enough.

Table A10 report results on growth from 1980 to 2000, following the same specifications. It is clear that tracts closer to 1970 industrial areas experience slower growth in total, manufacture, FIRE employment, median income, housing value and the number of college graduates in the subsequent two decades, and these negative growth effects are stronger if the relevant industrial areas are more polluted prior to 1980, and the tracts are more exposed to the pollutants emitted from these heavily-polluted industrial areas due to being downwind of them. Since from 1980 onward the air quality around the US is improving greatly, and the improvement is greater in areas that are more heavily polluted initially, the negative growth impact is not caused by worsening air pollution, but more of a result of self-reinforcing agglomeration forces. Tracts that have been able to attract more educated workforce and residents due to better environmental amenities are able to attract more if educated people would like to live and work near other educated people.

Table A1: Industrial Activities in the 1970s and 1970 Pollution

Variables	Monitor-level TSP: 1971-1979			Monitor-level PM10: 2000-2010		
Sample	Full Sample	Above-median polluted industrial area	Below-median polluted industrial area	Full Sample	Above-median polluted industrial area	Below-median polluted industrial area
$1(disind \in 0-1km)$	26.85***	22.36***	13.41***	7.159**	9.808**	2.899
	(2.528)	(1.876)	(1.588)	(2.980)	(4.125)	(3.520)
$1(disind \in 1-2km)$	14.79***	11.85***	8.491***	2.400	5.450**	0.154
	(2.922)	(4.352)	(1.760)	(2.385)	(2.230)	(3.463)
$1(disind \in 2-3km)$	6.910***	5.333	6.040***	2.679	0.158	-1.492
	(2.967)	(5.618)	(1.500)	(2.159)	(3.109)	(2.488)
$1(disind \in 3-4km)$	8.054***	1.323	8.837***	4.442**	4.917	3.523**
	(3.092)	(5.061)	(1.639)	(2.183)	(5.408)	(1.769)
$1(disind \in 0-1km)$*Same elevation	1.056	0.337	-0.575	-0.991	-3.664	2.097
	(2.468)	(3.660)	(1.408)	(2.858)	(3.881)	(2.452)
$1(disind \in 1-2km)$*Same elevation	1.637	0.869	0.622	2.724	0.103	1.319
	(2.793)	(3.949)	(1.422)	(2.536)	(2.255)	(3.598)
$1(disind \in 2-3km)$*Same elevation	4.732*	2.450	1.336	2.445	4.289	2.553
	(2.796)	(5.084)	(1.503)	(2.128)	(2.934)	(2.664)
$1(disind \in 3-4km)$*Same elevation	4.473	9.392**	-2.812	0.863	5.817	-4.458*
	(3.388)	(4.716)	(2.003)	(2.958)	(6.686)	(2.425)
Same elevation	5.629***	5.152	2.550**	5.206***	6.203**	3.045**
	(2.008)	(3.684)	(1.242)	(1.227)	(2.596)	(1.416)
Observations	4,968	2,471	2,422	1,893	821	842
R-squared				0.312	0.521	0.435
p-value (Sum of interactions=0)	0.0757	0.0213	0.5604	0.3421	0.8766	0.9539
p-value (Sum of elevation interacts=0)	0.0869	0.2029	0.5972	0.3069	0.4226	0.8452

Notes: Dependent variables are the average measure of TSP ambient concentration from 1970 to 1979 collected at each TSP monitor with positive reading during this period. $1(disind \in a-bkm)$ is an indicator of the distance of a TSP monitor from the closest 1970s industrial area being within a and b km. Downwind is an indicator of whether or not the TSP monitor is located downwind of the industrial area. Same elevation is a dummy that takes one if the TSP monitor is at the same elevation or less than 100m lower than its closest industrial area, and not obstructed by other areas in between. Controls include CBSA fixed effects, the distance from each monitor to transportation lines, natural amenities and the CBD. Robust standard errors are clustered at CBSA level.

Table A2: Economic outcomes in 2000: Same elevation dummy

	Outcomes in 2000: Above-median pollution level							
	By place of work			By place of residence				
VARIABLES	Employ density	Highskill ratio	Median earnings	Employ density	Highskill ratio	Median earnings	Housing value	Coll ratio
$1(disind \in 0-1km)$*same elevation	0.654***	0.00869	-0.0169	0.586***	-0.00737	-0.0982***	-0.148***	-0.04
	(0.161)	(0.00736)	(0.0190)	(0.140)	(0.00753)	(0.0276)	(0.0407)	(0.0
$1(disind \in 1-2km)$*same elevation	0.415***	0.0167**	-0.00766	0.358***	-0.00211	-0.0328	-0.0649	-0.02
	(0.157)	(0.00680)	(0.0215)	(0.123)	(0.00567)	(0.0301)	(0.0498)	(0.0
$1(disind \in 2-3km)$*same elevation	0.463**	-0.00397	0.000573	0.426***	-0.00493	-0.0613	-0.0496	-0.02
	(0.180)	(0.0112)	(0.0284)	(0.145)	(0.00853)	(0.0513)	(0.0503)	(0.0
$1(disind \in 3-4km)$*same elevation	0.144	-0.0261	-0.0252	0.0818	-0.00868	-0.0916**	-0.145***	-0.04
	(0.178)	(0.0264)	(0.0309)	(0.184)	(0.0120)	(0.0400)	(0.0472)	(0.0
$1(disind > 4km)$*same elevation	0.345**	0.00589	-0.0540**	0.345**	-0.00260	0.0324	0.0478	0.01
	(0.157)	(0.00813)	(0.0251)	(0.151)	(0.00882)	(0.0406)	(0.0332)	(0.0
$1(disind \in 0-1km)$	-0.344	-0.0417**	-0.00219	-0.298	-0.0277***	-0.0609	-0.0732	-0.05
	(0.249)	(0.0161)	(0.0291)	(0.257)	(0.00878)	(0.0470)	(0.0743)	(0.0
$1(disind \in 1-2km)$	-0.0571	-0.0417***	-0.0348	0.0424	-0.0212**	-0.0873	-0.0857	-0.04
	(0.218)	(0.0149)	(0.0305)	(0.202)	(0.00818)	(0.0577)	(0.0772)	(0.0
$1(disind \in 2-3km)$	-0.174	-0.0153	-0.0633*	-0.0281	-0.0229**	-0.0342	-0.0444	-0.02
	(0.221)	(0.0170)	(0.0331)	(0.188)	(0.0106)	(0.0767)	(0.0640)	(0.0
$1(disind \in 3-4km)$	-0.120	0.00414	0.0111	-0.0281	0.00132	0.0278	0.0665	0.01
	(0.209)	(0.00934)	(0.0287)	(0.166)	(0.00584)	(0.0361)	(0.0510)	(0.0
Observations	8,255	8,255	8,111	8,282	8,282	8,107	7,997	8,25
R-squared	0.401	0.060	0.191	0.583	0.273	0.246	0.550	0.16
p-value (Sum of elevation interacts=0)	0.0002	0.7616	0.2014	0.0002	0.1759	0.0113	0.0189	0.04

	Outcomes in 2000: Below-median pollution level							
	By place of work			By place of residence				
VARIABLES	Employ density	Highskill ratio	Median earnings	Employ density	Highskill ratio	Median earnings	Housing value	Coll ratio
$1(disind \in 0-1km)$*same elevation	-0.0974	-0.00862	-0.0113	-0.161	-0.00463	-0.0158	-0.102	-0.0
	(0.190)	(0.0160)	(0.0301)	(0.199)	(0.0214)	(0.0635)	(0.0909)	(0.04
$1(disind \in 1-2km)$*same elevation	-0.143	-0.0435**	-0.0373	-0.252	-0.000998	-0.0302	-0.0747	-0.02
	(0.191)	(0.0189)	(0.0316)	(0.155)	(0.0173)	(0.0512)	(0.0731)	(0.0
$1(disind \in 2-3km)$*same elevation	-0.234	-0.0262	-0.0245	-0.189	-0.00997	0.00329	-0.0269	-0.0
	(0.199)	(0.0165)	(0.0419)	(0.149)	(0.0123)	(0.0619)	(0.0694)	(0.0
$1(disind \in 3-4km)$*same elevation	-0.293	-0.0320*	-0.0308	-0.239	-0.0122	-0.0368	-0.0995	-0.0
	(0.202)	(0.0173)	(0.0341)	(0.193)	(0.0180)	(0.0768)	(0.0833)	(0.0
$1(disind > 4km)$*same elevation	0.942***	0.0310**	0.0265	0.898***	0.00551	-0.0498	0.0187	-0.00
	(0.158)	(0.0128)	(0.0165)	(0.139)	(0.0155)	(0.0513)	(0.0629)	(0.0
$1(disind \in 0-1km)$	0.748***	0.000832	0.0201	0.543***	-0.0241	-0.115*	-0.0928	-0.04
	(0.208)	(0.0201)	(0.0332)	(0.205)	(0.0233)	(0.0669)	(0.0984)	(0.04
$1(disind \in 1-2km)$	0.662***	0.0354*	0.0214	0.641***	-0.0192	-0.0785	-0.0800	-0.0
	(0.199)	(0.0200)	(0.0312)	(0.160)	(0.0179)	(0.0520)	(0.0743)	(0.0
$1(disind \in 2-3km)$	0.564***	0.0252	0.0252	0.435***	-0.00150	-0.0685	-0.0592	-0.00
	(0.175)	(0.0153)	(0.0379)	(0.139)	(0.0113)	(0.0621)	(0.0633)	(0.0
$1(disind \in 3-4km)$	0.534***	0.0317**	0.0226	0.442**	0.0114	0.0252	0.0768	0.00
	(0.195)	(0.0152)	(0.0341)	(0.185)	(0.0163)	(0.0727)	(0.0749)	(0.0
Observations	7,870	7,870	7,371	7,869	7,869	7,367	7,803	7,87
R-squared	0.373	0.111	0.287	0.451	0.330	0.351	0.562	0.27
p-value (Sum of elevation interacts=0)	0.5421	0.0666	0.3677	0.8676	0.6487	0.2519	0.4629	0.36

Notes: Results from the upper panel are obtained from a sample of census tracts that are closest to 1970 industrial areas with above-median pollution intensity measure, and those from the lower panel from those closest to industrial areas with below-median pollution level. Dependent variables of the first three columns are employment density, the ratio of FIRE, IT and professional services in total employment, and median earnings, counted at the place of work. The last five columns report results on employment density, the ratio of FIRE, IT and professional services in total employment, median earnings, median housing value and ratio of college graduates among the population over 25 years old, counted at their place of residence. $1(disind \in x-ykm)$ is an indicator of whether or not the distance from a tract to its closest industrial area is within x to y km. Same elevation is a dummy that takes one if the PM10 monitor is at the same elevation or less than 100m lower than its closest industrial area, and not obstructed by other areas in between. Controls include CBSA fixed effects, the distance from each monitor to transportation lines, natural amenities and the CBD. Robust standard errors are clustered at CBSA level.

Table A3: Growth from 1980-2000: Same elevation dummy

	Growth from 1980 to 2000: Above-median pollution level					
VARIABLES	Employment	Manufacture employment	FIRE employment	Median income	Housing value	College graduates
$1(disind \in 0-1km)$*Same Elevation	0.0436	0.0837	0.0918	-0.0347	-0.0152	0.0118
	(0.0374)	(0.0640)	(0.0573)	(0.0260)	(0.0341)	(0.0540)
$1(disind \in 1-2km)$*Same Elevation	-0.0712	-0.0541	-0.0625	0.0147	-0.0222	-0.0282
	(0.0691)	(0.102)	(0.0712)	(0.0198)	(0.0222)	(0.0823)
$1(disind \in 2-3km)$*Same Elevation	-0.162**	-0.208*	-0.168*	-0.0683***	-0.0437	-0.245**
	(0.0816)	(0.109)	(0.0997)	(0.0237)	(0.0312)	(0.105)
$1(disind \in 3-4km)$*Same Elevation	-0.0441	-0.164	-0.123	-0.0264	-0.0547	-0.0945
	(0.114)	(0.123)	(0.134)	(0.0323)	(0.0377)	(0.160)
$1(disind > 4km)$*Same Elevation	0.165**	0.141	0.122	-0.0358	-0.0932**	0.0677
	(0.0736)	(0.0948)	(0.0929)	(0.0231)	(0.0365)	(0.0863)
$1(disind \in 0-1km)$	0.272*	0.335*	0.178	0.0275	0.0416	0.337
	(0.162)	(0.189)	(0.217)	(0.0410)	(0.0661)	(0.248)
$1(disind \in 1-2km)$	0.346**	0.394**	0.295	-0.0201	0.0211	0.291
	(0.159)	(0.189)	(0.203)	(0.0323)	(0.0598)	(0.231)
$1(disind \in 2-3km)$	0.378**	0.443**	0.297*	0.0421	0.0243	0.426**
	(0.155)	(0.189)	(0.175)	(0.0304)	(0.0328)	(0.191)
$1(disind \in 3-4km)$	0.227**	0.330*	0.228*	0.0257	0.0300	0.230*
	(0.111)	(0.170)	(0.136)	(0.0318)	(0.0351)	(0.138)
Observations	6,563	5,154	6,329	6,598	6,423	6,532
R-squared	0.371	0.344	0.324	0.245	0.504	0.257
p-value (Sum of elevation interacts=0)	0.2596	0.2594	0.2796	0.0363	0.1296	0.2050

	Growth from 1980 to 2000: Below-median pollution level					
VARIABLES	Employment	Manufacture employment	FIRE employment	Median income	Housing value	College graduates
$1(disind \in 0-1km)$*Same Elevation	-0.0809	-0.0419	-0.0742	-0.0278	0.0368	-0.0295
	(0.0758)	(0.0860)	(0.0734)	(0.0256)	(0.0329)	(0.0768)
$1(disind \in 1-2km)$*Same Elevation	-0.0649	-0.0693	-0.0935	-0.0251	0.0676*	-0.0678
	(0.0767)	(0.0859)	(0.105)	(0.0307)	(0.0398)	(0.0959)
$1(disind \in 2-3km)$*Same Elevation	-0.0103	-0.178	-0.124	-0.0417	0.00760	-0.0613
	(0.0866)	(0.147)	(0.143)	(0.0285)	(0.0323)	(0.0891)
$1(disind \in 3-4km)$*Same Elevation	0.145***	0.116	-0.0548	0.0229	0.00848	0.0923
	(0.0479)	(0.0739)	(0.0904)	(0.0577)	(0.0495)	(0.0693)
$1(disind > 4km)$*Same Elevation	0.0936	0.0999	-0.0446	0.00326	-0.0014	0.0543
	(0.114)	(0.131)	(0.110)	(0.0226)	(0.029)	(0.113)
$1(disind \in 0-1km)$	0.0946	0.109	-0.0318	0.00174	-0.0167	-0.0507
	(0.0822)	(0.114)	(0.0917)	(0.0345)	(0.0397)	(0.0988)
$1(disind \in 1-2km)$	0.0477	0.0717	-0.0321	-0.00237	-0.0505	-0.0451
	(0.0871)	(0.109)	(0.121)	(0.0377)	(0.0458)	(0.110)
$1(disind \in 2-3km)$	0.00140	0.157	0.0289	-0.00280	-0.00229	-0.0273
	(0.0935)	(0.154)	(0.138)	(0.0316)	(0.0360)	(0.0996)
$1(disind \in 3-4km)$	-0.118**	-0.0878	0.00896	-0.0453	-0.000972	-0.135*
	(0.0478)	(0.0799)	(0.0768)	(0.0581)	(0.0461)	(0.0746)
Observations	6,563	5,154	6,329	6,598	6,423	6,532
R-squared	0.371	0.344	0.324	0.245	0.504	0.257
p-value (Sum of elevation interacts=0)	0.9307	0.3793	0.1403	0.4668	0.2417	0.7083

Notes: Results from the upper panel are obtained from a sample of census tracts that are closest to 1970 industrial areas with above-median pollution intensity measure, and those from the lower panel from those closest to industrial areas with below-median pollution level.Dependent variables are growth in total, manufacture, FIRE employment, median household income, median housing value and the number of college graduates from 1980 to 2000, all counted at the place of residence. $1(disind \in x-ykm)$ is an indicator of whether or not the distance from a tract to its closest industrial area is within x to y km. Same elevation is a dummy that takes one if the tract is at the same elevation or less than 100m lower than its closest industrial area, and not obstructed by other areas in between. Controls include CBSA fixed effects, the distance from each monitor to transportation lines, natural amenities and the CBD. Robust standard errors are clustered at CBSA level.

Table A4: Exclude coastal cities: economic outcomes in 2000

Outcomes in 2000: Tracts closest to above-median pollution level								
	By place of work			By place of residence				
VARIABLES	Employ density	Highskill ratio	Median earnings	Employ density	Highskill ratio	Median earnings	Housing value	College ratio
$1(disind \in 0-1km)$*Downwind	0.0124	0.00335	0.000700	-0.0235	-0.00416	0.00980	-0.0625**	-0.0214*
	(0.0872)	(0.00791)	(0.0126)	(0.0568)	(0.00413)	(0.0231)	(0.0297)	(0.0121)
$1(disind \in 1-2km)$*Downwind	0.0946	-0.0103	-0.0157	0.101	-0.0112***	-0.0391**	-0.0881***	-0.0370***
	(0.0819)	(0.00846)	(0.0147)	(0.0681)	(0.00378)	(0.0182)	(0.0319)	(0.0131)
$1(disind \in 2-3km)$*Downwind	-0.129	-0.00844	-0.0382*	-0.139	-0.0148***	-0.0280	-0.0193	-0.0212
	(0.114)	(0.00848)	(0.0205)	(0.0998)	(0.00502)	(0.0250)	(0.0362)	(0.0163)
$1(disind \in 3-4km)$*Downwind	-0.174	-0.0152	-0.0169	-0.135	-0.0127*	-0.0241	-0.0608	-0.0451*
	(0.177)	(0.0109)	(0.0291)	(0.105)	(0.00695)	(0.0323)	(0.0622)	(0.0238)
$1(disind > 4km)$*Downwind	0.178	0.0144	0.0269	0.192	0.00580	-0.0103	0.0338	-0.0120
	(0.198)	(0.0109)	(0.0215)	(0.152)	(0.00778)	(0.0359)	(0.0463)	(0.0153)
$1(disind \in 0-1km)$	-0.261	-0.0297*	0.0759**	-0.365**	-0.0188**	-0.143***	-0.220***	-0.0878***
	(0.189)	(0.0169)	(0.0294)	(0.180)	(0.00888)	(0.0462)	(0.0667)	(0.0240)
$1(disind \in 1-2km)$	-0.310*	-0.0217	0.0446*	-0.268*	-0.0105	-0.0808*	-0.136**	-0.0548**
	(0.173)	(0.0150)	(0.0253)	(0.147)	(0.00841)	(0.0429)	(0.0583)	(0.0210)
$1(disind \in 2-3km)$	-0.242*	-0.0131	0.0405	-0.118	-0.00468	-0.0338	-0.0700	-0.0282
	(0.128)	(0.0120)	(0.0270)	(0.113)	(0.00692)	(0.0391)	(0.0506)	(0.0192)
$1(disind \in 3-4km)$	-0.0822	-0.00912	0.0334	-0.0377	0.00103	-0.0160	0.0116	-0.00140
	(0.113)	(0.00761)	(0.0233)	(0.0814)	(0.00580)	(0.0304)	(0.0422)	(0.0165)
Observations	4,907	4,907	4,767	4,914	4,914	4,765	4,865	4,908
R-squared	0.292	0.108	0.247	0.384	0.345	0.281	0.487	0.232

Outcomes in 2000: Tracts closest to below-median pollution level								
	By place of work			By place of residence				
VARIABLES	Employ density	Highskill ratio	Median earnings	Employ density	Highskill ratio	Median earnings	Housing value	College ratio
$1(disind \in 0-1km)$*Downwind	-0.101	-0.00410	-0.00445	-0.0452	-0.00164	0.0153	-0.0206	-0.0186*
	(0.0953)	(0.00747)	(0.0126)	(0.0725)	(0.00387)	(0.0154)	(0.0289)	(0.0102)
$1(disind \in 1-2km)$*Downwind	-0.107	0.000508	0.00197	-0.0518	-0.00281	0.0138	-0.0482	-0.0123
	(0.0877)	(0.00623)	(0.0180)	(0.0612)	(0.00348)	(0.0189)	(0.0302)	(0.0109)
$1(disind \in 2-3km)$*Downwind	-0.159	-0.00585	0.0238	-0.225*	-0.00112	0.0347	-0.0252	-0.00551
	(0.135)	(0.00962)	(0.0202)	(0.116)	(0.00557)	(0.0263)	(0.0385)	(0.0148)
$1(disind \in 3-4km)$*Downwind	-0.113	0.000953	-0.0246	0.00298	-0.00641	-0.00991	-0.0255	0.00417
	(0.132)	(0.0152)	(0.0242)	(0.110)	(0.00770)	(0.0297)	(0.0586)	(0.0210)
$1(disind > 4km)$*Downwind	-0.124	0.00372	-0.0250**	0.0450	0.00452	0.0236	0.0446	0.0195
	(0.121)	(0.00656)	(0.0125)	(0.108)	(0.0104)	(0.0413)	(0.0641)	(0.0262)
$1(disind \in 0-1km)$	0.472**	-0.0292***	-0.00467	0.285*	-0.0239***	-0.150***	-0.229***	-0.0842***
	(0.189)	(0.0109)	(0.0243)	(0.170)	(0.00656)	(0.0273)	(0.0394)	(0.0185)
$1(disind \in 1-2km)$	0.431**	-0.0231**	-0.0341	0.330**	-0.0183***	-0.131***	-0.180***	-0.0639***
	(0.168)	(0.00894)	(0.0211)	(0.150)	(0.00634)	(0.0237)	(0.0370)	(0.0164)
$1(disind \in 2-3km)$	0.204	-0.0101	-0.0249	0.221*	-0.00925	-0.0780***	-0.0996**	-0.0339**
	(0.140)	(0.00795)	(0.0181)	(0.115)	(0.00663)	(0.0283)	(0.0386)	(0.0161)
$1(disind \in 3-4km)$	0.153	-0.00573	-0.0148	0.155	-0.00319	-0.0170	-0.0375	-0.0201
	(0.122)	(0.00689)	(0.0151)	(0.101)	(0.00540)	(0.0234)	(0.0326)	(0.0149)
Observations	5,878	5,878	5,558	5,875	5,875	5,554	5,835	5,878
R-squared	0.324	0.126	0.260	0.393	0.430	0.322	0.505	0.283

Notes: Results from the upper panel are obtained from a sample of census tracts that are closest to 1970 industrial areas with above-median pollution intensity measure, and those from the lower panel from those closest to industrial areas with below-median pollution level, in non-coastal cities. Dependent variables of the first three columns are employment density, the ratio of FIRE, IT and professional services in total employment, and median earnings, counted at the place of work. The last five columns report results on employment density, the ratio of FIRE, IT and professional services in total employment, median earnings, median housing value and ratio of college graduates among the population over 25 years old, counted at their place of residence. $1(disind \in x-ykm)$ is an indicator of whether or not the distance from a tract to its closest industrial area is within x to y km. Downwind is an indicator of whether or not the TSP monitor is located downwind of the industrial area. Controls include CBSA fixed effects, the distance from each monitor to transportation lines, natural amenities and the CBD. Robust standard errors are clustered at CBSA level.

Table A5: Determinants of local amenity perceived by workers from different sectors

VARIABLES	FIRE	IT	Edu Med	Professional	Public admin	Art entertain
$1(industrialarea)$	-0.0867***	-0.0741***	-0.0778***	-0.0825***	-0.0681***	-0.0471***
	(0.00968)	(0.0111)	(0.00785)	(0.00855)	(0.00823)	(0.00797)
disind<1km	-0.0911***	-0.0728***	-0.0824***	-0.0884***	-0.0819***	-0.0575***
	(0.0106)	(0.0122)	(0.00788)	(0.00935)	(0.0108)	(0.0101)
disind$\in 1-2km$	-0.0653***	-0.0520***	-0.0645***	-0.0677***	-0.0678***	-0.0467***
	(0.0109)	(0.0113)	(0.00830)	(0.0103)	(0.0108)	(0.0108)
disind$\in 2-3km$	-0.0435***	-0.0399***	-0.0457***	-0.0508***	-0.0462***	-0.0358***
	(0.00998)	(0.0113)	(0.00841)	(0.0100)	(0.00991)	(0.00955)
disind$\in 3-4km$	-0.0149**	-0.0176**	-0.0249***	-0.0185**	-0.0268***	-0.0147***
	(0.00609)	(0.00801)	(0.00556)	(0.00713)	(0.00570)	(0.00539)
Public school	0.0116***	0.00867***	0.0147***	0.0104***	0.0149***	0.0124***
	(0.00206)	(0.00253)	(0.00162)	(0.00216)	(0.00161)	(0.00153)
Distance to highway	-7.87e-07	3.56e-07	-8.38e-07	-1.02e-06	6.22e-07	5.39e-08
	(7.69e-07)	(8.25e-07)	(7.28e-07)	(6.99e-07)	(8.07e-07)	(6.56e-07)
Distance to the CBD	0.00604***	0.00534***	0.00640***	0.00557***	0.00681***	0.00591***
	(0.000504)	(0.000438)	(0.000633)	(0.000519)	(0.000467)	(0.000427)
Beach	0.0764***	0.0485*	0.0552*	0.0425	0.0787***	0.0626**
	(0.0287)	(0.0268)	(0.0281)	(0.0278)	(0.0245)	(0.0252)
Distance to railway	0.00610***	0.00514***	0.00471***	0.00502***	0.00268*	0.00379***
	(0.00156)	(0.00157)	(0.00131)	(0.00133)	(0.00152)	(0.00102)
Distance to water	1.22e-08	2.82e-07	6.03e-07	8.08e-07	2.72e-07	1.28e-07
	(5.12e-07)	(7.27e-07)	(3.96e-07)	(6.08e-07)	(6.92e-07)	(5.06e-07)
Observations	27,529	26,986	27,639	27,625	27,431	27,621
R-squared	0.781	0.832	0.870	0.785	0.830	0.918

VARIABLES	Manufacture	Wholesale	Retail	Farming	Construction	Utility
$1(industrialarea)$	-0.0336***	-0.0569***	-0.0559***	0.00204	-0.0554***	-0.0505***
	(0.00778)	(0.00875)	(0.00808)	(0.0173)	(0.00859)	(0.00686)
disind<1km	-0.0540***	-0.0678***	-0.0623***	0.0210	-0.0727***	-0.0625***
	(0.00659)	(0.0112)	(0.00878)	(0.0189)	(0.00875)	(0.00775)
disind$\in 1-2km$	-0.0498***	-0.0535***	-0.0489***	-0.00524	-0.0626***	-0.0442***
	(0.00680)	(0.0129)	(0.00937)	(0.0235)	(0.0103)	(0.00835)
disind$\in 2-3km$	-0.0408***	-0.0439***	-0.0370***	0.0219	-0.0546***	-0.0428***
	(0.00684)	(0.0109)	(0.00930)	(0.0224)	(0.00906)	(0.00722)
disind$\in 3-4km$	-0.0186***	-0.0198***	-0.0213***	0.0229	-0.0126***	-0.0174***
	(0.00626)	(0.00547)	(0.00545)	(0.0250)	(0.00613)	(0.00604)
Public school	0.0157***	0.0138***	0.0138***	-0.0192***	0.0158***	0.0151***
	(0.00188)	(0.00160)	(0.00177)	(0.00467)	(0.00159)	(0.00143)
Distance to highway	-1.11e-08	-2.55e-07	-5.41e-07	-1.03e-06	-1.47e-07	2.58e-08
	(8.29e-07)	(6.89e-07)	(7.18e-07)	(7.72e-07)	(8.06e-07)	(7.30e-07)
Distance to the CBD	0.00762***	0.00686***	0.00672***	-0.00210***	0.00777***	0.00618***
	(0.000500)	(0.000618)	(0.000537)	(0.000469)	(0.000666)	(0.000418)
Beach	0.0531*	0.0317	0.0786***	-0.0297	0.0815**	0.0855***
	(0.0289)	(0.0269)	(0.0292)	(0.0577)	(0.0325)	(0.0317)
Distance to railway	0.00165	0.00374**	0.00374**	0.00674**	0.00306**	0.00320**
	(0.00128)	(0.00169)	(0.00157)	(0.00309)	(0.00144)	(0.00139)
Distance to water	4.66e-07	-4.99e-08	6.16e-07*	-4.13e-07	1.99e-07	3.89e-07
	(4.51e-07)	(5.46e-07)	(3.57e-07)	(3.02e-07)	(4.93e-07)	(4.72e-07)
Observations	27,586	27,230	27,628	9,218	27,490	27,522
R-squared	0.844	0.816	0.901	0.120	0.844	0.797

Notes: Dependent variable is estimated amenity level in 2000 perceived by workers from different sectors, defined in equation (22), with H_{Ris} standing for the number of workers from sector s at place of residence and $\tilde{w_{ps}}$ standing for adjusted wages of sector s at place of work. CBSA fixed effects are controlled for. SE clustered at CBSA level.

Table A6: Determinants of local amenity perceived by different sectors

VARIABLES	FIRE	IT	Edu Med	Professional	Public admin
violent crime rate	-0.00506***	-0.00465***	-0.00482***	-0.00446***	-0.00494***
	(0.000665)	(0.000580)	(0.000515)	(0.000514)	(0.000581)
$1(industrialarea)$	-0.0998***	-0.122***	-0.0914***	-0.104***	-0.0538***
	(0.0179)	(0.0261)	(0.0135)	(0.0164)	(0.0109)
disind<1km	-0.0836***	-0.0968***	-0.0724***	-0.0843***	-0.0494***
	(0.0178)	(0.0242)	(0.0131)	(0.0162)	(0.0106)
disind$\in 1-2km$	-0.0783***	-0.0886***	-0.0751***	-0.0846***	-0.0472***
	(0.0176)	(0.0222)	(0.0134)	(0.0151)	(0.0116)
disind$\in 2-3km$	-0.0635***	-0.0796***	-0.0484***	-0.0696***	-0.0341***
	(0.0164)	(0.0199)	(0.0126)	(0.0159)	(0.0119)
disind$\in 3-4km$	-0.0499***	-0.0644***	-0.0363**	-0.0449***	-0.0253**
	(0.0158)	(0.0161)	(0.0138)	(0.0126)	(0.0104)
Public school	0.00194	0.000695	0.00635*	0.000994	0.00658*
	(0.00281)	(0.00364)	(0.00322)	(0.00328)	(0.00336)
Distance to highway	2.58e-06	1.55e-06	4.85e-06*	1.46e-06	7.00e-06**
	(2.11e-06)	(2.91e-06)	(2.50e-06)	(2.62e-06)	(2.84e-06)
Distance to the CBD	0.0113***	0.00545***	0.00659***	0.00896***	0.00654***
	(0.000701)	(0.00137)	(0.00128)	(0.000925)	(0.000951)
Dostance to railway	0.0146***	0.0161***	0.0102***	0.00752***	0.0106***
	(0.00317)	(0.00266)	(0.00194)	(0.00225)	(0.00376)
Distance to water	5.64e-07	-2.38e-06***	7.20e-07	-1.60e-06*	-5.90e-07
	(1.39e-06)	(7.49e-07)	(8.79e-07)	(8.03e-07)	(1.23e-06)
Beach	-0.0829**	0.0137	0.0324***	0.0204	-0.0358**
	(0.0345)	(0.0411)	(0.00987)	(0.0281)	(0.0135)
Observations	3,446	3,369	3,465	3,464	3,424
	0.831	0.823	0.868	0.825	0.835

VARIABLES	Art Entertain	Manufacture	Wholesale	Retail	Farming
violate crime rate	-0.00349***	-0.00318***	-0.00358***	-0.00394***	-0.00120
	(0.000427)	(0.000493)	(0.000583)	(0.000566)	(0.00200)
$1(industrialarea)$	-0.0360*	-0.0153	-0.0179	-0.0649***	-0.0821
	(0.0194)	(0.0104)	(0.0140)	(0.0178)	(0.124)
disind<1km	-0.0292*	-0.0286***	-0.0244*	-0.0643***	0.00171
	(0.0167)	(0.00908)	(0.0125)	(0.0156)	(0.115)
disind$\in 1-2km$	-0.0280	-0.0356***	-0.0290**	-0.0632***	-0.0270
	(0.0174)	(0.0100)	(0.0111)	(0.0146)	(0.124)
disind$\in 2-3km$	-0.0209	-0.0282***	-0.0374***	-0.0488***	0.00518
	(0.0148)	(0.00974)	(0.0106)	(0.0114)	(0.128)
disind$\in 3-4km$	-0.00341	-0.00899	-0.0219**	-0.0268**	-0.0581
	(0.0184)	(0.00907)	(0.0107)	(0.0122)	(0.130)
Public school	0.00863***	0.00631	0.00841*	0.00571	0.00519
	(0.00301)	(0.00539)	(0.00455)	(0.00405)	(0.00471)
Distance to highway	-1.86e-06	8.48e-07	3.36e-07	2.08e-06	6.00e-06
	(3.57e-06)	(2.48e-06)	(2.31e-06)	(2.81e-06)	(1.66e-05)
Distance to the CBD	0.00453**	0.00914***	0.00729***	0.00680***	-0.00454
	(0.00202)	(0.00134)	(0.00203)	(0.00154)	(0.00637)
Dostance to railway	0.00615***	0.00301	0.0114***	0.00753**	0.0181
	(0.00201)	(0.00315)	(0.00315)	(0.00296)	(0.0202)
Distance to water	-7.66e-07	4.72e-07	-3.22e-06***	-9.44e-07	3.34e-06
	(1.25e-06)	(1.53e-06)	(9.79e-07)	(1.38e-06)	(3.98e-06)
beach	0.0790***	-0.0148	-0.0388**	0.000829	-0.395*
	(0.0142)	(0.0330)	(0.0152)	(0.0155)	(0.208)
Observations	3,461	3,453	3,375	3,463	618
	0.892	0.825	0.747	0.908	0.138

Notes: Dependent variable is estimated amenity level in 2000 perceived by different sectors, defined in equation (22). Only tracts with 2000 crime rate data are kept. CBSA fixed effects are controlled for. SE clustered at CBSA level.

Table A7: Determinants of local amenity perceived by different sectors

VARIABLES	FIRE	IT	Edu Med	Professional	Public admin
PM10 2000	-0.00418**	-0.00418**	-0.00469***	-0.00308**	-0.00606***
	(0.00163)	(0.00184)	(0.00145)	(0.00144)	(0.00117)
$1(industrialarea)$	-0.101***	-0.0719***	-0.0984***	-0.0994***	-0.0705***
	(0.0287)	(0.0267)	(0.0203)	(0.0251)	(0.0247)
disind<1km	-0.0887***	-0.0500*	-0.0917***	-0.0971***	-0.0666***
	(0.0282)	(0.0273)	(0.0207)	(0.0262)	(0.0256)
disind$\in 1-2km$	-0.0714***	-0.0412	-0.0813***	-0.0877***	-0.0668***
	(0.0267)	(0.0260)	(0.0196)	(0.0242)	(0.0238)
disind$\in 2-3km$	-0.0572*	-0.0275	-0.0576***	-0.0761***	-0.0473*
	(0.0294)	(0.0282)	(0.0205)	(0.0264)	(0.0250)
disind$\in 3-4km$	-0.00755	0.00766	-0.0184	-0.00944	4.05e-05
	(0.0217)	(0.0202)	(0.0145)	(0.0195)	(0.0190)
Public school	0.00703**	0.00427	0.0111***	0.00680**	0.00825***
	(0.00321)	(0.00347)	(0.00292)	(0.00303)	(0.00298)
Distance to highway	6.21e-07	2.76e-06	-2.34e-08	1.05e-06	2.81e-06*
	(1.11e-06)	(2.59e-06)	(1.19e-06)	(1.66e-06)	(1.54e-06)
Distance to the CBD	0.00554***	0.00434***	0.00558***	0.00548***	0.00573***
	(0.000658)	(0.000778)	(0.000671)	(0.000598)	(0.000720)
Distance to railway	0.00856	0.00648	0.00666*	0.00512	0.000507
	(0.00529)	(0.00565)	(0.00391)	(0.00466)	(0.00530)
Distance to water	-8.69e-07	-3.79e-07	3.84e-08	3.26e-07	-7.08e-07
	(8.23e-07)	(1.20e-06)	(7.08e-07)	(6.22e-07)	(8.97e-07)
Observations	4,636	4,487	4,662	4,659	4,597
R-squared	0.822	0.866	0.903	0.847	0.856

VARIABLES	Art Entertain	Manufacture	Wholesale	Retail	Farming
PM10 2000	-0.00231	-0.00331	-0.00387**	-0.00308***	-0.00383*
	(0.00165)	(0.00206)	(0.00174)	(0.00145)	(0.00399)
$1(industrialarea)$	-0.0591***	-0.0358*	-0.0371	-0.0406*	0.0499
	(0.0190)	(0.0185)	(0.0248)	(0.0207)	(0.0819)
disind<1km	-0.0598***	-0.0517***	-0.0452*	-0.0474**	0.0452
	(0.0204)	(0.0180)	(0.0254)	(0.0208)	(0.0795)
disind$\in 1-2km$	-0.0632***	-0.0688***	-0.0498**	-0.0487**	-0.0409
	(0.0177)	(0.0174)	(0.0252)	(0.0191)	(0.0775)
disind$\in 2-3km$	-0.0461**	-0.0462***	-0.0273	-0.0281	0.113
	(0.0181)	(0.0169)	(0.0237)	(0.0177)	(0.123)
disind$\in 3-4km$	-0.0143	-0.0256	-0.0128	0.00632	0.0246
	(0.0142)	(0.0160)	(0.0186)	(0.0146)	(0.121)
Public school	0.00978***	0.00751***	0.00672**	0.00908***	-0.0119
	(0.00258)	(0.00286)	(0.00286)	(0.00295)	(0.0147)
Distance to highway	4.51e-06**	3.24e-06	3.36e-06	3.28e-06	6.19e-06***
	(2.09e-06)	(2.24e-06)	(2.39e-06)	(2.00e-06)	(2.17e-06)
Distance to the CBD	0.00497***	0.00704***	0.00628***	0.00637***	-0.00198
	(0.000687)	(0.00106)	(0.000814)	(0.000820)	(0.00163)
Distance to railway	0.00642**	0.000235	0.00506	0.00700	0.0159
	(0.00313)	(0.00423)	(0.00388)	(0.00477)	(0.0166)
Distance to water	4.24e-07	4.15e-07	-7.17e-07	7.16e-08	-1.85e-06
	(1.04e-06)	(5.57e-07)	(1.01e-06)	(7.92e-07)	(1.28e-06)
Observations	4,658	4,647	4,552	4,658	1,298
R-squared	0.941	0.862	0.854	0.921	0.203

Note Dependent variable is estimated amenity level in 2000 perceived by different sectors, defined in equation (22). Only tracts within 2 km to a PM10 monitor station are kept in the sample. CBSA fixed effects are controlled for. SE clustered at CBSA level.

Table A8: Determinants of local productivity by different sectors

VARIABLES	FIRE	IT	Edu Med	Professional	Public admin	Art entertain
$1(industrial area)$	-0.0412***	-0.0315**	-0.0231***	-0.0209**	-0.00424	-0.0222**
	(0.00912)	(0.0138)	(0.00589)	(0.00822)	(0.0103)	(0.00955)
disind<1km	-0.0464***	-0.0316***	-0.0148**	-0.0328***	0.00119	-0.0282***
	(0.00920)	(0.0136)	(0.00610)	(0.00769)	(0.0101)	(0.0101)
disind$\in 1-2km$	-0.0252***	-0.0312**	-0.0114*	-0.0284***	0.00436	-0.0162
	(0.00948)	(0.0150)	(0.00608)	(0.00838)	(0.0114)	(0.0115)
disind$\in 2-3km$	-0.0382***	-0.0315*	-0.00321	0.00165	0.0118	-0.00166
	(0.0116)	(0.0165)	(0.00683)	(0.00824)	(0.0104)	(0.0116)
disind$\in 3-4km$	0.00140	0.0270*	0.00251	0.00635	0.0228	0.00575
	(0.0130)	(0.0159)	(0.00717)	(0.0127)	(0.0145)	(0.0115)
Public School	-0.00536***	-0.00681**	0.0284***	-0.00455**	-0.00718***	-0.0177***
	(0.00197)	(0.00295)	(0.00154)	(0.00195)	(0.00204)	(0.00206)
Distance to highway	-2.37e-06***	-3.93e-06***	-1.76e-06***	-2.31e-06***	-3.08e-06***	-2.59e-07
	(5.42e-07)	(9.79e-07)	(3.95e-07)	(6.26e-07)	(6.65e-07)	(5.26e-07)
Distance to CBD	-0.00110***	-0.00123***	-0.00146***	-0.000914***	-0.00206***	-0.00184***
	(0.000358)	(0.000357)	(0.000228)	(0.000344)	(0.000404)	(0.000230)
Beach	-0.0463	0.0207	-0.0535	0.0861**	0.0149	-0.0193
	(0.0500)	(0.0687)	(0.0376)	(0.0341)	(0.0469)	(0.0492)
Distance to railway	0.00820***	0.00741***	0.00439***	0.00670***	0.00612***	0.00567***
	(0.00147)	(0.00213)	(0.00118)	(0.00142)	(0.00177)	(0.00127)
Distance to water	-4.26e-08	-4.03e-07**	-1.29e-07	-6.29e-08	-1.53e-07	1.90e-07
	(2.15e-07)	(1.66e-07)	(1.46e-07)	(1.87e-07)	(1.95e-07)	(1.98e-07)
Observations	24,345	17,497	27,210	25,455	19,301	22,864
R-squared	0.136	0.104	0.199	0.138	0.185	0.138

VARIABLES	Manufacture	Wholesale	Retail	Farming	Construction	Utility
$1(industrial area)$	-0.0111	-0.0222	-0.0195**	0.00204	-0.0105	-0.00346
	(0.0101)	(0.0118)	(0.00818)	(0.0173)	(0.00722)	(0.0101)
disind<1km	-0.0116	-0.0449***	-0.0256***	0.0210	-0.0206**	-0.0159
	(0.0126)	(0.0116)	(0.00941)	(0.0189)	(0.00807)	(0.0128)
disind$\in 1-2km$	-0.0291**	-0.0170	-0.0227**	-0.00524	-0.0188**	-0.00631
	(0.0118)	(0.0132)	(0.00990)	(0.0235)	(0.00872)	(0.0138)
disind$\in 2-3km$	-0.0119	-0.0153	-0.0351***	0.0219	-0.0117	-0.00835
	(0.00997)	(0.0140)	(0.0114)	(0.0224)	(0.00932)	(0.0156)
disind$\in 3-4km$	-0.00487	0.0124	-0.00669	0.0306	0.0148	0.00184
	(0.0133)	(0.0158)	(0.0108)	(0.0250)	(0.0107)	(0.0134)
Publich School	-0.0133***	-0.00862***	-0.0182***	-0.0192***	-0.00722***	-0.00604***
	(0.00277)	(0.00307)	(0.00204)	(0.00467)	(0.00226)	(0.00224)
Distance to highway	-3.45e-06***	-2.95e-06***	-2.32e-06***	-1.03e-06	-3.07e-06***	-2.40e-06***
	(7.29e-07)	(6.94e-07)	(5.16e-07)	(7.72e-07)	(5.53e-07)	(5.75e-07)
Distance to CBD	0.000173	-0.000388	-0.000884**	-0.00210***	-0.000495	-0.000471
	(0.000612)	(0.000489)	(0.000352)	(0.000469)	(0.000333)	(0.000353)
Beach	0.0117	0.0405	0.00136	-0.0297	-0.00680	0.0127
	(0.0582)	(0.0532)	(0.0455)	(0.0577)	(0.0277)	(0.0357)
Distance to railway	0.00633***	0.00599***	0.00358**	0.00674**	0.00448***	0.00570***
	(0.00161)	(0.00207)	(0.00155)	(0.00309)	(0.00138)	(0.00142)
Distance to water	1.03e-07	-2.66e-07	7.67e-08	-4.13e-07	-1.69e-07	-1.57e-07
	(2.39e-07)	(3.62e-07)	(2.66e-07)	(3.02e-07)	(2.43e-07)	(1.98e-07)
Observations	24,273	21,643	25,640	9,218	25,537	21,310
R-squared	0.115	0.095	0.129	0.120	0.192	0.084

Notes: Dependent variable is estimated productivity of different sectors in 2000, defined in equation (17). CBSA fixed effects are controlled for. SE clustered at CBSA level.

Table A9: Triple-differences specification: 2000 outcomes

	By place of work				By place of residence				
VARIABLES	1971-79 log TSP	Employ density	Highskill ratio	Median earnings	Employ density	Highskill ratio	Median earnings	Housing value	Colle ratio
IND1*TSP7179*Downwind	0.00904	-0.0890	-0.0158	-0.0183	0.0297	-0.0186**	-0.0712**	-0.108**	-0.03
	(0.0316)	(0.101)	(0.0162)	(0.0226)	(0.0594)	(0.00836)	(0.0286)	(0.0495)	(0.01
IND2*TSP7179*Downwind	0.0355	-0.114	-0.0179	-0.0172	-0.0645	-0.0200*	-0.0620***	-0.146***	-0.04
	(0.0359)	(0.162)	(0.0226)	(0.0198)	(0.0707)	(0.0109)	(0.0238)	(0.0460)	(0.01
IND3*TSP7179*Downwind	0.0303	-0.133	-0.0334*	-0.0319*	0.0240	-0.0180	-0.0165	-0.0750*	-0.02
	(0.0325)	(0.114)	(0.0171)	(0.0193)	(0.0824)	(0.0111)	(0.0268)	(0.0445)	(0.01
IND4*TSP7179*Downwind	0.115**	-0.0250	0.0448	-0.103**	-0.00473	-0.0283	-0.0996	-0.0778	-0.04
	(0.0561)	(0.370)	(0.0335)	(0.0430)	(0.274)	(0.0211)	(0.108)	(0.0957)	(0.03
IND1*TSP7179	0.136***	-0.101**	-0.00708	0.00169	-0.124***	-0.00846**	0.00612	0.0151	0.004
	(0.0108)	(0.0454)	(0.00514)	(0.00673)	(0.0357)	(0.00396)	(0.0156)	(0.0239)	(0.01
IND2*TSP7179	0.104***	-0.0627	0.00211	-0.00322	-0.0855**	-0.00424	0.00767	0.0250	0.006
	(0.0236)	(0.0572)	(0.00459)	(0.00753)	(0.0395)	(0.00288)	(0.0170)	(0.0252)	(0.01
IND3*TSP7179	0.101***	0.0153	0.00255	-0.00322	-0.00939	-0.00982*	-0.0191	-0.0113	-0.00
	(0.0175)	(0.0489)	(0.00608)	(0.00665)	(0.0442)	(0.00547)	(0.0194)	(0.0255)	(0.00
IND4*TSP7179	0.0439*	-0.0998	-0.000489	0.00727	-0.104*	-0.000947	0.0207	0.0410	0.017
	(0.0263)	(0.0676)	(0.00461)	(0.00797)	(0.0542)	(0.00197)	(0.0184)	(0.0253)	(0.01
IND1*Downwind	0.0186	-0.00178	0.0235	-0.00679	0.167	0.00714	0.0945**	-0.0150	0.013
	(0.204)	(0.204)	(0.0241)	(0.0293)	(0.151)	(0.0143)	(0.0475)	(0.0689)	(0.02
IND2*Downwind	-0.102	-0.115	0.0230	-0.0453	0.155	0.00363	0.0844*	0.0106	0.016
	(0.221)	(0.220)	(0.0251)	(0.0313)	(0.154)	(0.0167)	(0.0455)	(0.0747)	(0.02
IND3*Downwind	-0.0248	0.00929	0.0242	0.00107	-0.0207	0.000126	0.0132	-0.0663	-0.00
	(0.257)	(0.255)	(0.0296)	(0.0364)	(0.187)	(0.0183)	(0.0544)	(0.0784)	(0.03
IND4*Downwind	0.525*	0.642**	0.0332	-0.0566	0.391**	0.000686	-0.0516	0.0193	0.036
	(0.278)	(0.301)	(0.0276)	(0.0356)	(0.172)	(0.0160)	(0.0674)	(0.0716)	(0.02
IND1	0.0584**	0.516**	-0.0186	0.0168	0.239	-0.0412***	-0.205***	-0.255***	-0.10
	(0.0292)	(0.220)	(0.0168)	(0.0240)	(0.193)	(0.00942)	(0.0329)	(0.0529)	(0.01
IND2	0.0435*	0.455**	-0.00562	-0.00315	0.287	-0.0259***	-0.165***	-0.201***	-0.08
	(0.0257)	(0.195)	(0.0144)	(0.0179)	(0.184)	(0.00871)	(0.0311)	(0.0432)	(0.01
IND3	0.00560	0.217	-0.000269	-0.00773	0.181	-0.0137*	-0.122***	-0.115**	-0.05
	(0.0194)	(0.135)	(0.0118)	(0.0183)	(0.132)	(0.00736)	(0.0324)	(0.0480)	(0.01
IND4	0.00913	0.107	-0.00136	-0.0138	0.0890	-0.00464	-0.0519**	-0.0487	-0.02
	(0.0157)	(0.0953)	(0.0111)	(0.0149)	(0.0897)	(0.00556)	(0.0224)	(0.0417)	(0.01
Observations	10,565	10,508	10,508	10,176	10,532	10,532	10,170	10,251	10,51
R-squared	0.635	0.369	0.072	0.228	0.521	0.296	0.291	0.552	0.225

Notes: Dependent variable is estimated productivity of different sectors in 2000, defined in equation (17). CBSA fixed effects are controlled for. SE clustered at CBSA level.

Table A10: Triple-differences specification: Growth from 1980 to 2000

	Growth from 1980 to 2000					
VARIABLES	Employment	Manufacture employment	FIRE employment	Median income	Housing value	College graduates
IND1*TSP7179*Downwind	-0.0161	-0.00800	-0.0363	-0.0214**	0.0128	-0.0131
	(0.0233)	(0.0297)	(0.0332)	(0.0101)	(0.0159)	(0.0302)
IND2*TSP7179*Downwind	-0.0503*	-0.0547	-0.0645*	0.00935	0.00900	-0.0446
	(0.0277)	(0.0523)	(0.0332)	(0.0135)	(0.0146)	(0.0329)
IND3*TSP7179*Downwind	-0.0441	-0.0686	0.000227	-0.00646	0.00956	-0.0639
	(0.0327)	(0.0439)	(0.0466)	(0.0110)	(0.0179)	(0.0500)
IND4*TSP7179*Downwind	-0.0481	-0.0177	-0.0659	-0.0399*	-0.0167	-0.0922
	(0.0548)	(0.0696)	(0.0550)	(0.0241)	(0.0189)	(0.0694)
IND1*TSP7179	-0.0157	-0.00949	-0.0197	0.000921	-0.00178	-0.0100
	(0.0127)	(0.0184)	(0.0182)	(0.00438)	(0.00545)	(0.0144)
IND2*TSP7179	-0.000302	-0.00225	0.00494	-0.00663	0.00356	-0.00855
	(0.0110)	(0.0139)	(0.0138)	(0.00546)	(0.00745)	(0.0144)
IND3*TSP7179	0.00303	0.0328	-0.0556	-0.0171*	-0.00678	-0.0175
	(0.0299)	(0.0377)	(0.0361)	(0.00998)	(0.0113)	(0.0365)
IND4*TSP7179	-0.0224	-0.0290*	-0.0387	0.00320	-0.000243	-0.0176
	(0.0138)	(0.0166)	(0.0248)	(0.00400)	(0.00339)	(0.0190)
IND1*Downwind	-0.00143	-0.0430	-0.0169	-0.0122	0.0221	-0.00299
	(0.0547)	(0.0532)	(0.0705)	(0.0142)	(0.0356)	(0.0541)
IND2*Downwind	0.0355	0.0203	0.0584	-0.0312***	-0.0233	0.0181
	(0.0390)	(0.0560)	(0.0451)	(0.0114)	(0.0231)	(0.0396)
IND3*Downwind	0.00124	0.00544	-0.0193	-0.0135	-0.0143	0.00977
	(0.0480)	(0.0588)	(0.0536)	(0.0193)	(0.0234)	(0.0581)
IND4*Downwind	-0.0342	-0.0200	-0.104*	-0.0222	-0.0378	-0.0206
	(0.0557)	(0.0726)	(0.0618)	(0.0226)	(0.0248)	(0.0696)
IND1	-0.0196	-0.0737***	-0.0208	0.0224***	0.0207**	-0.0678**
	(0.0223)	(0.0279)	(0.0295)	(0.00818)	(0.00925)	(0.0291)
IND2	-0.0443**	-0.0973***	-0.0576*	0.0133	0.0127	-0.0927***
	(0.0215)	(0.0265)	(0.0300)	(0.00873)	(0.0117)	(0.0251)
IND3	-0.0355	-0.115***	-0.0687*	-0.00702	0.00616	-0.0867**
	(0.0320)	(0.0332)	(0.0404)	(0.0126)	(0.0178)	(0.0407)
IND4	-0.0170	-0.0719**	-0.0117	0.0126	0.0238	-0.0566
	(0.0268)	(0.0312)	(0.0365)	(0.0123)	(0.0229)	(0.0416)
Observations	12,499	9,602	11,889	12,607	12,200	12,420
R-squared	0.306	0.305	0.244	0.233	0.502	0.184

Table A11: IV estimation results

	IV: downwind*8 distance buffers					
	By place of work		By place of residence			
VARIABLES	Highskill ratio	Median earnings	Highskill ratio	Median earnings	Housing value	College ratio
$\Delta TSP1971-1979$	-0.00151	-0.00182	-0.00360	-0.00437	-0.00261	-0.00518
	(0.00143)	(0.00200)	(0.00371)	(0.00413)	(0.00518)	(0.00378)
Observations	6,903	6,782	6,926	6,777	6,629	6,905
$F_{1stStage}$	4.179	4.201	4.098	4.075	4.123	4.255
R-squared	0.044	0.177	0.132	0.252	0.553	0.078

	IV: downwind*1 distance buffer					
	By place of work		By place of residence			
VARIABLES	Highskill ratio	Median earnings	Highskill ratio	Median earnings	Housing value	College ratio
$\Delta TSP1971-1979$	-0.00178	-0.00252	-0.00497	-0.00663	-0.00235	-0.00667*
	(0.00176)	(0.00206)	(0.00480)	(0.00468)	(0.00588)	(0.00406)
Observations	6,903	6,782	6,926	6,777	6,629	6,905
$F_{1stStage}$	5.827	5.881	5.515	5.498	6.672	5.903
R-squared	0.027	0.079	0.080	0.220	0.553	-0.031

Notes: Dependent variables are the average measure of TSP ambient concentration from 1971 to 1979 collected at each TSP monitor with positive reading during this period. $1(disind \in a-bkm)$ is an indicator of whether or not the distance from a TSP monitor to the closest 1970s industrial area is within a and b km. Downwind is an indicator of whether or not the TSP monitor is located downwind of the industrial area, which is defined by Figure 1.1. Controls include CBSA fixed effects, the distance from each monitor to transportation lines, natural amenities and the CBD. Robust standard errors are clustered at CBSA level.

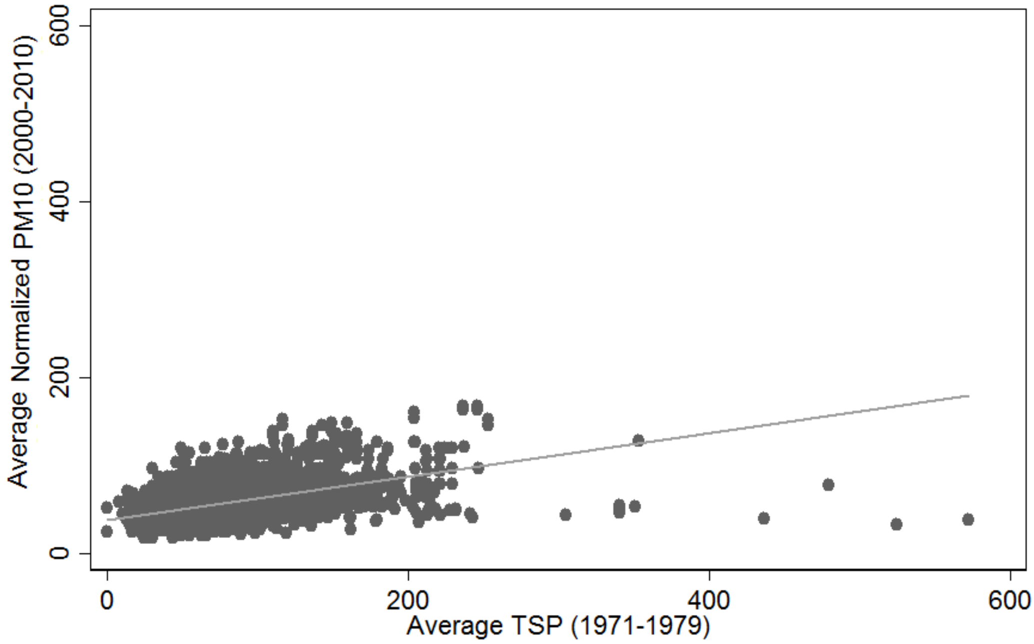
Table A12: Placebo Checks: Split sample

VARIABLES	%College graduates40	%College graduates50	log income 1950	Manager share50	Professional share50
$1(disind \in 0-1km)$*Downwind	-0.00943	-0.0103	0.0851**	-0.000378	-0.00769
	(0.00645)	(0.00880)	(0.0409)	(0.00462)	(0.00701)
$1(disind \in 1-2km)$*Downwind	0.0250*	0.0219**	0.0731	0.0163*	0.00924
	(0.0122)	(0.00995)	(0.0454)	(0.00854)	(0.00734)
$1(disind \in 2-3km)$*Downwind	0.00584	0.0408	0.107	-0.0303	0.0350**
	(0.0592)	(0.0417)	(0.0973)	(0.0296)	(0.0171)
$1(disind \in 3-4km)$*Downwind	-0.0224	0.0253	0.159	-0.0377	0.203***
	(0.0675)	(0.0436)	(0.126)	(0.0438)	(0.0545)
Downwind	0.00854	-0.0383	-0.114	0.0134	-0.0188
	(0.0613)	(0.0402)	(0.0938)	(0.0307)	(0.0174)
$1(disind \in 0-1km)$	-0.0575	-0.0502	-0.178	-0.0111	-0.00714
	(0.0596)	(0.0443)	(0.124)	(0.0226)	(0.0290)
$1(disind \in 1-2km)$	-0.0485	-0.0454	-0.152	-0.0114	-0.00645
	(0.0564)	(0.0401)	(0.110)	(0.0212)	(0.0243)
$1(disind \in 2-3km)$	-0.0286	-0.0513	-0.110	-0.00355	-0.0182
	(0.0491)	(0.0352)	(0.0932)	(0.0203)	(0.0212)
$1(disind \in 3-4km)$	-0.00909	-0.0328	-0.0598	-0.00271	-0.0159
	(0.0645)	(0.0384)	(0.106)	(0.0207)	(0.0175)
Observations	1,012	2,241	1,987	2,244	2,244
R-squared	0.167	0.189	0.273	0.236	0.121

VARIABLES	%College graduates40	%College graduates50	log income 1950	Manager share50	Professional share50
$1(disind \in 0-1km)$*Downwind	-0.0287**	-0.0332**	0.0330	-0.0222***	-0.0147**
	(0.0140)	(0.0157)	(0.0462)	(0.00671)	(0.00658)
$1(disind \in 1-2km)$*Downwind	-0.00382	0.0702	0.0278	0.00287	0.0454
	(0.0315)	(0.0644)	(0.0664)	(0.0181)	(0.0395)
$1(disind \in 2-3km)$*Downwind	-0.00574	-0.0108	-0.158	-0.0336**	0.0209
	(0.0262)	(0.0540)	(0.165)	(0.0154)	(0.0225)
$1(disind \in 3-4km)$*Downwind	0.142*	-0.0143	-0.0660	-0.0469**	-0.00900
	(0.0701)	(0.104)	(0.350)	(0.0230)	(0.0508)
Downwind	-0.0534	-0.0694	0.0681	0.000358	-0.000425
	(0.148)	(0.106)	(0.101)	(0.0173)	(0.0720)
$1(disind \in 0-1km)$	-0.0593	0.0202	-0.187*	-0.0190	0.0139
	(0.0424)	(0.0377)	(0.103)	(0.0192)	(0.0196)
$1(disind \in 1-2km)$	-0.0605*	0.00327	-0.110	-0.0154	0.00331
	(0.0347)	(0.0288)	(0.0821)	(0.0156)	(0.0149)
$1(disind \in 2-3km)$	-0.0332	0.00154	-0.104*	-0.0140	0.00721
	(0.0242)	(0.0211)	(0.0541)	(0.0114)	(0.0102)
$1(disind \in 3-4km)$	-0.0200	0.00533	0.00339	-0.000348	-0.00247
	(0.0193)	(0.0129)	(0.0374)	(0.00967)	(0.0102)
Observations	1,023	2,071	1,713	2,076	2,076
R-squared	0.358	0.281	0.413	0.230	0.172

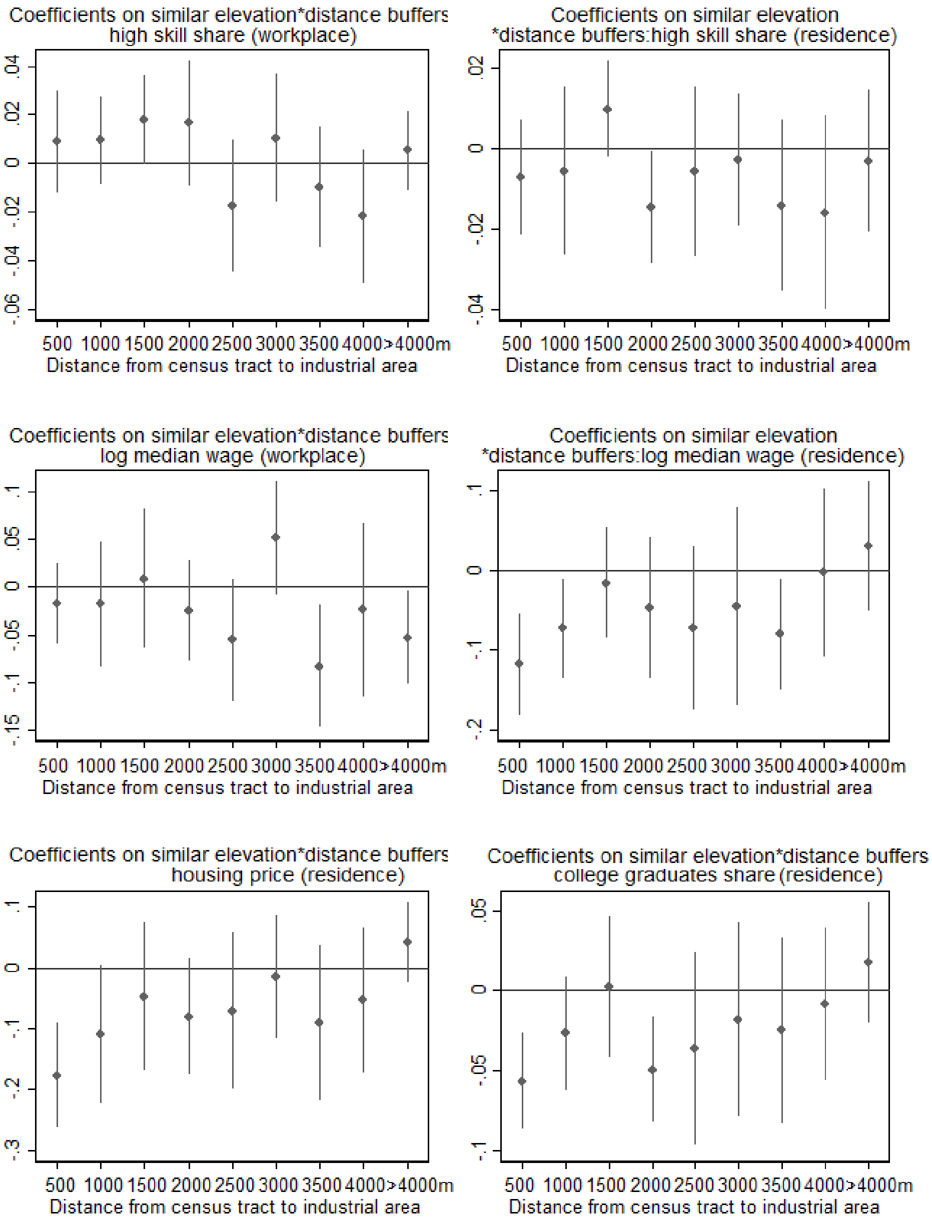
Notes: Results from the upper panel are obtained from a sample of 1950 census tracts that are closest to industrial areas with above-median pollution level in the 1970s, and those from the lower panel from those closest to industrial areas with below-median pollution level. Dependent variables are the share of college graduates in 1940 and 1950, log median income in 1950, the share of managers and professional/technical occuptions in total employment, $1(disind \in x-ykm)$ is an indicator of whether or not the distance from a tract to its closest industrial area is within x to y km. Downwind is an indicator of whether or not the TSP monitor is located downwind of the industrial area. Controls include CBSA fixed effects, the distance from each monitor to transportation lines, natural amenities and the CBD. Robust standard errors are clustered at CBSA level.

Figure A1: Pollution in the 1970s and 2000s



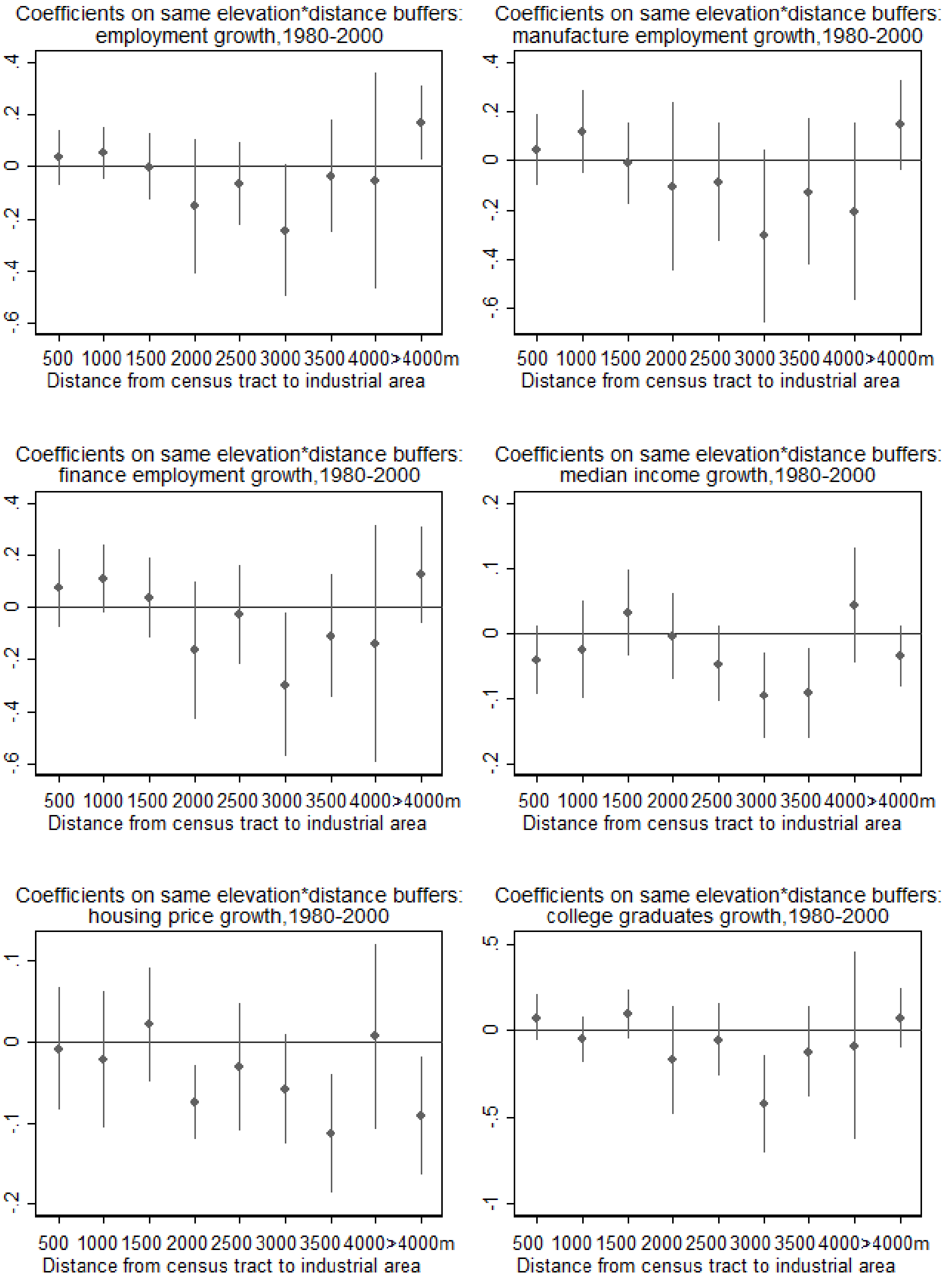
Notes: The graph plots the scatterplots of tract-level PM10 measures (y axis) in the 2000s against TSP measures (x axis) in the 1970s. For tracts with more than one monitor, an arithmetic average is taken. PM10 levels are normalized to TSP levels for better comparison according to the average TSP/PM10 ratio in a sample of monitor-year observations with both TSP and PM10 readings. Both axises are adjusted to the same scale.

Figure A2: Estimated coefficients on same elevation*distance buffers: Economic outcomes



Notes: The figures display the estimated coefficients and 95% confidence intervals in regressions where the dependent variables are housing price, college graduates share at residence, high skilled employment share at workplace and residence, median wage at workplace and residence. The independent variables are dummies of distance buffers of 500 metres are obtained from a sample of census tracts that are closest to industrial areas with above-median pollution level in the 1970s, and those from the lower panel from those closest to industrial areas with below-median pollution level. interval from each TSP monitor to the closest industrial area, interacted with a dummy of whether or not the TSP monitor is at the same elevation or less than 100 metres lower than the industrial area, and not obstructed by anything between them.All the coefficients are obtained in regressions on a sample of census tracts that are closest to above-median polluted industrial areas.

Figure A3: Estimated coefficients on same elevation*distance buffers: Economic outcomes



Notes: The figures display the estimated coefficients and 95% confidence intervals in regressions where the dependent variables are housing price, college graduates share at residence, high skilled employment share at workplace and residence, median wage at workplace and residence. The independent variables are dummies of distance buffers of 500 metres are obtained from a sample of census tracts that are closest to industrial areas with above-median pollution level in the 1970s, and those from the lower panel from those closest to industrial areas with below-median pollution level. interval from each TSP monitor to the closest industrial area, interacted with a dummy of whether or not the TSP monitor is at the same elevation or less than 100 metres lower than the industrial area, and not obstructed by anything between them.All the coefficients are obtained in regressions on a sample of census tracts that are closest to above-median polluted industrial areas.

Chapter 2

Travel Costs and Urban Specialization Patterns: Evidence from China's High Speed Railway System

How does intercity passenger transportation shape urban employment and specialization patterns? To shed light on this question I study China's High Speed Railway (HSR), an unprecedentedly large-scale network that connected 81 cities from 2003 to 2014 with trains running at speeds over 200 km/h. Using a difference-in-differences approach, I find that an HSR connection increases city-wide passenger flows by 10% and employment by 7%. To deal with the issues of endogenous railway placement and simultaneous public investments accompanying HSR connection, I examine the impact of a city's market access changes purely driven by the HSR connection of other cities. The estimates suggest that HSR-induced expansion in market access increases urban employment with an elasticity between 2 and 2.5. Further evidence on sectoral employment suggests that industries with a higher reliance on nonroutine cognitive skills benefit more from HSR-induced market access to other cities.

2.1 Introduction

Transportation costs play an important role in the location, agglomeration and evolution of economic activities. Yet, despite abundant research on the relationship between the cost of goods transportation and trade patterns, relatively little attention has been paid to the costs of passenger travel and their implications for labour markets. Reducing the cost of travel between cities not only removes obstacles to migration (Morten and Oliveira (2016), but also reduces the cost of face-to-face meeting across cities and allows remote sourcing of jobs. As airplanes or speedy trains make frequent day trips more feasible, firms may be more willing to locate their headquarters or R&D centres in centrally-located cities with large pools of talented employees, who can exert effective control over production plants in smaller cities with much lower urban costs. How significant are the benefits of infrastructure projects that dramatically increase the speed of intercity travelling, and how are these benefits distributed across sectors?

In this paper, I exploit the high speed railway (HSR) project in China, the largest in the world, as a natural experiment to study the benefits of improving passenger-dedicated transportation infrastructure. I examine the impacts of HSR connection through changes in its access to all the other cities, and focus on the one driven by indirect HSR connection to address endogeneity issues such as non-random route placements or simultaneous investments in other areas. The differential impacts of HSR on employment across sectors suggest that industries benefiting more from enhanced market access are either tourism-related or intensive in nonroutine cognitive skills. These findings highlight the role of improved passenger travel infrastructure in promoting the delivery of services across cities, facilitating cross-city labour sourcing and knowledge exchanges, and ultimately shifting the specialization pattern of connected cities towards skilled and communication intensive sectors.

The HSR expansion in China is an appropriate context for such a study. As of 2014, China had the world's longest and busiest HSR network with 15,399 km of track in service, connections between 81 cities, and an annual ridership of 857 million as of 2014.[1] HSR had a marked impact on people's travel patterns: after being connected to the HSR network, a city experiences an 18% increase in the number of passengers travelling by train and a 9.6% increase in the number of passengers travelling by any forms of transportation. The top-down rapid expansion of the HSR network also creates plausibly exogenous variation in each city's connectivity which does not depend on its own its own direct connection to the HSR system, since passengers from unconnected cities use HSR to travel to other cities by transferring at a nearby HSR hub.

In a difference-in-differences specification, I demonstrate that being connected to the HSR network leads to a significant increase in GDP and urban employment. As a first attempt at dealing with the problem of endogenous routes placement, I restrict my study sample to the cities that are either connected by the end of 2014 or will be connected by HSR according to the HSR plan of the Ministry of Railway (MOR)

[1]Table A6 and Figure A1

(MOR (2008)). Therefore, the analysis is identified through the timing of a city's connection to the HSR, which is affected by idiosyncratic factors such as the length of the line and engineering difficulties. I also test whether connected and unconnected cities have experienced differential trends in GDP or employment growth before the actual HSR connection by including the leads of the initial connection dummy. No HSR effects are found on GDP or employment before the actual connection.

To further explore the mechanisms at work, I adopt a "market access" approach similar to that in Donaldson and Hornbeck (2016), which is micro-founded by a model of cross-city labour sourcing. I measure how the expansion of the HSR network affects each city's "market access", and estimate the impacts of enhanced market access on the city's transportation and specialization patterns, as well as aggregate economic outcomes. A city's market access is approximated using an average of other cities' GDP inversely weighted by the bilateral costs of passenger travel. To account for the changes in cities' market access driven by improvements in different modes of transportation, including HSR, I assemble a network database of highways and railways in China from 2000 to 2014 to construct a time-varying travel cost matrix that takes account of the changes in time and fare costs brought by both highway and HSR expansions in China.

A major empirical advantage of this "market access" method is that it allows a city's market access to be affected by the HSR connections of other cities. Employees in non-HSR cities can travel to a nearby HSR town and transfer there for other destinations[2]. The separation of HSR-induced market access growth from direct HSR placements assists in dealing with the identification challenges of both endogenous infrastructure placement and simultaneous public spending and investments in relevant sectors. In my main specification, I use a measure of market access which deducts the increase in market access driven by a city's own connections. A one percent increase in this "non-connection-induced market access" (NCIMA) leads to an 8% increase in railway ridership and 2% increase in employment. The impacts are largest for tourism-related employment, followed by skilled employment including IT, finance, business services, education and scientific research, and smallest for other types of service and non-service employment.

Another benefit of this approach is that I can evaluate the impacts of alternative market access measures that capture different sources of improvement in intercity connectivity during this period. In particular, I evaluate the separate effects of improvements in HRS expansion, highway expansion, and parts of HSR development

[2] Ollivier et al. (2014) documents that 40% of the passengers taking the Wuhan-Guangzhou HSR line come from other railway lines or alternative modes of transportation.

that face little competition from air travel. This helps us to get a complete picture on the way various infrastructure improvements change transportation patterns and economic outcomes. HSR expansion does not reduce highway usage but highway expansion affects railway ridership significantly. A larger drop in air travel usage is observed in cities that benefit more from HSR over shorter distance trips. Over longer trips, HSR does not appear to steal passengers from civil aviation or road transportation, but it manages to attract new passengers. Regarding the economic outcomes, service employment is much more responsive to HSR-induced market access changes, particularly those over shorter distance trips, than highway-induced ones, while manufacturing employment and GDP responds more to highway-induced market access changes.

One of the most important findings in this paper is the implications of an improvement in intercity passenger transportation on cross-city specialization patterns. Conceptually, better intercity passenger transportation reduces the cost of face-to-face interactions across space, and should exert larger impacts on industries that are communication intensive. To test it, I begin by estimating the impacts of HSR across sixteen Chinese industries, and compare the estimated coefficients to the task contents of each industry, as per Autor et al. (2003). It is revealed that the benefits of better intercity passenger transportation increase in the industry-specific requirement of nonroutine cognitive tasks, and decrease in their reliance on manual or routine cognitive skills. On the contrary, the estimated impacts of highway expansion do not correlate with the nonroutine contents of industries. These results highlight the distinctive role of HSR in shifting the specialization patterns of cities towards interactive industries, compared to other forms of transportation.

This paper contributes to a growing literature on estimating the economic impacts of transportation infrastructure projects. Recent papers have studied the skill premia in local labour markets (Michaels (2008), long-term GDP growth (Banerjee et al. (2012), income volatility (Burgess and Donaldson (2010), gains from trade (Donaldson (Donaldson), and asymmetric effects on core and peripheral markets (Faber (2014). Also, papers in urban economics have explored the effects of urban transportation improvements on urban growth (Duranton and Turner (2012)) and urban form (Baum-Snow et al. (2016)). Relative to the existing literature, this article draws attention to a different type of transportation infrastructure, the inter-city passenger transportation and a new mechanism, the sourcing of labour across cities.

This paper also contributes to an extensive literature on urban growth and specialization. Previous research has emphasized the importance of agglomeration effects (Glaeser et al. (1992), amenities (Clark et al. (2002), human capital (Glaeser and Saiz (2004b) and intra-city transportation (Duranton and Turner (2012). This paper shows that better intercity passenger transportation can also act as a new engine of urban growth. Since I focus on how improved intercity transportation facilitates labour sourcing across cities and reshapes urban employment patterns, the paper is also related to the international trade literature on offshoring, such as Grossman and Rossi-Hansberg (2008), and Ottaviano et al. (2013).

Finally, this paper provides a rigorous empirical evaluation of the largest HSR project in the world up to now. Earlier evaluations of HSR projects in Japan and Europe have presented mixed evidence. Sasaki et al. (1997) suggests that the Shinkansen in Japan promoted local development and did not cause regional inequality. Bernard et al. (2014) bernard et al. (2014) further examines the response of Japanese firms to a particular Shinkansen line with a focus on supplier relationships. Behrens and Pels (2012) documents a significant change in passenger travel behavior along the Paris-London corridor after the high speed Eurostar was in operation. Ahlfeldt and Feddersen (2010) present evidence that the HSR line connecting Cologne and Frankfurt in Germany substantially increases the GDP of the regions that enjoy an increase in accessibility. While Albalate and Bel (2012) discovers that French HSR has neither accelerated industrial concentration nor promoted economic decentralization from Paris. The HSR project in China provides me with an excellent opportunity to evaluate the potential economic benefits of HSR due to its large scale. An earlier paper on Chinese HSR Zheng and Kahn (2013) finds that HSR connection boosts housing prices in China, using cross-sectional data. My paper looks at a wider range of economic outcomes with more detailed HSR data and different identification strategies.

2.2 Background and Discussion of HSR Usage in China

In 2008, the State Council in its revised Mid-to-Long Term Railway Development Plan set the goal of a national high-speed rail grid composed of four north-south corridors and four east-west corridors, with a budget of around 4,000 billion yuan (Council (2004). The construction costs of HSR range from 80-120 million RMB per km (US$13-20 million) excluding stations Bullock et al. (2012). The expansion of the HSR network in China from 2003 to 2014, is shown in Figure 2.1. Detailed information on the construction start date, opening date, distance and speed of all the operating lines is listed in Table A3 and A4. The objective of this HSR grid, as stated in the Plan, is to connect provincial capitals and other major cities with faster means of transportation. The placement of lines, according to the Ministry of Railway (MOR), should be based on a comprehensive consideration of the economic development, population and resource distribution, national security, environmental concerns and social stability of each region MOR (2008). Finally, HSR lines are expected to complement existing transportation networks as much as possible.

Before exploring the role of HSR in city specialization patterns, we first need to understand how it affects the way people travel. I pay attention to two key questions. First, how does HSR compete with other forms of transportation? Second, what are the socioeconomic characteristics of HSR passengers? I explore these questions using collected passenger surveys and official ridership data.

Table A5 reports results from three pieces of passenger surveys, conducted by Jianbin (2011), Wu et al. (2013) and Ollivier et al. (2014), respectively. Interviewees are drawn from four HSR lines, two short lines, and two long-haul ones. I collect answers to four kinds of question: (1) passengers' income; (2) purposes of travel; (3) means of transport to the HSR station and whether or not transfers on and off HSR are made, and (4) their alternative intercity travel choice before the introduction of HSR. From the second column, we learn that the average monthly income of HSR passengers ranges from ¥4300 to ¥6700, which roughly falls into the high-income group in China.[3]. A large proportion (25% to 40% along shorter HSR lines and 40% to 60% along longer HSR lines) of the passengers travel for business purposes. Regarding the substitution between the HSR and other forms of transport, as reported in column 4, none of the HSR passengers on the two short journeys preferred to fly to their destination before the advent of HSR, while 36% (Changchun-Jilin) and 61.5% (Beijing-Tianjin) of them listed conventional railway as their primary choice of transportation at that time, and 50% (Changchun-Jilin) and 32% (Beijing-Tianjin) of them had preferred road transport (including coaches/buses and private cars). Over the long-haul trip (Tianjin-Jinan), a large proportion (77%) mentioned air travel as their main choice before the HSR was in operation, followed by 18% who chose conventional rail travel, and almost none considered long-distance coach journeys. It is clear that HSR mainly competes with air travel for longer trips and with traditional railway/road travel over shorter lines.

I collect ridership[4] data on HSR and other forms of transportation from two

[3]According to the Chinese Statistics Bureau, in 2013, the average monthly income of urban residents is ¥2462, residents with average monthly income over ¥4700 are categorized into the high-income group.

[4]Ridership is defined by the National Statistics Bureau of China as the total number of trips made on a particular kind of transportation device. For road ridership, it only includes paid road trips (coaches, etc.), and excludes self-driving trips. For railway, the trips can start and end at any stops, including both destinations.

sources: the railway yearbook series and Department of Transportation reports. Table A6 reports selected ridership data on most of the HSR lines in service from 2009 to 2012, collected from China Railway Yearbooks. We observe a clear increasing trend in ridership for all the lines listed that have more than one year's ridership data, showing people's gradual acceptance of HSR as a new form of transportation. Among different lines, the most heavily used ones are median-to-short lines connecting two central cities, such as the Beijing-Tianjin, Shanghai-Nanjing, and Guangzhou-Shenzhen lines.

Turning to the total ridership data across different forms of transportation, presented in Figure A1. As noted, more than 85% of the trips were made by cars, buses, and coaches, followed by somewhat more than 10% using the railway. Air and water transport services were relatively less used, accounting for less than 2%. From 2010 to 2014, we observe a steady increase of HSR ridership, from 300 million in 2010 to 830 million in 2014. Over the course of HSR expansion, few changes can be seen in the percentage of passenger trips made by air or water. But we do observe a slight drop in the proportion of passengers carried by conventional railway from 8% to a little less than 7% and by road from 87% to 86%, although the number of railway and road ridership increases steadily. This evidence is consistent with the view that conventional railways face the strongest competition from HSR, most likely because a few of the services on some conventional railway lines are cut when the parallel HSR starts operation (Qin (2016)) and the fact that conventional railway and HSR are closer to each other in the fare/time-cost trade-off spectrum.

To put these numbers into perspective, we can compare the transportation pattern in China to that in the US. In 2014, the total HSR ridership in China was 830 million, more than the combined ridership by air and intercity rail/Amtrak of about 698 million in the US [5]. On the contrary, the total passenger-miles count on US highway (excluding private passenger cars) doubles its Chinese counterpart[6]. In a relative term, HSR appears to be twice as important as air and intercity railway combined in the US.

[5]The data on passenger travel in the US are obtained from the National Transportation Statistics by the Bureau of Transportation Statistics. In 2014, the total passenger-miles on air were 607,772 million, and the average length of travel by air is 1,440 miles in 2013, wich translates into 679 million trips made by air. Similarly, we obtain a ridership of 19 million on intercity rail/Amtrak.

[6]In 2014, the total passenger-miles on highways that excludes passenger cars are 1,492,801 miles in the US. In China, total road ridership is 19,082 million, with an average length of 39.14 miles, which adds up to 744,198 million. The information is obtained from the Department of Transportation annual report.

2.3 Conceptual framework

The main focus of this paper is to examine the impacts of improved cross-city transportation on specialization patterns across cities. Conceptually, when we think about the differential impacts of HSR connection on different industries, the reliance of these industries on face-to-face contact is essential. By facilitating face-to-face interactions of people, particularly skilled workers, HSR effectively reduces the unit cost of production in communication-intensive sectors, either through productivity boost brought by intense knowledge sharing or through enhanced remote sourcing.

To formalize the intuition on the general equilibrium effects of an improvement in cross-city passenger transportation, I begin with a simple model of labour sourcing based on Grossman and Rossi-Hansberg (2008)[7]. In this model, a tradable final good is produced using multiple tasks, and the production of a task can take place in one city while using the technology from another city, subject to communication costs between them. Therefore, a reduction in face-to-face communication cost through HSR connection allows a high productive city to source more tasks from other cities. Some quantitative predictions are generated from the model. The model predicts that the benefits of HSR connection work through enhancing the passenger access to other cities, which leads to employment growth proportional to the change in HSR-induced market access measure, approximated using an average of other cities' GDP inversely weighted by the bilateral costs of passenger travel. One could interpret an increase in this "passenger market access" not only as better chances in labour sourcing, but also as better access to other cities knowledge or customer pools.

The model features a single sector, where an improvement in market access translates directly into growth in aggregate city employment. To further evaluate the implications of HSR-induced market access on urban specialization, it could be extended to a multisector one. Think of the simplest case with two different industries, one interactive and one non-interactive; and two different sets of tasks, interactive and non-interactive. The production of the final good in the interactive industry depends on interactive tasks only, and vice versa. A reduction in passenger travel cost reduces the unit cost of interactive tasks production, to an extent that is proportional to the growth in passenger market access. With free final goods and labor mobility across cities and industries, HSR connection leads to a shift in comparative advantage towards the interactive industry, and ultimately to relatively higher growth in this industry in connected cities.[8]

[7]Details of the model are reported in Appendix A.

[8]It is noted that the aggregate cross-industry relocation of employment will depend on the equi-

Based on this line of thinking, I derive the following empirically testable hypothesis:

Hypothesis *HSR connection will lead to relatively higher growth in employment of communication-intensive industries in connected cities. And the benefits of HSR connection will appear as improved passenger access to other cities.*

To obtain a sensible measure of industry-specific dependence on face-to-face interactions, I consider the types of tasks required in each industry following Autor et al. (2003). They divided the required task contents in a particular industry into the following five categories: routine manual, routine cognitive, nonroutine manual, nonroutine analytical and nonroutine interactive; further, they came up with a measure of task intensity of these five types across 140 consistent census industries. According to their analysis, industries high in nonroutine analytical and interactive tasks involve more abstract thinking, problem-solving and complex communication activities. Naturally, a reduction in communication cost will more directly impact the cross-city employment patterns of these industries, other than the ones that focus more on manual or routine cognitive tasks.

The impacts of HSR on city economic outcomes are realized through improvements in the accessibility of connected cities to other cities as a result of a faster and more convenient means of transportation. Therefore, we should expect larger impacts for cities that are connected to a greater number of more prosperous destinations. In my subsequent empirical analysis, I will first check how direct connections to the HSR network lead to employment growth across different industries, before exploring the relationship between HSR-induced market access changes and urban specialization patterns.

2.4 Empirical Specification

2.4.1 Data

Prefecture-level socioeconomic data are drawn from China City Statistical Yearbooks from 2000 to 2013 and China Regional Economic Statistical Yearbooks from 2000 to 2011, since many of the variables are missing for years before 2000. The City Statistical Yearbook series report prefecture-level passenger ridership and volumes of goods transported by different modes of transportation, GDP, population, em-

librium relative price of the two final goods. Therefore, the model may have clear predictions over the absolute effects of HSR on the interactive industry employment in connected cities. However, to generate an expansion in interactive industry as a whole, we can introduce further assumptions on non-homothetic preferences or international markets to regulate the possible relative price changes.

ployment in 18 sectors[9], average wage, local government revenue and expenditure, total number and revenue of industrial firms, total number and sales of retail firms and a variety of city level infrastructure measures. The Regional Economic Statistical Yearbooks report statistics on prefecture-level housing prices. In my analysis, I only focus on prefecture-level cities, excluding prefecture-level autonomous regions. Haikou and Sanya are left out of my sample because they are cities on the Hainan Island and their accessibility to all the other cities cannot be changed by HSR connection easily. Some key socioeconomic data are missing along the time series for a few cities, which leaves me with 278 cities to work with throughout most of my analysis. Table A1 summarizes the source, year range, the total number of observations, number of cities with at least one year of observation and number of cities without missing values along the time series of all the outcome and control variables I use in this paper. Since I control for average city and provincial level GDP and population growth in the past three years throughout my analysis, I am effectively using only outcomes from 2003 to 2013 as my dependent variables.

The yearbooks also report ridership data on railway, road, air and water. Ridership is defined as the number of paid trips made on each form of transportation. Self-drive trips and trips on public transportation are not included in road ridership count. It should be noted here that ridership here is not limited to intercity ridership by definition. Road ridership could include coach trips made across towns or villages within the same prefecture city. But we have reasons to believe that most parts of the ridership come from intercity trips, especially for railway and air travel, as a city typically has only one main railway station and airport.

A prefecture-level city usually has an urban core (Shixiaqu) that consists mainly of urban residents and surrounding counties with a relatively larger proportion of rural population. For each variable, two separate statistics, one aggregated only to urban-ward (Shixiaqu) level and the other covering the whole area of the prefecture, are reported in the yearbooks. Throughout my analysis, I use the statistics counted at urban ward (Shixiaqu) level of prefecture cities since I am interested in the employment and resources flows across urban areas.[10]

High speed railway (HSR) lines are defined as railway lines running at an average speed of 250km/h or more, or passenger-dedicated-intercity-lines running at an average speed of 200km/h or more[11]. As the end of 2014, there were 43 HSR

[9]Detailed descriptions of these sectors, as well as the comparison with NAICS 2-digit industries, are reported in Table A2

[10]Although I stick to statistics counted at the urban core level. This level of aggregation is not available for some variables, such as ridership on variables modes of transport, and I have to use the one counted at the whole prefecture level.

[11]Major technical stipulation on railway, the Ministry of Railway, 2012.

lines in operation, with a total mileage of 11152 km (Table A3). Most information on the Chinese HSR system, including construction starting date, open date, length, designed speed and ridership on selected lines, is obtained from the Major Events, Finished and Ongoing Projects sections in the China Railway Yearbooks from 1999 to 2012. For a small proportion of lines that are opened in 2013 and 2014 and future HSR lines in plan, this information is not available from the most updated (2012) railway yearbook, so I have to rely on official news published on http://news.gaotie.cn as well as other online news sources[12] I doublechecked the information on the stops along each existing line from the official railway service website (www.12306.cn). Geo-referenced administrative unit data, as well as conventional railway routes, are obtained from the ACASIAN Data Center at Griffith University in Brisbane, Australia. Highway networks data of China in 2000, 2002, 2003, 2005, 2007 and 2010, from Baum-Snow et al. (2016) ,are kindly shared by the authors.

To check other mechanisms at work, I also bring in patent application data in China from the SIPO (State Intellectual Property Office) as a proxy for innovation activities within a city. In China, patents are divided into invention and utility-model patents, with the former category being more stringent and lasts longer.

2.4.2 Definition of market access variables

As briefly discussed in section 3, the benefits of HSR connection appear as improved passenger access to other cities. To better capture the treatment effects of HSR connection, I introduce measures of market access and examine the impacts of market access growth, induced by HSR or highway, on cross-city transportation and specialization patterns.

In practice, the market access of city k is defined as $MA_k = \sum_{j=1}^{N} \tau_{kj}^{-\theta} X_j$, where τ_{kj} is the travel cost between city k and j, and X_j is the GDP of city j [13]. Intuitively, a city enjoys higher market access if it boasts lower transportation cost to larger cities. τ_{kj} can be reduced through the expansion of transportation infrastructures, such as HSR or highways.

I use four market access measures throughout my analysis:

(1) Market access measure that captures both highway and HSR network expansion (MAall)

[12] Information on the construction starting date, designed speed and length for lines that started construction before 2012 is also included in the Railway yearbooks major events section. But I need to rely on online news sources for their exact opening date.

[13] This specification of market access can be derived from a labour sourcing model described in Appendix A.

(2) Market access measure that captures HSR network expansion only (MAHSR). Given the highway network of each year, the changes in market access purely driven by HSR network expansion are counted.

(3) Market access measure that captures highway network expansion only (MAhighway). Assuming that there is no change in HSR network, I calculate the changes in market access compared to the base year (2000) purely attributable to highway network changes.

(4) Market access measure that captures indirect HSR network expansion only (non-connection-induced market access, NCIMA): For a city i connected in year t, the market access changes for it using a counterfactual HSR network that bypasses city i but is otherwise the same as the real one. This measure considers only HSR expansion.

(5) Market access measure that captures the impacts of HSR on short distance trips (less than five hours) only (MAHSRless5). Only changes in bilateral travel costs for trips less than five hours are counted in this measure. This measure considers only HSR expansion.

These different measures allow me to examine differential impacts of different sources of improvements in intercity connectivity. We can look at the responses of ridership and economic outcomes to market access induced by highway and HSR connection separately.

The calculation of the market access variable requires the construction of a time-varying transportation cost matrix, τ_{km}, for each city pair. In my definition, I allow τ_{km} to incorporate both time and fare cost in travel. Conventional roads, highways and HSR represent different fare-time cost combinations and passengers face tradeoffs between fare and time cost. The upgrade from conventional roads to highways and to HSR changes τ_{km} by offering passengers alternative options to travel. The information on highway network is obtained from Baum-Snow et al. (2016), and the timing of HSR network expansion is shown in Table A3.

We have to rely on a few assumptions to construct τ_{km}, and the details are reported in Appendix E.

In Figure 2.2, I plot the distribution of log(MAHSR), the market access measure that considers only HSR expansion, both across cities and through time. The left graph presents the distribution of log(MAHSR) over all the observations. And the right graph shows the distribution of the residuals of log(MAHSR) conditional on city fixed effects. These graphs give us a basic idea on the pooled and within city variation of log market access. Meanwhile, the lower panel of Table 2.1 shows the mean and standard deviation of five logged market access measures from 2001 to

2014. The growth in HSR-driven market access during this period is roughly 2.4% and the growth in non-connection-induced market access is about 1.9%.

2.4.3 Difference in differences specification

In this section, I explore the aggregate impacts of HSR connection by regressing my outcome variables on a HSR connection dummy. The baseline estimation strategy is a difference in differences specification of the form:

$$ln(y_{it}) = \alpha_i + \beta_{rt} + \gamma * Connect_{it} + Controls_{it} + \epsilon_{it} \tag{2.1}$$

where y_{it} is an outcome of interest of city i within region r in year t, α_i is a city fixed effect, β_{rt} is a region[14] by year fixed effect, and $Connect_{it}$ is an indicator of whether city i of region r was connected to HSR in year t. The error term ϵ_{it} is clustered at the city level. Standard errors allow spatial dependence decaying in distance as in Conley (1999).

To test hypothesis 1, I divide the total employment into four groups: tourism-related employment, which includes hotels and catering services and wholesale and retail trade; skilled employment, which includes finance and insurance, real estate, information, business services, scientific research and technical services and education; other service employment, and other non-service employment. Apart from employment, I also look at other aggregate outcomes at city level, which include GDP, housing price, total fixed investments, retail sales, and total patents application.

The identifying assumption of difference-in-differences estimations is the parallel trends of outcomes between HSR-connected cities and the other cities should there be no HSR. However, if an HSR placement decision is based on past growth and expected future growth, this assumption may possibly not hold in reality. To mitigate this problem, I restrict my sample to 172 cities that are either connected or planned to be connected by HSR by 2020.

The primary identification challenge is not whether a city is connected to the HSR network, but rather what factors determine the timing of its connection. Several idiosyncratic factors appear to influence the opening time of HSR lines. First, as is evident from Table A3, the construction work on many of the existing lines (12 out of 45) began in 2005, following the passage of the Mid-to-Long Railway Plan in 2004. The timing of when each line opens is determined by the construction

[14] Cities are divided into 8 regions (Northeast, Northern Costal, Eastern Coastal, Southern Costal, Southwest, Northwest, the middle reaches of Yangtze River, the middle reaches of Yellow River.

progress, which largely depends on engineering difficulties. In Appendix F.1., I exploit variation in HSR connection timing purely driven by engineering difficulties and find similar effects on railway ridership, service and private employment as in the main regressions.

Another obstacle to identification is the possibility of simultaneous investments in other areas. Local governments may take HSR connection as a new engine for local economic growth and invest heavily in related projects. This government spending would create jobs that can be mistakenly interpreted as the direct impacts of HSR connection. At this stage, to deal with these concerns, I control for local government spending, other infrastructure measures such as the length of roads above a certain standard, the total area of new urban roads built, the area of green land, the number of public facilities such as theatres, hospitals, and public libraries. [15] Controlling for region-by-year fixed effects should be able to take care of any region-specific yearly shocks to local economic conditions. Apart from that, I also control for average city and province GDP and population growth for the past three years, as well as interactions of year dummies with distance to regional central cities for fear that the HSR connection decisions of a line connecting multiple nearby cities are correlated with temporary regional shocks at different levels, and that geographical centrality is both correlated with economic growth and connectivity improvement.

To check the parallel trend assumption, I run a variation of equation (1), controlling for the leads and lags of the initial connection dummy.

$$ln(y_{it}) = \alpha_i + \beta_{rt} + \sum_{m=1}^{3} \gamma_m FirstConnect_{i,t-m} + \sum_{n=0}^{4} \gamma_n FirstConnect_{i,t+n} + Controls_{it} + \epsilon_{it} \tag{2.2}$$

where $FirstConnect_{it}$ is a dummy variable indicating whether a city is first connected to the HSR network in year t. It switches to 1 only if the HSR line connecting city i is opened in year t. $FirstConnect_{i,t-m}$ is its m-th lag, and $FirstConnect_{i,t+n}$ is its n-th lead. Controlling for leads allows me to examine the pre-HSR effects of future railways as a placebo test and helps to disentangle anticipatory effects from actual connection effects. Controlling for lags enables me to trace the treatment effects in the years after initial connection. In reality, we expect anticipatory effects

[15] A potential problem with including these controls on the right-hand side of the regressions is that government spending and investment itself is endogenous—a government expecting better growth prospects invests more with or without HSR. However, if the correlations between the unobservable and government spending and that between HSR connection and government spending are both positive, then the coefficients on HSR connection or market access are underestimated by controlling for local government spending.

to be relevant for some outcome variables but not for others. For example, housing prices can respond positively to the HSR connection before the railway becomes operative because news about a future HSR station is usually known about five years before it opens and people make investment decisions based on this information. However, HSR-induced changes in employment are expected to be observed only after actual HSR lines are in operation since people can only benefit from HSR for travelling and commuting after it is in operation.

2.4.4 Market Access Approach

As briefly outlined in the conceptual framework, the benefits of HSR connection appear as improved passenger access to other cities. In this section, I use HSR-induced market access measures as the main independent variables to capture finer variations in the treatments of HSR connection.

One major empirical advantage of this market access approach is that it allows me to exploit the variation in a city's access to other cities driven by HSR expansion but has nothing to do with its own HSR placement. In short, an unconnected city can still benefit from HSR connection if a city close to it receives connection, which allows its passengers faster trips to other destinations through transfers. It effectively deals with the identification challenges of both endogenous route placement and simultaneous investments. In my main empirical specification, I examine the response of outcomes of interest to increase in a non-connection-induced market access for both connected[16] and unconnected cities.

Another benefit of this approach is that it enables me to examine differential impacts of different sources of improvements in intercity connectivity. We can look at the responses of ridership and economic outcomes to market access induced by highway and HSR connections separately. Another market access measure worth considering is the one that captures HSR impacts on shorter trips only. As I do not explicitly account for the changes in air travel costs during this period, I could be overestimating the effects of HSR on connectivity if passengers prefer air travel to HSR over longer-distance trips. Focusing on the impacts of HSR-induced market access changes for short distance trips only helps to alleviate the concern, as HSR holds strong advantages over air travel for shorter trips.

Throughout my analysis, my preferred specification is:

[16]For a connected city, its non-connection-induced market access is defined as its access to other cities through a hypothetical HSR network which is the same as the actual one except for bypassing this city.

$$ln(y_{it}) = \alpha_i + \beta_{rt} + \gamma * NCIMA_{it} + \theta * MAhighway_{it} + Controls_{it} + \epsilon_{it} \quad (2.3)$$

where y_{it} is an outcome of interest of city i within region r in year t, α_i is a city fixed effect, β_{rt} is a region by year fixed effect, $NCIMA_{it}$ and $MAhighway_{it}$ are non-connection-induced market access and highway access respectively.

Two concerns over identification using the non-connection-induced market access measure may arise here. First, apart from basing the HSR placement decision on the economic conditions of each city, the railway authority might design the network expansion plan to maximize the market access of connected cities intentionally and target at prosperous regions first. So "better located" cities that are close to other developed cities might experience larger increase in market access that may or may not depends on its own connection. Thus, throughout my analyses, I control for the past average city and provincial GDP as well as the interaction terms of year fixed effects and distance to provincial capital cities. Another challenge for identification is the common regional shocks that might jointly affect a city's economic outcome and the HSR placement of nearby cities. It is less of an issue after controlling for region-year effects and past city and provincial average GDP. Furthermore, I also adopt a specification similar to equation (2), by including the leads and lags of the increments in non-connection-induced market access (NCIMA) as main independent variables. The formal specification is:

$$ln(y_{it}) = \alpha_i + \beta_{rt} + \sum_{m=1}^{3} \gamma_m \Delta logNCIMA_{i,t-m} + \sum_{n=0}^{4} \gamma_n \Delta logNCIMA_{i,t+n} + Controls_{it} + \epsilon_{it} \quad (2.4)$$

Similarly, this specification helps us to check if the growth in the non-connection-induced-market-access (NCIMA) measure correlates with the trends of a variety of outcomes prior to the year when the actual increase in market access takes place. I plot the coefficients on the leads and lags of $\Delta logNCIMA$ in Figures 2.4 and 2.7.

2.4.5 Event-study for non-connection-market access

A potential challenge for identification using my non-connection-induced market access (NCIMA) measure is that some cities closer to each other may be able to drag HSR lines to their region, and their incentive and capacity of lobbying is correlated with their growth prospect. Controlling for region-specific year fixed effects may not be able to fully address this issue as the collective bargaining might occur at a smaller geographical scale. So apart from controlling for past provincial GDP, I also rerun my non-connection-induced market access regressions by including the leads and lags of increments in market access. If regions with better growth prospects are more likely to get HSR placement, then we should expect positive coefficients for leads of market access increments in the cities in these regions, no matter whether they are connected or not. The specification is:

The results are shown in Figures A2 and A3. We do not observe large differences in the pretrends across cities that experience higher and lower non-connection-induced market access growth, and there are clear trend breaks around the time when a city actually experiences changes in market access.

2.5 Estimation Results

2.5.1 Transportation patterns

I start by examining the effects of HSR connection on passenger transportation patterns before studying its implications on economic outcomes. The upper panel of Table 2.2 presents the difference-in-differences estimation results on the usage of different modes of transport. Controlling for city and region-by-year fixed effects, it is observed from Column (1) that HSR connection significantly increases railway ridership by 18%. The coefficient of total ridership on all forms of transportation (Column 4) is a smaller 9.6%) but still positive and significant, suggesting that HSR creates extra demands for transportation, and the poaching of ridership from other forms of transport is small. From Columns (2) and (3), it is clear that HSR connection leads to a small positive effects on road ridership, but a large 12% drop in the number of passengers who travel by air, which indicates that travelling by HSR is a close substitute for air travel, but not road transportation. It is also worth noting that no significant impacts of HSR are observed on the volume of goods transported by railway (Column 5), confirming our intuition that HSR is bringing changes to intercity passenger travel cost, but not goods trade cost.

A back-of-the-envelope calculation suggests that HSR brings in about 901 million extra passengers since its inception until 2013, taking into consideration the substitution and complementarity relation between HSR and other forms of transport [17]. To put this number in perspective, my HSR ridership data on the main existing lines (Table A6) suggests that the aggregate HSR ridership on 11 major lines (out

[17] My estimation suggests that cities with HSR connection experience 9.6% increase in total ridership. Using 2008 total ridership of all the cities with HSR connection by 2013 (9387 million) as a benchmark, this translates into 901 million new ridership.

of 28) in 2012 is 320 million, and the aggregate annual HSR ridership data reported by the Department of Transportation is 300 million in 2010, 420 in 2011, 470 in 2012 and 602 in 2013. It is observed that the estimated aggregate increase in total ridership is about half the size of the aggregate HSR ridership during this period. The difference could either be attributed to the substitution between HSR and other means of transport, most notably conventional railway and air travel, or due to the fact that in this simple difference-in-differences design, we assume cities with no HSR station to be 'untreated' while in reality, they could still benefit from HSR connection in other cities by transferring. But without a doubt, such sizable newly generated passenger flows, with a large proportion of business travellers, indicate large improvement in the communication and economic ties across Chinese cities, yielding large impacts on local economic outcomes possible.

The lower panel of Table 2.2 reports the responses of passenger travelling patterns to a city's non-connection-induced HSR and highway market access changes. The estimates are largely consistent with our intuition about the substitution and complementarity between different modes of transportation. One percent increase in HSR-induced market access[18] translates into an 8 percent increase in railway ridership, 2 percent drop in road transportation, and 7.4 percentage drop in air travel, while a one percent increase in highway-induced market access leads to 5.7% growth in road ridership and 4.2% drop in railway ridership and do not seem to affect air travel negatively. In sum, both HSR and highway expansions increase total passenger ridership but the estimated elasticity is larger for highway market access.

Table 2.3 reports how ridership on different forms of transportation responds to other market access measures, including the market access measure that incorporates both HSR and highway expansions (MAall), and that limited to short distance trips. For each variant of the market access measures apart from MAhighway, the regressions control for both the connection dummy and highway market access (MAhighway), to show that the main results on the relationship between market access and employment growth are not driven by direct connections *per se* or simultaneous highway expansion. It is worth noting that for the market access measure capturing only HSR impacts over short distance trips (MAHSRless5), the estimated elasticities are in general larger in magnitude, showing that connectivity improvements are more important when they decrease the travel costs between cities that are closer to each other. Interestingly, the estimated elasticity is six times larger in absolute value than the one in baseline for air travel ridership, but only less than

[18]Summary statistics on different measures of market access are reported in Table 2.1. From 2001 to 2014, the average NCIMA grows by 2.23% as a result of HSR expansion, and the average MAhighway grows by 1.86% following concurrent highway expansion.

doubles for railway ridership. A story consistent with this observation is that the substitution between HSR and air travel is strong over shorter-distance trips: HSR expansion at a closer range leads to a much larger drop in air travel ridership. For longer distance trips, civil aviation faces little competition from the newly available HSR service so the increase in HSR ridership is more likely to be newly-generated passenger flows. However, even for short distance trips where the substitution between HSR and air travel is stronger, the effect of a boost in market access on total ridership is large and positive, with an estimated elasticity of 8.2%. Piecing together the evidence, it is not hard to conclude that the estimated impacts of different aspects of market access changes echo with our intuition on the ways different modes of transport compete with each other and the qualitative evidence shown in section 2. Therefore, we should have more confidence in the validity of the "market access" approach and the estimation results on the impacts of HSR-led market access changes on economic outcomes of interest.

2.5.2 Specialization patterns

In my conceptual framework, I hypothesize that industries that rely more on communication and interpersonal skills will benefit more from HSR that reduces the cost of face-to-face interaction and communication across cities. In this section, I will directly test the implications of HSR connection on the specialization patterns across cities. I will first roughly divide total employment into skilled service employment, tourism-related service employment, other service and non-service employment and examine the impacts of HSR on them separately. I will then carry on a more detailed analysis on the effects of HSR on 16 industries and rank the estimated effects according to the cognitive content of each one.

The estimated coefficients on connection dummy are presented in Table 2.4. A city is estimated to experience a statistically significant 7% higher growth in employment after being connected to HSR compared to the unconnected. Columns (2) to (4) report the results on employment of four subcategories[19]. It is certain that the aggregate employment effects are largely driven by growth in tourism-related employment, which grows by 13% following the HSR connection. The effect on skilled employment (4.8%) is moderate, lower than that on non-service employment such as manufacturing, utility, and construction. One possible explanation is that the HSR connection effects on non-service sectors, such as manufacturing and con-

[19]Total employment is divided into tourism-related employment (retail and wholesale, hotel and restaurant), skilled employment (FIRE, IT, business services, research and technical services and education), non-service employment (agriculture, manufacture, construction, utilities) and other service employment.

struction, are driven by subsequent investments in public infrastructure and fixed assets. Market access results reported in the lower panel of Table 2.4 supports this interpretation. One percent increase in NCIMA leads to 2.5 and 3.9 percent growth in skilled and tourism-related employment, both higher than the estimated elasticity on aggregate employment, but exerts almost no effects on other non-service employment. It suggests that the benefits of better passenger access to other cities are mostly limited to skilled and tourism-related employment. Therefore, the significant positive impact of direct HSR connections on construction and manufacturing comes partly from subsequent investments accompanying HSR connections. On the contrary, highway-led market access growth leads to growth in non-service employment, most notably manufacturing, construction and utility, which is consistent with the idea that highway expansion has larger effects on trade cost, which benefits the tradable sectors more.

Table 2.5 reports the estimation results using alternative market access measures, controlling for both the direct HSR connection and highway market access expansion. The evidence presented here is consistent with the main effects: both HSR-driven and total market access growths have positive impacts on total employment, and the coefficients are largest for tourism-related sectors, followed by skilled sectors. Controlling for direct connection dummy does not take away much of the explanatory power of this non-connection-induced market access (NCIMA), compared to the results presented in Table 2.4, which suggests that this measure captures mostly market access improvements orthogonal to cities' own connections. The estimated elasticities are also larger in magnitude for the market access measure that captures only HSR impacts over short distance trips (MAHSRless5), which is consistent with the transportation patterns results. It shows that the connectivity boost across cities that are relatively close to each other plays a more important role in explaining both variation in railway ridership and employment growth.

Figure 2.4 shows event study results on employment of different categories, specified by equation (4). In these graphs, I plot the estimated elasticity on the leads and lags of the changes in non-connection-induced market access measure (NCIMA). It is clear that the coefficients on the leads of NCIMA are close to zero for all four categories of employment, which supports the parallel trends assumption: the boost in HSR-induced market access does not correlate with previous trends in employment across different sectors. Meanwhile, all four categories of employment grow gradually following HSR-induced market access boost, and the effects are largest for tourism-related and skilled service sectors. The treatment effects of HSR-induced market access are smallest in magnitudes for non-service employment, which is different from the results obtained using the connection dummy.

Subsequently, I present the estimated HSR effects across sixteen different industries in Table 2.6. It is clear that most industries respond positively to both non-connection-induced market access (NCIMA) and highway induced market access (MAhighway), but there exists considerable heterogeneity across sectors. For instance, manufacturing employment responds significantly to highway-led market access growth but little to HSR-induced one, while employment in finance, retail, medical services, etc. grow significantly as a result of better passenger access through HSR.

To establish a closer link between an industry's employment response to better intercity passenger transportation and its dependence on human interactions, I compare the industry-specific estimated coefficients to Autor et al. (2003)'s measures of the task requirement of each industry in nonroutine interactive activities, quantitative reasoning, routine cognitive skills, nonroutine and routine manual activities. In their context, routine tasks are defined as procedural, rule-based activities that are codifiable, and nonroutine tasks as those relying a lot on abstract skills such as problem-solving, intuition, persuasion, and creativity. We believe that industries with a focus on nonroutine tasks should benefit more from HSR connection as it improves face-to-face contacts of people across space.

It is evident from the left panel of Figure 2.5 that there exists a strong positive relationship between the importance of nonroutine interactive cognitive skills and the estimated impacts of better HSR-induced market access across all 16 industries. Industries that require a lot of intuition, creativity and human interactions such as FIRE, IT and business services both have higher measures of nonroutineness and respond more strongly to HSR-induced market access. On the contrary, the correlation between the impacts of highway-induced market access and the industry-specific task interactive task requirement drops to almost zero, as shown in the right panel. It suggests that HSR promotes specialization towards more interactive tasks much more than the highway and shows a clear distinctive role of passenger-dedicated transportation device in shaping cross-city specialization patterns. Similar contrasts can be drawn by comparing the sector-specific impacts of HSR and highway to the manual task requirement of each industry. Contrary to the patterns with interactive task content, we discover in Figure 2.6 that the estimated impacts of HSR decrease in the reliance of each industry on manual tasks, while the impacts of highway do not differ much across industry-specific manual task intensity. Analogous results on the other two dimensions of task intensity (routine cognitive skills and nonroutine quantitative reasoning skills) are presented in A2 and A3, with similar patterns: the estimated HSR impacts on sector-specific employment increase in nonroutine quantitative skills but decrease in routine cognitive skills.

Finally, I explore deeper into the impacts of HSR on tourism. It is clear that employment in catering and retailing experience huge boosts following the HSR connection. In the left graph of Figure 2.5, both the coefficients on retailing and catering employment lie above the fitted line, suggesting that there are factors other than the reliance on "people skills" that contribute to the observed impacts of HSR on these two industries. It would be interesting to account for the contribution of tourism to the observed patterns. In Table A7, I check if the responses of tourism-related economic activities to HSR connection is disproportionally larger for cities with better tourism resources. The China National Tourism Administration (CNTA) has divided major Chinese tourist attractions into five categories based on the code "Categories and Rating Standard of Tourist Attractions". [20] Here I define the indicator for tourism resources of a city as the number of 5A-class tourist attractions in that city. An additional 5A tourist attraction makes a city's total employment 26% more responsive to growth in market access compared to a city with none, but its retail/wholesale employment 130% more responsive.[21] Similar patterns show up if we compare the results across GDP and the total sales of wholesale/retail firms. The coefficient on the interaction term of tourism resources and market access is negative for GDP but is substantial and significant for retail sales. An additional 5A site makes the effect of HSR market access 56% stronger compared to the zero 5A site baseline on retail sales.

2.5.3 Broader Economic Impacts

In the previous sections, I have presented evidence on the impacts of HSR in shifting urban specialization patterns towards more communication-intensive and human-centred sectors. But it still remains a question whether or not HSR connections will exert any aggregate effects at the city level to justify the cost of HSR construction. In this section, I take a look at the impacts of HSR on other aggregate outcomes, such as GDP, investments and housing price.

The results are reported in the first column of Table 2.7. The direct HSR connection effect on GDP is estimated to be 3%, though not statistically significant,

[20] Chris Ryan, Gu Huimin and Fang Meng (2009). "Destination planning in China". In Chris Ryan and Gu Huimin. Tourism in China: Destination, Cultures, and Communities (1 ed.). pp. 11–37. ISBN 9780203886366.

[21] Both ratios are calculated by dividing the coefficients on the interaction terms of tourism resources and log(NCIMA) and the coefficients on log(NCIMA), yielding 0.571/2.187=0.26 and 2.826/2.171=1.3, respectively.

which is very similar to the estimates of 2.7% to 4.7% in Ahlfeldt and Feddersen (2010). Housing price also grow by 3% following the HSR connection. The patterns are similar with market access measures as the explanatory variables, reported in the lower panel of Table 2.7. One percent growth in non-connection-induced market access leads to significant 2.1%, 3.9% and 1.7% increases in aggregate employment, GDP and housing price. In the meantime, market access growth created by highway expansion only significantly improves GDP with an elasticity of 2.1%, with little effects on employment. It is consistent with the fact that the reduction in trade costs has larger impacts on goods market as opposed to the labour market, and consequently exerts more substantial effects on GDP other than employment.

To check if the GDP growth is largely driven by simultaneous investments, I observe the effects of HSR on the aggregate fixed investments of all the industrial firms[22] in China in column (4). The coefficient on connection dummy is negative and statistically significant, suggesting a decrease in fixed investments following HSR connections. I believe it is mostly due to the fact that the construction of HSR lines and stations leads to a boost in fixed investments before the actual operating of HSR lines. Considering the next row, the total fixed investments of all industrial firms do not respond to HSR or highway driven market access changes, which indicates that the observed GDP growth is not a result of expansion in fixed assets. On the contrary, HSR connection is associated with a significant 10% growth in total wholesale and retail sales, with an estimated elasticity in response to NCIMA growth being a significant 3.4. It confirms that a boost in retailing accounts for a significant part of the HSR-driven growth in GDP.

Finally, it is conceivable that HSR might generate growth in innovation through knowledge spillovers and enhanced scientific collaboration across cities. In the last column of Table 2.7, I check whether or not HSR connection leads to growth in patenting activities. The coefficient is positive but not statistically significant. However, in the last figure of Figure 2.7 and A4, I find a strong effect of direct HSR connections on patent applications after 2 to 3 lags, and a positive and significant coefficient on the third lag of NCIMA, as well. A plausible explanation is that innovation and patent application take time, and it is natural for the response in patents to take longer to show up.

[22]For industrial firms, the data include all the state-owned firms and all the private firms with industrial output larger than 5 million RMB per year. For wholesale/retail firms, the data include all the firms with annual sales greater than 5 million RMB

2.6 Robustness Checks

2.6.1 Robustness to alternative parameters and specifications

Table A9 reports the sensitivity analysis of the estimated market elasticity to alternative definitions of market access.

First of all, in constructing the travelling cost matrix, I assume that on average two trips per month are necessary for a task to be sourced, which is quite arbitrary. So I experiment with other required number of trips ranging from 0.5 time per month to 10 times per month. Given the same expansion in transportation networks, the percentage increase in market access induced by HSR should be mechanically larger for more frequent face-to-face contacts requirement. As shown in rows (1)-(4) of Table A9, most of the coefficients on ridership, employment, GDP and housing price remain positive and statistically significant at 0.01 level.

The calculation of market access also requires an approximate of θ, a parameter on the productivity distribution of cities. Following Donaldson (Donaldson), my baseline measure assumes a value of 3.6. For robustness, I adopt two other values of θ proposed by Eaton and Kortum (2002), 8.28 and 12.86. The signs and significance levels of most of the estimates are maintained with different θ.

2.6.2 Results on cities without an airport

As briefly mentioned, I do not incorporate air travel into my calculation of market access because air travel is less popular in China compared to railway and highway, and it is hard to get reliable information on the fare cost of planes as it fluctuates a lot across time. A potential problem with this oblivion is that I could be overestimating the impacts of HSR on city connectivity if planes and HSR are very close substitutes. Although I have shown that the substitutability of HSR and air travel is limited over mid-to-long distance trips, the increase in HSR ridership far more than compensate for the drop in civil aviation usage, and my results are robust to an alternative market access measure that captures only changes over short distance trips that face little competition from air travel, it makes sense to double-check my main results with a sample of cities that do not have an airport by the end of year 2013. The information on civil airports is obtained from the official website of CAAC (Civil Aviation Administration of China).

As shown in Table A8, The estimated coefficients on railway ridership and total ridership are larger in this small sample, and those on most economic outcomes are slightly lower than the baseline results. But the sign and significance of all the main results are preserved in the subsample of cities without airports.

2.7 Conclusion

This paper aims to make two primary contributions. First, I evaluate the impacts of one of the largest transportation infrastructure projects in the world, the high speed railway project in China, on local passenger travel patterns and economic outcomes. Second, I extend Donaldson and Hornbeck (2016)'s market access approach to evaluate different sources of intercity transportation improvements on cross-city sectoral employment patterns.

Despite the abundance of research on the cost of goods transportation and trade patterns, minimal attention has been paid to the cost of moving people or to the implications of it for the labour market. This paper exploits China's high speed railway (HSR) network expansion as a source of plausibly exogenous variation in passenger travel cost across Chinese cities over time. I find that an increase in HSR-induced market access leads to higher growth in industries with higher requirements in nonroutine cognitive skills rather than manual or routine cognitive skills, which is notably distinctive from the effects of highway development.

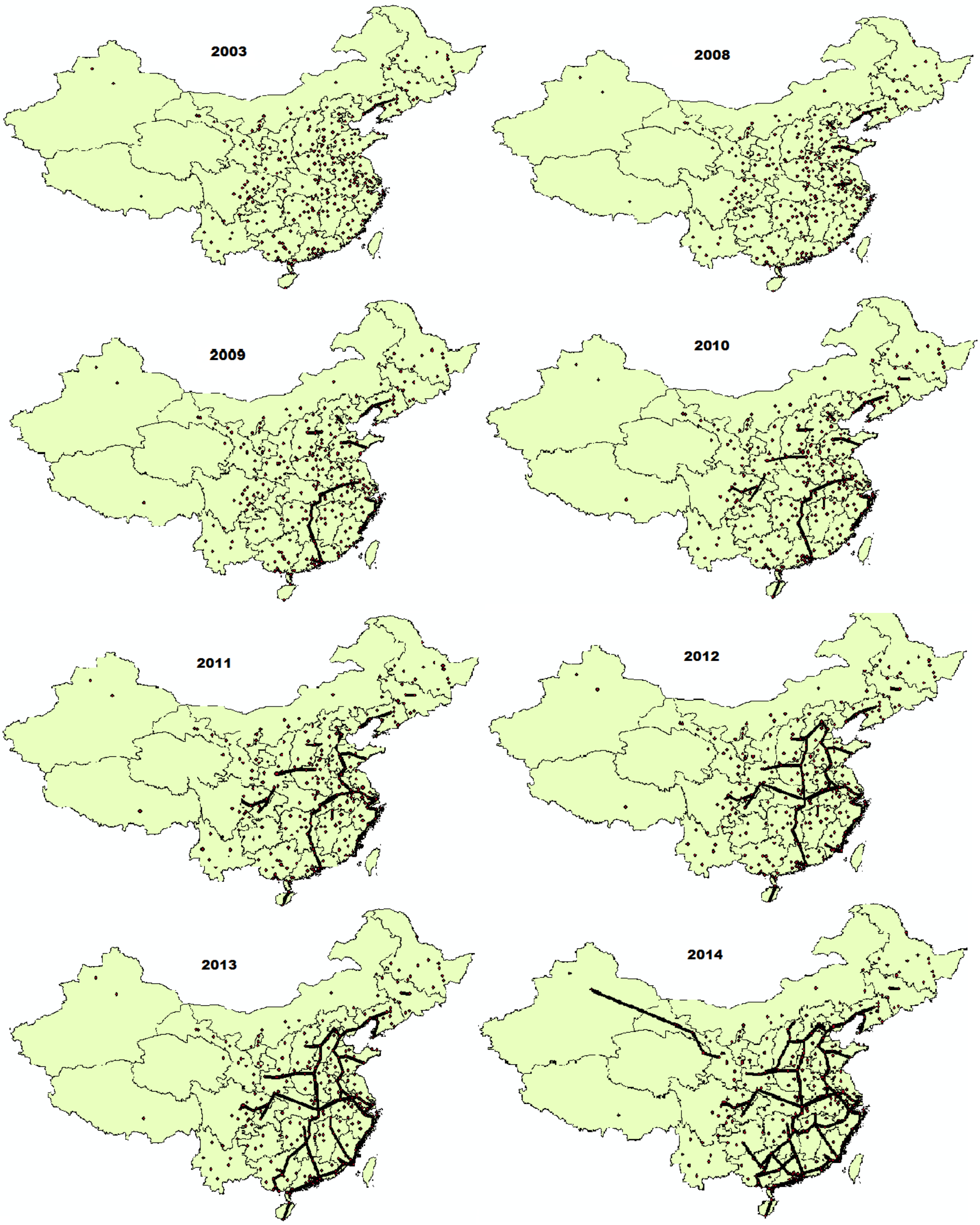


Figure 2.1: Evolution of HSR expansion from 2003 to 2014

Notes: These figures display the evolution of HSR expansion from year 2003 to 2014. The lines in bold black are lines in use by the end of that year. Each dot represents a prefecture-level city.

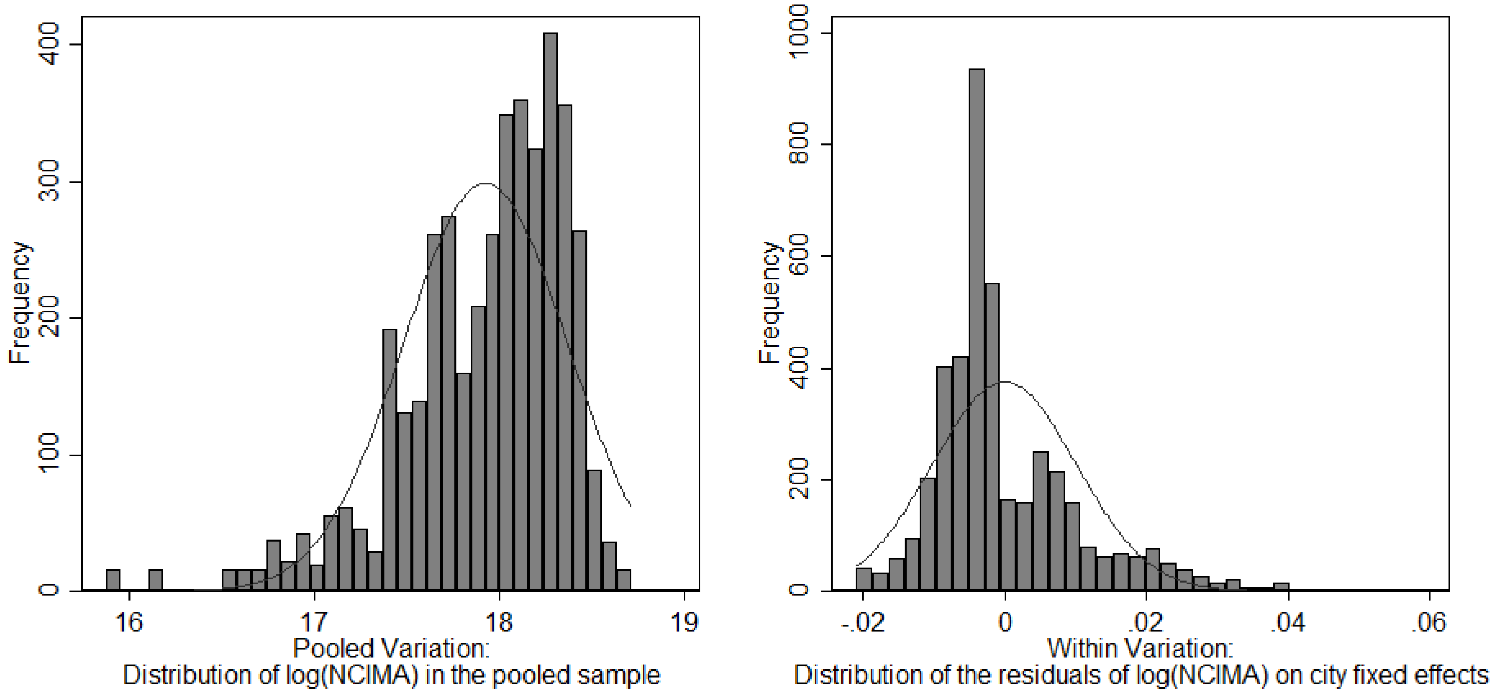
Table 2.1: Summary Statistics

	Connected		Non-connected	
	2001-2007	2007-2013	2001-2007	2007-2013
Employment growth	0.143	0.539	0.088	0.392
	(0.106)	(0.388)	0.(428	(0.361)
GDP growth	1.718	1.713	1.733	1.79
	(0.945)	(0.563)	(1.64)	(0.688)
Population growth	0.312	0.112	0.223	0.099
	(0.643)	(0.201)	(0.714)	(0.177)
	2001-2006	2006-2011	2001-2006	2006-2011
Housing price growth	0.810	1.399	0.695	1.330
	(0.453)	(0.464)	(0.432)	(0.338)
	2007	2013	2007	2013
Employment	43.71	67.59	14.6	20.04
	(7.16)	(11.79)	(1.41)	(2.05)
GDP	104.36	202.14	28.92	59.59
	(18.13)	(32.66)	(3.71)	(7.45)
Population	196.11	215.87	95.26	103.25
	(23.78)	(26.16)	(6.27)	(6.59)
	2006	2011	2006	2011
Housing price	2633.68	5577.08	1811.75	3712.23
	(1555.28)	(3421.74)	(700.96)	(1435.61)

	Growth in market access measures				
Year	2001	2003	2007	2010	2014
log(MAall)	17.9249	17.9264	17.9338	17.9422	17.9656
	(0.4431)	(0.4453)	(0.4439)	(0.4456)	(0.4420)
log(MAHSR)	17.9187	17.9193	17.9193	17.9274	17.9415
	(0.4452)	(0.4454)	(0.4454)	(0.4475)	(0.4433)
log(MAHSRless5)	17.9187	17.9187	17.9187	17.9239	17.9266
	(0.4453)	(0.4453)	(0.4453)	(0.4464)	(0.4464)
log(NCIMA)	17.9186	17.9191	17.9191	17.9261	17.9374
	(0.4453)	(0.4454)	((0.4454)	(0.4472)	(0.4428)
log(MAhighway)	17.9248	17.9257	17.9332	17.9335	17.9433
	(0.4443)	(0.4450)	(0.4438)	(0.4434)	(0.4439)

Notes: Unit: GDP: 1 billion RMB; Population and employment: 10000 people; Wage: RMB; Housing price: RMB/m^2; Market access: 1 billion RMB. A city is defined to be connected if it is connected by HSR by the end of year 2014. As housing price data only exist until 2011, I look at it over a slightly different time frame.

Figure 2.2: Pooled and within city variation of log(MAHSR)



Notes: Left: the distribution of log(MAHSR) over all the observations (city*year level). Right: the distribution of the residuals of log(MAHSR) netting out city fixed effects (within city variation)

Table 2.2: Impacts of HSR connection on transportation patterns

VARIABLES	log railway ridership	log road ridership	log air ridership	log total ridership	log railway goods
connect	0.186***	0.0283	-0.127**	0.0957***	0.106
	(0.0450)	(0.0344)	(0.0541)	(0.0261)	(0.0815)
Observations	1,637	1,746	731	1,773	1,642
R-squared	0.033	0.036	0.044	0.057	0.024
log(NCIMA)	8.096***	-2.095	-7.437**	1.244*	0.696
	(2.447)	(1.629)	(3.559)	(0.737)	(3.713)
log(MAhighway)	-4.213**	5.722**	1.958	3.312*	3.631
	(1.871)	(2.419)	(3.837)	(1.972)	(3.376)
Observations	2,480	2,791	1,175	2,808	2,544
R-squared	0.020	0.020	0.039	0.025	0.015

Notes: Data from the upper table are a panel of 172 Chinese prefecture cities annually from 2003 to 2013 that are connected or planned to be connected by HSR lines by the end of 2014. Connect is a dummy which is zero unless a city is connected to HSR before the end of that year, in which case it takes the value one. Data from the lower table are a panel of 278 Chinese prefecture cities annually from 2003 to 2013. NCIMA is the non-connection-induced-market-access measure. MAhighway accounts for market access changes by highway network expansion only. The dependent variables as listed are logs of railway passenger ridership (not restricted to high speed railway ridership), road ridership, air travel ridership, total ridership and volume of goods transported by railway (in tons). All outcome variables are counted at urban wards (shixiaqu) of prefecture cities. Cities are divided into 8 regions (Northeast, Northern Costal, Eastern Coastal, Southern Costal, Southwest, Northwest, the middle reaches of Yangtze River, the middle reaches of Yellow River. All regressions include city fixed effects and region-by-year fixed effects. Controls include government spending, other infrastructure measures, past city and provincial GDP, and interactions of year dummies with geographical centrality measures. Standard errors, reported in parentheses, are heteroskedasticity robust and clustered at the city level, allowing spatial dependence decaying in distance as in Conley (1999).* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 2.3: Impacts of HSR on transportation patterns: other market access variables

VARIABLES	log railway ridership	log road ridership	log air ridership	log total ridership	log railway goods
log(MAHSR)	5.610**	0.588	-5.300	1.883*	-1.370
	(2.408)	(1.767)	(3.862)	(1.059)	(3.264)
Observations	2,480	2,791	1,175	2,808	2,544
R-squared	0.024	0.016	0.020	0.017	0.014
log(NCIMA)	7.433***	-2.140	-6.435*	1.212	-0.280
	(2.604)	(1.746)	(3.535)	(0.803)	(3.853)
Observations	2,480	2,791	1,175	2,808	2,544
R-squared	0.026	0.016	0.021	0.017	0.013
log(MAall)	5.895**	0.847	-5.483	2.052*	-1.131
	(2.444)	(1.803)	(4.915)	(1.080)	(3.336)
Observations	2,480	2,791	1,175	2,808	2,544
R-squared	0.024	0.016	0.021	0.017	0.013
log(MAHSRless5)	12.80**	5.075	-41.42***	8.261**	3.072
	(6.346)	(4.787)	(15.96)	(3.832)	(11.28)
Observations	2,480	2,791	1,175	2,808	2,544
R-squared	0.022	0.016	0.029	0.018	0.013
log(MAhighway)	-2.653	4.781**	1.803	3.421**	1.664
	(1.949)	(2.537)	(4.191)	(1.976)	(3.440)
Observations	2,480	2,791	1,175	2,808	2,544
R-squared	0.020	0.020	0.039	0.025	0.015

Notes: Data are a panel of 278 Chinese prefecture cities annually from 2003 to 2013. Each cell corresponds to a separate regression. MAHSR is the market access measure accounting for HSR network changes only. NCIMA is the non-connection-induced-market-access measure. MAall is the market access measure accounting for both changes in HSR and highway networks. MAHSRless5 only accounts for market access changes by HSR when the bilateral travel time is less than 5 hours. MAhighway accounts for market access changes by highway network expansion only. For the regressions with log(MAall), log(MAHSR), log(NCIMA) and log(MAHSRless5) as independent variables, both the direct HSR connection dummy and log(MAhighway) are controlled for. The dependent variables as listed are logs of railway passenger ridership (not restricted to high speed railway ridership), road ridership, air travel ridership, total ridership and volume of goods transported by railway (in tons). All outcome variables are counted at urban wards (shixiaqu) of prefecture cities. All regressions include city fixed effects and region-by-year fixed effects. Controls include government spending, other infrastructure measures, past city and provincial GDP, and interactions of year dummies with geographical centrality measures. Standard errors, reported in parentheses, are heteroskedasticity robust and clustered at the city level, allowing spatial dependence decaying in distance as in Conley (1999).* significant at 10%; ** significant at 5%; *** significant at 1%.

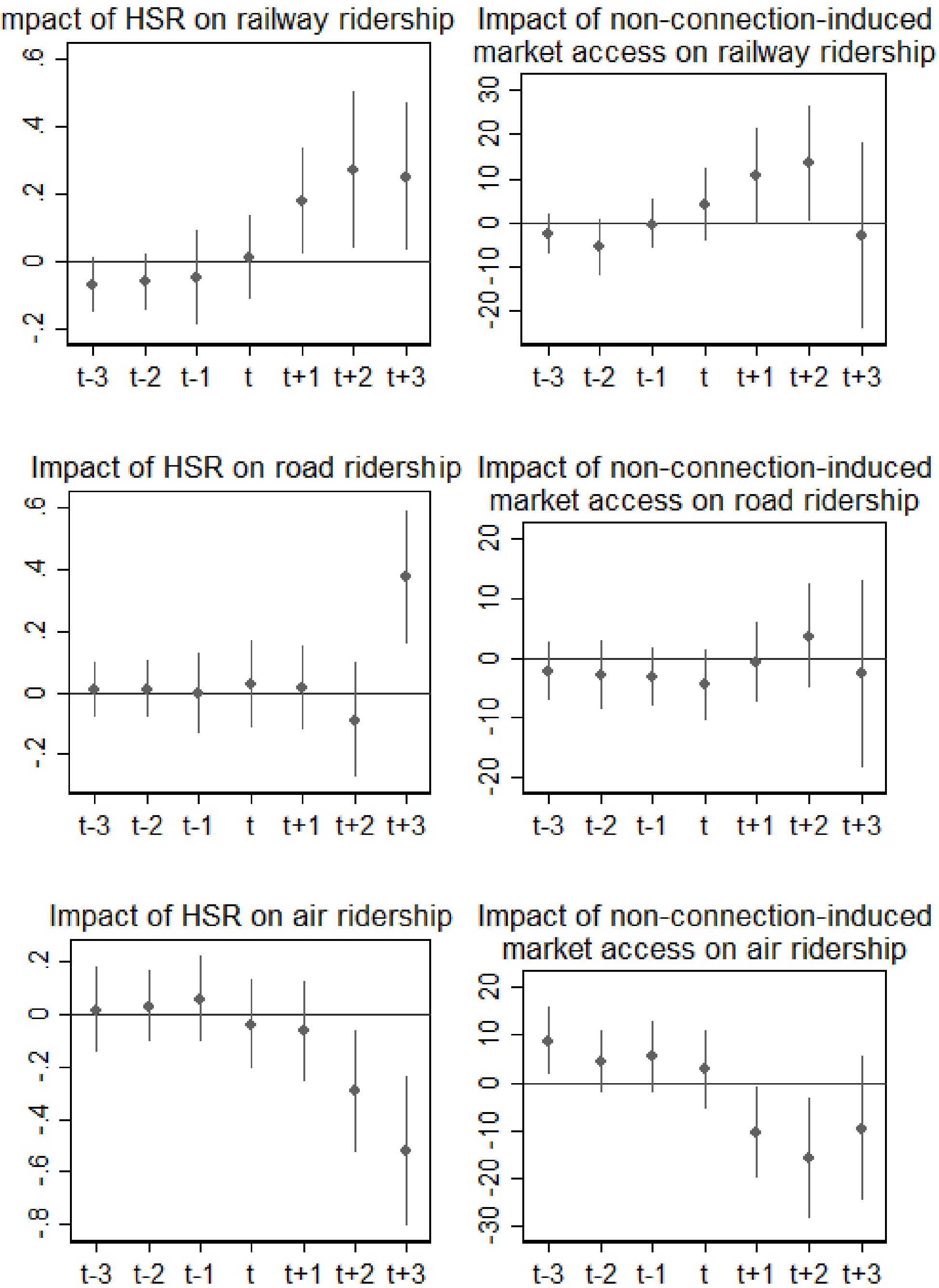


Figure 2.3: Event study: Transportation patterns

Notes: The figures display the estimated coefficients and 95% confidence intervals in regressions where dependent variables (from up to down) are the logs of railway ridership, road ridership and air travel ridership. The independent variables on the left column are the leads and lags of the initial connection dummy. The independent variables on the right column are the leads and lags of the increments in non-connection-induced market access (NCIMA). For both sets of regressions, the sample is a balanced panel from 2003-2011, as HSR connection information is available only until 2014, and the third lead is a missing value for observations after 2011. For the connection dummy, the panel includes 172 Chinese prefecture cities that are connected or planned to be connected by HSR lines by the end of 2014. For NCIMA, the panel includes 278 Chinese prefecture cities.

Table 2.4: Impacts of HSR on specialization patterns

VARIABLES	log employment	log skilled employment	log tourism employment	log other service employment	log other non-service employment
connect	0.0736***	0.0482***	0.131***	0.0225**	0.0650**
	(0.0123)	(0.0131)	(0.0247)	(0.0123)	(0.0283)
Observations	1,801	1,774	1,795	1,680	1,751
R-squared	0.035	0.030	0.031	0.011	0.029
log(NCIMA)	2.156***	2.466***	3.917***	1.032	0.551
	(0.677)	(0.515)	(1.165)	(0.718)	(1.937)
log(MAhighway)	-0.718	1.686***	-0.401	-0.346	6.393*
	(0.602)	(0.584)	(1.290)	(0.717)	(3.683)
Observations	2,877	2,813	2,847	2,604	2,788
R-squared	0.026	0.022	0.026	0.011	0.047

Notes:Data from the upper table are a panel of 172 Chinese prefecture cities annually from 2003 to 2013 that are connected or planned to be connected by HSR lines by the end of 2014. Connect is a dummy which is zero unless a city is connected to HSR before the end of that year, in which case it takes the value one. Data from the lower table are a panel of 278 Chinese prefecture cities annually from 2003 to 2013. NCIMA is the non-connection-induced-market-access measure. MAhighway accounts for market access changes by highway network expansion only. The dependent variables as listed are logs of total employment, skilled employment (includes IT, FIRE, education, business service and scientific research), tourism-related employment (includes wholesale and retail trade, hotels and catering service), other service and non-service employment. All outcome variables are counted at urban wards (shixiaqu) of prefecture cities. All regressions include city fixed effects and region-by-year fixed effects. Controls include government spending, other infrastructure measures, past city and provincial GDP, and interactions of year dummies with geographical centrality measures. Standard errors, reported in parentheses, are heteroskedasticity robust and clustered at the city level, allowing spatial dependence decaying in distance as in Conley (1999).* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 2.5: Impacts of HSR on specialization patterns: other market access variables

VARIABLES	log employment	log skilled employment	log tourism employment	log other service employment	log other non-service employment
log(MAall)	1.365**	2.328***	3.708***	1.150**	-0.182
	(0.640)	(0.521)	(1.112)	(0.560)	(1.127)
Observations	2,877	2,813	2,847	2,604	2,788
R-squared	0.028	0.027	0.029	0.012	0.015
log(MAHSR)	1.343**	2.281***	3.571***	1.120**	-0.217
	(0.628)	(0.511)	(1.096)	(0.547)	(1.108)
Observations	2,877	2,813	2,847	2,604	2,788
R-squared	0.028	0.027	0.029	0.012	0.015
log(NCIMA)	1.557**	1.969***	2.806**	1.434**	0.795
	(0.653)	(0.508)	(1.119)	(0.567)	(1.094)
Observations	2,877	2,813	2,847	2,604	2,788
R-squared	0.029	0.025	0.028	0.013	0.015
log(MAHSRless5)	9.022***	7.322***	13.46***	6.136	-0.144
	(2.012)	(1.919)	(3.543)	(3.770)	(3.345)
Observations	2,877	2,813	2,847	2,604	2,788
R-squared	0.033	0.026	0.029	0.014	0.015
log(MAhighway)	-0.500	1.937***	-0.0142	-0.258	6.451*
	(0.579)	(0.589)	(1.267)	(0.711)	(3.579)
Observations	2,877	2,813	2,847	2,604	2,788
R-squared	0.021	0.015	0.022	0.011	0.047

Notes: Data are a panel of 278 Chinese prefecture cities annually from 2003 to 2013. MAHSR is the market access measure accounting for HSR network changes only. NCIMA is the non-connection-induced-market-access measure. MAall is the market access measure accounting for both changes in HSR and highway networks. MAHSRless5 only accounts for market access changes by HSR when the bilateral travel time is less than 5 hours. MAhighway accounts for market access changes by highway network expansion only. For the regressions with log(MAall), log(MAHSR), log(NCIMA) and log(MAHSRless5) as independent variables, both the direct HSR connection dummy and log(MAhighway) are controlled for. The dependent variables as listed are logs of total employment, skilled employment (includes IT, FIRE, education, business service and scientific research), tourism-related employment (includes wholesale and retail trade, hotels and catering service), other service and non-service employment. All outcome variables are counted at urban wards (shixiaqu) of prefecture cities. All regressions include city fixed effects and region-by-year fixed effects. Controls include government spending, other infrastructure measures, past city and provincial GDP, and interactions of year dummies with geographical centrality measures. Standard errors, reported in parentheses, are heteroskedasticity robust and clustered at the city level, allowing spatial dependence decaying in distance as in Conley (1999).* significant at 10%; ** significant at 5%; *** significant at 1%.

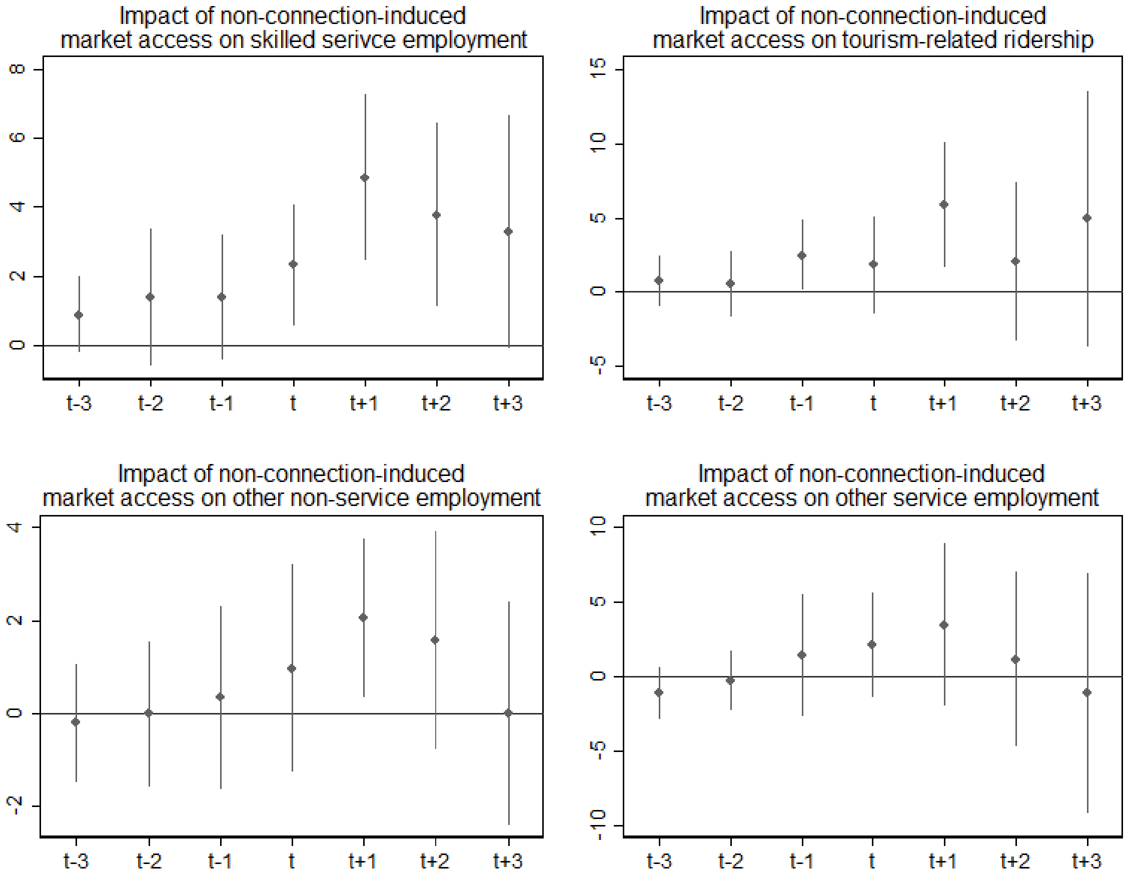


Figure 2.4: Event study: Specialization patterns

Notes:The figures display the estimated coefficients and 95% confidence intervals in regressions where the dependent variables are log skilled service employment, tourism-related service employment, other service and non-service employment. The independent variables are the leads and lags of the increments in non-connection-induced market access (NCIMA). For all the regressions, the sample is a balanced panel of 278 Chinese prefecture cities from 2003-2011, as HSR connection information is available only until 2014, and the third lead is a missing value for observations after 2011.

Table 2.6: Impacts of market access across industries

VARIABLES	log manufacture employment	log utility employment	log construct employment	log retail employment	log hotel/restaurant employment	log transport employment	log finance employment	log IT employment
log(NCIMA)	0.599	-0.250	-1.358	3.516**	1.991	-1.556	4.929***	1.794
	(3.541)	(1.515)	(1.771)	(1.425)	(1.536)	(1.240)	(1.077)	(1.927)
log(MAhighway)	12.13**	-0.925	-5.519***	-2.784*	2.027	-1.608	-1.552	-1.068
	(5.892)	(1.446)	(1.839)	(1.528)	(1.425)	(1.164)	(0.962)	(1.623)
Observations	2,796	2,350	2,354	2,621	2,363	2,618	2,362	2,338
R-squared	0.043	0.014	0.009	0.019	0.009	0.009	0.025	0.007

VARIABLES	log real estate employment	log research employment	log public employment	log medical employment	log education employment	log business service employment	log facility employment	log entertain employment
log(NCIMA)	2.327	2.138*	0.399	2.817***	1.628***	1.602	2.678*	2.645***
	(1.717)	(1.160)	(0.739)	(0.553)	(0.487)	(1.827)	(1.466)	(0.839)
log(MAhighway)	2.640	2.962**	0.231	0.477	2.831***	-0.769	1.676	-1.178
	(1.773)	(1.305)	(0.844)	(0.699)	(0.829)	(1.854)	(1.354)	(1.220)
Observations	2,847	2,864	2,269	2,363	2,363	2,557	2,358	2,354
R-squared	0.014	0.010	0.016	0.019	0.027	0.028	0.017	0.008

Notes The dependent variables are logs of sectoral employment across sixteen industries. Detailed definitions of these sectors are reported in Table A2. NCIMA is the non-connection-induced-market-access measure. MAhighway accounts for market access changes by highway network expansion only. All regressions include city fixed effects and region-by-year fixed effects. Controls include government spending, other infrastructure measures, past city and provincial GDP, and interactions of year dummies with geographical centrality measures. Standard errors, reported in parentheses, are heteroskedasticity robust and clustered at the city level, allowing spatial dependence decaying in distance as in Conley (1999).* significant at 10%; ** significant at 5%; *** significant at 1%.

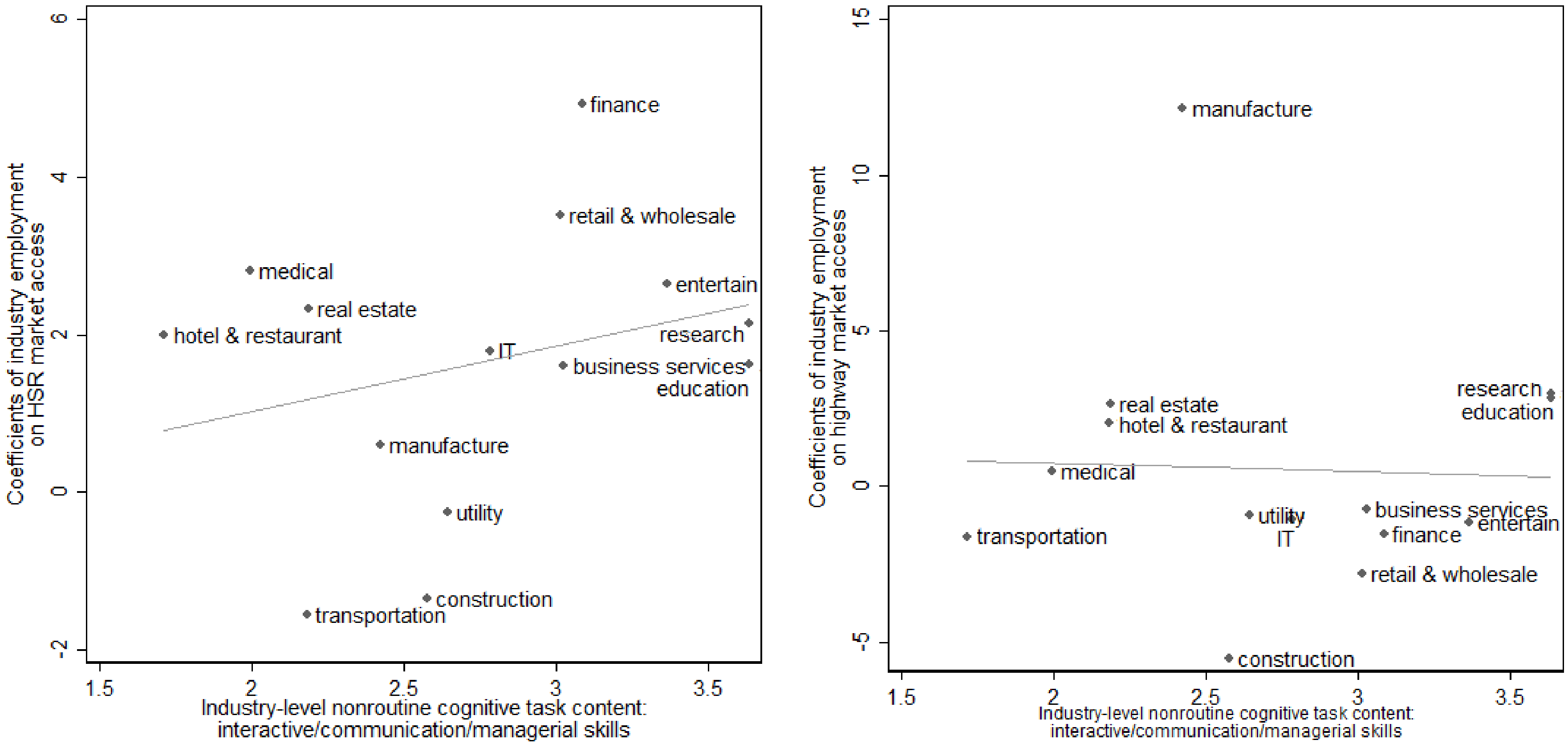


Figure 2.5: Sector-specific HSR and highway impacts and nonroutine interactive task intensity

Notes:The figures plot estimated coefficients on the impacts of NCIMA and highway-induced market access (MAhighway) on sectoral employment reported in Table 2.6 against sector-specific nonroutine interactive task intensity (interactive, communiction and managerial skills) reported in Autor et al. (2003). The matching between Chinese industries and the US industries is reported in Table A2.

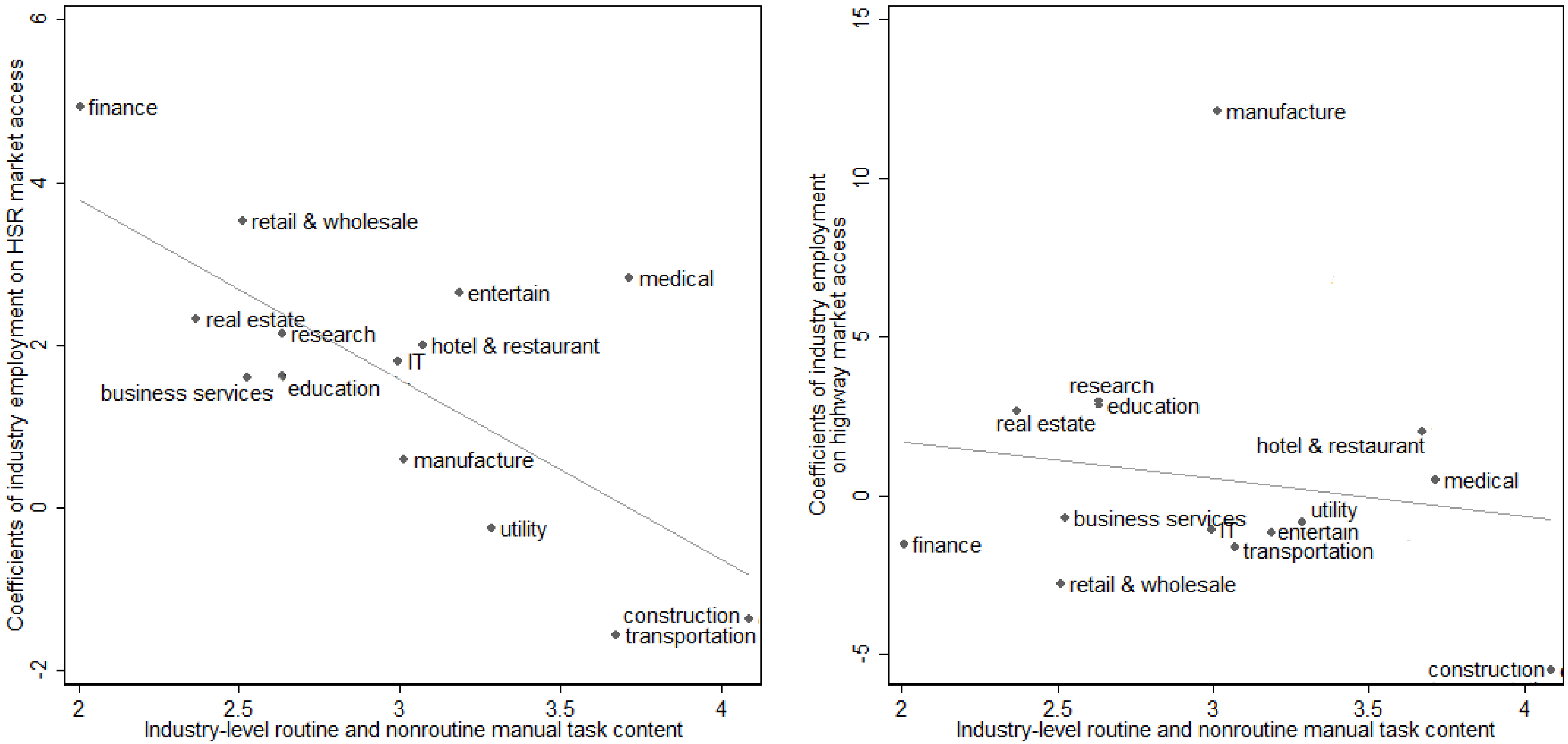


Figure 2.6: Sector-specific HSR and highway impacts and manual task intensity

Notes:The figures plot estimated coefficients on the impacts of NCIMA (left) and highway-induced market access (MAhighway) (right) on sectoral employment reported in Table 2.6 against sector-specific manual task intensity reported in Autor et al. (2003). The matching between Chinese industries and the US industries is reported in Table A2.

Table 2.7: Impacts of HSR on aggregate economic outcomes

VARIABLES	log employment	logGDP	log housing price	log industrial fixed investment	log retail sales	logpatent
connect	0.0736***	0.0354	0.0303**	-0.0846***	0.101**	0.0352
	(0.0123)	(0.0250)	(0.0152)	(0.0278)	(0.0423)	(0.0232)
Observations	1,801	1,801	1,468	1,315	1,803	1,563
R-squared	0.035	0.023	0.023	0.050	0.028	0.132
log(NCIMA)	2.156***	3.932***	1.669**	-1.348	3.431**	1.616
	(0.677)	(1.483)	(0.763)	(1.537)	(1.378)	(1.263)
log(MAhighway)	-0.718	2.141**	0.612	-0.0346	1.301	0.0414
	(0.602)	(0.913)	(0.801)	(1.280)	(2.149)	(0.598)
Observations	2,877	2,871	2,332	2,098	2,879	2,397
R-squared	0.026	0.051	0.074	0.053	0.022	0.079

Notes:Data from the upper table are a panel of 172 Chinese prefecture cities annually from 2003 to 2013 that are connected or planned to be connected by HSR lines by the end of 2014. Connect is a dummy which is zero unless a city is connected to HSR before the end of that year, in which case it takes the value one. Data from the lower table are a panel of 278 Chinese prefecture cities annually from 2003 to 2013. NCIMA is the non-connection-induced-market-access measure. MAhighway accounts for market access changes by highway network expansion only. The dependent variables as listed are logs of total employment, GDP, housing price, total fixed investment of industrial firms, sales of retail firms and total patents. All outcome variables are counted at urban wards (shixiaqu) of prefecture cities. All regressions include city fixed effects and region-by-year fixed effects. Controls include government spending, other infrastructure measures, past city and provincial GDP, and interactions of year dummies with geographical centrality measures. Standard errors, reported in parentheses, are heteroskedasticity robust and clustered at the city level, allowing spatial dependence decaying in distance as in Conley (1999).* significant at 10%; ** significant at 5%; *** significant at 1%.

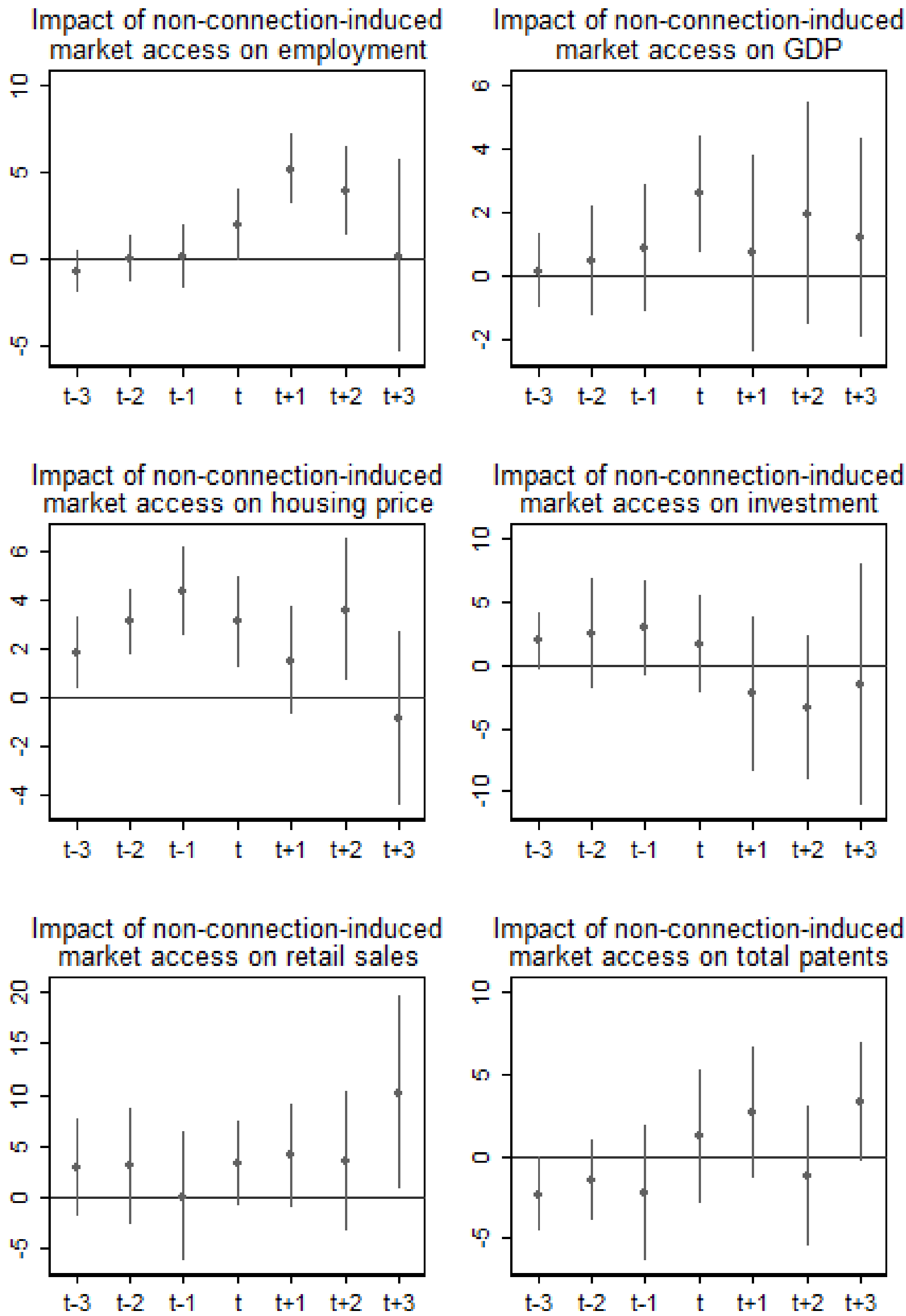


Figure 2.7: Event study: Aggregate economic outcomes

Notes:The figures display the estimated coefficients and 95% confidence intervals in regressions where the dependent variables are log total employment, GDP, housing private, fixed investments, total retail sales and patents. The independent variables are the leads and lags of the increments in non-connection-induced market access (NCIMA). For all the regressions, the sample is a balanced panel of 278 Chinese prefecture cities from 2003-2011, as HSR connection information is available only until 2014, and the third lead is a missing value for observations after 2011.

Appendices

Appendix B

Appendix of "Travel Costs and Urban Specialization: Evidence from China's High Speed Railway"

B.1 A Model of labour Sourcing

B.1.1 Preferences and Endowments

The economy consists of N cities and a single final good sector. Each city j is endowed with an inelastic supply of land ($\bar{H}_j$). The economy as a whole is endowed with an inelastic supply of workers who are perfectly mobile across cities. The representative consumer's preferences are defined over a final consumption good (C_j), housing (H_j) , and local-amenities (δ_j) taking the Cobb-Douglas form:

$$U_j = \delta_j C_j^{\alpha} H_j^{1-\alpha}, 0 < \alpha < 1 \tag{B.1}$$

Consumption good price is normalized to be 1. Housing rental rate is represented as q_j. The indirect utility function is:

$$V_j = b\delta_j w_j (q_j)^{\alpha-1} \tag{B.2}$$

where $b = \alpha^{\alpha}(1-\alpha)^{(1-\alpha)}$ is a constant.

For simplicity we assume that the final good is freely traded across cities. Cobb-Douglas utility function indicates that housing expenditures should take up $1-\alpha$ of total income. We assume that housing expenses are redistributed as a lump-sum to consumers, as in Helpman (1998). Therefore for city j, its total income can be represented as:

$$v_j L_j = w_j L_j + (1-\alpha) v_j L_j = w_j L_j / \alpha \tag{B.3}$$

Combined with housing market clearing condition, we obtain equilibrium housing rental rate:

$$q_j = \frac{1-\alpha}{\alpha} \frac{w_j L_j}{\bar{H}_j} \tag{B.4}$$

where $\bar{H}_j$ represents city j's inelastic housing supply.

B.1.2 Production

The final good y in city j is produced from a continuum of tasks $i \in [0,1]$ as in Grossman and Rossi-Hansberg (2008) under CES technology. Perfect competition and constant returns to scale apply.

$$y_j = [\int_0^1 x(i)^{(\sigma-1)/\sigma} dj]^{\sigma/(\sigma-1)} \tag{B.5}$$

Tasks are produced using labour according to a linear technology:

$$l_j(i) = \frac{x_j(i)}{z_j(i)} \tag{B.6}$$

A city j can remote source tasks from another city k. The productivity of producing task i at city k for city j is $z_{jk}(i)$. The origin city j's productivity on task i can be summarized by a vector $\mathbf{z_j(i)} = (z_{j1}(i), ...z_{jI}(i))$. I assume that the productivity vector for $i \in [0,1]$ and $k = 1, 2, ..., I$, is a random variable drawn independently across tasks and destination cities from a multivariate Frechet distribution with zero correlation across draws, $F_j(\mathbf{z_j(i)} = exp(-\sum_l T_j(z_{jl}(i))^{-\theta})$. The marginal distribution of $z_{jk}(i)$ is then $F_j(z_{jl})) = exp(-T_j(z_{jl}(i))^{-\theta})$.

Here I use the multivariate Frechet distribution instead of the single-variable version because I would like to have the ex-ante probabilities of a city j sourcing out its tasks to any cities to be positive.

By remote sourcing a task i from city k to city j, firms take advantage of city j's higher productivity $z_{jk}(i)$ and city k's lower labour cost w_k, subject to iceberg transportation costs, $\tau_{jk} > 1$. The iceberg cost can be interpreted as the loss of efficiency in management over longer distances.

When a task i is sourced from city k to city j, the labour requirement $l_{jk}(i)$ and total cost of production $g_{jk}(i)$ are:

$$l_{jk}(i) = \frac{\tau_{jk} x_j(i)}{z_{jk}(i)} \tag{B.7}$$

$$g_{jk}(i) = \frac{\tau_{jk} w_k(i)}{z_{jk}(i)} \tag{B.8}$$

Firms in city j that sources in task i looks for a lowest cost source of supply for that task. Following Eaton and Kortum (2002), the cost of sourcing task i from city k to city j has the distribution:

$$G_{jk}(g) = Pr[g_{jk}(i) \leq g] = Pr(\frac{w_k(i)}{z_{jk}(i)}\tau_{jk} \leq g) = 1 - exp((-T_j(w_k\tau_{jk})^{-\theta})g^{\theta}) \tag{B.9}$$

The cost of production in city j then has the distribution:

$$G_j(g) = Pr[g_{jk}(i) \leq g, k = 1, 2...N] = 1 - \prod_{k=1}^{N}[1 - G_{jk}(g)] = 1 - exp(-\Phi_j g^{\theta}) \tag{B.10}$$

where the parameter Φ_j of city j's cost distribution is:

$$\Phi_j = \sum_{k=1}^{N} T_j(w_k\tau_{jk})^{-\theta} \tag{B.11}$$

The actual unit cost of production of final good for the CES production function, assuming $\sigma < 1 + \theta$, is

$$c_j = \gamma\Phi_j^{-1/\theta} \tag{B.12}$$

B.1.3 Flows in Tasks Sourcing

Perfect competition implies that all firms receive zero profits. Prices of the final good are equalized across cities because of free trade. Therefore, unit cost of production of final goods should also be equalized across cities, that is:

$$c_j = \gamma\Phi_j^{-1/\theta} = c, \forall j \tag{B.13}$$

A direct implication is that $\Phi_j = \sum_{i=1}^{N} T_j(w_i\tau_{ji})^{-\theta}$ is also equalized across cities. In equilibrium, this cost equalization condition indicates that higher productivity in a central city usually goes with higher wages not only in the city itself, but also in surrounding cities. Additionally, a decrease in communication costs τ_{kj} between two cities is likely to drive up wages in these two cities relative to those in other cities.

Another important result derived following Eaton and Kortum (2002) describes X_{jk}, the total labour cost in city k of producing for city j, as:

$$X_{jk} = T_j(w_k\tau_{kj})^{-\theta}(\sum_{k=1}^{N} T_j(w_k\tau_{kj})^{-\theta})^{-1}X_j \tag{B.14}$$

where X_j is city j's total cost of labour production (domestically produced or sourced to other cities). It is obvious from the above equation that city j sources more tasks to city k if city k has a lower wage, or is closer to j.

Note here $\sum_{k=1}^{N} T_j(w_k d_k)^{-\theta} = \Phi$ is the same across all cities due to the unit labour cost equalization condition. And we have:

$$Y_k = \sum_{j=1}^{N} X_{jk} = \Phi^{-1}(w_k^{-\theta})\sum_{j=1}^{N} \tau_{kj}^{-\theta}X_jT_j \tag{B.15}$$

Here Y_k is the total labour income of city k, and can be represented as l_kw_k. To guide the empirical analysis, I define $MA_k = \sum_{j=1}^{N} \tau_{kj}^{-\theta}X_jT_j$ as the "market access" of k. Intuitively, the "market access" in city k is a weighted average of all other cities' "market size" scaled up by their productivity and scaled down by distances. It can be roughly thought to represent the "perceived" demand of k's labour from the whole country.

B.1.4 Spatial Equilibrium

Free labour mobility implies that indirect utility is equalized across all cities in equilibrium.

$$V_j = b\delta_jW_j(q_j)^{\alpha-1} = V, \forall i \tag{B.16}$$

Combined with the "market access" equation (18) and the market clearing condition (6), we are now able to characterize a similar spatial equilibrium with endogenous variables N_k, w_k, q_k, and exogenous variables $T_k, \delta_k, \bar{H}_k, MA_k$, as in Glaeser and Gottlieb (2009).

Given $\{T_k, \delta_k, \bar{H}_k, MA_k\}$, a spatial equilibrium is $\{N_k, W_k, q_k\}$ such that

1. $q_j = \frac{1-\alpha}{\alpha}\frac{w_jl_j}{\bar{H}_j}$ (Housing Market Clearing)
2. $V_j = b\delta_jw_j(q_j)^{\alpha-1} = V, \forall i$ (Workers Mobility)
3. $l_kw_k = \Phi^{-1}(w_k^{-\theta}) * MA_k$ (labour Income and Market Access)
4. $\sum_k l_k = N$ (Aggregate labour market clearing)
5. $\sum_k C_k = \sum_k y_k$ (Aggregate goods market clearing)

To summarize, my model has the following implications: First, a city's labour income $w_j l_j$ is increasing in its market access, defined as the access to all the other cities' technology. Second, housing prices increase following wages because of free labour mobility. Whether the increase in market access translates into higher employment or higher wage depends on the housing supply restrictions of the city.

B.2 Services and fares of HSR

Most of the HSR lines in China operate intensively, but there is a large variety in the number of services across lines: the busiest Nanjing-Shanghai line carries more than 160 trains one way per day while the number of services for the Western Chengdu-Dazhou line is only 7. Detailed information on the number of services for each line is listed in Table A4. A typical HSR train normally carries eight cars, with a total capacity of about 600.

There are two main categories of high speed trains in China. One set of trains runs at a designed speed of 350km/h, and their train numbers start with G. The other runs at a designed speed of 250km/h with their train numbers led by D, compared to a top speed of 120km/h for pre-HSR trains. There are national fare scales for the two speeds but in practice, HSR fares vary slightly from line to line. The price, travelling time and cost per kilometer for passengers are listed in Table A4. For the lines served by trains of different speeds I report the price and travelling time of the category starting with G (350km/h).

Undoubtedly, the introduction of HSR drastically decreases the travel time across cities. The travel time between Beijing and Shanghai has been shortened from 13 hours to 5 after the introduction of the Beijing-Shanghai HSR line in 2011, for a second-class price of 553 RMB (90 USD). By comparison, a flight from Beijing to Shanghai usually takes a bit more than two hours for a full price of 1290 RMB (210 USD, including 160 RMB fees and taxes). Even if discounts are usually available for flights, HSR still proves to be a good value of money if we consider the extra time to get to the airport and occasional delays by air.

B.3 Market Access: Construction of cost parameters

As briefly mentioned in section 4.2, the calculation of the market access variables requires the construction of a time-varying transportation cost matrix for each city pair. We have to rely on a few assumptions to construct τ_{km}.

First, I allow τ_{km} to incorporate both time and fare cost in travel. I generalize passenger choices among different modes of transportation into three distinct combinations of time-fare tradeoffs. (1) Travel by conventional roads (not highways) or slower railway at a speed of 60 km/h and monetary cost of 0.1 km/h (2) Travel by highways or faster conventional rail (K or T initials) at a speed of 100 km/h, with a fare/monetary cost of ¥0.23/km [1] (3) Travel by HSR at a speed of 220 km/h, with a fare cost of ¥0.43[2] per kilometer. I assume that all cities are connected by conventional roads had they not been connected by highway or HSR. The information on highway network is obtained from Baum-Snow et al. (2016). A city is defined to be connected by highway if the distance between the centroid of the city and the nearest highway is less than 30 km. [3]

Second, wage plays an essential role in passengers' fare-time cost tradeoff: people who earn more have a higher value of time and are more likely to take high speed trains at a higher cost. For a particular city pair, I assume that the value of time is determined by the average wage of these two cities. If the value of hours saved by taking HSR does not cover the extra cost in fares, I assume that the HSR connection between these two cities does not reduce bilateral travel cost. To avoid additional endogeneity, I use the wage data in 2007, the year before most HSR lines were opened, in all the calculations, since there are a few wage data missing from 2000.

Third, I restrict the maximum number of transfers to be two to better capture people's travelling behavior. It is therefore assumed that if passengers need to make more than two transfers travelling by HSR, they will stick to alternative forms of transportation and HSR does not change their travel cost.

Fourth, to capture the nature of travel cost facilitate intercity communication, we need an approximation of the required frequency of face-to-face contacts from one city to another for an intercity link to be successfully built. Without sensible estimates on labour sourcing cost, I assume that on average, two trips per month

[1] The fare of traditional railway is set according to a formula, adding up three parts: 0.06/km for seaters, 0.016/km extra for air-conditioning, 0.024/km extra for pre-HSR top speed, and usually scaled up by 50% for new air-conditioned trains, which are prevalent nowadays. The fare is also allowed to be scaled up by local railway authorities according to the operational costs; here I use an arbitrary scaling up rate of 50%. As for highways, the monetary cost usually consists of toll fees and fuel cost. The former is around ¥0.5/km for a small vehicle and the latter around 0.8/km for average fuel efficiency. Assuming on average four passengers occupy a small vehicle, the cost per person is 0.33/km. Taking long-distance coach would cost slightly less. For convenience I assume the fare and time cost are the same for conventional trains and road travel.

[2] Imputed from Table A3. For long-haul trains with a travel time above 10 hours, I scale up the average fare rate by a half to account for the extra cost of sleepers.

[3] Baum-Snow et al. (2016) only reports changes in highway network for the year 2000, 2002, 2003, 2005, 2007 and 2010. According to them, data on the year 2000, 2005 and 2010 are more accurate. For years without data, I assume the highway network follows the previous year.

are necessary for all kinds of tasks. In Section 8, I experiment with multiple values of the number of trips from 0.5 to 10 per month as robustness checks.

Finally, an estimation of market access as $MA_k = \sum_{m!=k}^{N} \tau_{km}^{-\theta} Y_m$, requires an estimation of the decay parameter θ. In my model, the parameter θ measures the dispersion of productivity. I follow Eaton and Kortum (2002) and Donaldson (2014) and try several different measures over a wide range (3.6, 8.28 and 12.86) for robustness, but stick to 3.6 from Donaldson (2014) in my main analysis. I find the results to be robust to the selection of θ, as reported in Table A9. To avoid endogeneity caused by geographically correlated local shocks, I use base year (2000) GDP in calculations of all the market access variables.

B.4 Additional robustness checks

B.4.1 Exploiting variation in HSR opening time due to variation engineering difficulties only

As mentioned in Section 3.2, one way to get around the concerns that there are economic and political reasons to push some lines to be finished earlier that are correlated with the expected growth prospects of the connected cities is to focus on a subset of connected cities where the construction of HSR starts at the same year and to exploit the variation in project completion solely driven by engineering difficulties. Specifically, I look at the cities where HSR construction work commences in 2005, following the passage of "Mid-to-long run HSR plan" immediately. I collect data on the construction duration, length, the number of stations, bridge and tunnel length of the eleven HSR lines that started construction in 2005. From an engineering point of view, an HSR line takes longer to be built if it is longer, has more stations and has a higher bridge and tunnel ratio. To implement the idea, I regress the duration of construction on HSR line length, the number of stations and bridge/tunnel ratio and get a predicted duration of construction of these lines as the fitted value of this regression. I then set the engineering-difficulty-determined opening date of each earlier HSR lines to be the construction starting date plus predicted duration of construction. The new engineering-difficulty-determined opening year changes for four lines (Wuhan-Guangzhou, Zhengzhou-Xi'an, Fujian-Xiamen, and Jiangmen-Xinhui). Although this operation leaves me with a very small sample (27 cities), I still find similar effects on railway ridership, service employment and private employment as in the main regressions (Table A10). However, the effects on manufacturing employment are negative in this case, which might be attributable to especially large construction effects as they are the earlier connectors, and the construction work might be more costly and requires more input with fewer experiences. It means that we should observe larger manufacture and construction employment increase before the actual connection of HSR, leaving the coefficient of connection dummy to be negative.

B.4.2 HSR impacts on peripheral areas

Table A11 presents the results on the impacts of HSR connection or HSR-induced market access on the peripheries of prefecture cities. I run separate regressions and replace the dependent variables as the outcomes of the whole prefecture city excluding its central urban area. I find small but insignificant negative impacts of direct HSR connection on GDP and employment in the peripheral areas of prefecture cities. It suggests that a small proportion of the positive impacts of HSR connection on urban-core employment is a result of an accelerating urbanization process. Direct HSR connection or HSR induced market access boost promote the attractiveness of central cities and draws people and economics activities from rural areas to urban centers.

Table A1: Data Appendix

Variables	Source	Year Range	Obs.	No. of cities	No. of cities with no missing values
GDP	City Statistical Yearbook	2000-2013	3888	282	261
Population	City Statistical Yearbook	2000-2013	3887	282	261
Employment	City Statistical Yearbook	2000-2013	3887	282	261
manufacturing employment	City Statistical Yearbook	2000-2013	3799	282	261
Service Employment	City Statistical Yearbook	2000-2013	3887	282	261
Wage	City Statistical Yearbook	2000-2013	3794	282	261
Housing Price	Regional Economic Statistical Yearbooks	2000-2011	3311	281	261
Railway Ridership	City Statistical Yearbook	2000-2013	3441	255	230
Road Ridership	City Statistical Yearbook	2000-2013	3940	284	268
Air Ridership	City Statistical Yearbook	2000-2013	1488	134	98
Total Ridership	City Statistical Yearbook	2000-2013	3960	284	268
Goods Transported by Railway	City Statistical Yearbook	2000-2013	3465	252	231
SOE employment	Regional Economic Statistical Yearbooks	2000-2011	3355	282	261
Utility employment	City Statistical Yearbook	2000-2013	3870	282	261
Construction employment	City Statistical Yearbook	2000-2013	3880	281	261
FIRE employment	City Statistical Yearbook	2000-2013	3705	282	261
IT employment	City Statistical Yearbook	2003-2013	3069	282	274
Business service employment	City Statistical Yearbook	2003-2013	3033	282	271
Retail employment	City Statistical Yearbook	2002-2013	3331	282	270
Catering employment	City Statistical Yearbook	2003-2013	3092	282	277
Educaion employment	City Statistical Yearbook	2003-2013	3101	282	277
Mining employment	City Statistical Yearbook	2000-2013	2887	206	198
Government Spending	City Statistical Yearbook	2000-2013	3641	282	261
Area of urban paved roads	City Statistical Yearbook	2002-2013	3423	282	277
Urban green land area	City Statistical Yearbook	2003-2013	3105	284	281
Number of theatres	City Statistical Yearbook	2000-2013	3856	282	262

Table A2: Match between Chinese industries and US two-digit industries

Code	Chinese industries	US industries	NAICS
A	Agriculture, forestry, animal production and hunting, fishing	Forestry, Fishing, Hunting and Agriculture Support	11
B	Mining and quarrying	Mining	21
C	Manufacturing	Manufacturing	31
D	Production and distribution of electricity heating power, gas and water	Utilities	22
E	Construction	Construction	23
F	Wholesale and retail trade	Wholesale and retail trade	42,43
G	Transportation, warehousing and post	Transportation and Warehousing	48
H	Hotels and catering services	Accommodation and Food Services	72
I	Information transmission, computer services and software	Information	51
J	Finance and Insurance	Finance and Insurance	52
K	Real estate	Real Estate and Rental and Leasing	53
L	Leasing and business sevices	Professional, Scientific, and Technical Services	54
M	Scientific research, technical services	Professional, Scientific, and Technical Services	54
N	Management of water conservancy, environment and public facilities		
O	Household services, repair and other services		
P	Education	Educational Services	61
Q	Health,social work	Health Care and Social Assistance	62
R	Culture, sports and entertainment	Arts, Entertainment, and Recreation	71

Table A3: HSR lines in use before the end of 2014

HSR lines	Construction Time	Openning Time	Length(km)
Qinhuangdao-Shenyang	01/01/1999	01/07/2003	405
Hefei-Nanjing	11/06/2005	19/04/2008	154
Beijing-Tianjin	07/04/2005	01/08/2008	120
Qingdao-Jinan	28/01/2007	20/12/2008	393
Shijiazhuang-Taiyuan	11/06/2005	01/04/2009	190
Hefei-Wuhan	01/08/2005	01/04/2009	351
Dazhou-Chengdu	01/05/2005	07/07/2009	148
Ningbo-Taizhou-Wenzhou	27/10/2005	28/09/2009	268
Wenzhou-Fuzhou	08/01/2005	28/09/2009	298
Wuhan-Guangzhou	23/06/2005	26/12/2009	968
Zhengzhou-Xian	01/09/2005	06/01/2010	455
Fuzhou-Xiamen	01/10/2005	26/04/2010	275
Chengdu-Dujiangyan	04/11/2008	12/05/2010	65
Shanghai-Nanjing	01/07/2008	01/07/2010	301
Nanchang-Jiujiang	28/06/2007	20/09/2010	131
Shanghai-Hangzhou	01/04/2009	26/11/2010	150
Yichang-Wanzhou	01/12/2003	22/12/2010	377
Wuhan-Yichang	17/09/2008	23/12/2010	293
Haikou-Sanya (Eastern Coastal line)	29/09/2007	30/12/2010	308
Changchun-Jilin	01/04/2008	30/12/2010	111
Jiangmen-Xinhui	18/12/2005	07/01/2011	27
Beijing-Shanghai	18/04/2008	30/06/2011	1433
Guangzhou-Shenzhen	20/08/2008	26/12/2011	116
Longyan-Xiamen	25/12/2006	01/07/2012	171
Zhengzhou-Wuhan	15/10/2008	28/09/2012	536
Hefei-Bengbu	20/05/2009	16/10/2012	132
Haerbin-Dalian	23/08/2007	01/12/2012	921
Beijing-Zhengzhou	26/12/2007	26/12/2012	693
Nanjing-Hangzhou	01/04/2009	01/07/2013	249
Panjin-Yingkou	31/05/2009	12/09/2013	89
Tianjin-Qinhuangdao	08/11/2008	01/12/2013	261
Xiamen-Shenzhen	23/11/2007	28/12/2013	502
Xi'an-Baoji	18/12/2009	28/12/2013	148

Guangxi Coastal (Nanning-Qinzhou-Beihai)	11/12/2008	28/12/2013	261
Liuzhou-Nanning	27/12/2008	28/12/2013	223
Wuhan-Xianning	26/03/2009	28/12/2013	90
Taiyuan-Xi'an	03/12/2009	01/07/2014	678
Nanchang-Changsha	26/02/2009	16/09/2014	344
Hangzhou-Nanchang	18/04/2010	10/12/2014	582
Lanzhou-Wulumuqi	01/01/2010	26/12/2014	1776
Guangzhou-Nanning	11/09/2008	18/04/2014	577
Huanggang-Wuhan -Huangshi	02/10/2009	18/06/2014	97

Notes: Data are obtained from the Major Events and the Finished and Ongoing Projects sections in the China Railway Yearbooks from 1999 to 2012.

Table A4: Service and fare information of HSR lines

HSR lines	Trains/day	Speed (km/h)	Price	Duration	Cost/km
Qinhuangdao-Shenyang	20	250	125	02:30	0.31
Hefei-Nanjing	57	250	60.5	00:44	0.36
Beijing-Tianjin	87	350	54.5	00:33	0.47
Qingdao-Jinan	35	250	116.8	02:28	0.32
Shijiangzhuang-Taiyuan	16	250	68	01:35	0.36
Dazhou-Chengdu	7	200	110	02:45	0.74
Ningbo-Taizhou	44	250	49	00:59	0.32
Taizhou-Wenzhou	48	250	35.5	00:45	0.31
Wenzhou-Fuzhou	37	250	88.5	02:05	0.3
Wuhan-Guangzhou	58	350	463.5	04:10	0.48
Zhengzhou-Xian	28	350	229	02:27	0.5
Hefei-Wuhan	41	250	105	02:02	0.3
Fujian-Xiamen	65	250	71.5	01:39	0.26
Chengdu-Dujiangyan	6	220	15	00:33	0.23
Shanghai-Nanjing	162	350	139.5	01:30	0.46
Nanchang-Jiujiang	14	250	42	01:05	0.32
Shanghai-Hangzhou	81	350	73	00:50	0.49
Yichang-Wanzhou	16	200	162	04:47	0.43
Haikou-Sanya (Eastern Coastal line)	26	250	83.5	01:50	0.27
Changchun-Jilin	42	250	31.5	00:50	0.28
Jiangmen-Xinhui	19	200	10	00:07	0.37
Beijing-Shanghai	37	350	553	05:05	0.31
Guangzhou-Shenzhen	127	350	74.5	00:33	0.64
Longyan-Xiamen	15	200	48.5	01:10	0.31
Zhengzhou-Wuhan	54	350	244	02:10	0.46
Hefei-Bengbu	29	350	70	00:57	0.53
Haerbin-Dalian	18	350	403.5	04:14	0.44
Beijing-Zhengzhou	55	350	309	03:22	0.45
Nanjing-Hangzhou	73	350	117.5	01:45	0.47
Panjin-Yingkou	5	350	27.5	00:26	0.31
Xiamen-Shenzhen	39	350	150.5	03:50	0.30
Xi'an-Baoji	18	350	51.5	00:59	0.35
Tianjin-Qinhuangdao	16	350	120	01:25	0.46

Guangxi Coastal (Nanning-Qinzhou-Beihai)	23	250	61	01:38	0.30
Liuzhou-Nanning	31	250	65.5	01:30	0.30
Wuhan-Xianning	48	250	24.5	00:24	0.27
Hangzhou-Nanchang	42	350	263.5	02:57	0.46
Nanchang-Changsha	45	350	157	01:42	0.46
Lanzhou-Wulumuqi	3	250	548.5	11:25	0.31
Taiyuan-Xi'an	15	250	178.5	03:47	0.28
Guangzhou-Nanning	34	250	169	00:39	0.30
Huanggang-Wuhan -Huangshi	127	250	74.5	00:33	0.64

Table A5: HSR passenger survey results

<table>
<tr><th>Line</th><th>Sample Size</th><th>Average monthly Income</th><th>Purpose of travel5</th><th>Connecting or transfer modes4</th><th colspan="3">Alternative modes 7</th></tr>
<tr><td rowspan="3">Changchun-Jilin (110km)1</td><td rowspan="3">1001</td><td rowspan="3">4300</td><td>Non-commute business: 26%</td><td>Car/Taxi: 48%</td><td colspan="3">Bus/Coach: 50%</td></tr>
<tr><td>Leisure: 46%</td><td rowspan="2">Public transport: 43%</td><td colspan="3">Ordinary train: 36%</td></tr>
<tr><td>Commuting: 19%</td><td colspan="3">Air: 0%</td></tr>
<tr><td rowspan="3">Beijing-Tianjin3 (120km)</td><td rowspan="3">1108</td><td rowspan="3"></td><td>Non-commuting Business: 39%</td><td>Car/Taxi: 34.7%</td><td colspan="3">Bus: 20.5%</td></tr>
<tr><td>Leisure: 33%</td><td rowspan="2">Public transport: 60.1%</td><td colspan="3">Other train: 61.5%</td></tr>
<tr><td>Commuting: around 15%</td><td colspan="3">Car: 11.60%</td></tr>
<tr><td rowspan="3">Wuhan-Guangzhou2 968km)</td><td rowspan="3">556</td><td rowspan="3">4500</td><td>Non-commute Business: 39%</td><td>No transfer: 49%</td><td colspan="3" rowspan="3"></td></tr>
<tr><td>Leisure: 33%</td><td>Transfer to road: 40%</td></tr>
<tr><td>Commuting: around 15%</td><td>Transfer to air:1.5%</td></tr>
<tr><td rowspan="4">Tianjin-Jinan1 (Part of Beijing-Shanghai 1318km)</td><td rowspan="4">1001</td><td rowspan="4">6700</td><td>Non-commute Business: 62%</td><td>Transfer to road 67%</td><td></td><td>Short trip 6</td><td>Long trip</td></tr>
<tr><td>Leisure: 28%</td><td rowspan="3">Other: 29%</td><td>Bus</td><td>32%</td><td>1%</td></tr>
<tr><td rowspan="2">Commuting: 0%</td><td>Other Train</td><td>40%</td><td>18%</td></tr>
<tr><td>Air</td><td>7%</td><td>77%</td></tr>
</table>

Notes: 1. Passenger survey results from Ollivier et al. (2014)
2. Passenger survey resultsJianbin (2011)
3. Passenger survey results Wu (2006)
4. The exact questions differ across surveys. For the survey on Wuhan-Guangzhou line, the question is "What are your transfer choices off HSR to your final destination, if any?". For the other surveys, the question is "What are your transportation choices to the HSR station?"
5. The percentages of business travel, leisure and commuting do not add up to 100% and the remaining respondents choose other purposes.
6. Short trips are defined as trips shorter than 300 km, and long trip longer than 300 km.
7. Passengers are asked "What are your preferred ways to travel before the introduction of HSR for similar trips?" in the surveys.

Table A6: Annual ridership for major HSR lines [1], (in millions passengers)

HSR lines	2009	2010	2011	2012
Qingdao-Jinan	21.22	23.95	26.74	29.63
Beijing-Tianjin	16.41	20.22	21.04	21.50
Changchun-Jilin			8.84	9.12
Shijiazhuang-Taiyuan [2]	4.64 (6.19)	7.46	8.48	8.83
Hefei-Wuhan [2]	1.62 (2.16)	2.97	3.49	4.57
Shanghai-Shenzhen [3]	2.17	20.89	35.06	
Zhengzhou-Xian	X[5]	3.74	3.82	6.37
Guangzhou-Shenzhen [4]	33.49	36.95	39.05	35.78
Guangzhou-Shenzhen -Hong Kong				10.56
Guangzhou-Zhuhai			14.06	16.31
Beijing-Guangzhou [3]		20.52	30.87	38.57
Beijing-Shanghai [2]			20.59 (35.29)	54.81
Shanghai-Nanjing			52.46	56.96
Shanghai-Hangzhou			16.29	18.52
Nanchang-Jiujiang			10.80	11.53
Dalian-Shenyang				1.01
Haikou-Sanya (Eastern Coastal line)			9.82	10.48
Hefei-Nanjing			1.37	2.83

Notes:1. The ridership data on Guangzhou-Shenzhen line are obtained from the annual report of the GUANGSHEN Railway Co.,Ltd. The other ridership information is obtained from the China Railway Yearbooks from 1999 to 2012.

2. The Hefei-Wuhan and the Shijiazhuang-Taiyuan lines were opened on April 1st, 2009. The Beijing-Shanghai line was opened on June 30th, 2011. Therefore the ridership is counted for only 9 months and 7 months, respectively. Projected ridership numbers of the whole year for these three lines in 2009 and 2011 are reported in the parentheses.

3. The Shanghai-Shenzhen line contains several segments: Ningbo-Taizhou-Wenzhou (open on 28/09/2009), Wenzhou-Fuzhou (open on 28/09/2009), Shanghai-Hangzhou (open on 26/10/2010), Fujian-Xiamen (26/04/2010), Hangzhou-Ningbo (in construction), Xiamen-Shenzhen (open on 26/12/2013). Similarly, the Beijing-Guangzhou line consists of three line segments: Wuhan-Guangzhou line (opened on 26/12/2009), Wuhan-Shijiazhuang line (opened on 28/09/2012) and Shijiazhuang-Beijing line(opened on 26/12/2012). Therefore the change in ridership reflects both the change for existing lines and the newly-generated ridership on new segments.

4. This line is an upgraded line (D-initial) opened in 2007 with a top speed of 200km/h, different from the newly constructed high speed line (Shenzhen-Guangzhou-Hong Kong) opened in 2011 with a top speed of 350km/h)

5. X denotes missing data. The blank indicates that the line is not opened in that year.

Table A7: Tourism

VARIABLES	log employment	log retail employment	log cater employment	log skilled employment	log other service employment	log other non-service employment	logGDP	log retail sales
log(NCIMA)	2.187***	2.171	-0.958	1.467***	-0.342	1.228	3.666**	2.426
	(0.648)	(1.655)	(1.728)	(0.533)	(0.767)	(1.791)	(1.538)	(1.515)
tourism resources* log(NCIMA)	0.571**	2.826***	5.021***	1.336***	1.656***	-0.350	-0.683**	1.363*
	(0.263)	(0.761)	(0.560)	(0.326)	(0.370)	(1.132)	(0.322)	(0.735)
Observations	2,805	2,550	2,547	2,747	2,546	2,717	2,800	2,807
R-squared	0.048	0.030	0.032	0.040	0.020	0.060	0.131	0.033

Notes: Data are a panel of 278 Chinese prefecture cities annually from 2003 to 2013. NCIMA is the non-connection-induced-market-access measure. log(NCIMA) is interacted with the number of 5A tourist attractions of each city. The dependent variables as listed are logs of total employment, retail and wholesale trade, hotel and restaurants employment, skilled employment (includes IT, FIRE, education, business service and scientific research), other service and non-service employment, total GDP and total retail sales. All outcome variables are counted at urban wards (shixiaqu) of prefecture cities. All regressions include city fixed effects and region-by-year fixed effects. Controls include government spending, other infrastructure measures, past city and provincial GDP, and interactions of year dummies with geographical centrality measures. Standard errors, reported in parentheses, are heteroskedasticity robust and clustered at the city level, allowing spatial dependence decaying in distance as in Conley (1999).* significant at 10%; ** significant at 5%; *** significant at 1%.

Table A8: Robustness: Limited to cities without an airport

VARIABLES	log railway ridership	log road ridership	log total ridership	log railway goods
connect	0.144**	-0.0450	0.142***	0.200*
	(0.0577)	(0.122)	(0.0438)	(0.102)
Observations	829	735	840	806
R-squared	0.056	0.012	0.069	0.031
log(NCIMA)	10.98***	-0.787	3.265*	-4.562
	(3.894)	(2.254)	(1.808)	(6.670)
log(MAhighway)	-2.159	5.784*	5.110**	-1.395
	(2.054)	(3.459)	(2.559)	(4.191)
Observations	1,207	1,329	1,341	1,229
R-squared	0.051	0.045	0.050	0.037

VARIABLES	log employment	logGDP	log housing price	log skilled employment	log tourism employment	log other service employment	log other non-service employment	logpatent
connect	0.0960***	0.0493**	0.0389*	0.0583***	0.118***	0.0575**	0.0709	0.00541
	(0.0192)	(0.0199)	(0.0198)	(0.0149)	(0.0315)	(0.0260)	(0.0560)	(0.0391)
Observations	840	841	685	835	839	796	817	754
R-squared	0.050	0.034	0.042	0.042	0.044	0.016	0.040	0.172
log(NCIMA)	1.393	3.949***	1.530	2.752***	0.5327	2.193**	1.065	-2.631*
	(1.041)	(1.269)	(0.961)	(0.612)	(2.1036)	(1.079)	(3.062)	(1.578)
log(MAhighway)	-0.136	3.490***	-1.304	1.937***	1.937***	1.632*	4.065	-0.513
	(0.862)	(0.903)	(1.028)	(0.715)	(1.7022)	(0.902)	(2.691)	(0.694)
Observations	1,341	1,339	1,087	1,326	1,320	1,226	1,303	1,141
R-squared	0.024	0.122	0.041	0.033	0.033	0.022	0.029	0.121

Notes:Data are a panel of 130 Chinese prefecture cities without any airports by the end of 2013 annually from 2003 to 2013. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A9: Robustness to Alternative Parameters and Specifications

VARIABLES	log railway ridership	log employment	log manufacture employment	log service employment	loggdp	log housing price
Set required trips to be 0.5 per month	19.31***	8.093***	8.530	9.363***	11.51**	4.801**
	(7.134)	(1.878)	(10.85)	(1.565)	(4.654)	(2.400)
Set required trips to be 1 per month	10.25***	4.386***	4.637	5.070***	6.149**	2.580**
	(3.800)	(1.015)	(5.833)	(0.839)	(2.514)	(1.304)
Set required trips to be 5 per month	3.013***	1.418***	1.470	1.612***	1.777**	0.773*
	(1.115)	(0.316)	(1.785)	(0.252)	(0.762)	(0.418)
Set required trips to be 10 per month	2.101***	1.041***	1.041	1.164***	1.185**	0.532*
	(0.768)	(0.224)	(1.254)	(0.175)	(0.521)	(0.302)
Set parameter θ to 8.28 in equation (19)	5.652**	2.442***	7.559***	2.382***	1.614	1.210
	(2.877)	(0.867)	(2.903)	(0.701)	(1.149)	(1.178)
Set parameter θ to 12.86 in equation (19)	5.667***	2.112***	3.767	2.838***	2.771***	1.589*
	(2.097)	(0.560)	(2.783)	(0.431)	(0.993)	(0.924)

Notes:Each row reports a set of estimates from the indicated specification, as discussed in the text (section 6.1). The relevant market access measure is the one that only captures the HSR-induced market access changes (MAHSR). Robust standard errors clustered at city level are reported in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A10: Exploiting variation in HSR connection timing purely due to engineering difficulty differences

VARIABLES	lograilpas	logemployment	logmanuemp	logserviceemp	logwage	loggdp	loghouseprice
connectalt	0.159**	-0.00342	-0.0701*	0.0476*	0.0518**	-0.0605**	0.00317
	(0.0680)	(0.0297)	(0.0392)	(0.0242)	(0.0251)	(0.0302)	(0.0417)
Observations	213	226	227	227	225	227	226
No. of cities	27	27	27	27	27	27	27
R-squared	0.323	0.297	0.299	0.220	0.253	0.545	0.121
VARIABLES	log SOE employment	log other service employment	log catering employment	log industrial firm numbers	log industrial revenue	log retail firm numbers	log retail sales
connectalt	0.0199*	0.0656**	-0.0772	0.00412	-0.00751	-0.0118	0.207***
	(0.0108)	(0.0256)	(0.0705)	(0.0424)	(0.127)	(0.0598)	(0.0667)
Observations	224	227	227	227	227	224	227
No. of cities	27	27	27	27	27	27	27
R-squared	0.147	0.265	0.289	0.288	0.215	0.200	0.213

Notes:Data are a panel of 27 Chinese prefecture cities annually from 2003 to 2013, whose HSR construction started in 2005. Connectalt is a dummy of an alternative connection measure. It is the projected year of HSR opening based purely on engineering difficulty driven construction duration predicted by the length of HSR lines and total bridge/tunnel percentage. Standard errors, reported in parentheses, are heteroskedasticity robust and clustered at the city level, allowing spatial dependence decaying in distance as in Conley (1999).* significant at 10%; ** significant at 5%; *** significant at 1%.

Table A11: Impacts of HSR on peripheral areas

VARIABLES	log suburban employment	log suburban population	log suburban GDP	log suburban employment	log suburban population	log suburban GDP
connect	-0.0325	0.00464	-0.0644			
	(0.0302)	(0.00677)	(0.0459)			
log(NCIMA)				-0.724	-0.428	-5.710*
				(1.254)	(0.375)	(3.439)
Observations	1,768	1,771	1,771	2,801	2,806	2,801
No. of cities	172	172	172	279	279	279
R-squared	0.043	0.022	0.036	0.032	0.020	0.034

Notes:Data are a panel of 278 Chinese prefecture cities annually from 2003 to 2013. The outcome variables are employment, population, and GDP at the peripheries (areas out of urban wards) of prefecture cities. Standard errors, reported in parentheses, are heteroskedasticity robust and clustered at the city level, allowing spatial dependence decaying in distance as in Conley (1999).* significant at 10%; ** significant at 5%; *** significant at 1%.

Figure A1: Aggregate ridership on different modes of transportation: 2010 to 2014

Notes:The figures display the aggregate number and percentage of ridership on high speed railway, conventional railway, road, air and water travel from 2010 to 2014. The unit is 100 million.

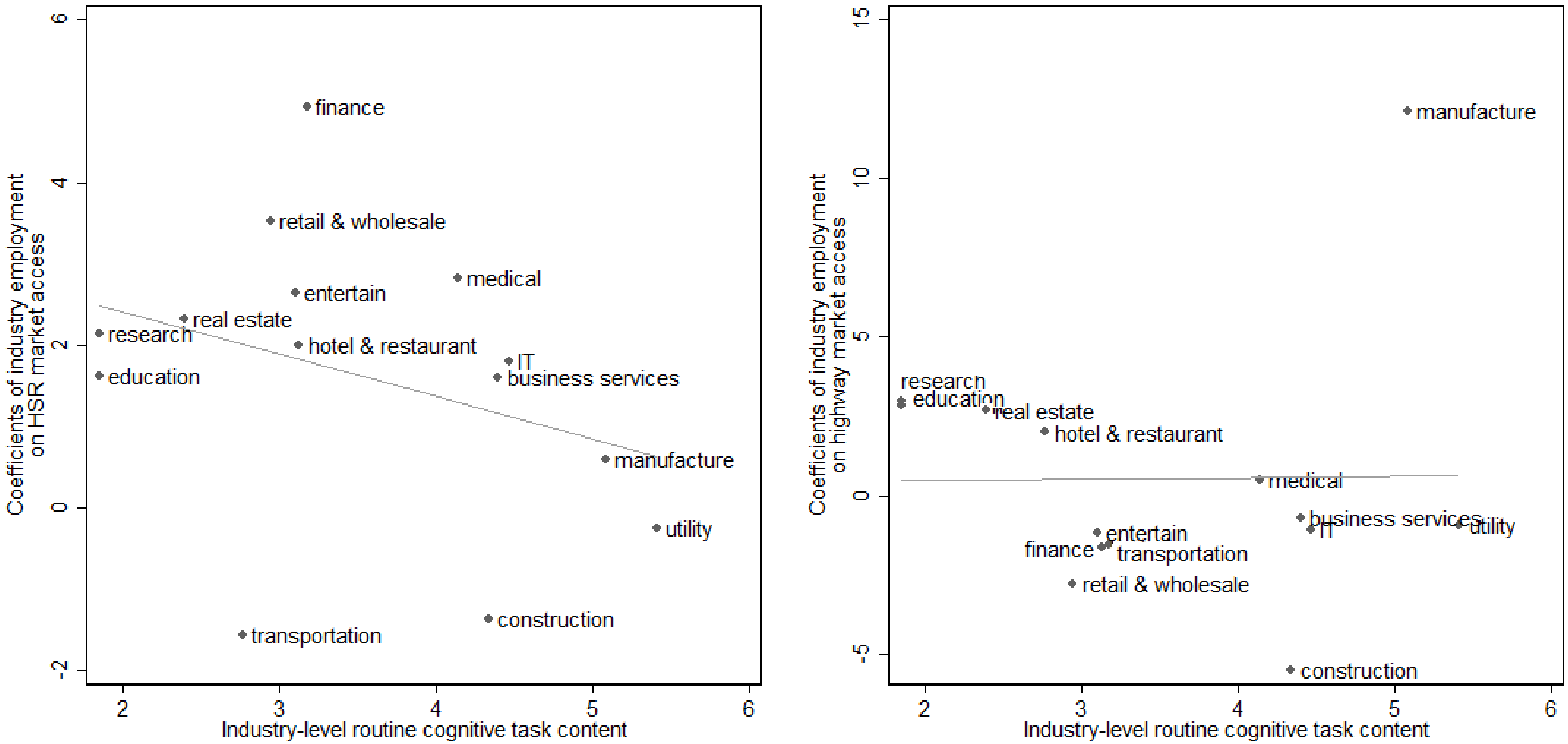


Figure A2: Sector-specific HSR and highway impacts and routine cognitive task intensity

Notes:The figures plot estimated coefficients on the impacts of NCIMA highway-induced market access (MAhighway) on sectoral employment reported in Table 2.6 against sector-specific routine task intensity reported in Autor et al. (2003). The matching between Chinese industries and the US industries is reported in Table A2.

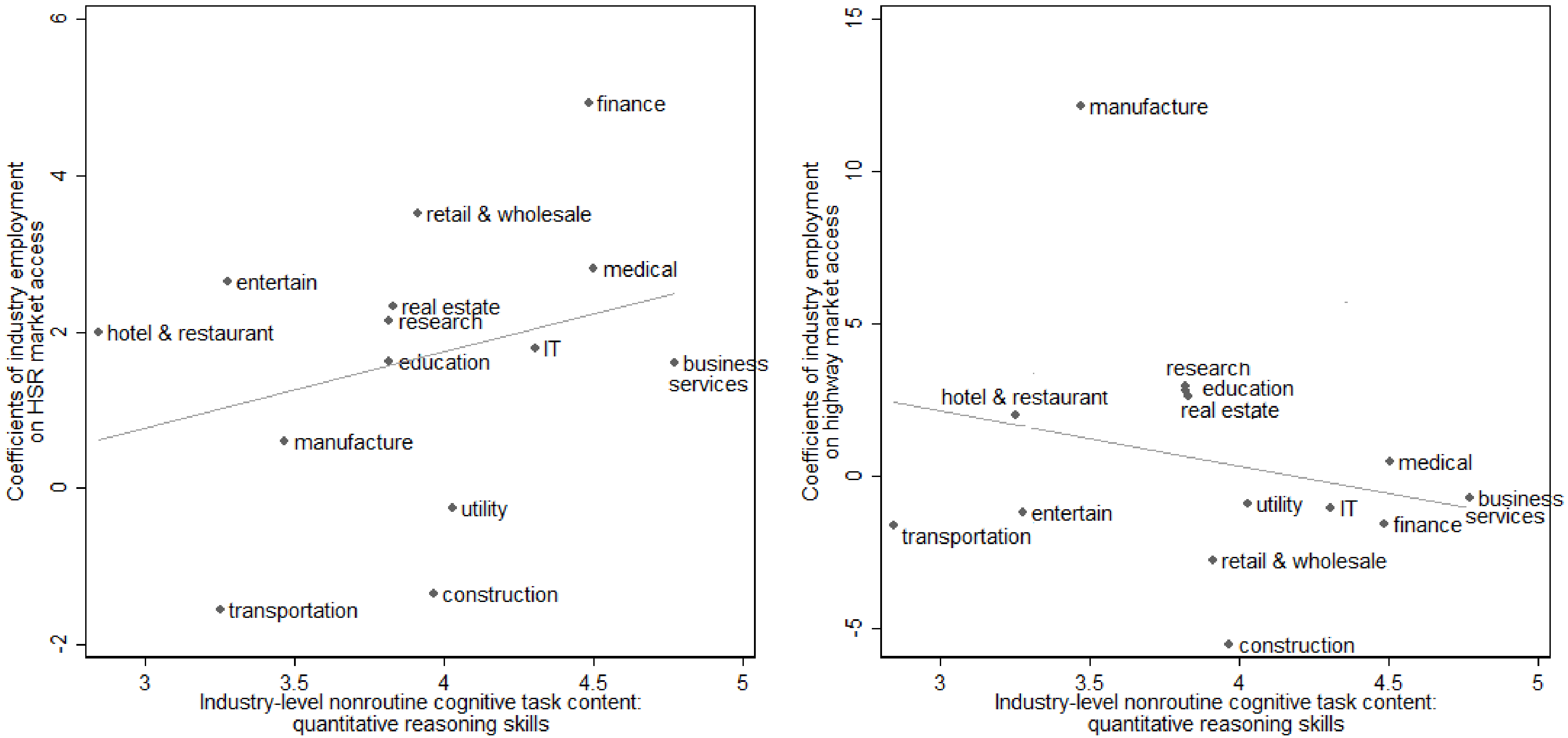
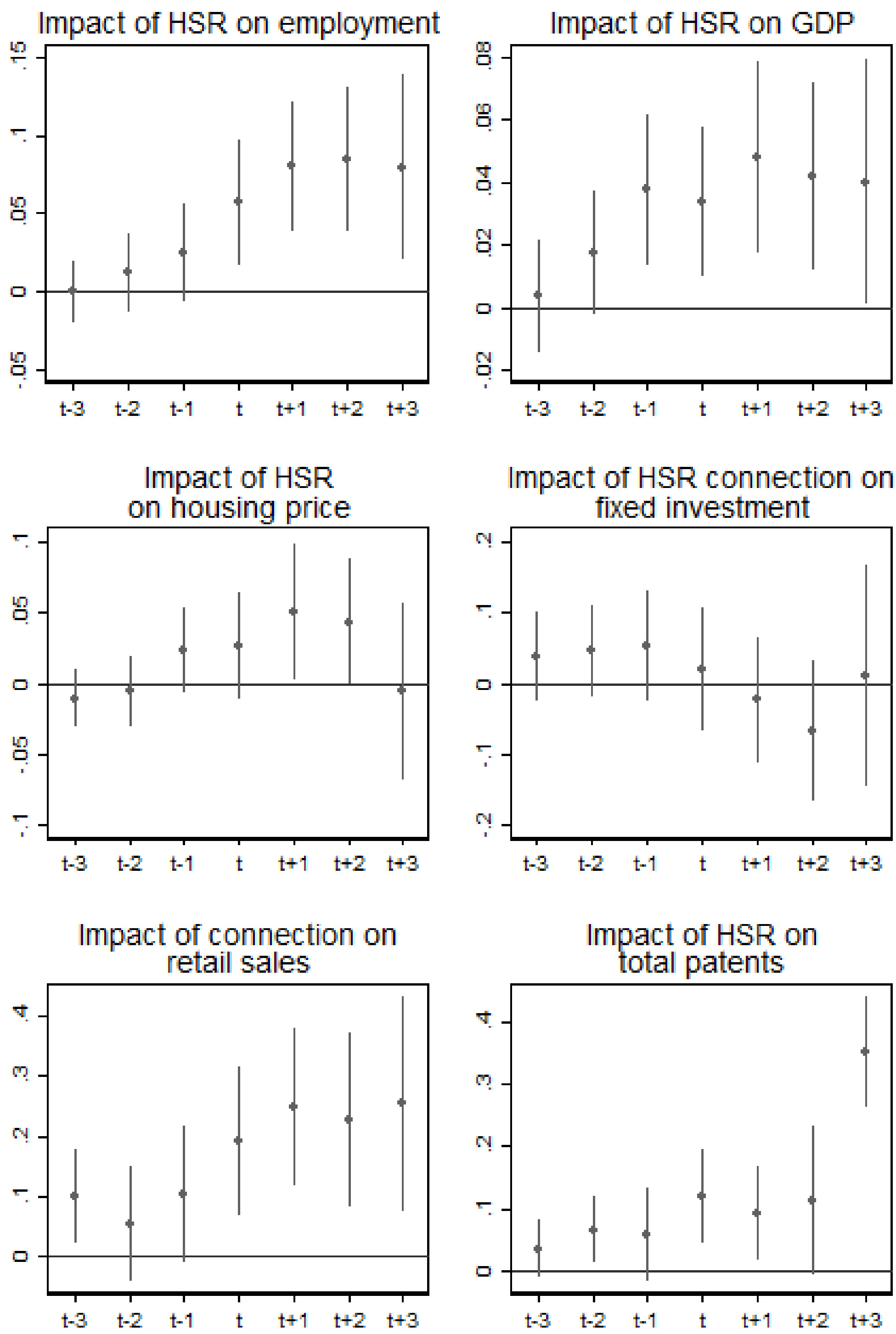


Figure A3: Sector-specific HSR and highway impacts and nonroutine analytical task intensity

Notes:The figures plot estimated coefficients on the impacts of NCIMA highway-induced market access (MAhighway) on sectoral employment reported in Table 2.6 against sector-specific nonroutine analytical task intensity reported in Autor et al. (2003). The matching between Chinese industries and the US industries is reported in Table A2.

Figure A4: Event study with connection dummy: Aggregate economic outcomes



Notes:The figures display the estimated coefficients and 95% confidence intervals in regressions where the dependent variables are log total employment, GDP, housing private, fixed investments, total retail sales and patents. The independent variables are the leads and lags of the initial connection dummy. For all the regressions, the sample is a balanced panel of 172 Chinese prefecture cities from 2003-2011, as HSR connection information is available only until 2014, and the third lead is a missing value for observations after 2011.

Chapter 3

Where does the Wind Blow? Green Preferences and Spatial Misallocation in the Renewable Energy Sector

Are “greener" investments less efficient? This paper looks at the location choices of wind power investors. I measure the efficiency loss in this sector due to location choices away from profit-maximization and explore the factors contributing to it. Using extensive information on wind resources, transmission, electricity prices and other restrictions are relevant for the siting choices of wind farms, I calculate the predicted profitability of wind power projects for all the possible places across the contiguous US, use the distribution of this profitability as a counterfactual for profit-maximizing wind power investments and compare it to the actual placement of wind farms. The average predicted profit of wind projects would have risen by 47.1% had the 1770 current projects in the continental US been moved to the best 1770 sites. It is also shown that 80% and 42% respectively of this observed deviation can be accounted for by within-state and even within-county distortions. I show further evidence that a large proportion of the within-state spatial misallocation is attributable to green investors' “conspicuous generation" behaviour: wind farms in more environmental-friendly counties are more likely to be financed by local and non-profit investors, are closer to cities, are much less responsive to local fundamentals and have worse performance ex-post. The implementation of state policies such as Renewable Portfolio Standard (RPS) and price-based subsidies are related to better within-state locational choices through attracting more for-profit investments to the "brown" counties, while lump-sum subsidies have the opposite or no effects.

My findings have salient implications for environmental and energy policy: policy makers should take account of the non-monetary incentives of renewable investors when determining the allocative efficiency of policies.

3.1 Introduction

Location is the most important determinant of some industries' productivity. Large economic loss can occur when plants are located in wrong places due to insufficient information on site suitability, distorting land use restrictions or certain place-based policies. Although spatial misallocation has widely acknowledged as an important source of inefficiency, it is hard to measure the exact loss in productivity caused by poor location for two reasons. First, a lot of the locational fundamentals that matter for specific industries are not observed by researchers. Second, various agglomeration and dispersion make the suitability of locations for firms in particular sectors dependent on the pretense of other firms from the same or related industries nearby. In this paper, I attempt to circumvent these problems by looking at the locational efficiency within the renewable energy sector, a sector where locational fundamentals are very important and largely observable, where agglomeration and dispersion forces are relatively weak, and where regional energy policies play a significant role. My research design also allows me to uncover factors that contribute to the mislocation-induced efficiency loss within the wind power sector.

Economic efficiency and environmental benefits of renewable energy sector, as suggested by Cullen (2013), Zivin et al. (2014) and Callaway et al. (2015), critically rely on the proper siting of these projects. As a general rule, wind turbines should be located in places with strong and stable winds, reasonably good access to electricity transmission, high wholesale electricity prices and no restrictions on wind farm development. In this paper, I adopt a novel method to directly compare the location choices of actual wind projects across the contiguous US to a profit-maximization counterfactual spatial allocation using rich information on local wind intensity, grid access, electricity price, as well as restrictions on wind power placement.

In practice, I divide the contiguous US into 75,147 10km*10km grid-cells and evaluate the profitability of placing wind power projects in each of these cells, subject to necessary exclusions. I then calculate the differences in predicted profits between existing cells and the best N cells in predicted profitability, where N is the number of grid-cells with wind power projects. I find a discrepancy of 47.1% in predicted profits: if we move all of the 1770 current wind projects to the best possible sites, the predicted average profit of these projects will grow by 47.1%. Moreover, I find the *within-state* misallocation alone accounts for an efficiency loss of 37.4%, over 80% of the total observed distortion. Even if we focus only on within county distortion, the most conservative loss in efficiency is still measured as 19.8%. Large efficiency gains are expected had wind power investors been better at picking sites within their own states or even within their own counties. In fact, equalizing state-level incentives for green energy is only able to boost aggregate efficiency by 1% to 5% [1] since cross-state distortion is not large in magnitude compared to within-state distortion.

The natural next step is to examine potential explanations for the significant observed locational inefficiency. A closer look at the data reveals that wind farms located in "greener" counties, measured by local support for the Democratic and Green Party in presidential elections, are located at less windy sites and perform worse ex-post. They also tend to be invested by non-profit and local investors. They are also more likely to be placed in urban areas, commonly thought to be not suitable for wind power projects[2] but obviously make these wind turbines more salient to the public. I further show that differences in green preferences are quantitatively important in accounting for the observed within-state and within-county distortion. Moving from 25 to 75 percentile in the local "greenness" measure leads to the location of wind farms 20% less responsive to local fundamentals and more than doubles the within-county distortion measure.

One possible explanation for this behaviour is that instead of doing a global search for the most productive sites, "greener" developers of renewable energy projects might prominently install wind turbines on their own properties or at least within their local counties as a demonstration of preferences for environmental protection. It could also be that green investors are smaller and unspecialized organizations with disproportionally higher search and monitoring costs. In either case, the existence of non-monetary motives for renewable energy is central to this particular locational misallocation, a phenomenon with novel and interesting policy implications. Policies that are ex-ante equivalent and are equally attractive to profit-maximizing investors might actually screen investors with different levels of green

[1] The aggregate efficiency gain from equalizing state-level incentives is calculated by estimating the policy treatment effects and generating the predicted wind capacity addition for each state by assuming the intensity of policies to the same across states, while keeping the aggregate treatment effects of wind capacity addition to be the same. Some assumptions are needed to evaluate the change in aggregate efficiency level in the counterfactual configuration. I assume the average efficiency level for each state under the counterfactual allocation is the same as the mean/median/max estimated profitability of occupied cells (before any renewable policies are applied). The estimated change ranges from 1% to 5% under these different assumptions.

[2] "Locations in narrow valleys and canyons, downwind of hills or obstructions, or in forested or urban areas are likely to have poor wind exposure.", by the National Renewable Energy Lab (NREL), http://www.bbc.co.uk/blogs/ethicalman/2009/12/why_micro_wind_turbines_dont.html

preferences differently, resulting in starkly different ex-post allocative efficiency.

Therefore, I further investigate the role of state-level renewable policies in changing allocative efficiency within-state and how it interacts with investors' green preferences. I collect information on these policies from DSIREUSA (Database of State Incentive for Renewables and Efficiency), and loosely divide them into three categories: quantity-based policies such as Renewable Portfolio Standard (RPS), per-unit-price-based (performance-based) policies such as feed-in-tariff and certain corporate tax breaks, and direct subsidies (non-performance-based) such as tax breaks on equipment and property. I try to aggregate several different policies into a single index of policy intensity under these three categories based on their impacts on the predicted profits of a typical wind farm project. In a simple difference-in-differences specification, I find RPS and price-based policies associated with significantly better locations of wind projects within-state. An important reason is that these policies are more attractive to pure profit-maximizing investors, who are adding capacities mainly in "brown" counties. Direct subsidies neither change within-state allocative efficiency nor have differential impacts on wind power capacity addition across counties with different environmental attitudes. For better identification, I restrict my sample to gridcells around state borders and check the dynamic effects of renewable policy incentives before and after their actual implementation. The key results are robust to these alternative specifications.

I then come up with a model of private provision of public good with heterogeneous green preferences, similar to Jacobsen et al. (2014). I introduce search costs for picking a suitable site for wind farms and assume that green investors derive additional utility from having wind farms in their local areas, rendering fewer benefits from searching. This model nicely accommodates all my key empirical findings. It also predicts that when comparing the extra public benefits of policies, direct subsidies are dominated by other performance-based or mandate-type policies with the existence of green preferences.

My empirical findings have several novel and important policy implications. The most important rationale of renewable support schemes is that they are the more politically-accepted way to internalize the public benefits generated by renewable electricity generation. Therefore, they should be designed in a way to realign public and private benefits/costs of renewable investments. However, one of the most important lessons we learn here is that we have paid too little attention to the importance of green preferences in green investors' private benefits, which is shown to be negatively correlated to the public benefits generated by a wind farm project given the same amount of private costs[3]. In light of this, non-performance-based renewable support schemes are clearly dominated since they tend to screen in greener but less efficient investments. The tradeoff between price-based instruments and renewable energy standard largely depends on to what extent the standard is able to incorporate each location's unique mix of electricity generation resources and other restrictions associated with the public benefits of renewable energy. On a related note, to engage agents with strong environmental preferences, promoting markets for green electricity where people can purchase electricity generated from renewable sources at a premium and get visible awards for it would be a better idea than encouraging them to invest in their own renewable energy projects.

This paper contributes to a burgeoning literature on green preferences and consumer behaviour. Kahn and Kok (2014) looks at the capitalization of green labels in California housing market. Sexton and Sexton (2014) attributes consumers' enthusiasm on Prius to "conspicuous conservation", a costly signalling of one's concerns for the environment. Bollinger and Gillingham (2012) underscores peer effects as the motives for people to install solar panels. Instead, I am exploring the importance of green preferences in steering investors' behaviour and it is somewhat surprising to notice that the importance of green preferences is also significant in this setting, where agents are perceived to be more "rational" and profit-oriented. A major distinction of this paper from the previous research is that I explicitly document and quantify the loss in efficiency due to this special "conspicuous generation" motive of green investors and examine the effects of financial incentives in partially offsetting it. It also relates to the intrinsic incentive crowding out topics in psychology and public economics literature, also from a very different angle. I show that extrinsic incentives such as renewable energy subsidies, albeit crowd out intrinsic motivation in green investments, encourage the investors to adopt a more "profit-maximizing" thinking, which could be desirable from the policy makers' perspective.

I also evaluate the impacts of renewable energy policies from an unusual angle. In my paper, I assess how the implementation of state-level renewable energy policies reshape the cross-state and within-state allocative efficiency of wind farms. Among plenty of papers that explicitly looked at the effectiveness of renewable energy policies (Bird et al. (2005),Yin and Powers (2010), Delmas and Montes-Sancho (2011)) systematically analyses the causes and effectiveness of typical US state-level policies in adding capacities. At a more micro level, Cook and Lin (2015) finds that Danish renewable incentives significantly impacted the timing of shutdown and upgrade

[3]In the United States, the correlation between environmental friendliness and local wind resources is negative. Moreover, the additional environmental benefits of wind power generation are smaller with a higher proportion of other renewable or clean energy in local energy mix.

decisions made by turbine owners.

This paper is also related to the broader practical question of second-best energy policy design in face of large multidimensional heterogeneity when the first-best is unattainable. I document an unintended source of distortion in this case: the tendency of some subsidies to attract environmentalism-inspired but less efficient investments. Other papers have looked at different policies at different scenarios. Ryan (2012) shows how regulation might hurt social welfare through increasing market power. Fowlie (2010) shows that heterogeneity in plant ownership structure largely affects the effectiveness of environmental regulation. Ito and Sallee (2015) discusses the pros and cons of attribute-based regulation, which helps to equalize compliance costs but brings in extra distortion.

Finally, this paper makes a contribution to the spatial economics literature by directly estimating the loss in aggregate productivity due to spatial misallocation. The particular setting of my problem allows me to quantify this kind of loss by directly comparing the actual location distribution to a well-established counterfactual using rich information specific to the industry, without relying on a structural model as in Bryan and Morten (2015) and Fajgelbaum et al. (2015). My findings underline the importance of investors' preferences in determining industrial location, consistent with a "jobs follow people" story.

The paper is structured as follows: section 2 prepares the readers with the background knowledge of US wind power industry and relevant renewable energy policies; section 3 introduces the data and methods to measure wind farm locational distortions; section 4 presents the main findings on different sources of distortion; section 5 shows evidence on the distorting roles of green preferences and counteracting policy effects; section 6 presents a simple model of private provision of public goods with green preferences that brings together all my empirical findings; section 7 concludes.

3.2 Background

3.2.1 Wind power in the US

Wind power in the United States expands quickly during the past several years and is taking up an increasingly important role in the energy mix of the US. As of the end of 2014, the total wind capacity was 65,879 MW, which generates 4.45% of the total electricity produced in the US. Over the past ten years, the US wind industry has had an average annual growth of 25.6% over the last 10 years.

Wind power is widely considered to be the most cost-effective type of renewable energy apart from hydropower and is therefore expected to grow even more in the future as the country relies more on renewable energy. A US Department of Energy report finds 35% wind penetration by 2050 is "plausible", in terms of grid reliability and cost, as well as the industry's ability to scale up.[4] And the EPA projects that renewables could rise to 28 percent of the electricity supply by 2030 with Clean Power Plan in place.[5] Therefore it is time for us to think about how efficiently have we been able to place existing wind projects and what can we do to improve the allocative efficiency of this sector. Removing the persistent distortion in this sector may prove to be as important as innovation in wind power generation and storage technology in bringing renewable energy to be cost-competitive with fossil fuels.

Figure 1 shows the distribution of wind farms across the US. Figure 2 looks deeper into the allocative efficiency of them. Figure 2.1 plots the density of wind farm distribution across different wind power classes. Wind power class is a measure of wind resources, where 7 stands for the strongest wind and 1 stands for the weakest. The NREL (National Renewable Energy Lab) suggests that only areas with WPC greater or equal to 3 are suitable for utility-scale wind turbine applications[6]. However, from figure 1.1 we can see that about 30% of the current US wind projects are located in areas with WPC smaller than 3. Figures 2.2 and 2.3 further show that the wind farms that are located in low wind areas (WPC=1 & 2) are not closer to electricity grid or are in areas with higher retail electricity prices than their counterparts in the middle range wind areas (WPC=3 & 4), suggestive of a significant amount of spatial misallocation of wind farms across the country. Finally, Figure 2.4 plots the average local environmental attitude measure[7] of wind farms across different wind classes. Quite interestingly, I find that the wind farms exposed to little wind are located in counties with higher preferences for environmental protection. Therefore strong green preferences of the investors could work against incentivizing them to look for sites that make the most economic sense. In section 5, I am going to explore these phenomena quantitatively.

3.3 Data

My analysis draws on three main sources of data: the database on the fundamentals of wind farm location, information on the distribution and performance of wind power projects, and a comprehensive dataset on state-level renewable energy incen-

[4] http://energy.gov/sites/prod/files/WindVision_Report_final.pdf

[5] http://www.vox.com/2015/8/4/9096903/clean-power-plan-explained

[6] http://www.nrel.gov/gis/wind_detail.html

[7] Local environmental attitude at county level is measured as a linear combination of average county income, college graduate share, votes share for democratic and Green Party in 2012 presidential election, similar to Allcott (2015)

tives. I will describe them in turn.

3.3.1 Locational fundamentals

To establish a reliable counterfactual of profit-maximizing wind farm distribution, we need information on the local fundamentals that are critical to the profitability of wind farms. I collect information on wind resources, electricity transmission lines, electricity prices, and the restrictions on wind placement. I then compile a database of 75,147 10km*10km gridcells covering the whole continental US and match all the locational fundamental attributes to each gridcell. It allows me to work out a single measure of potential wind power project profitability at the cell level.

Wind resources: The main wind resource data I use are drawn from the annual average wind resource data used in the Renewable Electricity Futures Study (http://www.nrel.gov/analysis/re_futures) from the National Renewable Energy Laboratory (NREL). The majority of the onshore wind data was modelled at a 50 m hub height and vertically adjusted to 80 m height to better represent current wind technology. Wind resources are divided into 7 categories with 1 representing the worst and 7 the best.

One particular drawback of using an annual average wind resource measure lies in the fact that there is a large variation in wind intensity from time to time, and the revenue generated from wind production largely depends on how the peak of wind power coincides with that of electricity demand. To deal with this issue, I obtain alternative simulated wind production data the National Renewable Energy Laboratory's (NREL) Eastern Wind dataset[8] and Western Wind dataset[9]. These datasets are created for energy integration studies by NREL and its partners. Simulated power production from hypothetical wind plants in every ten minutes from 2004 to 2006 is generated for 32,043 sites across the Western United States and 1,326 across the Eastern United States. Mapping these sites to my gridcells generates time-series wind production information for 5866 gridcells.

Electricity transmission: I draw the information on electricity transmission infrastructure from a GIS file on 2001 US main electricity transmission lines above 60KV.

Agricultural land value: The county-level agricultural land value for 2014 is collected from the United States Department of Agriculture (USDA) statistics service dataset [10].

[8]http://www.nrel.gov/electricity/transmission/eastern_wind_methodology.html

[9]http://www.nrel.gov/electricity/transmission/western_wind_methodology.html

[10]www.nass.usda.gov

Electricity prices: I obtain 2010 retail electricity rates by residential, commercial and industrial uses across over 4,000 pricing units from data reported by Energy Information Agency of the US (EIA). Rates were matched from EIA data and Ventyx (2010) territory shapes. I complement this data with hourly wholesale electricity rates from 2004-2010 across 24 pricing hubs gathered from Bloomberg.

Exclusions: I rely on the National Land Cover Database 2001 (NLCD2001) to eliminate places that are not suitable for wind power development. Incompatible land use includes urban, wetlands and perennial snow areas. Mountainous areas with a slope steeper than 20 degrees, calculated using the USGS national 90-meter spatial resolution National Elevation Dataset, are also excluded. Finally, I exclude Bureau of Land Management (BLM) and National Science Foundation (NSF) protected areas, brownfield, national parks, federally owned land, national trails and tribal lands, according to the BLM. GIS data on exclusion are matched to the gridcell database. A gridcell is defined as not suitable for wind power development if over 70% of its area is covered by excluded areas.

3.3.2 Wind power projects distribution and performance

I merge three datasets to get an as complete as possible picture of the characteristics and performances of current wind power projects across the continental US. US Geological Survey (USGS) gather information on the exact location, mode, operation date and owner wind farm of over 48,000 wind turbines in the US through March 2, 2014. Energy Information Agency (EIA) publishes annual reports on power plant generation (EIA-923) and generators (EIA-861) up to 2013, which includes information on capacity, generation, emission, interconnection and other characteristics of 821 wind power plants whose operation commenced before 2013. I also obtain detailed ownership, developer and operator information on over 1214 wind projects from Thewindpower (`www.thewindpower.net`). I merge these three datasets together by the names of the plant/project and year of operation. Over 70% plants in the EIA dataset are matched to both USGS turbine-level dataset and Thewindpower project-level dataset.

3.3.3 State Level Renewable Policies

There are various support schemes for renewables across the US implemented at different levels. At the federal level, we have the Production Tax Credit (PTC) and the Investment Tax Credit (ITC), which reduces federal income taxes for qualified tax-paying owners of renewable energy facilities based on either electrical output or capital investment in renewable energy projects.

At the state level, the most important policy is the Renewable Portfolio Standard (RPS), where utilities within the implementing states are required to source a given proportion of its electricity generation from renewable sources. Apart from it, there exists a number of different kinds of subsidies. I try to categorize them into performance-based and non-performance-based ones for my analysis. Support schemes can also be awarded by individual utilities or municipals, but many of them are direct responses to RPS. Therefore, throughout this paper, I am going to limit my attention on state-level policies only.

Information on state-level renewable energy policies and incentives is gathered from the Database of State Incentives for Renewables and Efficiency (DSIREUSA, www.dsireusa.org). Since there are so many different types of renewable policies and incentive schemes, I categorize them into three main groups and generate a single index of policy intensity within each group. I use a few criteria of exclusion to simplify my categorization. These three groups are:
(1) Direct fixed cost subsidies that cover parts of the fixed cost of wind projects and are not dependent on performances, including equipment sales tax exemption, property tax exemption, grants, interconnection cost subsidy, support on feasibility studies etc.;
(2) Price based subsides given to per unit electricity generated, hence depends on performances, including feed-in-tariff, performance-based rebates, and corporate tax credits;
(3) Quantity based policies that stipulate the minimum amount of renewable electricity generated, such as renewable portfolio standard (RPS).
I then apply the following rules to exclude policies that are not considered for my analysis.
1. I focus only on state-level policies. Policies on the federal or municipal level are not considered. Policies implemented by individual utilities are not included as well.
2. I exclude policies that cannot be categorized loosely into the aforementioned three groups. Policies like green power purchase options or loan programs are not counted.
3. I exclude policies that are not awarded directly to wind farm developers, such as industrial support for wind turbine and parts manufacturers.
4. I exclude policies with too restrictive size or ownership requirements. (Policies with a maximum limit over 10MW and a minimum limit under 100MW, or are dedicated to particular ownership groups (i.e. residential only) are excluded)

With these requirements in mind, I define the index for per-unit-price-based (performance-based) policies as the total amount of extra money given to per unit electricity generation, the index for direct subsidies (non-performance-based) as the estimated percentage of total upfront cost saved, and the index for quantity based policy (RPS) as the "real" measure of target stringency each year ($RPS_{st} = Norminal\, RPS_{st} - \frac{Renewable_{s,t-1}}{Total_{s,t-1}}$), where $Norminal\, RPS_{st}$ is the nominal RPS target on the minimum proportion of electricity sales from renewable sources, and $\frac{Renewable_{s,t-1}}{Total_{s,t-1}}$ is the actual proportion of electricity sales from renewable sources last year.

3.4 Measuring Locational Distortions

I follow three steps to obtain a systematic measure of locational misallocation at different levels in wind power industry. First, I evaluate the contribution of locational fundamentals to wind power plant performance. Second, I divide the continental US into 75,147 10km*10km gridcells and calculate the predicted profitability of each cell. Third, I define my distortion measure as the difference between the average profit of current wind projects and the average that can be attained should they be reallocated to the best gridcells.

To weigh the contribution of different locational factors such as local wind resources and transmission access to the general profitability of wind power projects, I define the location-varying predicted revenue per kW of wind capacity installed as:

$$Predicted\, Capacity\, Factor * (1 - \%Loss\, in\, Transmission) * AverageElectricityPrice/kW$$

As there are different measures on wind resources and electricity prices, I come up with multiple measures of revenue for robustness, which I will discuss later.

On the cost side, two of the most important sources of location-varying fixed cost are grid interconnection cost and land rental cost[11]. I subtract the location-varying fixed cost from the revenue function to get a profitability measure of wind farms. The interconnection cost is calculated based on the distance of wind farms to the closest electricity grid. EIA-861 series report interconnection costs for most of the wind power generating units in the US. Therefore I regress the actual interconnection

[11] A report by the European Wind Energy Association http://www.ewea.org/fileadmin/files/library/publications/reports/Economics_of_Wind_Energy.pdf shows that grid connection and land rent takes up 8.9% and 3.9% of the total setting up cost of a typical 2 MW wind turbine

cost on the distance to electricity transmission lines and the size of the power plant to and get a prediction of interconnection costs for each wind turbine installed in any of the 75,147 gridcells. I amortize these two sources of fixed cost over 15 years, the lifespan of a typical wind farm, with an annual interest rate of 3%.

To calculate predicted wind power production, capacity factor is a common measure in electrical engineering defined as the ratio between annual total electricity generation and the maximum amount of electricity generated at full capacity during one year. Since wind power is an intermittent energy source and wind turbines are not working when there is no wind, the capacity factor of a typical wind power plant usually ranges from 20% to 40%. I predict the capacity factor for a typical wind power plant in a given gridcell using information on the annual average wind speed of that gridcell. To obtain a reliable relationship between average wind speed and power plant capacity factor, I look into the National Renewable Energy Laboratory (NREL) Eastern and Western wind datasets, which reports hourly wind speed and simulated capacity factor over two years across 30,000 sites in the US. I regress the simulated capacity factor on yearly average wind speed to get a coefficient of the importance of wind resources to production efficiency.

A shortcoming of this method in predicting wind power generated revenue lies in the fact that there is a large variation in wind intensity from time to time, and the revenue generated from wind production largely depends on how the peak of wind power coincides with that of electricity demand. So as an alternative I also use the simulated capacity factor reported by the NREL Eastern and Western wind datasets directly. The advantage of the second source is that it provides us with detailed variation in simulated wind power production per hour for three years, and the disadvantage being this information is only available for only 5866 of my 75,147 gridcells. Among them, only 317 of the 1,770 occupied cells are covered. To avoid dropping too many occupied cells from my sample, for those occupied cells without detailed time-series wind production information, I use the information from the closest sites to them as a proxy as long as the distance between the cell and the observed simulation site is less than 30 km. This operation leaves me with 1128 occupied cells in the end.

I use both wholesale and retail electricity prices in my revenue calculation. Both methods have their respective pros and cons. Wholesale electricity prices are the prices faced by wind power plants and they are available at high frequency, allowing us to capture the fluctuation of electricity demand across different points of time within a day. But they are only observed at 24 trading hubs. Retail electricity prices are available at over 4000 price units across the US annually. But they are the prices faced by consumers and markups between retail and wholesale prices could be different across places. I use the retail prices for my main specification as I believe it could better capture the demand side differences but I use wholesale prices for robustness. I then factor in the loss in transmission and get an estimate of the amount of money received per unit electricity generated by the wind farm ($(1-\%Loss\, in\, Transmission) * average\, electricity\, price$). The loss in transmission depends, of course, on the type of prices I use. With retail prices, the loss depends on the distance to the distribution lines, and with wholesale prices, the loss depends on both the distance to the closest 375 kV electricity transmission lines and the distance to the electricity trading hub.

For robustness, I define four different profitability measures. On the production side, I use either the predicted wind power production based on annual average wind speed, or the simulated wind power production hourly data that are available only for a subset of gridcells. As for the price, I use either retail or wholesale price data. For simplicity, to generate the profitability measure using hourly simulated data and wholesale prices, I define off-peak time to be 12:00 p.m. to 8:00 a.m. next morning, and peak time to be the rest. I then aggregate both wholesale electricity price data and simulated wind production data to peak and off-peak ones and generate the total predicted revenue. The combination of two production estimation methods and two price sources produces four profitability measures. The baseline one is the one that uses annual average wind speed and retail electricity price. Table 3.1 reports the correlation across these four measures. As Eastern and Western Wind datasets use a different methodology in simulation. I split them into two separate samples and report the correlation separately. It is clear that the correlation between them are quite high.

With a reliable measure of potential profitability of wind power projects across the continental US, I define the total loss in wind farm spatial misallocation as:

$$\frac{Average\, profit\, of\; 1770\; best\, cells\, nationwide - Average\, profit\, of\, 1770\; built\, up\, cells\, nationwide}{Average\, profit\, of\; N\; built\, up\, cells\, nationwide} \tag{3.1}$$

Over concerns about grid stability, I impose a restriction on the upper bound of wind penetration: in the optimal allocation, the proportion electricity coming from the wind should not be more than 30% of the total generation for any states.

Similarly, I am able to produce a within-state(county) measure of mis-location loss:

$$\frac{Average\, profit\, of\, the\, N\, best\, cells\, in\, state(county) - Average\, profit\, of\, N\, built\, up\, cells}{Average\, profit\, of\, N\, built\, up\, cells} \tag{3.2}$$

As mentioned, for robustness I generate four different measures of wind power profitability. Accordingly, I come up four distortion measures. Table 3.2 report these spatial misallocation measures at national level. The baseline reveals a total efficiency loss measure of 47.1%. That being said, the average profit of 1,770 continental US wind farms will increase by 47.1% should they be moved to the best 1,770 gridcells in the US. The measured distortion (43.6%) is slightly smaller if we are using wholesale other than retail electricity prices. Because the simulation method is different across the Eastern and Western datasets, I generate the distortion measures for Eastern/Western US separately. So the distortion measure from Row 3 to Row 6 can be interpreted as the change in average profit by moving the current wind farms to the best cells in Eastern/Western US. Since the simulation data are only available for a subset of gridcells (4,661 for the Western US, 2,003 for the Eastern US.), distortion measures based on them are more likely to be underestimated, and the extent of underestimation is larger for the Eastern subsample with less alternative gridcell's. They report alternative distortion measures from 11% to 37%.

Within-state allocative efficiency loss for different states is reported in Table 3.3. The first column and the second column report the distortion measures based on profitability measures using predicted production data based on annual average wind speed. The third and fourth columns report distortion measure based on the simulated production measures that take account of fluctuation in wind resources across time. It is clear that these four within-state distortion measures are highly correlated. For the rest of my analysis, I stick to the first column as my baseline.

We can see that there is a large variation in the current allocation efficiency across the US states. In Iowa, the observed efficiency loss due to the mislocation of wind projects is less than 10%, while in Maine, the average profit of wind power plants can go up by 110% if they are placed optimally. Weighting state-specific within-state efficiency loss with the total wind capacity of each state gives us a 37.4% efficiency loss driven purely by suboptimal wind farm siting choices within state. Even more surprisingly, the measured efficiency loss remains at 19.8% even if we only consider within-county distortion, which should be relatively free of most concerns on scheduled electricity demand and supply at the state level. It means that instead of placing wind farms in the wrong states, we should worry more about wind power investors not choosing the right sites within their own states or even within their own counties. To take a closer look at what might drive within-county misallocation, Figure 2 plots the relationship between measured within-county misallocation and the support for the Democratic Party at the county level, revealing a significant and negative relationship. Moving from the 25th to 75th percentile in the local "greenness" measure more than doubles the within-county distortion measure.

As a more rigorous attempt to examine the factors that contribute to this observed within-county spatial misallocation, I turn to regression analysis. Table 3.4 reports the correlation between a normalized distortion measure and county-level greenness measures, the mean and standard deviation of profitability within-county, the percentage of cells that are not suitable for wind power placement, as well as a variety of demographic and social economic measures. It is shown that the support for the democratic and Green Party is positively correlated with the county-level distortion measures. Apart from that, the distortion measure is also increasing in the college graduates share in some specifications. Other county characteristics do not seem to correlate with this within-county distortion measure in a systematic way.

It is noted that the calculation of profitability inevitably relies on some assumptions. Several factors not considered in my calculation might still play a role in the decisions of wind power investors, such as the economics of scale for maintenance, unobservable local regulatory constraints and concerns for grid resilience. If these factors make the current locations of wind farms better for wind development than my profitability measure predicts, then I overestimate the aggregate misallocation based on this measure. Another question is whether or not some unobservable factors also account for the correlation of local environmental friendliness and measured misallocation. One possibility is that environmental friendly communities also care about the "side effects" of wind power development more, such as bird strikes, noises and degradation of scenic views. In that sense, deviation from profit maximization is not necessarily suboptimal among these "green" communities. However, previous empirical research on the hedonic effects of wind turbines reveal that the negative impacts of wind turbine construction on local property prices are modest in the US (Heintzelman and Tuttle (2012)). Given the size of measured profitability loss, and the fact that a large proportion of wind farms actually locates in places with little wind, I believe measurement errors stemming from omission of these unobservable factors do not take away the main message of misallocation measures.

3.5 Green Preferences and Spatial Misallocation

In this section, I attempt to evaluate the efficiency loss from the suboptimal siting choices made by those who invest in wind power out of environmental concerns. My main hypothesis is that for either inner satisfaction or a demonstration of the pro-social behaviour, wind farm developers who invest out of environmental concerns display stronger local bias: instead of surveying more sites to place their wind turbines, they tend to have them in their backyards. This behaviour can be interpreted as a particular way to signal one's "greenness" through producing their own electricity, a phenomenon we term "conspicuous generation". Previous papers have documented this kind of behaviour by comparing solar panel placement patterns across "green" and "brown" communities. (Bollinger and Gillingham (2012)) I will focus more on the potential efficiency loss stemming from this phenomenon. Moreover, I will further explore how the implementation of renewable energy policies interact with these intrinsic motives and shifts the overall allocative efficiency level in particular ways.

In the following sections I document the following empirical findings:

(1) Wind farms in "greener" counties are located in less profitable sites, are less responsive to local fundamentals and perform worse ex-post. The negative relationship between wind farm performance and county level environmental attitude only exists in a sample of wind farms that are invested by local investors, but not those invested by national or international developers.

(2) Wind farms in "greener" counties are more likely to be invested by non-profit organizations and local investors, and located in cities.

(3) Performance-based renewable energy policies improve the within-state allocation of wind farms, partly through attracting more wind investments to "brown" counties.

3.5.1 "Green" wind farm performance

Here I use the merged plant-level data to look for any significant disparities in ex-ante location choices and ex-post performances of wind farms across counties of different environmental attitude. The baseline specification is:

$$y_{it} = \alpha * demrate_c + \beta_{state} + \gamma_t + \epsilon_{it} \tag{3.3}$$

The sample is the plant-level dataset with 774 plants (out of 821 in total) fully matched to the project level dataset. y_{it} are characteristics of wind power plant

i that starts operating in year t, including capacity factor (productive efficiency), predicted profitability based on locational fundamentals, actual profit calculated using capacity factor and retail electricity price, ownership type, indicators of whether or not the investor is local and located in cities. I control for state and operation year fixed effects for the first three variables in linear regressions and only year fixed effects for the latter three in logit regressions. Standard errors are clustered at state level. $demrate_c$ is the votes share for the Democratic Party in 2012 presidential election.

The results are shown in the upper panel of Table 3.5. Column 1-3 indicate that wind farms located in greener counties are placed in worse location ex-ante and perform worse in terms of productivity and revenue ex-post. Column 4 shows that their investors are more likely to be non-profit, such as governments, public organizations (NGOs and universities), municipal and cooperative utilities, which reveals significant differences in the nature of renewable investments across counties with different green preferences. Column 5 shows that wind power projects in greener counties are also more likely to be set up by local investors, defined as investors whose investments are limited within their own state, contrary to international or national developers such as EDF Renewables or GE energy, who spread their projects in various states. Column 6 indicates that the wind farms from "greener" counties are more likely to be located in urban areas, defined by the US Census Bureau. Having wind farms in urban areas is usually considered suboptimal because of more obstruction to incoming winds, higher land price and more restrictions on production due to noises and other potential disturbance of wind turbine operation to human activities. As a result locating wind farms closer to cities is likely to serve other purposes for green investors: they could be signaling their concerns for environmental protection to people who can easily see their wind turbines working; or as non-profit organizations, they are less efficient in monitoring and maintaining wind farms due to the lack of specialized personnel, which forces them to have their properties closer to where they are.

Another plausible interpretation of the worse site choice and ex-post performances for wind farms located in greener counties is that these counties are more welcoming to renewables and set lower entry barriers for wind farm investors. I address this issue in the lower panel of Table 3.5 by splitting the sample of wind farms into a local subsample that contains only wind farms whose investors only invest within-state and a non-local one. It is quite clearly that the worse ex-ante site choice and ex-post profitability of wind farms in greener counties are almost purely driven by those owned by local investors, which is contradictory to what we would expect if the lower entry barrier of wind farms to greener counties is the main story behind my findings.

To check this hypothesis from another perspective, I further leverage the gridcell level data to see if the placement of wind farms are less responsive to local fundamentals in greener counties. A major advantage of this specification is that it takes account of the possibilities that more environmentally friendly counties might start with systematic differences in terms of wind resources or restrictions in wind farm development, and evaluate the efficiency of location choices given these constraints. My main specification takes the form of:

$$Capacity_{it} = \alpha * profitability_i + \beta * demrate_c + \gamma * profitability_i * demrate_c + \theta_s + \delta_t + controls_i + \epsilon_{it} \tag{3.4}$$

$Capacity_{it}$ is the wind capacity added to cell i in year t, $profitability_{it}$ is a measure of predicted wind farm profitability at cell i; $demrate_c$ represents the green preferences of county c, measured as the votes share for Democratic Party in the 2012 presidential election of that county. γ shows how the responsiveness of wind power placement to profitability varies across "green" and "brown" counties. State and year fixed effects, as well as the interactions between profitability and polynomials of year trends are controlled for.

Since my dependent variable is lower-bounded by zero, linear regression might not be the most suitable specification. For robustness, I try different estimation methods that pay extra attention to the zeros in the left-hand side variables. Due to the censorship nature of this problem, I employ panel data Tobit estimation for all the regressions involving gridcell-level data. I follow Honoré (1992) practice to consistently estimate the coefficients in a panel Tobit setting with fixed effects. Since the distribution of the wind capacity added to each gridcell per year is highly dispersed with a large proportion (99.99%) of it clustered at zero, for the sake of computational convenience, I assemble a new sample with information on all the cell with wind farm placement, as well as a 10% sample randomly drawn from the remaining cells, keeping the panel structure.

Table 3.5 shows that profitability indeed plays a lessor role in the location choices of wind farms in greener counties, mostly because they are less prone to be placed in windier places. One standard deviation in the greenness measure makes the placement of wind farms 12% less responsive to the profitability of potential sites. The results hold with an alternative measure of environmental friendliness, such as votes share of the Green Party.

The above results not only show that wind farms located in greener counties perform worse, they also indicate that the deviation of wind farm placement from the optimum *within-county* is larger for more environmentally friendly counties. So it is not just that greener counties set lower entry barriers for green energy investments, but their investors are actually worse in placing given the amount of wind capacities within counties. The fact that green investors in renewable energy make worse location decisions grant us with novel and interesting policy implications: to maximize the impact of subsidies on renewable in generating public benefits, policy makers should focus on bringing more "brown" but profit-maximizing investors into the market instead of encouraging green and utility-maximizing agents to produce more. In the next part, I evaluate the role of three different types of renewable energy policies in correcting or exacerbating this green-preferences-related suboptimal misallocation within-state.

3.5.2 Renewable policies and allocative efficiency

As has been shown in section 4, most of the observed efficiency loss due to suboptimal siting of wind farms can be accounted for by spatial misallocation within-state, or even within-county. In this part, I manage to check if renewable policies affect the within-state allocation of wind farms. It is worth noting that in theory, if all the existing wind power investments are outcomes of profit-maximization and the search cost for better sites is a fixed cost, then only price-based subsidies should be effective in improving within-state allocation since it increases the benefit of conducting a more thorough search. Even if that is the case, we should not expect any differences in the policy impacts between "green" and "brown" counties, under the assumption that the only differences between investors from "green" and "brown" counties lie in their entry standards. The other two types of policies should only change participation constraints and attract less profitable projects. The baseline specification is:

$$Capacity_{it} = \Sigma_p^3 \beta_p * policies_{pst} + \Sigma_p^3 \gamma_p * profitability_i * policies_{pst} + \theta_i + \delta_t + \epsilon_{it} \quad (3.5)$$

$Capacity_{it}$ stands for wind capacity added to gridcell i in year t, $profitability_i$ is the measure of the predicted distant-varying profit for a typical wind farm operating in gridcell i, $policies_{pst}$ is the intensity of policy p implemented in state s in year t, where p indicates which group (per-unit-price-based, direct subsidies, RPS) does the policy index belong to. Cell and year fixed effects are controlled and the standard errors are clustered at the state level. I am also controlling for the interactions between profitability and polynomials of year trend in case there is a year trend governing the response of wind power placement to profitability.

Coefficients on $policies_{pst}$ are the estimates of the treatment effects of renewable energy policies on wind capacity addition in a basic difference in differences setting. The identification assumption is that conditional on the gridcell-level predicted profitability, as well as cell and year, fixed effects, the growth in wind capacity addition should follow a parallel trend across different states in absence of any policies. These assumptions are challenged if there are active business groups pushing for certain policies and they are also investing more heavily in local renewable energy programs, which could well be true in reality. I try some other measures to sharpen my identification in my robustness checks. First, I restrict my sample to only cells around state borders only, where they are much more similar to each other apart from the timing and intensity of state-level renewable policies. However, there is still the concern that apart from the policies I am examining there might be other unobservable policies or change of rules implemented at the same time. So furthermore, I restrict my sample to gridcells in states that have implemented at least one of the policies so their effects are identified through the variation of *when* the policies are implemented and how significant these policies are, instead of which states manage to implement policies. Finally, since the lobbying usually takes time and for most of the policies and there is usually a time gap between the enacting and implementation of policies, if the concern for avid green investors pushing policies is valid, we should be able to see the capacity addition diverges across treated and control states even before the implementation of policies. So as another robustness check, I look at the leads and lags of incentive changes to trace the dynamic impacts of policies before and after their actual implementation. There seem to be no discernible differences in pretrends across treated and control cells. The results on these extra specifications are reported in the appendix and the main results are largely robust.

The coefficients of the interaction terms, $profitability_i * policies_{st}$, measure how the implementation of policies changes the responsiveness of wind farm placement to profitability. Positive coefficients indicate that with renewable energy policies in place, the placement of wind farms follows local fundamentals better. Even if we believe that the identification of the treatment effects of policies on wind capacity addition is plagued with concerns about policy endogeneity and anticipation effects, it is hard to think about an alternative explanation on why should the *responsiveness* of wind farm placement to profitability would change hand in hand with renewable policies.

As shown in Table 3.7, both RPS and price-based subsidies improve the within state allocative efficiency of wind power projects. The magnitude is quite large: one standard deviation increase in the intensity of RPS increases the responsiveness of wind placement decision to profitability by 42% and one SD increase in the intensity of price-based policies improves that by 57%. Direct subsidies that are not performance-based do not seem to change the within-state allocation of wind farms quite significantly after we control for the cross terms of profitability and year fixed effects. Results from Tobit estimation are shown in the last two columns and the signs and significance of coefficients largely hold.

Needless to say, the interpretation of our results on the estimated coefficient of responsiveness γ_p largely depends on the distribution of cells by their measured potential profitability in different states. Suppose the states that implement renewable energy policies have larger dispersion in the higher end of the wind resources distribution, then even if both treated and control states experience the same trend that moves the placement of wind farms up to more profitable gridcells by the same percentiles, our estimates will pick up some improvement in allocative efficiency attributable to policies. Therefore it is crucial to adopt an alternative specification that looks at the role of renewable policies in shifting the placement of wind farms within the distribution of gridcells by potential profitability in each state. This specification will also help us to know if subsidies lift efficiency level through reducing the number of worst located projects or attracting the best ones.To implement the idea, I adopt the expected profitability distribution of the occupied cells for each state before any renewable subsidies are placed as a benchmark, divide all the cells into different groups according to their places in the benchmark and check the differential impacts of policies across different groups. Specifically, I divide the cells within each state into three groups: the ones above the 75th percentile of pre-subsidy occupied cells, the ones below the 25th percentile and the ones in between. A particular type of renewable policy that significantly improves the efficiency level of wind projects may work through either increasing the number of projects in the first group, decreasing the number of projects in the second group, or both. I interact the indicators for these three groups with the intensity measures of renewable energy policies $policies_{pst}$ to examine the impacts effects of different kinds of renewable policies on shifting the profitability distribution of occupied cells.

The results are shown in Table 3.7. As can be seen, price-based performance subsidies are most effective in reducing the probability of bad project placement in cells with expected profitability lower than the 25th percentile of pre-subsidy occupied cells, while quantity-based renewable portfolio standard (RPS) appears to be both reducing the occurrence of bad project placement and adding capacities to the good cells at the same time, Consistent with our intuition, non-performance-based fixed subsidies have similar effects in adding wind capacities in cells across different profitability groups.

3.5.3 Renewable policies and green preferences

The significance and magnitude of the previous results on the impacts of renewable policies on within-state allocative efficiency present a stark contrast to what we should expect if the investors had been following constrained profit maximization in making location decisions before the policies are put in place. Combined with the evidence on the characteristics of wind power plants in environmentally friendly counties, it is reasonable to argue that the improvement of within-state allocation of wind farms could come from the fact that these policies manage to counteract some pre-existing distortions: the local bias of green investors in choosing project sites seems to be a salient and prominent one.

We have reasons to believe that extra financial benefits related to wind power investments might incentivize "green" and "brown" investors differentially. A quick look at the incentive scheme of our three groups of policies suggests that direct subsidies should be equally attractive to all kinds of investors while profit-maximizing investors prefer price-based subsidies since they are getting more with higher production efficiency. Under RPS, all the utilities within the implementing states are required to source a given proportion of its electricity sold from renewable sources. To comply with this requirement, utilities are either investing in their own wind farms or trying to encourage efficient and stable sources of supply from private investors. Given its mandate nature, a utility serving mainly "brown" counties with less pre-existing green investments is required to expand its renewable energy supply much more aggressively than their "green" counterparts. Also, extra capacities invested by utilities as a purpose to meet the mandate are more likely to follow where the wind is in order to maximize the amount of "dirty" electricity replaced.

Therefore, we expect RPS and price-based subsidies to be more effective in adding capacities to "brown" counties with better wind resources as they have been under-targeted by previous wind power investments driven by environmental concerns. To sum up, assuming the existence of green preferences, there are two sources of potential gains in within-state allocative efficiency following the implementation of renewable energy policies. First, performance-based financial incentives and possibly RPS increase the returns to better site choice and encourage project developers to invest more in searching for better sites. Second and more interestingly, there exists a relocation effect: these policies are shifting new wind capacities from "green" counties to "brown" counties, where renewable investments are more profit-oriented and follow fundamentals more strongly. It is likely to be the result of more for-profit investors with less local biases entering the markets in response to renewable energy incentives.

I check it with the following simple regression:

$$Capacity_{it} = \Sigma_p^3 \beta_p * policies_{pst} + \Sigma_p^3 \gamma_p * demrate_c * policies_{pst} + \theta_i + \delta_t + \epsilon_{it} \quad (3.6)$$

This regression aims at checking if certain renewable energy policies that are proven to be effectively improving the within-state allocation of wind farms also manage to shift new capacities from "greener" but less efficiently located places to "browner" and more profit-oriented ones. From Table 3.9, we see that both RPS and price-based subsidies are adding more wind capacities to "brown" counties disproportionally. On the contrary, direct subsidies are adding disproportionally more wind capacities to more environmentally-friendly counties, most likely due to the fact that their non-performance-based nature ensures the same amount of payments to different kinds of projects, and green but less efficient investors are not punished by their worse performances. This could be one of the reasons why RPS and price-based subsidies work better in improving the within-state responsiveness of wind farm placement to profitability while fixed subsidies do not.

To account for the importance of this relocation effect in explaining the policy-induced improved within-state allocation, I adopt a slight variation of specification (6) by replacing the greenness index with a dummy that switches to 1 for counties above the 75th percentile of the continuous greenness index. We find green counties under this metric to be 40% less responsive to profitability and RPS/price-based policies seem to be adding capacities to brown counties only. These results indicate that the pure relocation of new capacities to brown counties by RPS and price-based subsidies is going to increase the responsiveness of wind farm placement to profitability by 10% and is hence able to explain about 25%-30% of the improved responsiveness due to RPS and price-based policies.

Another way to check the differential effects of different policies in screening investors is to check how the ownership types of wind power projects respond differently to these three types of renewable incentives. Obviously, RPS should be more effective in attracting utility-invested wind capacities as it directly applies to utilities. Meanwhile, private and for-profit investments should respond more to price subsidies than fixed subsidies. I test these hypotheses in Table 3.10. It is clear that the results are largely consistent with the mechanism examined in this paper previously, with RPS adding more utility-invested wind capacity and price subsidies more effective in adding private capacity.

3.6 Model

In this section, I will present a very simple model on the private provision of public goods. A distinctive feature of this model is that providers in public goods differ in their environmental attitudes.Those with green preferences display local biases when choosing sites, which decreases their incentives to search. Performance-based subsidies not only increase the return to site searching but also relax the participation constraint of for-profit investors more. Therefore the efficiency gains act through both intensive and extensive margins.

3.6.1 Wind power production

In the model, I assume that the production of renewable energy is solely determined by the locational fundamentals $x_i \in (0,1)$ of location i. To model the location choices of wind farm investors, I assume that an investor based at i can search for better sites by paying a search cost s. By searching, she moves closer to the best spot for wind power production. The profit function for her is thus defined as:

$$\pi_i = x_i + s(1-x_i)^{1/2} - s^2 - F \tag{3.7}$$

where x_i represents the local fundamentals at the investor's original place, s is the search cost and F is other fixed costs in setting up a wind power plant.

For a pure profit-maximizing agent, $s^* = \frac{(1-x_i)^{1/2}}{2}$, indicating that conditional on participation, the wind power investors coming from places with worse fundamentals search more.

3.6.2 Green preferences

We then start with a simple model of utility over a numeraire private good, c and the pleasure derived from supplying public goods. We assume there are two dimensions of heterogeneity for investor i: her local fundamentals for wind power development $x_i \in (0,1)$ and her environmental preference $b_i \in (0,\bar{b})$. The pleasure from supplying public goods is proportional to her environmental preference b_i. In the meantime, investors display local biases to varying extent. My previous empirical evidence reveals that wind projects are more likely to be locally invested in "greener" counties, suggesting that more environmentally friendly investors might display stronger local biases, due to either demonstration effects or the fact that green investors are worse at searching for an ideal site. In the model, I assume that the dis-utility from locating a project further away from the investor's original place is an increasing function of her green preference b_i and the difference between the local fundamentals of her original and final location.

The final utility function is defined as:

$$\begin{aligned} U_i &= c_i - sb_i(1-x_i)^{1/2} + b_i \\ s.t.\ c_i &= \pi_i = x_i + s(1-x_i)^{1/2} - s^2 - F \end{aligned}$$

The electricity price is normalized as 1.

The public benefit generated from a wind project is the amount of greenhouse gas emission reduction, thus should be proportional to the total amount of electricity it produces, $x_i + s(1-x_i)^{1/2}$.

The utility-maximizing search effort can be solved as $s^* = \frac{(1-b_i)(1-x_i)^{1/2}}{2}$, the utility derived from wind investment is therefore $U^* = x_i + \frac{(1-b_i)^2(1-x_i)}{4} + b_i - F$, and the public benefit generated is $e^* = x_i + \frac{(1-b_i)(1-x_i)}{2}$. It is easy to see that the optimal search effort is decreasing in b_i, a direct consequence from green investors reluctance to locate their wind farms away.

Without subsidies, only investors with $U^* > 0$ invest in wind projects. Given b_i, the cutoff in local fundamental x_i is $\bar{x} = \frac{4F-(b_i+1)^2}{2b_i-b_i^2+3}$, where only investors located in places with local fundamentals $x_i > \bar{x}$ choose to invest.

Lemma 1. When $b_i < 1$, $\bar{x}$ is decreasing in b, greener investors are more likely to invest

Proof: $\frac{\partial \bar{x}}{\partial b_i} < 0$ when $b_i < 1$.

3.6.3 Policy choices

In this section, I discuss how the introduction of different types of renewable subsidies affect investors' search efforts and the participation constraints for wind power development.

Here I focus on two types of renewable energy policies. Performance-based subsidy changes the electricity price received by investors to be $p > 1$. With performance-based subsidy, the expected profit for Direct subsidy takes f from the fixed cost F. Therefore, the profit function becomes $\pi_i = p*(x_i+s(1-x_i)^{1/2})-s^2-F$ under performance-based subsidy and $\pi_i = x_i + s(1-x_i)^{1/2} - s^2 - F + f$ under fixed subsidy.

Corollary 1. Performance-based policies increase search efforts for all the investors. The effects do not differ across investors with different environmental attitudes.

Proof $s^* = \frac{(p-b_i)(1-x_i)^{1/2}}{2}$, $\frac{\partial s^*}{\partial p} > 0$, $\frac{\partial s^*}{\partial b_i \partial p} = 0$

Proposition 1. For sufficiently small b_i and reasonable restrictions on value of parameters F and p, in response to performance-based subsidies, the cutoff in x_i drops more for smaller b_i, in other words, performance-based policy is going to add more wind capacity to areas with less environmental oriented investors.

Proof: $\bar{x} = \frac{4F-(b_i+p)^2}{4-(b_i-p)^2}$, $\frac{\partial \bar{x}}{\partial b_i \partial p} > 0$ if $F > 0, 1 < p < 2$ and $0 < b_i < 2.71$

Proposition 2. In response to direct subsidies, the cutoff in x_i drops by the same proportion for investors with different b_i, in other words, direct subsidies add same amount of capacity to areas with different environmental orientation, conditional on local fundamentals,

Proof: $\frac{\partial \bar{x}}{\partial b_i \partial f} = 0$

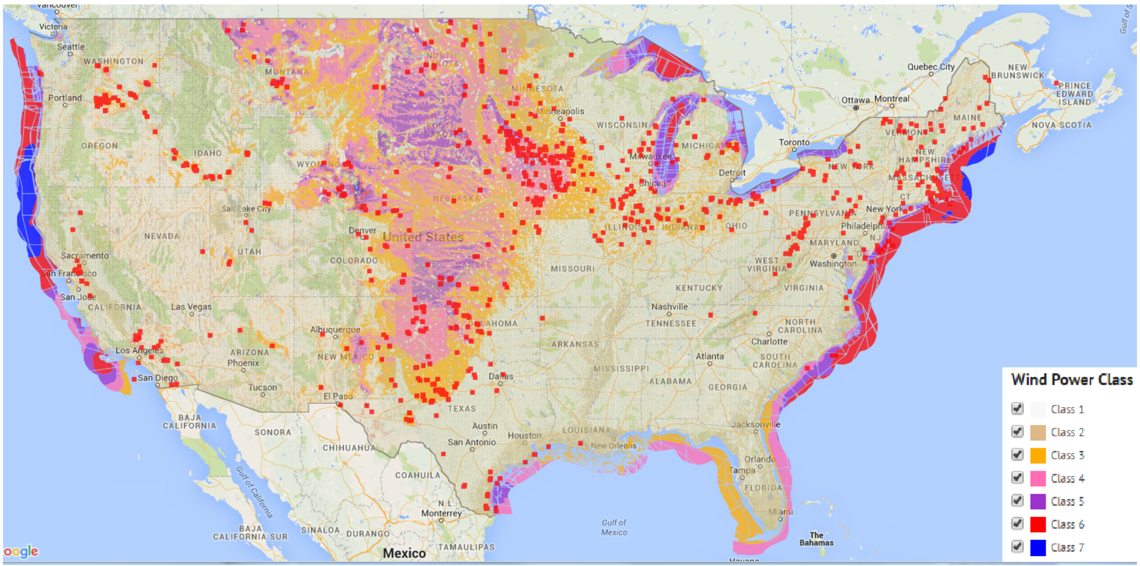
This simple stylized model could accommodate the following empirical findings I have documented in the previous sections. First, "greener" investors are less responsive to fundamentals because they search less. Second, "greener" investors are more likely to invest in renewables. Third, performance-based policies are going to improve the allocative efficiency through inducing more wind capacity added by less environmental-friendly but more profit-oriented investors.

3.7 Concluding Remarks

This paper aims to make two primary contributions. First, I quantify the efficiency loss in the renewable energy sector due to spatial misallocation of wind farms and decompose it into a within-state and cross-state components. These measures are important for us to understand some special characteristics of this industry and to think about the potential impacts of alternative policies on its overall efficiency. Second, I manage to link a significant proportion of the observed within-state distortion to green investors' "conspicuous generation" behavior, namely placing their wind turbines close to where they are instead of locating them in places that make more economic sense. I then come forward to evaluate the role of certain renewable energy policies in partially offsetting the efficiency loss in this way. In short, apart from the heterogeneity in the physical cost of producing GHG free energy, heterogeneity in people's green preferences is also important in determining the public benefits of renewable energy investments. Therefore policy makers should bear in mind the screening effects of policies on investors' non-pecuniary incentives in making a comparison across different types of incentive schemes that are equivalent in other dimensions. In light of this, to encourage people's involvement in supporting renewable energy, extra efforts should be made to create and promote a market for green electricity where people concerned with environmental protection can buy renewable electricity at a premium and possibly awarded in a visible way, instead of encouraging individual households to generate their own clean electricity. Advocates of gird-free distributed energy generation and "home-energy independence" should not only look at the positive side of distributed generation on grid stability but also pay due attention to the potential gains from trade and economics of scale abandoned in this movement.

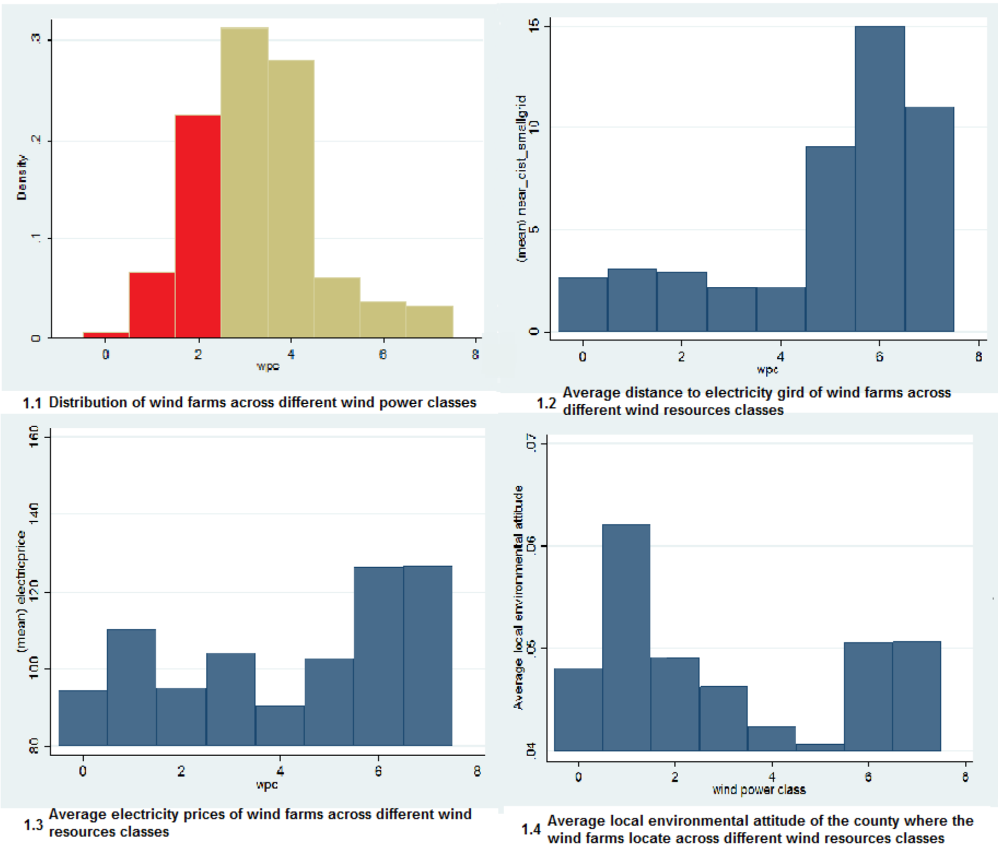
My next step is to quantify the effectiveness of different types of policies in (1) Adding renewable energy capacity; (2) Improving the efficiency of renewable investments; with the existence of large heterogeneity in green preferences across investors in a more structural way. Given the importance and observability of fundamentals in renewable energy sector, it would be interesting to know how much information on the profitability of projects in different locations would policy makers be able to incorporate into their decisions.

Figure 3.1: Wind resources and wind farm distribution



Notes: Each red dot represents a wind farm. WPC (wind power class) is a categorical measure of wind resources on a 1-7 scale, 7 being the strongest. Each wind power class is represented by a color, as shown in the legend. Data visualization courtesy of The Wind Prospector - NREL.

Figure 3.2: Distribution of wind farms across different wind power classes



1.1 Distribution of wind farms across different wind power classes

1.2 Average distance to electricity gird of wind farms across different wind resources classes

1.3 Average electricity prices of wind farms across different wind resources classes

1.4 Average local environmental attitude of the county where the wind farms locate across different wind resources classes

Notes: WPC (wind power class) is a categorical measure of wind resources on a 1-7 scale, 7 being the strongest. Figure 1.1 plots the density of wind farms across WPC. Figure 1.2 and 1.3 shows the average distance to the electricity grid and the average local retail electricity prices of wind farms across different WPC, respectively. Figure 1.4 plots the average local environmental attitude of the county where the wind farms locate across different WPC classes.

Table 3.1: Correlation across different measure of wind farm profitability

NREL Western Wind Dataset Sample

	Measure 1	Measure 2	Measure 3	Measure 4
# of gridcells	4661	4661	75147	75147
Correlation	Measure 1	Measure 2	Measure 3	Measure 4
Measure 1	1.0000			
Measure 2	0.8519	1.0000		
Measure 3	0.6558	0.3671	1.0000	
Measure 4	0.3814	0.4856	0.6662	1.0000

NREL Eastern Wind Dataset Sample

	Measure 1	Measure 2	Measure 3	Measure 4
# of gridcells	2003	2003	75147	75147
Correlation	Measure 1	Measure 2	Measure 3	Measure 4
Measure 1	1			
Measure 2	0.601	1		
Measure 3	0.6091	0.1176	1	
Measure 4	0.2598	0.1594	0.7979	1

Notes: I report the correlation of four different wind power profitability measures. Measure 1 is the baseline measure that combines predicted production based on wind speed data with retail electricity price. Measure 2 is generated with predicted production based on wind speed data with Bloomberg wholesale price data. Measure 3 and 4 takes account of the variation in wind power production. Measure 3 is generated using Eastern/Western Wind datasets wind power simulated production data and Bloomberg wholesale price data. Measure 4 uses Eastern/Western Wind datasets wind power simulated production data and average retail electricity data, under the assumption that offpeak electricity price is 0.63 of peak electricity price. As the methodology in simulating wind power production is different for the Eastern and Western wind datasets, I split the sample into two (Eastern and Western US) and report the correlation separately for them.

Table 3.2: Alternative measures of aggregate spatial misallocation

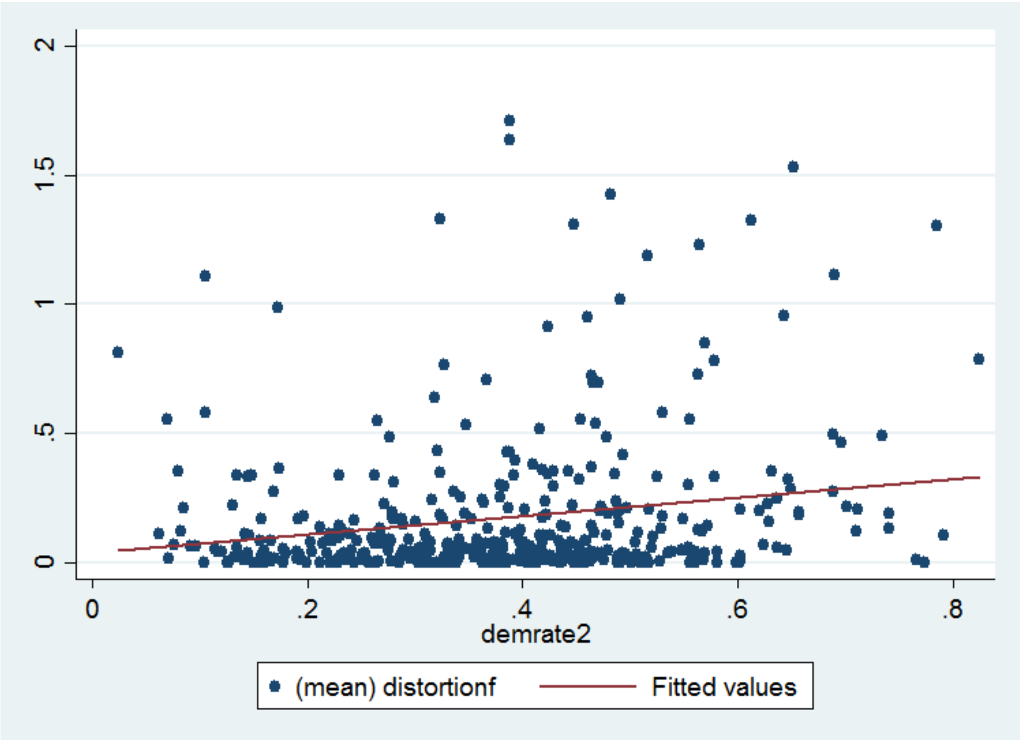
Specification	Sample	Measure
NREL wind power class data & retail electricity price (baseline)	Full sample	0.4719
NREL wind power class data & wholesale electricity price	Full sample	0.4366
NREL Eastern/Western datasets & retail electricity price	Eastern sample	0.1846
NREL Eastern/Western datasets & retail electricity price	Western sample	0.3761
NREL Eastern/Western datasets & wholesale electricity price	Eastern sample	0.1177
NREL Eastern/Western datasets & wholesale electricity price	Western sample	0.2376

Table 3.3: Within-state locational misallocation

State	Measure 1	Measure 2	Measure 3	Measure 4
IA	0.140439	0.199903	0.049872	0.072414
IN	0.143414	0.14436	0.236253	0.2098
IL	0.148962	0.291946	0.074424	0.168646
WV	0.209326	0.173536	0.059405	0.003158
ND	0.221021	0.189135	0.073325	0.071608
WA	0.23381	0.135259	0.121944	0.129677
CO	0.256589	0.19782	0.252514	0.256425
NH	0.281912	0.308476	0.098451	0.128257
ID	0.305632	0.31702	0.158998	0.150273
VT	0.316129	0.332688	0.089663	0.113455
OR	0.321848	0.289195	0.316987	0.337921
KS	0.363492	0.145481	0.164593	0.17919
OH	0.385767	0.41451	0.410419	0.72385
NE	0.404587	0.177351	0.143211	0.237728
OK	0.416231	0.155157	0.167612	0.056249
NC	0.433151	0.799545	0.175103	0.15487
CA	0.460538	0.701351	0.289593	0.280029
MN	0.485057	0.23437	0.339635	0.206461
MD	0.565588	0.520721	0.213071	0.259811
MO	0.577686	0.405438	0.192134	0.183553
MT	0.633305	0.259725	0.377029	0.375105
SD	0.407496	0.219858	0.297603	0.414985
TX	0.373097	0.345533	0.149471	0.152742
WY	0.392716	0.367695	0.339811	0.338902
NM	0.822967	0.281664	0.425106	0.274453
PA	0.892974	0.959326	0.281608	0.182485
MI	0.908642	0.639816	0.210498	0.145349
WI	0.95731	1.229424	0.210885	0.505282
ME	1.194982	1.283586	0.30433	0.652842
NY	1.218893	1.353304	0.307033	0.18198
US	.37402001	.3447733	.1855745	.1846892

Notes: Distortion is the measure of within-state distortion in wind farm placement calculated from (2). Four different measures of distortion are reported. The first is the baseline measure that combines predicted production based on wind speed data with retail electricity price. The second one is generated with predicted production based on wind speed data with Bloomberg wholesale price data. The third one and fourth take account of the variation in wind power production. The third one is generated using Eastern/Western Wind datasets wind power simulated production data and Bloomberg wholesale price data. The fourth one uses Eastern/Western Wind datasets wind power simulated production data and average retail electricity data, under the assumption that offpeak electricity price is 0.63 of peak electricity price. For the whole US, the distortion measure is a weighted average of within-state distortion by total capacity.

Figure 3.3: Within-county distortion and county level green preferences



Notes: The construction of within-county measure of locational distortion in wind farm placement is described in section 4. Demrate is the votes share for Democratic Party in the 2012 presidential election of that county. The slope of fitted line is 0.35 (standard error 0.088).

Table 3.4: Within-county distortion and county characteristics

VARIABLES	distortion	distortion	distortion	distortion
demrate	1.185*	1.024**		
	(0.694)	(0.518)		
greenrate			59.70	90.29**
			(57.29)	(45.46)
SD(profitability)	0.00590	-0.00501	0.0144	0.000617
	(0.0262)	(0.0348)	(0.0284)	(0.0371)
Mean(profitability)	0.0115	0.0241**	0.0117	0.0264**
	(0.0118)	(0.0106)	(0.0123)	(0.0115)
No. of wind farms in county	-0.0239	-0.0118	-0.0262	-0.0150
	(0.0245)	(0.0257)	(0.0245)	(0.0257)
% of non-suitable cells	-0.779***	-0.756***	-0.798***	-0.697***
	(0.279)	(0.222)	(0.298)	(0.249)
Median household income	9.11e-05	0.000104**	7.35e-05	8.97e-05
	(6.97e-05)	(5.24e-05)	(7.34e-05)	(5.85e-05)
Building permits	-0.000162	-0.000451*	-0.000158	-0.000490*
	(0.000122)	(0.000271)	(0.000124)	(0.000269)
Retail sales pc	-1.01e-05	1.31e-05	-6.78e-06	1.60e-05
	(1.08e-05)	(1.70e-05)	(1.09e-05)	(1.75e-05)
% of college graduates	0.0393**	0.0120	0.0364*	0.0110
	(0.0196)	(0.0192)	(0.0196)	(0.0191)
% of high school graduates	-0.0234	-0.00878	-0.0101	-0.00358
	(0.0265)	(0.0211)	(0.0236)	(0.0201)
% female	0.00829	-0.0298	0.00872	-0.0309
	(0.0295)	(0.0346)	(0.0295)	(0.0349)
% while alone	0.0259	0.0112	0.0277	0.0106
	(0.0295)	(0.0312)	(0.0291)	(0.0307)
% African alone	-0.0117	-0.0368	-0.00742	-0.0406
	(0.0406)	(0.0464)	(0.0404)	(0.0469)
% Asian alone	0.0168	0.00311	0.0235	0.00584
	(0.0309)	(0.0331)	(0.0302)	(0.0328)
Mean travel time to work	0.00838	0.0109	0.00774	0.0103
	(0.0185)	(0.0205)	(0.0187)	(0.0206)
Housing units	1.02e-06	4.45e-06	1.27e-06	4.31e-06
	(2.12e-06)	(4.64e-06)	(2.17e-06)	(4.70e-06)
Homeownership rate	0.0167	0.00144	0.0142	-0.00203
	(0.0163)	(0.0110)	(0.0162)	(0.0114)
Median housing value	5.63e-07	-9.52e-07	6.99e-07	-9.24e-07
	(1.26e-06)	(2.93e-06)	(1.26e-06)	(2.97e-06)
No. of firms	-2.04e-06	-7.70e-06	-2.92e-06	-5.99e-06
	(7.75e-06)	(1.75e-05)	(7.93e-06)	(1.75e-05)
Observations	398	262	398	262
R-squared	0.075	0.082	0.068	0.080

Notes: Distortion is the normalized measure of deviation from the optimal level at the county level, defined as the ratio between the percentage gain in average profitability should current projects be placed at the best positions and the percentage gain from a random allocation to the optimal allocation. Robust clustered standard error at the state level. I exclude the counties that have only one gridcell occupied from the sample and report the regression results in column 2 and 4.

Table 3.5: Characteristics and performances of wind farms in Counties with different level of greenness

Method	Linear regression			Logit		
VARIABLES	CF	wpc	Revenue	1(nonprofit)	1(local)	1(urban)
Demrate	-0.049**	-1.516**	-1.290*	1.483*	1.385*	1.974**
	(0.023)	(0.667)	(0.732)	(0.806)	(0.720)	(0.957)
Observations	756	760	756	774	774	774
R-squared	0.433	0.509	0.368	0.0123	0.00981	0.0159
STATE FE	YES	YES	YES	NO	NO	NO
YEAR FE	YES	YES	YES	YES	YES	YES

Sample	Non-local			Local		
Method	Linear regression					
VARIABLES	CF	wpc	nonprofit	CF	wpc	nonprofit
Demrate	-0.0071	-0.918	-0.008	-0.144*	-2.771**	0.246*
	(0.034)	(0.744)	(0.005)	(0.071)	(1.0681)	(0.122)
Observations	414	417	418	354	355	355
R-squared	0.473	0.561	0.027	0.445	0.61	0.207
State FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Notes: In the upper panel, the sample is a matched wind power plants data. Column 1-3 show results for linear regressions with state and operating year fixed effects. CF is the capacity factor of the power plant (total electricity produced/maximum electricity production at full capacity). WPC stands for the wind resource category measure of where the plant is. Revenue is the product of the capacity factor and wholesale electricity price (deducting transmission loss). Column 4-6 are logit regressions where the dependent variable is a dummy on whether or not the power plant is invested by non-profit investors, by local investors and located in urban areas. In the lower panel, I split the full sample into a local and a non-local subsamples. The local subsample includes only wind farms whose investors only invest within the state. The non-local one contains the wind farms whose investors have wind power projects in more than one state. Robust clustered standard error at the state level. * significant at the 0.1 level; ** significant at the 0.05 level; *** significant at the 0.01 level.

Table 3.6: Responsiveness to fundamentals of wind farms with different level of greenness: linear regressions

Variables	capacity	capacity	capacity	capacity	capacity	capacity
wind speed	0.000240***	0.000245***	4.62E-05	7.66E-05		
	(8.60E-05)	(8.60E-05)	(7.30E-05)	(6.92E-05)		
distgrid	-2.27e-06*	-2.46e-06*	-2.02e-06**	-2.19e-06**		
	(1.22E-06)	(1.30E-06)	(8.75E-07)	(9.69E-07)		
urban			0.00015	-3.37E-05		
			(0.00012)	(0.00013)		
profitability					0.000043***	0.000041*
					(8.14e-06)	(6.80e-06
wind speed*greenrate	-0.0135***		-0.0109***			
	(0.00386)		(0.00343)			
distgrid*greenrate	0.000737***		0.000327			
	(0.00025)		(0.00024)			
urban*greenrate			0.0209*			
			(0.0128)			
profitability*greenrate					-0.00183**	
					(0.000798)	
wind speed*demrate		-0.000143*		-0.000103**		
		(8.20E-05)		(4.44E-05)		
distgrid*demrate		1.58E-05		1.24E-06		
		(1.80E-05)		(9.84E-06)		
urban*demrate				0.000314*		
				(0.00022)		
profitability*demrate						-0.000016
						(9.30e-06
Observations	1,421,225	1,421,225	2,464,110	2,464,110	1,421,225	1,421,22
R-squared	0.002	0.002	0.001	0.001	0.002	0.002
State FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Notes: Sample is gridcell level panel data. The dependent variable is the amount of wind capacity installed per km^2 to a gridcell in a year. Wind speed is the average wind speed of the gridcell calculated according to NREL wind resrouces categorization. distgrid is the distance from the gridcell to the closest main electricity grid (in km). Profitability is the distance-varying profit measure of the gridcell. Urban is a dummy on whether or not the gridcell is inside urban areas. The first two and last two columns report results on a sample leaving out the cells that are considered to be not suitable for wind power development, while results on the middle two columns are estimated on the full sample. State and year fixed effects, as well as state-specific year trends are controlled. Robust clustered standard error at the state level. * significant at the 0.1 level; ** significant at the 0.05 level; *** significant at the 0.01 level.

Table 3.7: Renewable policies and within-state allocation: profitability measure

VARIABLES	Linear capacity	Linear 1(plant)	Tobit capacity	Logit 1(plant)
profitability	2.09e-05***	4.95e-05***	0.0702***	0.0343***
	(5.32E-06)	(7.54E-06)	(0.0160)	(0.0133)
RPS	-0.0306**	-0.0778***	-10.07	-3.238
	(0.012)	(0.0267)	(8.909)	(3.065)
fixsubsidy	-0.00235	-0.00678	2.215	0.797
	(0.00664)	(0.0237)	(2.473)	(1.075)
pricesub	-0.0251	-0.0553	-38.85***	-14.44***
	(0.0153)	(0.0354)	(13.88)	(4.839)
profitability*rps	0.00294***	0.00732***	0.649***	0.202***
	(0.0008)	(0.00186)	(0.221)	(0.0749)
profitability*fixsub	0.000735*	0.00304*	-0.0954	-0.0198
	(0.00038)	(0.00129)	(0.132)	(0.0564)
profitability*pricesub	0.00463**	0.00908*	2.419***	1.045***
	(0.00183)	(0.00472)	(0.695)	(0.283)
Observations	2,464,143	2,464,143	254,331	254,331
State FE	NO	NO	YES	YES
Gridcell FE	YES	YES	NO	NO
Year FE	YES	YES	YES	YES
R-squared	0.002	0.004	0.012	0.015

Notes:Sample is gridcell level panel data. Dependent variables in column (1) and (3) are the amount of wind capacity installed per km^2 to a gridcell in a year. Dependent variables in column (2) and (4) are dummies on whether or not a wind power plant is built at a gridcell in a year. RPS is the real stringency of Renewable Portfolio Standard for implementing states, defined in section 3.3. Pricesub is the amount of subsidy given to per unit electricity generation. Fixsub is the proportion of total upfront cost reduced by subsidies. Robust clustered standard error at the state level. * significant at the 0.1 level; ** significant at the 0.05 level; *** significant at the 0.01 level.

Table 3.8: Renewable policies and within-state allocation: Change in wind farm profitability distribution

VARIABLES	Linear capacity	Linear 1(plant)	Tobit capacity	Logit 1(plant)
1(above 75th pct)*pricesub	0.0147	0.0078	6.891	11.9
	(0.00201)	(0.00428)	(4.84)	(7.98)
1(75th-25th pct)*pricesub	0.0249**	0.0493***	1.419	-2.092
	(0.0117)	(0.0123)	(3.193)	(7.993)
1(below 25th pct)*pricesub	-0.00303***	-0.00740**	-3.916	-14.62
	(0.00087)	(0.00323)	(4.476)	(13.7)
1(above 75th pct)*fixsub	0.00936	0.0368	1.257	2.665
	(0.0178)	(0.0511)	(0.917)	(1.962)
1(75th-25th pct)*fixsub	0.0151	0.0127	0.293	0.897
	(0.0126)	(0.0527)	(0.304)	(0.752)
1(below 25th pct)*fixsub	0.00152	0.0186	1.426	3.279
	(0.00327)	(0.0137)	(0.962)	(2.011)
1(above 75th pct)*RPS	0.0347***	0.110***	3.555*	9.559*
	(0.0121)	(0.031)	(1.902)	(5.196)
1(75th-25th pct)*RPS	0.0153	0.032	0.562	1.453
	(0.0126)	(0.0271)	(2.387)	(6.05)
1(below 25th pct)*RPS	-0.00739***	-0.0131***	-3.19	-5.714
	(0.00174)	(0.00429)	(2.939)	(7.976)
Observations	1,937,364	1,937,364	224,532	210,924
State FE	NO	NO	YES	YES
Gridcell FE	YES	YES	NO	NO
Year FE	YES	YES	YES	YES
R-squared	0.002	0.004	0.012	0.015

Notes:Sample is gridcell level panel data. Dependent variables in column (1) and (3) are the amount of wind capacity installed per km^2 to a gridcell in a year. Dependent variables in column (2) and (4) are dummies on whether or not a wind power plant is built at a gridcell in a year. 1(above 75th pct) is a dummy that switches to one if expected profitability of the cell is higher than the 75th percentile of existing wind projects within the state before any renewable subsidies are applied. 1((75th-25th pct) is the indicator of whether or not the profitability of the cell falls into the 75th and 25th percentile of existing wind projects within the state before any renewable subsidies are applied, while 1(below 25th pct) indicates whether or not the profitability of cell is lower than the 25th percentile of existing pre-subsidy wind projects. RPS is the real stringency of Renewable Portfolio Standard for implementing states, defined in section 3.3. Pricesub is the amount of subsidy given to per unit electricity generation. Fixsub is the proportion of total upfront cost reduced by subsidies. Robust clustered standard error at the state level. * significant at the 0.1 level; ** significant at the 0.05 level; *** significant at the 0.01 level.

Table 3.9: Differential Impacts of Policies on wind capacity across "green" and "brown" counties

VARIABLES	Linear capacity	Linear capacity	Tobit capacity	Logit capacity
RPS	0.014	0.00202*	2.281	1.573
	(0.0168)	(0.0167)	(3.828)	(3.981)
pricesub	0.0241*	0.0153**	0.103	-3.721
	(0.0131)	(0.0073)	(7.409)	(6.054)
fixsubsidy	-0.0175**	-0.0009	-1.604**	-1.543
	(0.00628)	(0.00406)	(0.784)	(0.839)
greenrate*RPS	-0.388		-45.994	
	(0.284)		(66.886)	
greenrate*pricesub	-1.885**		-10.516	
	(0.329)		(153.610)	
greenrate*fixsub	1.732		47.015***	
	(1.118)		(14.999)	
demrate*RPS		-0.0248		-2.861
		(0.0294)		(7.304)
demrate*pricesub		-0.0397**		-7.302
		(0.0198)		(11.025)
demrate*fixsub		0.0542***		4.556***
		(0.0165)		(1.266)
Observations	2,464,110	2,464,110	284658	284658
R-squared	0.002	0.002	0.14	0.14
State FE	NO	NO	YES	YES
Gridcell FE	YES	YES	NO	NO
Year FE	YES	YES	YES	YES

Notes:Sample is gridcell level panel data. The dependent variable is the amount of wind capacity installed per km^2 to a gridcell in a year. The first two columns report results from linear regression in the full grid-cell sample and the last two columns report results from Tobit estimation in a sample with all the built-up cells and 10% of the other cells. Demrate and greenrate are the democratic and Green Party votes share in 2012 presidential election of the county. RPS is the increment in RPS requirement for implementing states. Pricesub is the amount of subsidy given to per unit electricity generation. Fixsub is the proportion of total upfront cost reduced by subsidies. Robust clustered standard error at the state level. * significant at the 0.1 level; ** significant at the 0.05 level; *** significant at the 0.01 level.

Table 3.10: Differential impacts of policies on wind different types of investments

VARIABLES	Utility capacity(MW)	Nonprofit capacity(MW)	Private capacity(MW)	1(Utility)	1(Nonprofit)	1(Private)
RPS	0.00295*	-0.000059	0.010546	0.00184	0.000134	0.018026
	(0.00178)	(0.000078)	(0.007972)	(0.00126)	(0.00034)	(0.01532)
pricesub	0.00254	0.000081	0.008287	0.0016	0.000303	0.023791***
	(0.0029)	(0.000145)	(0.006103)	(0.00179)	(0.000434)	(0.011433)
fixsub	0.00026	0.000089	0.004007***	0.0004	0.000123	0.012869***
	(0.00065)	(0.000021)	(0.001464)	(0.00067)	(0.000159)	(0.004166)
Obs	142,3073	142,3073	142,3073	142,3073	142,3073	142,3073
Gridcell FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
R-squared	0.0002	0.0001	0.0013	0.0002	0.0001	0.0026

Notes: Sample is gridcell level panel data. Gridcells that are considered not suitable for wind power development are dropped The dependent variables of the first three columns are the amount of wind capacity installed per km^2 by utilities, nonprofit investors and private profit-oriented investors to a gridcell in a year. The dependent variables of the last three columns are whether or not a gridcell has wind power capacity installed by utilities, nonprofit investors, and private profit-oriented investors. RPS is the increment in RPS requirement for implementing states. Pricesub is the amount of subsidy given to per unit electricity generation. Fixsub is the proportion of total upfront cost reduced by subsidies. Robust clustered standard error at the state level. * significant at the 0.1 level; ** significant at the 0.05 level; *** significant at the 0.01 level.

Appendices

Appendix C

Appendix of "Where does the Wind Blow? Green Preferences and Spatial Misallocation in the Renewable Energy Sector"

C.1 Robustness

C.1.1 Dynamic impacts of renewable subsidies

As mentioned in section 5.2, as an extra robustness check, I manage to trace the dynamic impacts of renewable policies before and after their actual implementation, based on the idea that if these policies are seriously endogenous, their “treatment effects" might show up even before the actual implementation of them. In practice, I adopt the following specification:

$$Capacity_{it} = \alpha_i + \beta_t + \sum_{m=1}^{3} \gamma_{mp} * \Delta policies_{p,s,t-m} + \sum_{n=0}^{2} \gamma_{np} \Delta policies_{p,s,t-m} + Controls_{it} + \epsilon_{it} \tag{C.1}$$

where $\Delta policies_{p,s,t}$ is the increment in the intensity of policy p implemented in state s in year t, while $\Delta policies_{p,s,t-m}$ and $\Delta policies_{p,s,t+n}$ are the m-th lead and n-th lag of the variable. The estimated coefficients are reported in Figure A1. I interact $\Delta policies_{p,s,t-m}$ and $\Delta policies_{p,s,t+n}$ with cell-level profitability to check if the changes in responsiveness to profitability also go hand in hand with the actual implementation of policies.The exact specification is:

$$Capacity_{it} = \alpha_i + \beta_t + \sum_{m=1}^{3} \gamma_{mp} * \Delta policies_{p,s,t-m} + \sum_{n=0}^{2} \gamma_{np} \Delta policies_{p,s,t-m} + \sum_{m=1}^{3} \zeta_{mp} * \Delta policies_{p,s,t-m} * profitability_i + \sum_{n=0}^{2} \zeta_{np} \Delta policies_{p,s,t-m} * profitability_i + Controls_{it} + \epsilon_{it} \tag{C.2}$$

Similarly, I interact them with support for the Democratic Party at county level following specification (6) to see if different kinds of policies add capacities to counties with different environmental attitude. The coefficients on the interaction terms, as well as their 95% confidence intervals, are shown in Figure A2 and A3.

C.1.2 Responses of other projects attributes to renewable policies

Another related question is whether of not the observed response of location choices of wind farms to changes in renewable policies is just a proxy of other responses. The investment of a wind farm involves a series of joint decisions, including the choices of project size, turbine type, and location. These choices depend on each other in different ways. For instance, a fixed non-performance-based subsidy might help the project with upfront costs, inducing the investor to pursue larger projects and more advanced turbine types. In the meantime, large projects have a higher land requirement, resulting in different location choices that might be more or less efficient depending on the context. Although these explanations will not invalidate my main story directly, as they are also examples of the selection effects of financial incentives. It would be interesting to check if other attributes of the wind projects other than location also respond to renewable energy subsidies, and if so, to which direction.

In this section, I look at two other project attributes: the size of the project, measured in total capacity installed, and the characteristics of the turbines, measured by turbine height and blade length. It is generally believed that higher turbine and longer blade makes use of wind resources more efficiently. [1] I check how they correlate with the local fundamentals and respond to renewable subsidies.

Table A2 shows the results. The upper panel reports the regression results on

[1] http://www.siemens.com/innovation/en/home/pictures-of-the-future/energy-and-efficiency/sustainable-power-generation-windpower-hexcrete-tower.html; http://cleantechnica.com/2015/03/23/us-energy-dept-prowl-bigger-longer-wind-turbine-blades/

the relationship between various project characteristics and local wind resources, and the lower panel reports results on how these characteristics respond to state renewable energy policies and differ across counties with different environmental attitudes. The analysis on project size is carried out with plant level data and that on turbine height and blade size uses turbine level data. It is clear from the upper panel that there's no strong correlation between all these three project attributes to local wind conditions, suggesting that location decision is probably made relatively independent from project size and turbine type choices, or at least the latter decisions does not seem to push the relevant project to a place with definite better or worse wind conditions. It is also not the case that less than desirable location choices are compensated by more powerful wind turbines.

Results reported from the lower panel of Table A2 suggest that the introduction of price-based subsidies and RPS do not lead to significant changes in project size and the quality of wind turbines. However, larger non-performance-based fixed subsidies do seem to encourage larger projects. A possible explanation is that larger fixed subsidies paid out upfront help the wind power investors overcome financial constraints that prevent them from building larger wind farms. The right three columns show the relationship between wind farm characteristics and local environmental attitudes and there is no significant correlation between green preferences and the wind arm attributes that we are interested in.

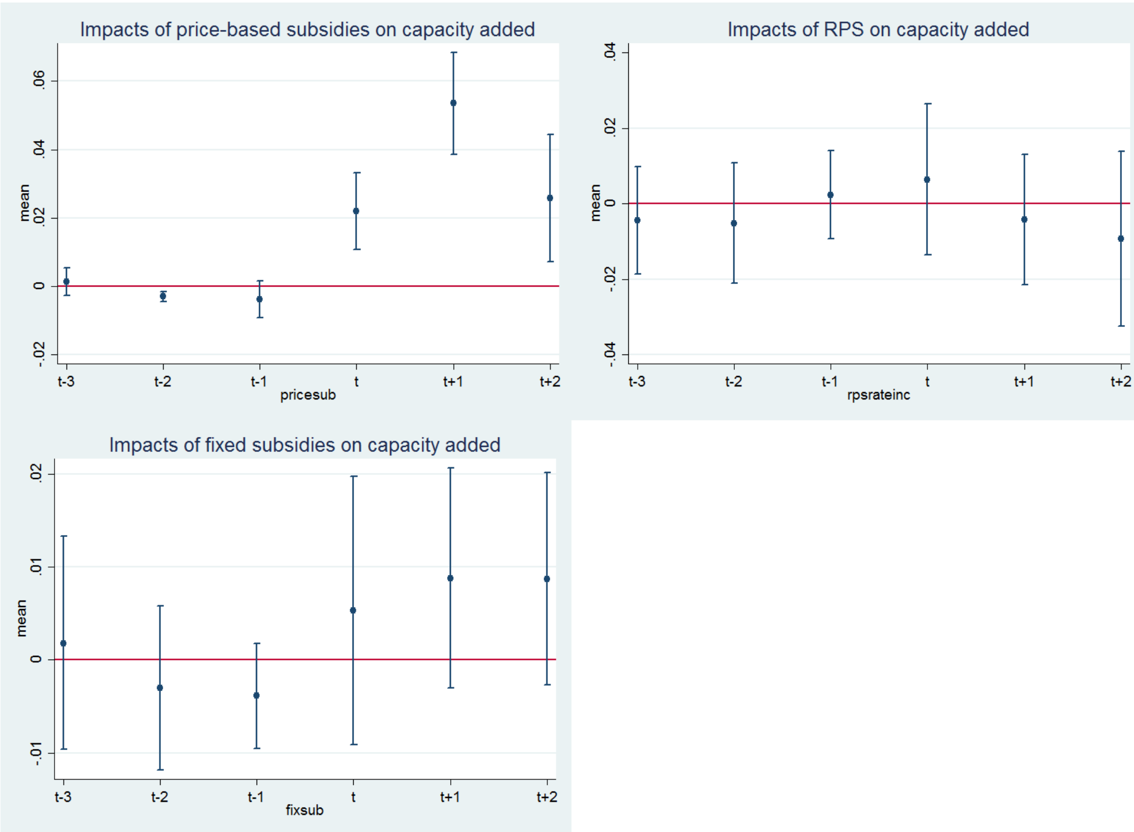
Therefore one conclusion we can draw from the previous analysis is that the robust relationship between renewable energy policies and improved efficiency of wind farms documented in the paper is most likely capturing the direct responses of wind farm site choices to financial incentives instead of proxies of other responses regarding other aspects of the wind farm projects.

Table A1: Robustness: Cells on state borders

VARIABLES	capacity	capacity	capacity
RPS	-0.00675	0.00836	0.00861
	(0.00638)	(0.0146)	(0.013)
pricesub	-0.0132	0.107***	0.126***
	(0.0149)	(0.0297)	(0.0282)
fixsubsidy	-0.0244***	-0.0184**	-0.0076
	(0.00891)	(0.00769)	(0.00932)
profitability*RPS	0.000803		
	(0.0005)		
profitability*pricesub	0.00450***		
	(0.00136)		
profitability*fixsubsidy	0.00074		
	(0.00067)		
demrate*RPS		-0.00755	
		(0.0195)	
demrate*pricesub		-0.162***	
		(0.0508)	
demrate*fixsubsidy		0.0452**	
		(0.0182)	
greenrate*RPS			-0.0806
			(0.152)
greenrate*pricesub			-2.130***
			(0.491)
greenfix			0.209
			(0.141)
Observations	748,869	748,869	748,869
R-squared	0.002	0.002	0.002
Number of cells	22,693	22,693	22,693

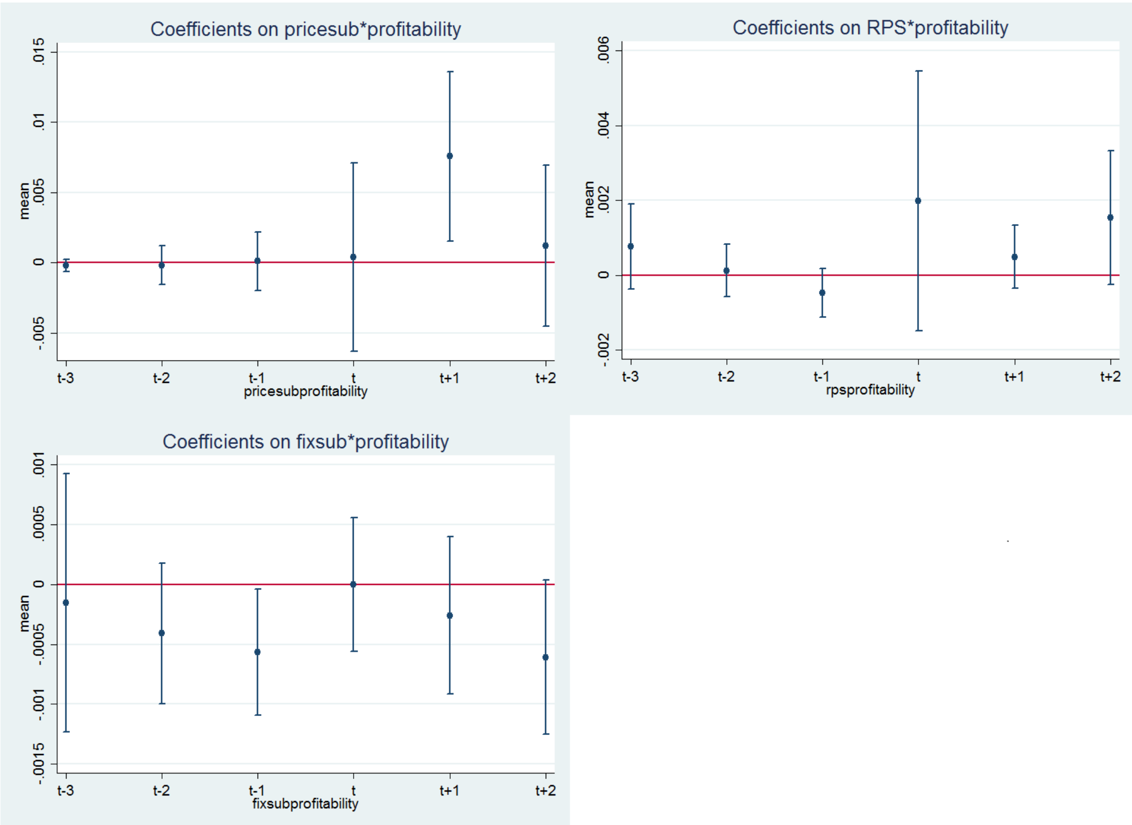
Notes:Sample is gridcell level panel data, limited to gridcells within 25 kilometers distance from state borders. The dependent variable is the amount of wind capacity installed per km^2 to a gridcell in a year. Robust clustered standard error at the state level. RPS is the real stringency of Renewable Portfolio Standard for implementing states, defined in section 3.3. Pricesub is the amount of subsidy given to per unit electricity generation. Fixsub is the proportion of total upfront cost reduced by subsidies. Profitability is the predicted profitability of a typical wind farm at the gridcell. Demrate and greenrate are county level votes shares for the democratic and Green Party at 2012 presidential election. Cell FE and year FE are all controlled. * significant at the 0.1 level; ** significant at the 0.05 level; *** significant at the 0.01 level.

Figure A1: Coefficients on the leads and lags of renewable policy intensity



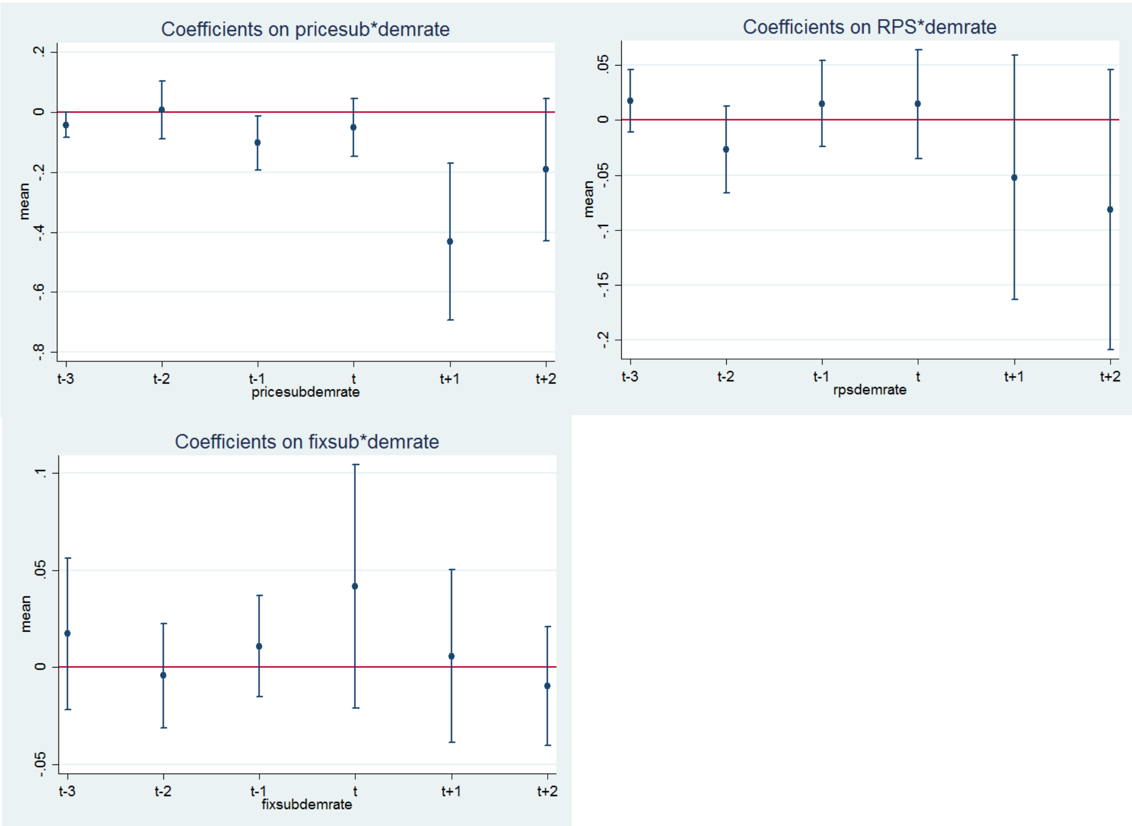
Notes: This graph plots the coefficients and 95% CI on the leads and lags of renewable energy policy intensity, as specified in equation (10).

Figure A2: Coefficients on the interactiosn of cell profitability and policy intensity leads/lags



Notes: This graph shows the coefficients and 95% CI on the interactions of the leads and lags of renewable energy policy intensity and cell level profitability measure, in a specification includes both leads/lags, profitability and their interaction terms.

Figure A3: Coefficients on the interactions of Democratic Party support and policy intensity leads/lags



Notes: This graph shows the coefficients and 95% CI on the interactions of the leads and lags of renewable energy policy intensity and county level support for Democratic Party, in a specification includes both leads/lags, democratic support and their interaction terms.

Table A2: Responses of other project attributes

VARIABLES	projectsize	bladelength	towerheight	projectsize	bladelength	towerheight
WPC	-1.738	-0.774	-1.648	0.646	0.391	0.743
	(7.177)	(0.521)	(0.99)	(6.23)	(0.301)	(0.595)
Observations	817	39,718	39,574	817	39,718	39,574
R-squared	0.291	0.949	0.913	0.311	0.941	0.907
State FE	NO	NO	NO	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

VARIABLES	projectsize	bladelength	towerheight	projectsize	bladelength	towerheight
Pricesub	-18.4	1.07	-3.644			
	(100.1)	(10.46)	(39.05)			
Fixsub	885.7**	54.85	-41.36			
	(388.5)	(43.04)	(59.92)			
RPS	-232.5	3.161	7.58			
	(273.8)	(9.211)	(35.99)			
Demrate				35.23	0.886	-2.18
				(37.99)	(1.957)	(2.398)
Observations	817	39,718	39,574	767	39,500	39,356
R-squared	0.319	0.919	0.95	0.311	0.92	0.952
State FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Notes:The dependent variables as listed are the size of wind project measured by total MW installed, turbine blade length and turbine tower height. The sample with wind project size as the dependent variable is matched plant-project level data and the sample with turbine blade length and tower height as dependent variables is turbine level data. WPC (wind power class) is a categorical measure of wind resources on a 1-7 scale, 7 being the strongest. RPS is the real stringency of Renewable Portfolio Standard for implementing states, defined in section 3.3. Pricesub is the amount of subsidy given to per unit electricity generation. Fixsub is the proportion of total upfront cost reduced by subsidies. Demrate is the county level votes shares for a Democratic Party at 2012 presidential election. Robust clustered standard error at the state level. * significant at the 0.1 level; ** significant at the 0.05 level; *** significant at the 0.01 level.

Bibliography

Ahlfeldt, G. M. and A. Feddersen (2010). From periphery to core: economic adjustments to high speed rail. *Documents de treball IEB* (38), 1.

Ahlfeldt, G. M., S. J. Redding, D. M. Sturm, and N. Wolf (2015). The economics of density: Evidence from the berlin wall. *Econometrica 83*(6), 2127–2189.

Albalate, D. and G. Bel (2012). High-speed rail: Lessons for policy makers from experiences abroad. *Public Administration Review 72*(3), 336–349.

Albouy, D. (2008). Are big cities bad places to live? estimating quality of life across metropolitan areas. Technical report, National Bureau of Economic Research.

Alder, S., D. Lagakos, and L. Ohanian (2014). Competitive pressure and the decline of the rust belt: A macroeconomic analysis. Technical report, National Bureau of Economic Research.

Allcott, H. (2015). Site selection bias in program evaluation. *The Quarterly Journal of Economics*, qjv015.

Allen, T., C. Arkolakis, and X. Li (2015). Optimal city structure. *Yale University, mimeograph*.

Ambrus, A., E. Field, and R. Gonzalez (2015). Loss in the time of cholera: Long-run impact of a disease epidemic on the urban landscape. Technical report, Citeseer.

Auffhammer, M., A. M. Bento, and S. E. Lowe (2009). Measuring the effects of the clean air act amendments on ambient pm10 concentrations: The critical importance of a spatially disaggregated analysis. *Journal of Environmental Economics and Management 58*(1), 15–26.

Autor, D., F. Levy, and R. J. Murnane (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics 118*(4), 1279–1333.

Banerjee, A., E. Duflo, and N. Qian (2012). On the road: Access to transportation infrastructure and economic growth in china. Technical report, National Bureau of Economic Research.

Banzhaf, S. H. and R. P. Walsh (2008). Do people vote with their feet? an empirical test of tiebout's mechanism. *The American Economic Review 98*(3), 843–863.

Baum-Snow, N., L. Brandt, J. V. Henderson, M. A. Turner, and Q. Zhang (2016). Roads, railroads and decentralization of chinese cities. *The Review of Economics and Statistics*.

Baum-Snow, N. and D. A. Hartley (2016). Accounting for central neighborhood change, 1980-2010.

Bayer, P., N. Keohane, and C. Timmins (2009). Migration and hedonic valuation: The case of air quality. *Journal of Environmental Economics and Management 58*(1), 1–14.

Behrens, C. and E. Pels (2012). Intermodal competition in the london–paris passenger market: High-speed rail and air transport. *Journal of Urban Economics 71*(3), 278–288.

Behrens, K., G. Duranton, and F. Robert-Nicoud (2014). Productive cities: Sorting, selection, and agglomeration. *Journal of Political Economy 122*(3), 507–553.

Bento, A., M. Freedman, and C. Lang (2015). Who benefits from environmental regulation? evidence from the clean air act amendments. *Review of Economics and Statistics 97*(3), 610–622.

bernard, A. B., A. MOXNES, and S. Y. Umeno (2014). Geography and firm performance in the japanese production network. *Geography*.

Bird, L., M. Bolinger, T. Gagliano, R. Wiser, M. Brown, and B. Parsons (2005). Policies and market factors driving wind power development in the united states. *Energy Policy 33*(11), 1397–1407.

Blomquist, G. C., M. C. Berger, and J. P. Hoehn (1988). New estimates of quality of life in urban areas. *The American Economic Review*, 89–107.

Bollinger, B. and K. Gillingham (2012). Peer effects in the diffusion of solar photovoltaic panels. *Marketing Science 31*(6), 900–912.

Bryan, G. and M. Morten (2015). Economic development and the spatial allocation of labor: Evidence from indonesia. *Manuscript, London School of Economics and Stanford University*.

Bullock, R. H., A. Salzberg, and Y. Jin (2012). High-speed rail–the first three years: Taking the pulse of china's emerging program. Technical report, World Bank.

Burgess, R. and D. Donaldson (2010). Can openness mitigate the effects of weather shocks? evidence from india's famine era. *The American economic review*, 449–453.

Callaway, D., M. Fowlie, and G. McCormick (2015). Location, location, location? what drives variation in the marginal benefits of renewable energy and demand-side efficiency? *URL http://nature. berkeley. edu/˜ fowlie/CFM. pdf. Manuscript: UC Berkeley*.

Chay, K. Y. and M. Greenstone (2003). Air quality, infant mortality, and the clean air act of 1970. Technical report, National Bureau of Economic Research.

Ciccone, A. and G. Peri (2006). Identifying human-capital externalities: Theory with applications. *The Review of Economic Studies 73*(2), 381–412.

Clark, T. N., R. Lloyd, K. K. Wong, and P. Jain (2002). Amenities drive urban growth. *Journal of urban affairs 24*(5), 493–515.

Combes, P.-P., G. Duranton, and L. Gobillon (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics 63*(2), 723–742.

Conley, T. G. (1999). Gmm estimation with cross sectional dependence. *Journal of econometrics 92*(1), 1–45.

Cook, J. A. and C.-Y. C. Lin (2015). Wind turbine shutdowns and upgrades in denmark: Timing decisions and the impact of government policy.

Council, S. (2004). zhong chang qi tie lu wang gui hua. Technical report, State Council.

Couture, V. and J. Handbury (2015). Urban revival in america, 2000 to 2010. In *American Economic Association annual meeting. https://www. dropbox. com/s/hlmnbjsx32eaafb/UrbanizationPatterns_ vCurrentPublic. pdf.*

Cullen, J. (2013). Measuring the environmental benefits of wind-generated electricity. *American Economic Journal: Economic Policy 5*(4), 107–133.

Currie, J., L. Davis, M. Greenstone, and R. Walker (2015). Environmental health risks and housing values: evidence from 1,600 toxic plant openings and closings. *The American economic review 105*(2), 678–709.

Davis, D. R. and J. I. Dingel (2014). The comparative advantage of cities. Technical report, National Bureau of Economic Research.

De la Roca, J., G. I. Ottaviano, and D. Puga (2014). City of dreams.

Delmas, M. A. and M. J. Montes-Sancho (2011). Us state policies for renewable energy: Context and effectiveness. *Energy Policy 39*(5), 2273–2288.

Dericks, G. and H. R. Koster (2016). Density restrictions and the urban spatial structure: Evidence from blitz bombings. Technical report, Mimeo.

Diamond, R. (2016). The determinants and welfare implications of us workers' diverging location choices by skill: 1980–2000. *The American Economic Review 106*(3), 479–524.

Donaldson, D. Railroads of the raj: Estimating the impact of transportation infrastructure. *American Economic Review*.

Donaldson, D. and R. Hornbeck (2016). Railroads and american economic growth: A "market access" approach. *The Quarterly Journal of Economics*, qjw002.

Duranton, G. and M. A. Turner (2012). Urban growth and transportation. *The Review of Economic Studies 79*(4), 1407–1440.

Eaton, J. and S. Kortum (2002). Technology, geography, and trade. *Econometrica 70*(5), 1741–1779.

Epple, D. and H. Sieg (1999). Estimating equilibrium models of local jurisdictions. *Journal of political economy 107*(4), 645–681.

Faber, B. (2014). Trade integration, market size, and industrialization: evidence from china's national trunk highway system. *The Review of Economic Studies*, rdu010.

Fajgelbaum, P. D., E. Morales, J. C. S. Serrato, and O. M. Zidar (2015). State taxes and spatial misallocation. Technical report, National Bureau of Economic Research.

Fee, K., D. Hartley, et al. (2012). The relationship between city center density and urban growth or decline. Technical report.

Feyrer, J., B. Sacerdote, A. D. Stern, A. Saiz, and W. C. Strange (2007). Did the rust belt become shiny? a study of cities and counties that lost steel and auto jobs in the 1980s [with comments]. *Brookings-Wharton Papers on Urban Affairs*, 41–102.

Fowlie, M. (2010). Emissions trading, electricity restructuring, and investment in pollution abatement. *The American Economic Review 100*(3), 837–869.

Glaeser, E. L. and J. D. Gottlieb (2009). The wealth of cities: Agglomeration economies and spatial equilibrium in the united states. *Journal of Economic Literature 47*(4), 983–1028.

Glaeser, E. L., H. D. Kallal, J. A. Scheinkman, and A. Shleifer (1992). Growth in cities. *Journal of political economy 100*(6), 1126–1152.

Glaeser, E. L., S. P. Kerr, and W. R. Kerr (2015). Entrepreneurship and urban growth: An empirical assessment with historical mines. *Review of Economics and Statistics 97*(2), 498–520.

Glaeser, E. L. and G. A. Ponzetto (2007). Did the death of distance hurt detroit and help new york? Technical report, National Bureau of Economic Research.

Glaeser, E. L. and A. Saiz (2004a). The rise of the skilled city. *Brookings-Wharton Papers on Urban Affairs 2004*(1), 47–105.

Glaeser, E. L. and A. Saiz (2004b). The rise of the skilled city. *Brookings-Wharton Papers on Urban Affairs 2004*(1), 47–105.

Greenstone, M., J. A. List, and C. Syverson (2012). The effects of environmental regulation on the competitiveness of us manufacturing. Technical report, National Bureau of Economic Research.

Grossman, G. M. and E. Rossi-Hansberg (2008). Trading tasks: A simple theory of offshoring. *The American Economic Review 98*(5), 1978–1997.

Guerrieri, V., D. Hartley, and E. Hurst (2013). Endogenous gentrification and housing price dynamics. *Journal of Public Economics 100*, 45–60.

Hanna, R. and P. Oliva (2015). The effect of pollution on labor supply: Evidence from a natural experiment in mexico city. *Journal of Public Economics 122*, 68–79.

Heblich, S., A. Trew, and Y. Zylberberg (2016). East side story: Historical pollution and neighborhood segregation. Technical report, Mimeo.

Heintzelman, M. D. and C. M. Tuttle (2012). Values in the wind: a hedonic analysis of wind power facilities. *Land Economics 88*(3), 571–588.

Helpman, E. (1998). The size of regions. *Topics in public economics: Theoretical and applied analysis*, 33–54.

Honoré, B. E. (1992). Trimmed lad and least squares estimation of truncated and censored regression models with fixed effects. *Econometrica: journal of the Econometric Society*, 533–565.

Hornbeck, R. and D. Keniston (2014). Creative destruction: Barriers to urban growth and the great boston fire of 1872. Technical report, National Bureau of Economic Research.

Ito, K. and J. M. Sallee (2015). The economics of attribute-based regulation: Theory and evidence from fuel-economy standards.

Jacobsen, M., J. LaRiviere, and M. Price (2014). Public goods provision in the presence of heterogeneous green preferences. Technical report, National Bureau of Economic Research.

Jianbin, L. (2011). Survey and analysis on passenger travel characteristics and distribution features of wuhan-guangzhou high speed railway [j]. *Railway Standard Design 11*, 003.

Kahn, M. E. and N. Kok (2014). The capitalization of green labels in the california housing market. *Regional Science and Urban Economics 47*, 25–34.

Kuminoff, N. V., V. K. Smith, and C. Timmins (2013). The new economics of equilibrium sorting and policy evaluation using housing markets. *Journal of Economic Literature 51*(4), 1007–1062.

Lee, S. and J. Lin (2015). Natural amenities, neighborhood dynamics, and persistence in the spatial distribution of income.

Logan, J. R., Z. Xu, and B. J. Stults (2014). Interpolating us decennial census tract data from as early as 1970 to 2010: A longitudinal tract database. *The Professional Geographer 66*(3), 412–420.

Michaels, G. (2008). The effect of trade on the demand for skill: Evidence from the interstate highway system. *The Review of Economics and Statistics 90*(4), 683–701.

MOR (2008). Mid-to-long railway plan. the ministry of railway of china. Technical report.

Moretti, E. (2004). Workers' education, spillovers, and productivity: evidence from plant-level production functions. *The American Economic Review 94*(3), 656–690.

Moretti, E. (2012). *The new geography of jobs.* Houghton Mifflin Harcourt.

Morten, M. and J. Oliveira (2016). Migration, roads and labor market integration: Evidence from a planned capital city.

Ollivier, G., R. Bullock, Y. Jin, and N. Zhou (2014). High-speed railways in china: A look at traffic. *China Transport Topic 11.*

Ottaviano, G. I., G. Peri, and G. C. Wright (2013). Immigration, offshoring, and american jobs. *The American Economic Review 103*(5), 1925–1959.

Price, C. V., N. Nakagaki, K. J. Hitt, and R. M. Clawges (2007). Enhanced historical land-use and land-cover data sets of the us geological survey. Technical report, Geological Survey (US).

Qin, Y. (2016). 'no county left behind?'the distributional impact of high-speed rail upgrades in china. *Journal of Economic Geography*, lbw013.

Redding, S. J. and D. M. Sturm (2016). Estimating neighborhood effects: Evidence from war-time destruction in london. *Princeton University, mimeograph*.

Ryan, S. P. (2012). The costs of environmental regulation in a concentrated industry. *Econometrica 80*(3), 1019–1061.

Sasaki, K., T. Ohashi, and A. Ando (1997). High-speed rail transit impact on regional systems: does the shinkansen contribute to dispersion? *The annals of regional science 31*(1), 77–98.

Sexton, S. E. and A. L. Sexton (2014). Conspicuous conservation: The prius halo and willingness to pay for environmental bona fides. *Journal of Environmental Economics and Management 67*(3), 303–317.

Siodla, J. (2015). Razing san francisco: The 1906 disaster as a natural experiment in urban redevelopment. *Journal of Urban Economics 89*, 48–61.

Sullivan, D. M. (2016). The true cost of air pollution: Evidence from house prices and migration.

Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of political economy 64*(5), 416–424.

Walker, W. R. (2013). The transitional costs of sectoral reallocation: Evidence from the clean air act and the workforce. *The Quarterly journal of economics*, qjt022.

Wilson, W. E. and H. H. Suh (1997). Fine particles and coarse particles: concentration relationships relevant to epidemiologic studies. *Journal of the Air & Waste Management Association 47*(12), 1238–1249.

Wu, J. (2006). Environmental amenities, urban sprawl, and community characteristics. *Journal of Environmental Economics and Management 52*(2), 527–547.

Wu, K., C. Fang, M. Zhao, et al. (2013). The intercity space of flow influenced by high-speed rail: A case study for the rail transit passenger behavior between beijing and tianjin. *Acta Geographica Sinica 68*(2), 159–174.

Yin, H. and N. Powers (2010). Do state renewable portfolio standards promote in-state renewable generation. *Energy Policy 38*(2), 1140–1149.

Zheng, S. and M. E. Kahn (2013). China's bullet trains facilitate market integration and mitigate the cost of megacity growth. *Proceedings of the National Academy of Sciences 110*(14), E1248–E1253.

Zivin, J. G. and M. Neidell (2012). The impact of pollution on worker productivity. *The American economic review 102*(7), 3652–3673.

Zivin, J. S. G., M. J. Kotchen, and E. T. Mansur (2014). Spatial and temporal heterogeneity of marginal emissions: Implications for electric cars and other electricity-shifting policies. *Journal of Economic Behavior & Organization 107*, 248–268.