Title
Essays on Environmental Economics

Author
Tang, Qu

Publication Date
2015

Peer reviewed|Thesis/dissertation

Essays on Environmental Economics

by

Qu Tang

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy
in
Agricultural and Resource Economics
in the
Graduate Division
of the
University of California, Berkeley

Committee in charge:

Professor Gordon C. Rausser, Chair
Associate Professor Meredith Fowlie
Professor Catherine D. Wolfram

Spring 2015

Essays on Environmental Economics

by

Qu Tang

Abstract

Essays on Environmental Economics

by

Qu Tang

Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Professor Gordon C. Rausser, Chair

This dissertation is comprised of three essays that apply microeconomics theory and econometric methods to study important issues in environmental economics.

In the first essay, I investigate the impacts of imposing inter-state trade restrictions on the compliance costs of coal-fired electric generating units (EGUs) in the context of a U.S. SO_2 emissions trading program (the Acid Rain Program). Over the past decade, tremendous efforts have been devoted to modifying emissions trading programs to address cross-state air pollution problems. The modification involves imposing more restrictions on emissions trading across geographical areas. The empirical question is how severe trade restrictions affect the regulated firms' compliance costs. Using rich data from the Acid Rain Program, this essay developed a discrete-continuous model to estimate electric generating units' compliance strategies and marginal abatement costs associated with the nationwide uniform emissions trading as the program was implemented in practice. Based on the estimation results, this essay then simulated units' compliance behaviors and the corresponding compliance costs if interstate trading had been prohibited. The results show that the aggregate compliance costs would increase more than one and a half times for the same emissions reduction goal due to the narrower trading markets in the counterfactual policy design with trade restrictions, and the costs would vary dramatically across space. Combined with the analysis on the benefit side, the results of this essay could be used to predict welfare impacts associated with trade restrictions at both national level and state level. And it may shed light on the future modification and implementation of EPA's cross-state air pollution regulations.

The second essay applies an equilibrium sorting model to a brand-new housing market in Beijing, China to estimate household preferences for neighborhood public goods provision, including public transportation services, public primary schools, and environmental amenities. The equilibrium sorting model is based on a discrete choice model of household residential location decisions. Relying on a unique, detailed data set on housing location, price, and other household characteristics, I estimate the model following the two-step BLP method, taking into account the heterogeneity of household preferences, incorporating neighborhood-specific

unobservable characteristics, and addressing the endogeneity of housing prices using instrumental variables. The results suggest that in general, lower housing price, better environmental amenities, and being closer to job centers will increase the choice opportunity of a neighborhood, and public transportation systems play a more important role in the neighborhoods far away from urban centers. Moreover, different households show varying preferences for these public goods. A distinct fact is that in addition to income, people's preferences vary greatly with generation (head age of households) and job type (whether there are public employees), which reveal the significant differences between generations and illustrate the welfare for public employees within the context of the transitional economy in China. This preference heterogeneity implies that future policies should be more geographically asymmetric, locally targeted and tailored based on specific socio-economic characteristics.

The third essay estimates the impact of climate change on the crop yields in China. I use a 11-year county-level panel data set covering more than 1,000 counties to estimate the effects of random year-to-year variation in weather on three major crops yields, including rice, wheat, and corn. Because it is not easy for small-scale farmers to adapt to climate change quickly in short time, these estimates could be used to plausibly predict the short to medium-run impacts of climate change on crop yields in China. The essay finds that over the period 2040-2060, projected climate change would reduce rice yield by 1.18% under a comparatively high emission scenario and by 0.08% under a medium-low scenario, reduce corn yield by 2.21% and 1.64% under the two emission scenarios, respectively, and increase wheat yield by 6.68% and 5.48% under the two emission scenarios, respectively. These findings may shed light on future policy designs to enhance the adaptive capacity of agriculture in China and thus ensure food security in the context of climate change.

Contents

Acknowledgements

I am incredibly indebted to my advisor, Gordon Rausser, for his invaluable guidance and continuous support through my graduate studies. I am also deeply grateful to Meredith Fowlie for her insightful advice and encouragement. I thank Sofia Villas-Boas, Catherine Wolfram, Denis Nekipelov, Peter Berck, Jeff Perloff, Jaremy Magruder, Ben Handel, Kenneth Train, and seminar participants at UC Berkeley for their comments and suggestions. Thanks to all of my wonderful classmates and colleagues in the ARE department who have made my graduate life so much more enjoyable. The staff in the ARE department provided invaluable administrative help. In particular, I thank Gail Vawter, Diana Lazo, and Carmen Karahalios for the professionalism and friendliness. I would also like to express special gratitude to my parents for their love and support. Most of all, thank you to my beloved husband Yizhen, for always being there with your love, patience, and understanding, and to our dear son Timothy, for filling our life with courage, hope, and happiness. Without you, I can't get through those difficult times.

Chapter 1. Introduction

This dissertation is comprised of three essays that apply microeconomics theory and econometric methods to study important issues in environmental economics.

In the first essay, I investigate the impacts of changing the geographic scope of emissions trading on affected firms' abatement costs in the United States. More specifically, the empirical question asked in this essay is how costly it would be if a single, nation-wide emissions trading market were geographically separated, and emissions trading were unlimited within a state but prohibited across states. This research question is relevant to a complex challenge for environmental regulators to address interstate transport of air pollution. It is motivated by the controversial "Good Neighbor" provision in the Clean Air Act, which requires each state to take responsibility for the pollution generated within the state boundary but significantly contributing to air quality problems in other states. Several times over the past decade, the U.S. EPA has attempted to interpret this provision by imposing restrictions on emissions trading across states.

I conduct the analysis in the context of the Acid Rain Program (ARP), which created a single nation-wide emissions trading market in the United States for SO_2. I first econometrically estimate the ex post compliance costs under the ARP as implemented in practice. With the benefit of the abundant data collected since the implementation of the ARP from 1995, I estimate a discrete-continuous compliance choice model using unit level data from the U.S. Energy Information Administration (EIA) form EIA-767, Federal Energy Regulatory Commission (FERC) form FERC-423, and EPA's continuous emission monitoring system (CEMS). In the model, an electric generating unit (EGU) is assumed to make a discrete choice between two major SO_2 abatement strategies, installing flue gas desulfurization equipment ("scrubber") or switching to/blending with low-sulfur coal (non-scrubbing) in the first step. Conditional on the chosen abatement strategy, the continuous emission rate of the unit is determined in the second step. Using the estimates from the discrete-continuous choice model, I then simulate how much EGUs' compliance costs would increase and how they would be distributed among states associated with the counterfactual policy design, which assume that the nationwide allowances trading market had been geographically separated and interstate trading had been prohibited.

The results show that under the ARP as implemented, the ex post estimated SO_2 allowance prices were 228$/ton and 510$/ton to achieve the 1995 ARP cap and 2000 ARP cap, respectively, leading to aggregate compliance costs of $\$107.46\times10^6$ in 1995 and $\$1136.43\times10^6$ in 2000. If the ARP had been designed allowing only intrastate allowance trading, the aggregate compliance cost would increase to $\$533.62\times10^6$ in 1995 and $\$1772.54\times10^6$ in 2000. Also, these costs vary dramatically across states, with a range of the allowance prices from zero in some states to more than $1000/ton in other states.

This essay contributes to the literature by combining discrete abatement strategy choices and continuous emission rate decisions using a discrete-continuous model. This model fits the observed units' choice data better than a simple discrete choice model, and leads to more plausible estimates for marginal abatement costs. Moreover, the results show substantive cost increase in the counterfactual policy design, and the cost increase is very unevenly distributed across states. Combined with the analysis on the benefit side, the results of this essay could be used to predict welfare impacts associated with trade restrictions at both national level and state level. And it may shed light on the future modification and implementation of EPA's cross-state air pollution regulations.

The second essay applies an equilibrium sorting model to investigate the residential location choice in a brand-new housing market in Beijing, China. Since 1998, the Chinese central government started to abolish the long-established welfare housing system, and in Beijing, it was until 2002 that such a system had been totally abandoned. From then on, households had more and more opportunities to sort themselves into neighborhoods within the city according to their heterogeneous preferences for neighborhood attributes. Relying on a unique, detailed data set collected from Beijing Household Transportation Survey in 2005, this essay provides the first empirical study of individual household's heterogeneous preference for neighborhood public goods provision, including public transportation services, public primary schools, government-owned parks and non-park green space.

In the model, a household's utility from choosing a particular neighborhood is decomposed into three parts. First is the mean indirect utility provided by the neighborhood, which is determined by all observed and unobserved neighborhood attributes including average housing price, and is common to all households. The second part is the portion of utility unique to each household through interacting household characteristics with neighborhood attributes, which represents individual household's heterogeneous preference. And the third part is an idiosyncratic error term. To identify the parameters of a household's utility function, I follow a two-step procedure discussed in Berry et al. (1995). By including unobserved neighborhood attributes into the mean indirect utility, the parameters for the interaction terms can be identified in the first step using a multinomial logit model. The second step recovers the remaining parameters by decomposing the mean indirect utility into its observable and unobservable components. The endogenous housing price in each neighborhood is instrumented using attributes in other neighborhoods in the second step.

The estimation results suggest that people do take into consideration public goods provision when they choose their residences. In general, lower housing price, better environmental amenities, and being closer to job centers will increase the choice opportunity of a neighborhood. Public transportation systems play a more important role in the neighborhoods far away from urban centers. Moreover, different households show varying preferences for these public goods. A distinct fact is that in addition to income, people's preferences vary greatly with generation (head age of households) and job type (whether there are public employees), which reveals the significant differences between generations and illustrates the welfare for public employees within the context of transitional economy. These findings can be used in evaluating the effects of government policies on social welfare, and also can shed light on future policy designs.

The third essay investigates the effects of climate change on the yields of major crops in China, including rice, wheat, and corn. Understanding the precise link between climate variables and crop yields could help to enhance the adaptive capacity of agriculture and thus ensure food security in the context of climate change.

Over the past 100 years, the average temperature in China has risen by 0.5~0.8 ℃, and the annual precipitation has decreased significantly in most of northern China, eastern part of the northwest, and northeastern China. By 2050, the average temperature in China is projected to increase by 2.3~3.3 ℃ as compared with that in 2000, the nationwide precipitation is projected to increase of 5~7%, and extreme weather/climate events are expected to occur more frequently.

Such climate change will bring unprecedented challenges to China's agriculture sector because of the sensitivity of agricultural production to climate variables.

This essay provides empirical estimates of climate change on major crops yields in China in a panel setting. Following Dischênes and Greenstone (2007), using a detailed county-level panel dataset in China from 2000 to 2010, I regress the annual yields of rice, wheat, and corn against two annual weather parameters, namely, temperature and precipitation, conditional on city fixed effects and province-by-year fixed effects. Therefore, the weather parameters are identified from the county-specific deviations in yearly weather from the average climate within the same city after adjusting for shocks common to all counties in the same province. Since the variations in weather are plausibly exogenous and random after controlling for time-invariant idiosyncratic features of the city and time-variant shocks of the province, the effects of weather on crop yields are therefore clearly identified in the panel data setting.

The estimation results show that the relationship between all three crops yields and temperature is inverted U-shaped, and the relationship between rice and corn yields and precipitation also follows the inverted U-shape pattern, but wheat yield is decreasing with respect to precipitation during the research period. Using the estimation results, the essay then predicts the impacts of future climate change on the three crops yields under two emission scenarios (RCP8.5 and RCP4.5) in China. Overall, the impacts of climate change over the short to medium term (2040-2060) will be negative on rice and corn yields, but positive on wheat yield for the two scenarios. Specifically, the average rice yield is expected to decrease by 0.08-1.18%, depending on different emission scenarios, the average wheat yield would increase by 5.48-6.68%, and the average corn yield would have the largest decrease by 1.64-2.21%. These yield effects imply an economic loss of $\$343\times10^6$ to $\$524\times10^6$ in the rice sector, and $\$361\times10^6$ to $\$487\times10^6$ in the corn sector, while an economic gain of $\$783\times10^6$ to $\$955\times10^6$ in the wheat sector.

Chapter 2. Costs to Trade Restrictions: Evidence from Limiting Interstate SO_2 Emissions Trading in the United States[1]

[1] I would like to thank comments from Gordon Rausser, Meredith Fowlie, Sofia Villas-Boas, Peter Berck, Catherine Wolfram, and Shihe Fu. All errors are my own.

2.1 Introduction to Chapter

For non-uniformly mixed pollutants, such as sulfur oxide (SOx) and nitrogen oxide (NOx), pollution damage varies across space due to the location of emissions sources and the complexity of dispersion conditions. Achieving ideal economic efficiency requires emission regulations to be differentiated, that is, the marginal abatement cost should be set equal to the marginal damage across sources (Tietenberg, 1995). Theoretically, differentiation could be incorporated into an emissions trading system either by setting non-uniform trading ratios between regulated sources according to their marginal damages (Montgomery, 1972; Krupnick et al., 2000; Farrow et al., 2005; Muller and Mendelsohn, 2009; Fowlie and Muller, 2013), or by dividing the control area into multiple submarkets with uniform one-for-one permit trading within a submarket and limited trading between submarkets (Tietenberg, 1995; Mendelsohn, 1986). However, the practical application of these differentiated policy designs is far too complicated and most of non-uniformly mixed pollutants are currently subject to undifferentiated regulations.

Over the past decade, the Environmental Protection Agency (EPA) has made tremendous efforts to address cross-state air pollution problems by attempting to modify the emissions trading programs. The modifications involve two main aspects. One is to set emissions caps at a finer geographic level, considering regional variation in emission damages and abatement costs. Another is to impose more restrictions on emissions trading across space. After repeated lawsuits, rehearing, and revisions, EPA's most recent attempt--the Cross State Air Pollution Rule (CSAPR)--was upheld by the Supreme Court on April 29, 2014. The CSAPR, issued on August 8, 2011, was designed to curtail SO_2 and NOx emissions in upwind states but affecting downwind states by setting state-specific caps, allowing only intrastate emissions trading, and prohibiting interstate trading. Challenges still exist, including debates over EPA's modeling method for determining state-specific emissions reduction responsibilities, and how geographically flexible the emissions trading should be allowed. For EPA to take further steps to implement a differentiated emissions trading program, a more sophisticated analysis will be needed.

This essay aims to investigate the impacts of imposing trade restrictions on the compliance costs of coal-fired electric generating units (EGUs) in the context of the Acid Rain Program (ARP). As the first large-scale, long-term emissions trading program, the ARP has achieved remarkable success in terms of SO_2 emissions reduction and abatement costs saving (Schmalensee and Stavins, 2013). Under the ARP, SO_2 emissions permits ("allowances") were allocated to affected EGUs and traded on a one-for-one basis across the whole country, which gave the EGUs great flexibility to achieve the SO_2 emissions reduction goal at a cost considerably lower than that of command-and-control regulations (Carlson et al., 2000). In this essay, I first econometrically estimate the ex post compliance costs under the ARP as implemented in practice, then I simulate how much these costs would increase and how they would be distributed among states if the nationwide allowances trading market under the ARP had been geographically separated and interstate trading had been prohibited, as proposed in the CSAPR.

A number of empirical studies have investigated the compliance strategies of power plants under the ARP. A unit has two major options to comply with the ARP. It can install a

capital-intensive flue gas desulfurization equipment ("scrubber") and continue to burn cheap high-sulfur coal. Alternatively, it can use more expensive low-sulfur coal with almost no large capital investment. Previous studies usually assume that a unit is making a 0-1 decision, either scrubbing or completely switching to low-sulfur coal, so a discrete choice model is applied. For example, using a probit model of the choice between scrubbing and switching to low-sulfur coal under different policy instruments, Keohane(2002) finds that the change in the average cost of scrubbing or coal-switching has a greater influence on the probability of scrubbing for units under the ARP than for units under a uniform emissions rate standard. Arimura (2002) estimates a probit model of EGUs' compliance decision making between coal-switching and purchasing allowances under the ARP in Phase I (from 1995 to 1999) and applies this model to investigate the effects of public utility commission (PUC) regulations on the abatement choices of Phase I units. Fullerton, McDermott and Caulkins (1997) also analyze EGUs' compliance choices using a discrete choice model and numerically estimate the effects of PUC regulations. Carlson et al. (2000) do take into account continuous blending with low-sulfur coal and empirically estimate marginal abatement cost functions for coal blending at all EGUs without scrubbers using a translog cost function, however, they ignore the discrete choice decision between scrubbing and non-scrubbing.

With the benefit of the abundant data collected since the implementation of the ARP from 1995, I estimate a discrete-continuous compliance choice model using unit level data from the U.S. Energy Information Administration (EIA) form EIA-767, Federal Energy Regulatory Commission (FERC) form FERC-423, and EPA's continuous emission monitoring system (CEMS). In the model, an electric generating unit is assumed to make a discrete choice between scrubbing or switching to/blending with low-sulfur coal (non-scrubbing) in the first step. Conditional on the chosen abatement strategy, the continuous emission rate of the unit is determined in the second step. Using the estimates from the discrete-continuous choice model, I then simulate EGUs' compliance choices and the corresponding abatement costs associated with counterfactual policy designs.

The findings show that under the ARP as implemented, the ex post estimated SO_2 allowance prices were 228\$/ton and 510\$/ton to achieve the 1995 ARP cap and 2000 ARP cap, respectively, leading to aggregate compliance costs of $\$107.46\times10^6$ in 1995 and $\$1136.43\times10^6$ in 2000. If the ARP had been designed allowing only intrastate allowance trading, the aggregate compliance cost would increase to $\$533.62\times10^6$ in 1995 and $\$1172.54\times10^6$ in 2000. Also, these costs vary dramatically across states, with a range of the allowance prices from 0\$/ton in some states (Alabama, Georgia, Kansas, Pennsylvania, and Tennessee in 1995; Arizona, Colorado, Delaware, Kansas, Michigan, Nebraska, Nevada, Oklahoma, Utah, and Washington in 2000) to more than 1000\$/ton in other states (Michigan and Wisconsin in 1995; Alabama, Georgia, New Hampshire, New Mexico, South Carolina, South Dakota, and Virginia in 2000).

This essay contributes to the literature by combining discrete abatement strategy choices and continuous emission rate decisions using a discrete-continuous model developed by Dubin and McFadden (1984). This model fits the observed units' choice data better than a simple discrete choice model. Figure 2.1 shows histogram for the heat input ratio of low-sulfur coal in all non-scrubbing power plants. It is clear that to comply with the ARP, most non-scrubbing power plants chose partially blending with low-sulfur coal (as the heat input ratio of low-sulfur

coal is less than one) rather than totally switching to low-sulfur coal (in which case the heat input ratio of low-sulfur coal would equal one). Furthermore, Figure 2.2 shows that the highest blending ratio of low-sulfur coal in most non-scrubbing plants could reach one. This, to some extent, demonstrates that the partially blending decision is not determined by technological constraints but affected by variables such as the price premium of low-sulfur coal, unit capacity, unit heat rate, and so on. I estimate these variables in the continuous step model. Moreover, this essay is among the first to empirically evaluate the costs of imposing geographical restrictions on SO_2 emissions trading. The results show substantive cost increase in the counterfactual policy design, and the cost increase is very unevenly distributed across states. Combined with the analysis on the benefit side, the results of this essay could be used to predict welfare impacts associated with trade restrictions at both national level and state level. And it may shed light on the future modification and implementation of EPA's cross-state air pollution regulations, such as the CSAPR.

The rest of the chapter is organized as follows. Section 2.2 introduces the background of SO_2 emission regulations in the United States. In Section 2.3, a discrete-continuous model is developed for estimating the marginal abatement costs of SO_2. Section 2.4 summarizes the data, the analysis procedure, and the identification strategies. Estimation and simulation results are provided in Section 2.5 and Section 2.6, respectively. Section 2.7 concludes.

2.2 Background on the U.S. SO_2 Emission Regulations

2.2.1 The Acid Rain Program (ARP)

The Acid Rain Program (ARP), established under Title IV of the 1990 Clean Air Act Amendments (1990 CAAA), was the first large-scale, long-term application of an emission permits (called "allowances" in the 1990 legislation) trading system to control emissions.[2] The overall goal of the program was to achieve significant environmental and public health benefits through reductions of sulfur dioxide (SO_2) and nitrogen oxide (NOx) emissions. In particular, the program required annual SO_2 emissions reductions of 10 million tons below 1980 levels (around 17 million tons). The program was implemented in two phases: Phase I from 1995 to 1999, and Phase II from 2000 beyond.

Phase I affected 263 of the dirtiest electric generating units (Table A units) at 110 mostly coal-fired generating plants located in 21 eastern and midwestern states. An additional 182 units voluntarily participated in Phase I of the program as substitution or compensating units, bringing the total number of Phase I units to 445.[3] Based on the historical fuel consumption (generally the unit's 1985-1987 average) and a specific emissions rate of around 2.5 pounds of SO_2 per million

---

[2] Detailed information about the ARP is provided on EPA's website http://www.epa.gov/airmarkets/progsregs/arp/, and in EPA Acid Rain Program Progress Reports.

[3] Montero (1999) examines the welfare implications of the voluntary opt-in provision of the ARP, and finds that this provision does not significantly affect the price and trading volume of the SO_2 allowance market, though most voluntarily participating units have counterfactual emissions below their allowance allocations. In this essay, I ignore the opt-in decision of these additional units and treats them exactly the same as Table A units.

Btus of heat input, the 263 Table A units were allocated annual tradable emissions allowances of roughly 5.6 million tons from 1995, and the total allowances issued to the 445 Phase I units were about 8.7 million tons. An allowance authorized a unit to emit one ton of SO_2 in the year of issuance or any year thereafter. Once allocated, allowances could be bought, sold, traded, or banked for use in future years.

Phase II, which began in 2000, covered all fossil-fueled electric generating units with an output capacity of greater than 25 megawatts, encompassing over 2,000 units in all. The allowance allocation rule in Phase II was similar to that of Phase I, but with a tighter emission rate of around 1.2 pounds of SO_2 per million Btus of heat input.

Figure 2.3 shows the annual SO_2 allowances available to the ARP units, as well as the actual emissions from these units. Significant SO_2 emission reductions were achieved. In 1995, the first compliance year of the program, the SO_2 emissions of regulated Phase I units were 5.3 million tons, a 44% reduction from 1980 emissions level. In 2000, the first year of Phase II of the program, the SO_2 emissions of all regulated units were 11.2 million tons, a reduction of 35% from the 1980 emissions level. It is worth noting that there was significant "over-compliance" during Phase I because of the banking provisions of the program. Anticipating more stringent regulations in Phase II, units chose to bank allowances for future use.

SO_2 allowance prices ranged between 100$/ton and 200$/ton during Phase I and the early years of Phase II (Figure 2.4), reflecting the low cost of implementing the ARP. Since 2004, SO_2 allowance prices have increased dramatically due to regulatory changes discussed in the following section.

2.2.2 EPA's Recent Modifications on the SO_2 Emissions Trading Programs

To eliminate the significant contribution of upwind states' SO_2 emissions to the nonattainment status of the national ambient air quality standard (NAAQS) for fine particles (PM2.5) in downwind states, EPA issued the Clean Air Interstate Rule (CAIR) in May, 2005, applying a more stringent emission cap on 27 eastern states and the District of Columbia.[4] The CAIR established an emissions trading program for SO_2 based on the ARP, but in which more than one allowance was surrendered for each ton of SO_2 emission in the regulated states. In particular, SO_2 allowances of vintage (i.e., the year in which the allowance was issued) 2009 and earlier offset one ton of SO_2 emissions, vintages 2010 through 2014 offset 0.5 tons of emissions, and vintages 2015 and beyond offset 0.35 tons of emissions (FR, 70(91), pp. 25274). Because the SO_2 allowances of the ARP could still be used for compliance with the more stringent CAIR, the allowance prices increased dramatically from 273$/ton in EPA's 2004 auction to 703$/ton in the 2005 auction, as shown in Figure 2.4 (Shmalensee and Stavins, 2013).

In the 60 days after EPA published the CAIR, North Carolina and a number of utilities filed petitions for review in the Circuit Court of Appeals for the District of Columbia. On July 11, 2008, the Court responded to the petitions by vacating the CAIR in its entirety and remanding it back to EPA. One important reason for the Court's vacating CAIR was that the region-wide SO_2 cap for upwind states was "arbitrary and capricious" (*State of North Carolina v.*

[4] For more information on the CAIR, see EPA's website http://www.epa.gov/cleanairinterstaterule/.

Environmental Protection Agency, No. 05-1244, 2008, District of Columbia Court of Appeals). The Court ruling stated that "[The CAIR] must actually require elimination of emissions from sources that contribute significantly and interfere with maintenance in downwind nonattainment areas. To do so, [EPA] must measure each state's 'significant contribution' to downwind nonattainment..." On December 23, 2008, a Court ruling allowed the CAIR to remain in effect temporarily until EPA issued a valid replacement. The nationwide, unlimited interstate SO_2 allowance trading system created by the ARP was invalidated by the Court's decision on the CAIR (Schmalensee and Stavins, 2013), and as a result, the SO_2 allowance prices decreased dramatically after 2005, as shown in Figure 2.4.

On July 6, 2011, responding to the Court's concerns on the CAIR, EPA finalized the Cross-State Air Pollution Rule (CSAPR) to replace the CAIR.[5] Under the CSAPR, 23 states were required to reduce annual SO_2 and NO_x emissions from 2012 to help downwind areas attain the 24-Hour and/or Annual PM2.5 NAAQS, and 16 of them were required to make additional reductions from 2014. By 2014, the CSAPR, combined with other state and EPA actions, would reduce power plant SO_2 emissions by 73% (around 6.4 million tons) from 2005 levels.[6] To achieve the emission reductions target and ensure each state eliminates its own significant contribution to downwind pollution, state-specific emission caps were set and more constraints were added to the trading scheme by allowing only intrastate allowances trading and prohibiting interstate trading. The ARP SO_2 allowances were no longer applicable to the CSAPR, which resulted in the collapse of the nationwide SO_2 market. Allowance prices fell to record low levels, as shown in Figure 2.4 (Schmalensee and Stavins, 2013). After repeated lawsuits, rehearing, and revisions, the U.S. Supreme Court finally upheld the CSAPR on April 29, 2014, but challenges still exist regarding its implementation.

2.2.3 Additional Historical Background for the Federal Control of SO2 Emissions[7]

As the first significant federal air pollution legislation, the 1970 Clean Air Act Amendments (CAAA) established national maximum standards for ambient concentrations of SO2. Under the 1970 Amendments, emissions from new power plants constructed after 1971 were subject to a new source performance standard (NSPS), with an emissions rate not allowed to exceed 1.2 pounds of SO_2 per million Btus of heat input, while pre-1970 plants remained subject to controls under State Implementation Plans (SIPs), which were required by the amendments to comply with national air quality standards. Both new and old emission sources could choose how to meet the emission standards during this period.

Amended in 1977, the Clean Air Act Amendments required new coal-fired power plants built after 1978 not only to meet the 1.2 lbs SO_2/mmBtu emissions rate constraint, but also to operate with scrubbers even if they burned low-sulfur coal. Therefore, new emission sources lacked flexibility in choosing their abatement strategies after the 1977 Amendments.

---

[5] For more information on the CSAPR, see EPA's website http://www.epa.gov/airtransport/CSAPR/.

[6] EPA FACT SHEET. The Cross-State Air Pollution Rule: Reducing the Interstate Transport of Fine Particulate Matter and Ozone http://www.epa.gov/crossstaterule/pdfs/CSAPRFactsheet.pdf.

[7] This section draws on Joskow and Schmalensee (1998).

2.3 The Econometric Model

In this section, I first investigate an individual unit's choice of SO_2 abatement strategies under the Acid Rain Program using a discrete-continuous model based on the Dubin and McFadden (1984) two-step estimation method. Using this model and the estimation results, I further simulate the compliance costs associated with the counterfactual SO_2 emissions trading design in Section 2.6.

2.3.1 Discrete-Continuous Model

An individual unit n faces J discrete choices of SO_2 abatement strategies, indexed by $j = 1,...,J$. I consider two major strategies in this essay, with $j = 1$ being installing a scrubber to reduce SO_2 emissions, and $j = 0$ non-scrubbing, which includes totally switching to low-sulfur coal or partially blending with low-sulfur coal and purchasing SO_2 allowances. Conditional on choice j, the unit continuously chooses input quantities of high-sulfur coal and/or low-sulfur coal to minimize the electricity production cost in time period t, subject to an electricity output constraint. The unit's cost minimization problem is expressed as follows (time indexes are suppressed for simplicity):

$$min_{Q_{nj}^h, Q_{nj}^l} C_{nj} = w_{nj}^h \cdot Q_{nj}^h + w_{nj}^l \cdot Q_{nj}^l + K_{nj}$$

subject to (2.1)

$$f(Q_{nj}^h, Q_{nj}^l) \geq y_n$$

where w_{nj}^h and w_{nj}^l are the allowance-price-included fuel price for high-sulfur coal and low-sulfur coal, respectively (cent/btu); Q_{nj}^h and Q_{nj}^l are heat inputs for high-sulfur coal and low-sulfur coal, respectively (mmbtu); K_{nj} is the other fixed costs of choice j (cent); and y_n is the electricity output (MW), which is assumed not to change with respect to the compliance strategy choice j.

Solving the minimization problem (2.1) gives the conditional cost function

$$C_{nj} = C(j, y_n, w_{nj}^h, w_{nj}^l, K_{nj}, \xi_j) \tag{2.2}$$

In the first step, the unit will choose scrubbing ($j = 1$) or non-scrubbing ($j = 0$) to abate SO_2 emissions. The probability of choosing alternative j is

$$Pr_j = Prob\{(\xi_j, \xi_i) : C(j, y_n, w_{nj}^h, w_{nj}^l, K_{nj}, \xi_j) < C(i, y_n, w_{ni}^h, w_{ni}^l, K_{ni}, \xi_i) \; for \; i \neq j\} \tag{2.3}$$

In the second step, the unit will choose the consumption levels (mmbtu) of high-sulfur coal Q_{nj}^h and/or low-sulfur coal Q_{nj}^l conditional on the abatement strategy choice j. $j = 1$ and $Q_{nj}^l = 0$ imply that the unit chooses scrubbing as its SO_2 abatement strategy; $j = 0$ and $Q_{nj}^h = 0$ totally switching to low-sulfur coal; and $j = 0$ and $Q_{nj}^h > 0$ partially blending with low-sulfur coal. Therefore, the two abatement strategies, switching to low-sulfur coal and blending with low-sulfur coal, are distinguished in this step. By Shephard's lemma,

$$Q_{nj}^{h} = \frac{\partial C(j, y_n, w_{nj}^{h}, w_{nj}^{l}, K_{nj}, \xi_j)}{\partial w_{nj}^{h}} \tag{2.4}$$

$$Q_{nj}^{l} = \frac{\partial C(j, y_n, w_{nj}^{h}, w_{nj}^{l}, K_{nj}, \xi_j)}{\partial w_{nj}^{l}} \tag{2.5}$$

From equation (2.4) and (2.5), I can further derive the actual emission rate (lb/mmBtu) of unit n conditional on the choice j:

$$e_{nj} = e(j, y_n, w_{nj}^{h}, w_{nj}^{l}, K_{nj}, \xi_j) \tag{2.6}$$

2.3.2 Econometric Specification

To estimate the unit's choice of SO_2 emission abatement strategies and coal consumption, I specify a reduced form of the general model represented by (2.2) and (2.6), following Newell and Pizer (2008). For the cost function, I take a commonly used translog form, as in Carlson et al. (2000):

$$\begin{aligned} ln(C_{nj}) = {} & \alpha_1 + \gamma_1 \cdot ln(K_{nj}) + \gamma_2 \cdot ln(w_{nj}^{h}) + \gamma_3 \cdot ln(w_{nj}^{l}) + \gamma_4 \cdot ln(y_n) \\ & + \frac{1}{2}\gamma_5 \cdot (ln(w_{nj}^{h}))^2 + \frac{1}{2}\gamma_6 \cdot (ln(w_{nj}^{l}))^2 + \frac{1}{2}\gamma_7 \cdot (ln(y_n))^2 + \frac{1}{2}\gamma_8 \cdot ln(w_{nj}^{h}) \cdot ln(w_{nj}^{l}) \\ & + \frac{1}{2}\gamma_9 \cdot ln(w_{nj}^{h}) \cdot ln(y_n) + \frac{1}{2}\gamma_{10} \cdot ln(w_{nj}^{l}) \cdot ln(y_n) + \xi_{nj} \end{aligned} \tag{2.7}$$

Consider $j = 1$, the condition when the unit has chosen scrubbing. $ln(K_{n1})$ represents the unit's scrubbing capital cost.[8] I assume that only high sulfur coal is burned for scrubbing units, so all terms with w_{n1}^{l} are dropped off from the cost equation (2.7).[9] Moreover, in a SO_2 allowance market, the allowance-price-included high-sulfur coal price for a scrubbing unit could be written as

$$w_{n1}^{h} = \tau \cdot \mu_n^{h} \cdot s_n + p_n^{h} \tag{2.8}$$

where τ is the SO_2 allowance price (\$/ton), μ_n^{h} is the sulfur content of high-sulfur coal (lb/mmbtu), $s_n = 1 - scrubbing\ efficiency$ is the residual of sulfur after scrubbing, and p_n^{h} is the market price of high-sulfur coal.

Therefore, the conditional cost function of scrubbing is specified as

---

[8] Here I ignore the operating cost of scrubbing, as it is far smaller than the capital cost. In the data, the average reported capital cost for scrubbing units is $\$119.3 \times 10^6$, which is 30 times higher than the average operating cost.

[9] The average sulfur content for scrubbing units in the data is 1.7 lbs sulfur/mmbtu. 80% of scrubbing units burn high-sulfur coal with a sulfur content higher than 0.6 lbs sulfur/mmbtu, resulting in SO_2 emission rates higher than 1.2 lbs SO2/mmbtu. 65% of scrubbing units burn high-sulfur coal with a sulfur content higher than 1.25 lbs sulfur/mmbtu, resulting in SO_2 emission rates higher than 2.5 lbs SO2/mmbtu.

$$ln(C_{n1}) = \alpha_1 + \gamma_1 \cdot ln(K_{n1}) + \gamma_2 \cdot ln(w_{n1}^h) + \gamma_4 \cdot ln(y_n) + \frac{1}{2}\gamma_5 \cdot (ln(w_{n1}^h))^2 + \frac{1}{2}\gamma_7 \cdot (ln(y_n))^2 + \frac{1}{2}\gamma_9 \cdot ln(w_{n1}^h) \cdot ln(y_n) + \xi_{n1} \tag{2.9}$$

I assume that $\xi_{n1} \sim N(0,\sigma_1^2)$ and is independent across units.

Consider $j = 0$, the condition when the unit has chosen non-scrubbing. The non-scrubbing units could comply with the ARP by switching to or blending with low-sulfur coal, and/or purchasing SO_2 allowances from the allowance market. I assume that there is zero fixed cost for non-scrubbing, so $ln(K_{nj})$ is dropped off from the cost equation (2.7).[10] In a SO_2 allowance market, the allowance-price-included prices for high-sulfur coal and low-sulfur coal for a non-scrubbing unit could be written as

$$w_{n0}^h = \tau \cdot \mu_n^h + p_n^h \tag{2.10}$$

$$w_{n0}^l = \tau \cdot \mu_n^l + p_n^l \tag{2.11}$$

where τ, μ_n^h, and p_n^h are defined the same as in equation (2.8); μ_n^l is the sulfur content of low-sulfur coal; and p_n^l is the price of low-sulfur coal in the coal market.

Therefore, the conditional cost function for non-scrubbing is specified as

$$ln(C_{n0}) = \alpha_0 + \gamma_2 \cdot ln(w_{n0}^h) + \gamma_3 \cdot ln(w_{n0}^l) + \gamma_4 \cdot ln(y_n) + \frac{1}{2}\gamma_5 \cdot (ln(w_{n0}^h))^2 + \frac{1}{2}\gamma_6 \cdot (ln(w_{n0}^l))^2 + \frac{1}{2}\gamma_7 \cdot (ln(y_n))^2 + \frac{1}{2}\gamma_8 \cdot ln(w_{n0}^h) \cdot ln(w_{n0}^l) + \frac{1}{2}\gamma_9 \cdot ln(w_{n0}^h) \cdot ln(y_n) + \frac{1}{2}\gamma_{10} \cdot ln(w_{n0}^l) \cdot ln(y_n) + \xi_{n0} \tag{2.12}$$

Similarly, I assume that $\xi_{n0} \sim N(0,\sigma_0^2)$ and is independent across units.

Units might make the decision not only by minimizing costs but also by responding to other regulation factors (Carlson et al., 2000; Fowlie, 2010). Therefore, in the first step, I re-write the criteria for unit n choosing non-scrubbing as

$$C_{n0} < C_{n1} \cdot exp\{\eta \cdot R_n\} \tag{2.13}$$

where R_n represents other factors affecting unit decisions, including age_by90_n, boiler age by 1990, and a dummy variable $PROTECT_n = 1$ if the unit (plant) is located in the states as Kentucky, Illinois, Indiana, Ohio, or Pennsylvania, where the burning of high-sulfur coal is protected (Arimura, 2002).

[10] Keohane (2002) used 10$/KW as the fixed cost for switching to eastern low-sulfur coals and 50$/KW for switching to western low-sulfur coals based on surveys performed by Denny Ellerman and his colleagues. These are much smaller than the fixed cost for scrubbing in the data, which is around 225$/KW.

Therefore, in the first discrete choice step, the probability of unit n choosing non-scrubbing is

$$\begin{aligned} Pr_{n0} &= Prob(ln(C_{n1}) - ln(C_{n0}) + \eta_1 \cdot PROTECT_n + \eta_2 \cdot age\_by90_n > 0\} \\ &= Prob\{(\alpha_1 - \alpha_0) + \gamma_1 \cdot ln(K_{n1}) + \gamma_2 \cdot \Delta ln(w_n^h) - \gamma_3 \cdot ln(w_{n0}^l) \\ &\quad + \frac{1}{2}\gamma_5 \cdot \Delta(ln(w_n^h))^2 - \frac{1}{2}\gamma_6 \cdot (ln(w_{n0}^l))^2 - \frac{1}{2}\gamma_8 \cdot ln(w_{n0}^h) \cdot ln(w_{n0}^l) \\ &\quad + \frac{1}{2}\gamma_9 \cdot \Delta ln(w_n^h) \cdot ln(y_n) - \frac{1}{2}\gamma_{10} \cdot ln(w_{n0}^l) \cdot ln(y_n) + \eta_1 \cdot PROTECT_n \\ &\quad + \eta_2 \cdot age\_by90_n + (\xi_{n1} - \xi_{n0}) > 0\} \end{aligned} \tag{2.14}$$

where $\Delta ln(w_n^h) = ln(w_{n1}^h) - ln(w_{n0}^h)$ and $\Delta(ln(w_n^h))^2 = (ln(w_{n1}^h))^2 - (ln(w_{n0}^h))^2$.

Once a unit chooses scrubbing, only high-sulfur coal is consumed and the emission rate is determined by the sulfur content of high-sulfur coal burnt in the unit and the residual of sulfur after scrubbing as

$$e_{n1} = \mu_n^h \cdot s_n \tag{2.15}$$

If a unit chooses non-scrubbing, I directly specify a reduced form of the emission rate without satisfying the theoretical integrability constraints:[11]

$$e_{n0} = \theta_0 + \theta_1 \cdot (ln(w_{n0}^l) - ln(w_{n0}^h)) + \theta_2 \cdot ln(y_n) + \sum_{m=1}^{M} \theta_{2+m} \cdot X_{nm} + \zeta_{n0} \tag{2.16}$$

where $(ln(w_{n0}^l) - ln(w_{n0}^h))$ is the allowance-price-included price premium of low-sulfur coal for non-scrubbing units, y_n is electricity output, X_{nm} is a vector of unit characteristics, and $\zeta_{n0} \sim N(0, \sigma_\zeta^2)$ is allowed to be correlated with ξ_{n0}.

2.4 Data and Estimation Procedure

2.4.1 Data

2.4.1.1 General Data Description

The primary data source used in this essay is the U.S. Energy Information Administration (EIA) form EIA-767 from 1985 to 2003, which collects annual survey data from all U.S. steam-electric

[11] The reason for this reduced form specification is threefold. First, the derived input demand functions from a translog cost specification are too complicated to estimate. Second, the unit level input cost data necessary to estimate the linear-log input cost share equations from a translog cost function is lacking in the essay. Third, the ultimate goal of the essay is to simulate EGUs' SO_2 emissions with respect to different allowance prices, which requires the estimation of the emission rate.

generating plants with a nameplate capacity of 10 megawatts or larger.[12] Specific unit-level information contained in the data includes location, in-service and retirement date, capacity, total electricity output, capital and operating and maintenance costs for scrubbing units, and average heat content, sulfur content, and ash content of all coal types. Only coal-fired units are considered in this essay. Because coal-fired units constructed after 1978 were required to operate with scrubbers and had no flexibility to choose SO_2 abatement strategies according to the 1977 CAAA, these units are also excluded from the sample. Therefore, the sample consists of two types of units: type "D" units constructed between 1971 and 1978 and subject to a new source performance standard (NSPS) based on the 1970 CAAA, and type "N" units not covered by the NSPS. I also exclude units of these two types that installed scrubbers to comply with other SO_2 emission regulations before the establishment of the ARP in 1990. The unit-level EIA 767 data are further combined with plant-level fuel consumption data from FERC Form 423 to obtain detailed coal consumption information, including heat content, sulfur content, and delivery contract/spot market price for each type of coal.

Using the EPA Continuous Emission Monitoring System (CEMS) data, these sample units are divided into two groups based on their regulation phases under the ARP. Phase I units include the 263 dirtiest generating units (Table A units) which were required by the ARP to reduce their SO_2 emissions from 1995, and some additional units that voluntarily chose to join Phase I of the Program (Opt-In units, Compensating units, and Substitution units in the CEMS data). Phase II units, which began to reduce their SO_2 emissions in 2000, include almost all coal-fired units with an output capacity of greater than 25 megawatts.

After excluding observations with missing data, the final sample includes 872 units, of which 380 are Phase I units and 492 are Phase II units. Of them, 39 Phase I units and 25 Phase II installed scrubbers (Table 2.1).

2.4.1.2 Decision Period

EGUs' compliance choices between scrubbing and non-scrubbing are modeled in this essay as a static decision process. Given that it usually takes at least one year for a unit to install a scrubber, I assume that Phase I units made their scrubbing decisions between 1990 and 1994 to comply with the SO_2 emissions reduction requirement in the first year (1995) of Phase I, and Phase II units made their scrubbing decisions between 1996 and 1999 in order to comply with the SO_2 emissions reduction requirement in the first year (2000) of Phase II. This assumption is consistent with the data. As shown in Table 2.2, 34 of 39 Phase I scrubbing units started to scrub between 1992 and 1996, and 20 of 25 Phase II scrubbing units started to scrub between 1997 and 2001.

2.4.1.3 Data Sources, Limitations, and Summary Statistics

Scrubbing capital costs. For scrubbing units, the capital costs are reported directly in the EIA-767 data. All costs are converted to 1996 dollars using the Handy Whitman Electric Annual Construction Cost Indexes (U.S. Census Bureau). The average scrubbing capital cost was around

[12] Data after 2003 are not used in the estimation because SO_2 allowance prices increased dramatically during 2004-2005 due to the announcement of the CAIR, and then decreased sharply after 2005 because of the vacating of the CAIR.

$\$119.3 \times 10^6$ for scrubbing units in the sample. However, 10 of the 64 scrubbing units reported scrubbing capital costs lower than $\$2 \times 10^6$. I generate a dummy variable $extreme = 1$ to denote these units. Capital costs are then divided by unit capacity to generate average scrubbing capital costs. As shown in Table 2.3, the average capital costs for Phase I units were around 276.36$/KW, and for Phase II units, they decreased to 163.14$/KW due to technical progress. The scrubbing capital costs for non-scrubbing units are not observable in the data, and will be econometrically estimated as a function of a number of unit characteristics. The estimation process will be discussed more in the next subsection.

Sulfur content. FERC Form 423 provides plant level sulfur content data for each type of coal. I use a sulfur content of 0.6 pounds per million Btu as a threshold to distinguish low-sulfur coal from high-sulfur coal. Coal with a sulfur content less than 0.6 lbs/mmbtu is defined as low-sulfur coal, and otherwise high-sulfur coal. This threshold meets the average Phase II emission standards and is used in the related literature (Carlson et al., 2000). Based on the decision period assumption, I use the average sulfur contents of high-sulfur coal and low-sulfur coal over 1990 to 1995 for Phase I units and 1996 to 2000 for Phase II units in the discrete step estimation.[13] For units with missing plant level sulfur content data, I use state level data. In general, units choosing to scrub used to burn coal with higher sulfur contents, and the average sulfur content of high-sulfur coal burned in Phase I units was higher than that in Phase II units (Table 2.3).

Coal price. Plant level coal prices are also drawn from FERC Form 423. Similar to the sulfur content data, I use the average delivery contract coal prices between 1990 and 1995 for Phase I units and between 1996 and 2000 for Phase II units in the discrete step estimation. All prices are converted to 1996 dollars using the producer price index for intermediate materials, supplies, and components (U.S. Bureau of Labor Statistics). Following Carlson et al. (2000), for units burning no low-sulfur coal (e.g., scrubbing units) or with missing low-sulfur coal data, I approximate the price as the product of the plant's observed high-sulfur coal price and the ratio of low- to high-sulfur coal prices in the state in which the plant is located. This approximation also applies to units with missing high-sulfur coal prices. As shown in Table 2.3, units choosing to scrub usually have higher price premiums for low-sulfur coal, and both the high-sulfur coal and the low-sulfur coal prices for Phase II units are generally lower than those for Phase I units.

Allowance price. In addition to annual allocations, EPA auctions approximately 2.8% of the total annual allowances once a year. These auctions provide important information allowing power plants to forecast allowance prices. I use the average spot auction prices for 1995 allowances in 1993 and 1994 as the allowance price for Phase I units, which is 141.35$/ton in 1996 dollars. For Phase II units, I use the average seven-year advance auction prices for 2000 in 1993 and 1994 as the allowance price, which is 127.43$/ton in 1996 dollars.

Baseline unit characteristics. I use 1985-1987 as the baseline period because an individual unit's initial SO_2 allowances were generally allocated based on the unit's average heat input during this period. I collect each unit's baseline information (1985-1987 average) from EIA 767 data, including sulfur content of coal burned in the unit, heat input, boiler capacity, and

[13] I extend the time period to 1995 for Phase I units, because some Phase I units started to burn low-sulfur coal in 1995 to comply with the ARP, and for these units no sulfur content data exist for low-sulfur coal during 1990-1994. Similarly, I extend the time period to 2000 for Phase II units.

electricity generation. I then divide heat input by electricity generation to generate the heat rate for each unit, which measures the efficiency of the unit. Table 2.4 provides summary statistics of unit characteristics by compliance strategy and ARP Phase. In general, scrubbing units have larger capacities, lower heat rates, and are younger than non-scrubbing units.

Variables for the conditional emission rate equation. The Continuous Emissions Monitoring System (CEMS) provides unit level annual SO_2 emission data measured in tons of SO_2 emitted, which is then divided by the unit heat input to get the emission rate measured in pounds of SO_2 per mmbtu. One of the important independent variables is the allowance-price-included price premium of low-sulfur coal for non-scrubbing units $(ln(w_{n0}^{l}) - ln(w_{n0}^{h}))$, where $ln(w_{n0}^{h})$ and $ln(w_{n0}^{l})$ are expressed in equation (2.10) and (2.11) and the variables for calculating these two prices are explained above. Electricity generation is obtained from EIA-767. I also include baseline unit characteristics in the equation, such as unit capacity, sulfur content of coal burned in the unit, unit age, and heat rate. Conditional on the choice of scrubbing or non-scrubbing, Phase I units started to comply with the ARP from 1995, and Phase II units from 2000. Therefore, I use data from 1995 to 2003 for Phase I units and data from 2000 to 2003 for Phase II units as an unbalanced panel data set to estimate the conditional emission rate equation. Table 2.5 provides summary statistics of these variables. The price premium is negative because the allowance price is included as expressed in equation (2.10) and (2.11).

2.4.2 Analysis Procedures and Identification

The econometric objective of the essay is to estimate equation (2.14), the probability of a unit's non-scrubbing decision, and equation (2.16), the conditional emission rate if the unit chooses non-scrubbing. I conduct the analysis in the following steps.

2.4.2.1 Estimate the Scrubbing Capital Cost for Non-Scrubbing Units

For non-scrubbing units, the scrubbing capital cost $ln(K_{n1})$ is not observable in the data, so I cannot directly estimate equation (2.14) without first generating estimates of $ln(K_{n1})$ for non-scrubbing units. Following Keohane (2002), I start by econometrically estimating $ln(K_{n1})$ as a function of L unit characteristics as

$$ln(K_{n1}) = \beta_0 + \sum_{l=1}^{L} \beta_l \cdot ln(X_{nl}) + \varepsilon_n \tag{2.17}$$

where X_{nl} represents a vector of unit characteristics, and $\varepsilon_n \sim N(0, \sigma_\varepsilon^2)$. Guided by engineering models (EPA IPM 2010; EPA IAPCS 1999), I choose baseline unit nameplate capacity, coal sulfur content, and heat rate as X_{nl} in equation (2.17).

Equation (2.17) can be estimated using OLS. However, if units with lower scrubbing costs are more likely to choose to scrub, OLS estimates will be biased. Therefore, I use Heckman's two-step method to correct for this selection bias in estimating the scrubbing capital cost equation (2.17). A unit chooses to install a scrubber and therefore its scrubbing capital cost is observed, if the anticipated total generating cost with a scrubber is smaller than the cost with coal switching/blending under the impacts of certain regulation factors, that is,

$$C_{n0} > C_{n1} \cdot exp\{\eta \cdot R_n\} \tag{2.18}$$

which is the reverse of equation (2.13).

The valid excluded variables used for identifying the coefficients in the scrubbing capital cost equation (2.17) are the prices of high-sulfur coal and low-sulfur coal, which significantly affect the selection of scrubbing, but do not directly affect the scrubbing capital cost. More estimation details are provided in Appendix A.

2.4.2.2 Estimate Units' Discrete Compliance Choice Between Scrubbing and Non-Scrubbing

Using the predicted scrubbing capital costs $ln(\widehat{K_{n1}})$ in the previous step for all units and other observed variables, I then estimate the probit model (2.14) for units' discrete compliance choice between scrubbing and non-scrubbing.

2.4.2.3 Calculate/Estimate Units' Continuous Decisions on Emission Rate

If a unit chooses scrubbing, the emission rate is simply the sulfur content of the high-sulfur coal burnt in the unit multiplied by the residual of sulfur after scrubbing, as expressed in equation (2.15).

If a unit chooses blending with or switching to low-sulfur coal rather than scrubbing, the unit's emission rate is estimated using equation (2.16). In the discrete-continuous choice model, the error term ζ_{n0} in the continuous emission rate equation is assumed to depend in general on the discrete choice $j = 0$, that is, ζ_{n0} is correlated with the error term ξ_{n0} in the non-scrubbing cost equation. In fact, this discrete-continuous choice model can be considered as a sample selection model (Newey, 2007; Newell and Pizer, 2008). Therefore, to estimate the emission rate equation (2.16), I first generate the inverse Mills ratio of the probit model in the previous discrete step, and then estimate the emission rate equation with the inverse Mills ratio added on the right-hand side to account for the selection bias. The variables excluded for identifying the coefficients in the continuous emission rate equation are the characteristics of scrubbing, including the scrubbing capital costs and scrubbing efficiency, which significantly affect the choice of abatement strategies, but do not directly affect the emission rates associated with coal-switching/coal-blending.

Using an unbalanced panel data set with data from 1995 to 2003 for Phase I units and from 2000 to 2003 for Phase II units, I run a random-effects model with selection correction to estimate the emission rate equation. I cannot run a fixed-effects model because the selection correction term derived from the discrete step is constant within unit.

Additionally, the price premium of low-sulfur coal may be endogenous. For example, coal suppliers might charge a higher price for low-sulfur coal when the demand for low-sulfur coal from non-scrubbing units increases to lower emission rates and comply with the requirement of the ARP. Therefore, I use the distance to the Powder River Basin (PRB), the major low-sulfur coal production region in the U.S., to instrument the price premium.

2.4.2.4 Simulate Units' Compliance Choices and Emissions Under the Implemented and Counterfactual Policy Designs

Based on the estimated results of the discrete-continuous model, I simulate units' compliance choices with different SO_2 prices and then calculate the corresponding aggregate SO_2 emissions using the following equation

$$Aggregate\ Emission = \sum_{n=1}^{N} (H_n \cdot e_{n0} \cdot Pr_{n0} + H_n \cdot e_{n1} \cdot (1 - Pr_{n0})) \tag{2.19}$$

where H_n is the heat input of unit n , e_{n0} is the emission rate if the unit chooses non-scrubbing and e_{n1} the emission rate if the unit chooses scrubbing, and Pr_{n0} is the probability of non-scrubbing and $1 - Pr_{n0}$ the probability of scrubbing. Given a price of SO2 emission allowance τ (\$/ton), Pr_{n0} can be estimated in the first discrete choice step, e_{n0} can be estimated in the second continuous step, and e_{n1} is directly calculated using the sulfur content of high-sulfur coal and the residual sulfur after scrubbing, as presented in equation (2.15). Repeating these steps for a range of SO_2 allowance prices gives the marginal abatement cost curve, which is used to evaluate the compliance costs associated with different policy designs. For the observed nationwide uniform trading scheme under the ARP, the marginal abatement cost curve is estimated by aggregating the emissions of all units, and for the counterfactual policy design with only interstate emissions trading, the marginal abatement cost curves are estimated state by state by aggregating emissions of only units within the same state.

2.5 Estimation Results

2.5.1 Estimates of the Scrubbing Capital Cost

I first estimate the scrubbing capital cost equation (2.17) using Heckman's two-step method. Table 2.6 presents the estimated scrubbing capital costs from the second-step OLS regression with selection bias correction, as well as the results estimated using a simple OLS regression. The Heckman estimation results show that the coefficient of *ln*(*capacity*) is negative and significant at the 1% level, and the coefficient of *ln*(*heat rate*) is positive and significant at the 5% level. Therefore, the average scrubbing capital cost decreases with unit capacity, and increases with unit heat rate. The coefficient for period 2 is negative and significant at the 5% level, implying that units covered in Phase II of the ARP usually have lower average scrubbing costs. The coefficient on the selection correction term is negative and significant at the 5% level, which confirms that units choosing to scrub have systematically lower scrubbing costs than non-scrubbing units.[14]

2.5.2 Estimates of the Non-Scrubbing Discrete Choice Model

Using predicted scrubbing capital costs, I now estimate the objective equation (2.14), the probit model of units' non-scrubbing discrete choice. The results are presented in Table 2.7. The

[14] Note that 10 of 64 scrubbing units reported extremely low scrubbing capital costs. If I exclude these 10 observations, the coefficient on the selection correction term remains negative but is not significant.

coefficient of $ln(K_{n1})$ is positive and significantly different from zero at the 1% level, suggesting that units with higher scrubbing capital costs are less likely to install scrubbers. The coefficient of $-ln(w_{n0}^{l})$ is positive and significantly different from zero at the 1% level, suggesting that units facing lower low-sulfur coal prices tend to choose non-scrubbing. The coefficients for boiler age and the ARP Phase II are both positive and significantly different from zero at the 1% level. This shows that older units and Phase II units are more likely to choose non-scrubbing.

I use the estimated probit model to predict unit compliance choices. Table 2.8 compares predicted choices with observed choices to test the performance of the model. A unit is predicted to switch to/blend with low-sulfur coal if the estimated probability of non-scrubbing is greater than 0.5 or 0.68. The latter cutoff value, 0.68, is the one that results in the same number of predicted choices as actual choices in the data, as used in Keohane (2002). Using the latter cutoff, the model correctly predicts 43 out of 64 (67.2%) scrubbing units, and 787 out of 808 (97.4%) non-scrubbing units.

2.5.3 Estimates of the Continuous Emission Rate Equation Conditional On Non-Scrubbing

Given the results of the compliance choice probit model, I then estimate the continuous emission rate with selection bias correction conditional on choosing non-scrubbing. The estimation results of this step are presented in Table 2.9.

The sign of $ln(price\ premium)$ is expected to be positive since units facing more expensive low-sulfur coal tend to burn less low-sulfur coal and thus to emit more SO_2. However, the simple OLS and regular Heckman estimation results give negative coefficients for the price premium. One possible reason for this counter-intuitive sign is the endogeneity of the price premium of low-sulfur coal discussed above. So I use distance to the Powder River Basin (PRB) to instrument the price premium. The results of the Heckman IV estimation are shown in the third column of Table 2.9, and more details on the IV estimation are provided in Appendix B. With the instrument used, the coefficient of $ln(price\ premium)$ is significantly positive. Also, the emission rate increases with electricity generation and the original sulfur content, and decreases with unit capacity. The coefficient on the selection correction term is not significant, which might indicate that the selection model has already included the main drivers of choosing non-scrubbing. Another possible reason is that I estimate emission rate rather than direct demand of low-sulfur coal, so there might be other factors offset the selection effects, such as the sulfur content of high-sulfur coal.

2.6 Simulation

2.6.1 Marginal Abatement Costs under the ARP

Based on the estimated results of the discrete-continuous model, I first simulate the compliance choices and aggregate SO_2 emissions of the sample units associated with different SO_2 prices under the ARP when unlimited interstate trading was allowed. Figure 2.5 depicts the simulation results, with the horizontal axis representing aggregate SO_2 emissions (1,000tons) and the

vertical axis the price of SO_2 allowances ($/ton). The two vertical lines in Figure 2.5 represent the imposed emissions caps, which are set equal to the annual SO_2 allowances allocated to the regulated units under the ARP in 1995 and 2000, respectively. The simulated equilibrium allowance prices associated with the 1995 emissions cap (5353.95 thousand tons) and the 2000 emissions cap (6414.17 thousand tons) are around 228 $/ton and 510 $/ton, respectively, which are higher than the actual market equilibrium prices under the ARP in 1995 (131 $/ton) and 2000 (123 $/ton). One reason is that the model underestimates the probability of units' choosing scrubbing. Another important reason for the relatively high simulated allowance price in 2000 is that banking is not considered. If the imposed cap in 2000 is set equal to the actual emissions in 2000 (9186.79 thousand tons), the simulated allowance price is around 189 $/ton, much closer to the actual market price.

Figure 2.6 shows the corresponding marginal abatement cost curves, where the horizontal axis is aggregate SO_2 emission abatements (1000tons), the vertical axis measures marginal abatement costs ($/ton), and the two vertical lines represent the SO_2 emissions abatements associated with the SO_2 emissions caps in 1995 and 2000, respectively. Integrating under the MAC curves and left to the abatement lines, I obtain the simulated abatement costs for sample units under the ARP, summarized in Table 2.10. For Phase I units to achieve the ARP cap in 1995, the total compliance costs are around $\$107.46\times10^6$. For all sample units to achieve the ARP cap in 2000, the total compliance costs are around $\$1136.43\times10^6$.

Another point of interest is the effect of coal price changes on EGUs' SO_2 emission reductions. Figure 2.7 shows that low-sulfur coal price premium decreased greatly due to railroad deregulation in 1980 (Busse and Keohane, 2007). In Figure 2.5 and Figure 2.6, I use the average coal price between 1991 and 1995 for Phase I units and between 1996 and 2000 for Phase II units to simulate aggregate emissions and abatements. To determine the effect of coal price changes, I also simulate aggregate SO_2 emissions using average coal prices during the baseline period (1985 to 1987), as shown in Figure 2.8. For Phase I units, the effect of the coal price decrease, represented by the vertical distance between the two curves in the left panel of Figure 2.8, accounts for more than 300 $/ton of the fall in the marginal abatement costs. For Phase II units, the effect of the coal price decrease, represented by the vertical distance between the two curves in the right panel of Figure 2.8, accounts for more than 600 $/ton of the fall in the marginal abatement costs.

2.6.2 Simulation Results under the Counterfactual Policy Design

If the ARP had been designed to prohibit any interstate trading, while the state-specific caps were set equal to the sum of initial SO_2 allowances allocated to the units within the state under the ARP, the total costs to achieve the 1995 cap and the 2000 cap would have increased to $\$533.62\times10^6$ and $\$1772.54\times10^6$, respectively.[15] The market clearing SO_2 allowance prices range from 0 $/ton to 3618 $/ton in 1995, and from 0 $/ton to 1482 $/ton in 2000. Table 2.10 summarizes the simulation results. The narrower markets and less trading opportunities under the

[15] If the state-specific caps had been redesigned based on the ex-ant cost-benefit analysis of the scenario with no interstate trading, the cost increase associated with trade restrictions would be lower. Therefore, the simulated results in this essay could be interpreted as upper bounds of the cost increase.

trade restrictions would be one of the main reasons leading to the cost increase. Another reason that might drive up compliance costs but is not included in the current simulation is that large units are more likely to exert market power in a relatively smaller market with fewer emission sources.

Furthermore, as summarized in Table 2.11 and Table 2.12, the compliance costs vary dramatically across states due to the difference in cap stringency and heterogeneous marginal abatement costs. In some states, the state-specific caps are not binding and therefore the SO_2 allowance prices are zero (Alabama, Georgia, Kansas, Pennsylvania, and Tennessee in 1995; Arizona, Colorado, Delaware, Kansas, Michigan, Nebraska, Nevada, Oklahoma, Utah, and Washington in 2000), while in some states the extreme stringent state-specific caps result in the SO_2 prices higher than 1000 $/ton (Michigan and Wisconsin in 1995; Alabama, Georgia, New Hampshire, New Mexico, South Carolina, South Dakota, and Virginia in 2000).

2.7 Conclusions

To address cross-state air pollution problems, one important way is to modify emissions trading programs by imposing geographical trade restrictions. Such trade restrictions would affect the compliance costs of regulated firms. The empirical question asked in this essay is how costly it would be if a single, nation-wide emissions trading market were geographically separated, and emissions trading were unlimited within a state but prohibited across states. To answer this question, I first establish a discrete-continuous model to estimate electric generating units' compliance strategies and marginal abatement costs associated with the nationwide uniform SO_2 emissions trading under the Acid Rain Program (ARP). Based on the estimation results, I then simulate units' compliance behavior and the corresponding compliance costs if interstate trading had been prohibited.

Under the nationwide, uniform SO2 allowance trading market established by the ARP, the ex post estimated aggregate compliance cost is around $\$107.46\times10^6$ for achieving the 1995 cap and $\$1136.43\times10^6$ for the 2000 cap, and the market clearing allowance prices are 228 $/ton and 510 $/ton in 1995 and 2000, respectively.

The simulation results show substantive cost increase in the counterfactual policy design, and the cost increase is very unevenly distributed across states. If only intrastate trading is allowed and interstate trading is completely forbidden, the aggregate compliance cost will increase to $\$533.62\times10^6$ in 1995 and $\$1772.54\times10^6$ in 2000. For some states with extremely tight caps, allowance prices are higher than 1000$/ton in this scenario. For some states, the state specific emissions caps are not binding, leading to 0$/ton allowance prices. These results, combined with the analysis on the benefit side, could be used to predict welfare impacts associated with trade restrictions at both national and state level. And it also may shed light on the future modification and implementation of EPA's cross-state air pollution regulations.

Finally, EGUs' compliance choice between scrubbing and non-scrubbing is modeled as a static decision process in this essay. Also, EGUs' allowance banking decision is not included in the current model. To better understand EGUs' compliance decisions, a more complicated dynamic model is needed.

2.8 Tables and Figures

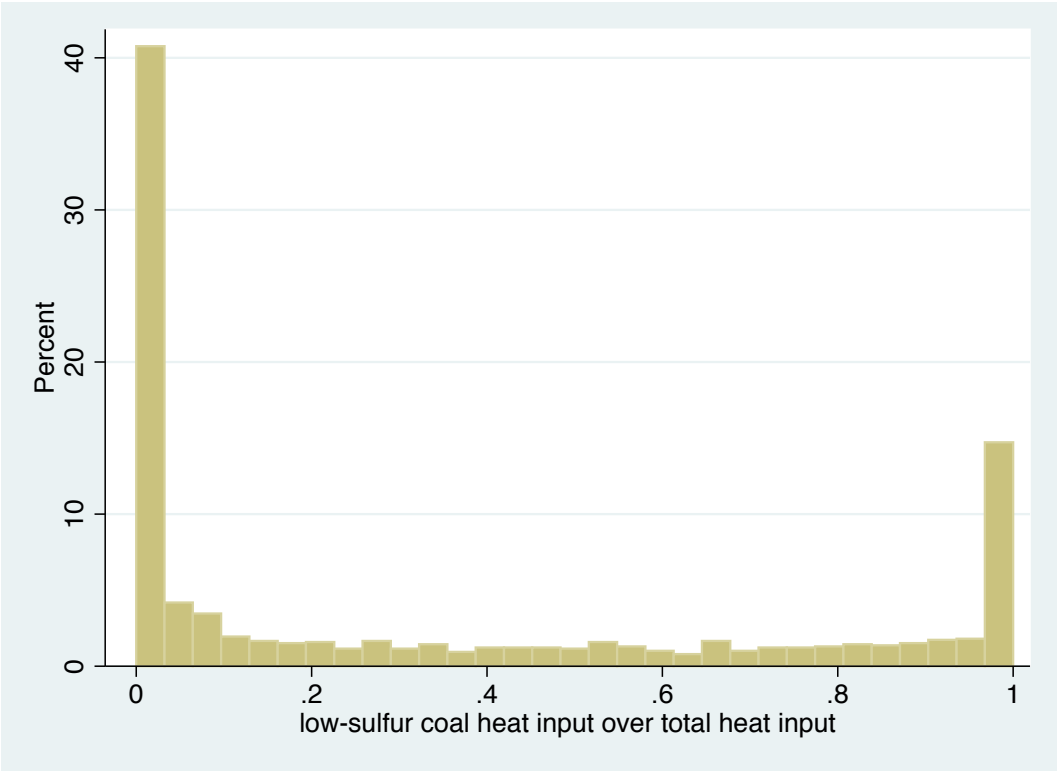

Figure 2.1: Histogram of Heat Input Ratio of Low-Sulfur Coal in All Non-Scrubbing Plants

Notes: Figure generated using data of all non-scrubbing Phase I units after 1995 and all non-scrubbing Phase II units after 2000. Heat input ratio of low-sulfur coal is calculated by dividing plant level low-sulfur coal heat input by total heat input. The heat input ratio of low-sulfur coal in more than 80% of non-scrubbing power plants is smaller than 1, implying that most non-scrubbing power plants chose partially blending with low-sulfur coal to comply with the ARP.

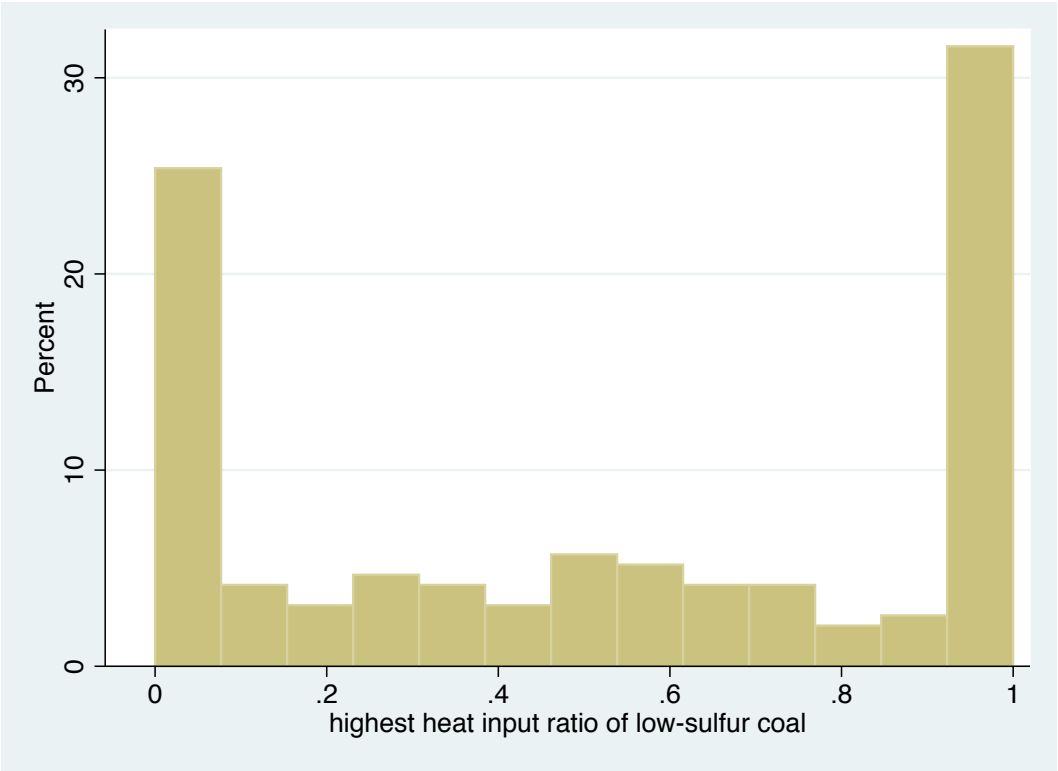

Figure 2.2: Histogram of the Highest Heat Input Ratio of low-Sulfur Coal in All Non-Scrubbing Plants

Notes: Figure generated using the highest low-sulfur coal heat input ratio over 1995-2003 for all non-scrubbing Phase I units and over 2000-2003 for all non-scrubbing Phase II units. More than 30% of non-scrubbing units could reach one. This, to some extent, demonstrates that the partially blending decision is not determined by technological constraints but affected by variables such as the price premium of low-sulfur coal, unit capacity, unit heat rate, and so on.

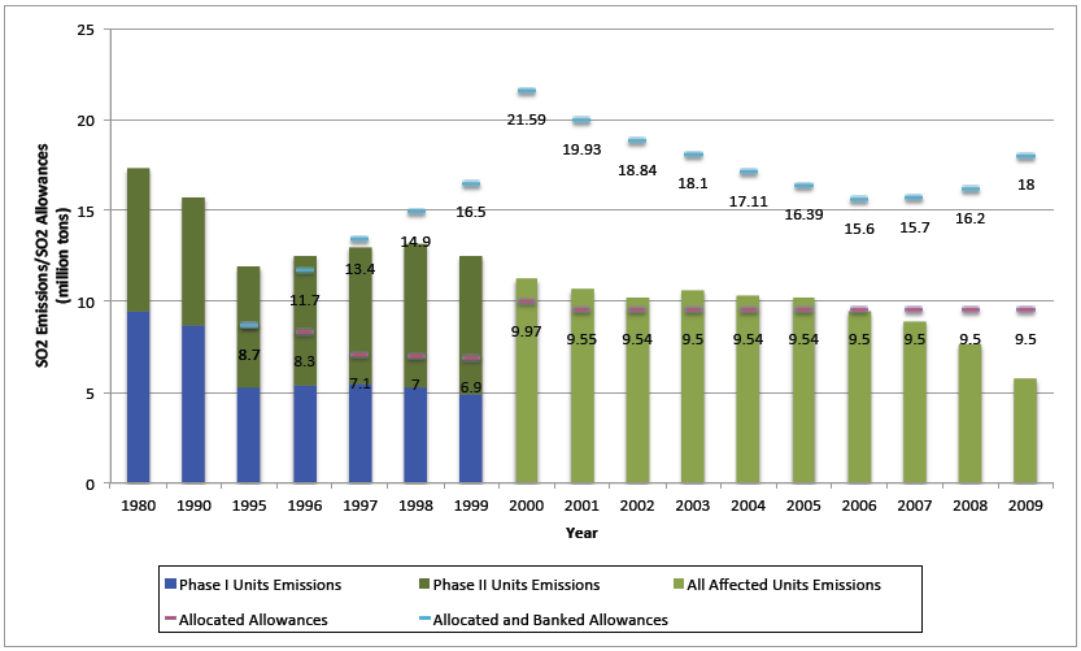

Figure 2.3: SO2 Allowances for and Emissions from ARP Units

Notes: Figure generated using data of all units regulated by the Acid Rain Program from 1980 to 2009. Data source: EPA Acid Rain Program Progress Reports 1995 to 2009.

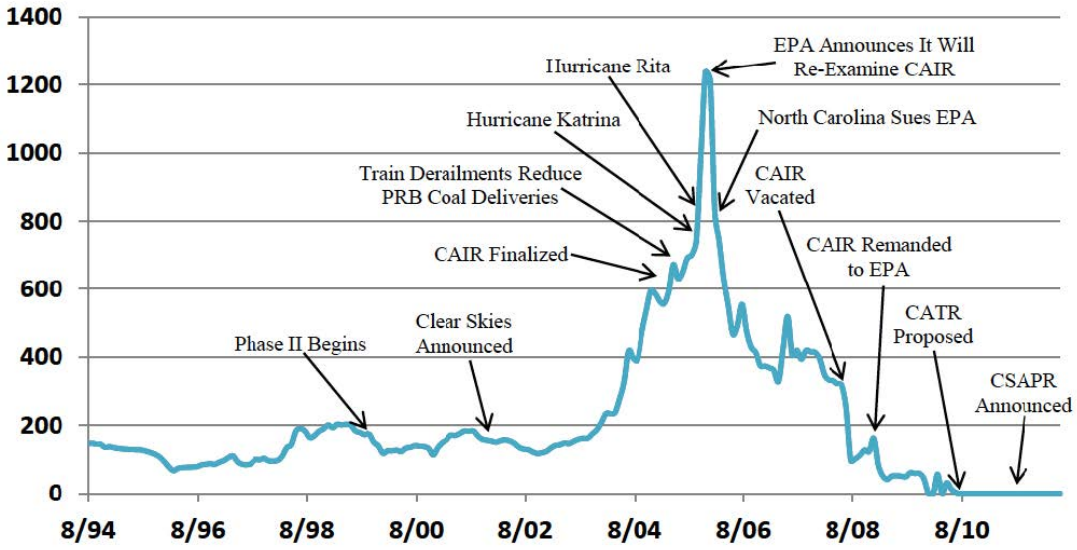

Figure 2.4: SO2 Allowance Prices and the Regulatory Environment, 1994-2012 (1995$ per ton)

Notes: Reprinted of Figure 2 from "The SO_2 Allowance Trading System: The Ironic History ", by Richard Schmalensee and Robert N. Stavins, 2013, *Journal of Economic Perspectives*. Copyright 2013 by American Economic Association. Reprinted with permission.

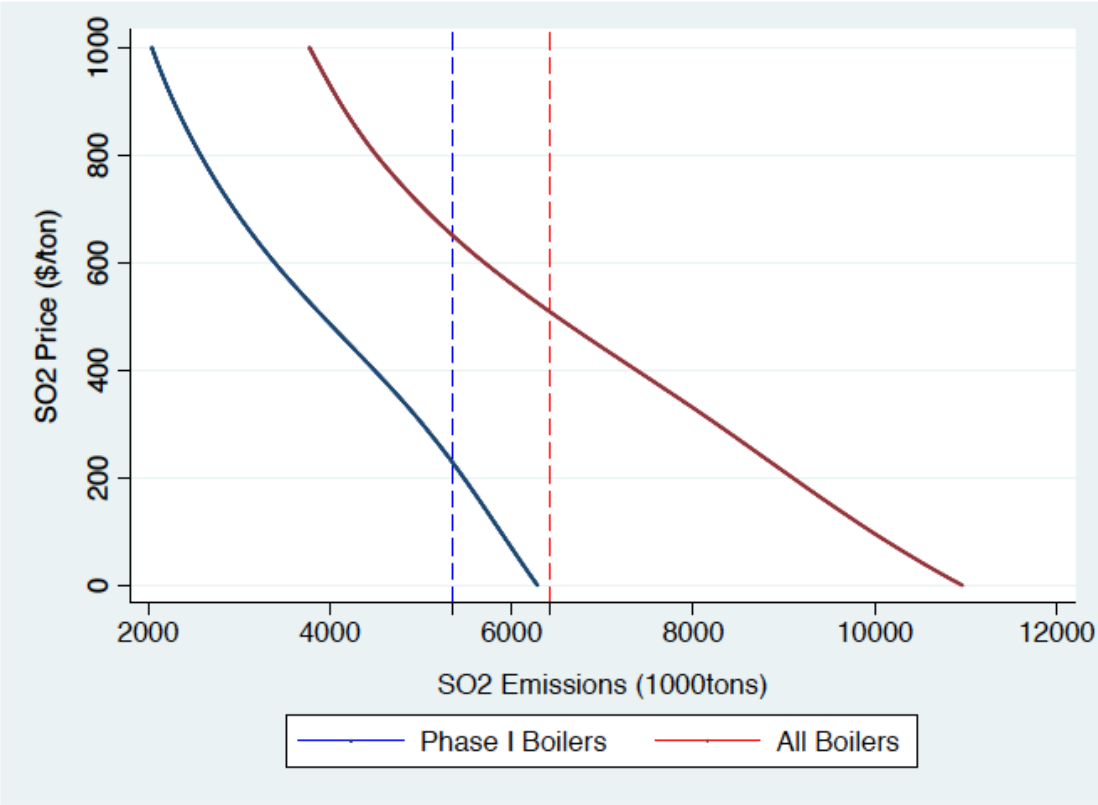

Figure 2.5: Simulated Aggregate SO2 Emissions of Sample Boilers under the ARP

Notes: This figure depicts the relationship between SO_2 allowance price and aggregate SO_2 emissions of the sample units under the ARP when unlimited interstate trading was allowed, by simulating units compliance choices using the estimates of the discrete-continuous model. The blue and red dashed vertical lines represent the imposed caps, which are set equal to the annual SO_2 allowances allocated to the regulated units under the ARP in 1995 (5353.95 thousand tons) and 2000 (6414.17 thousand tons), respectively.

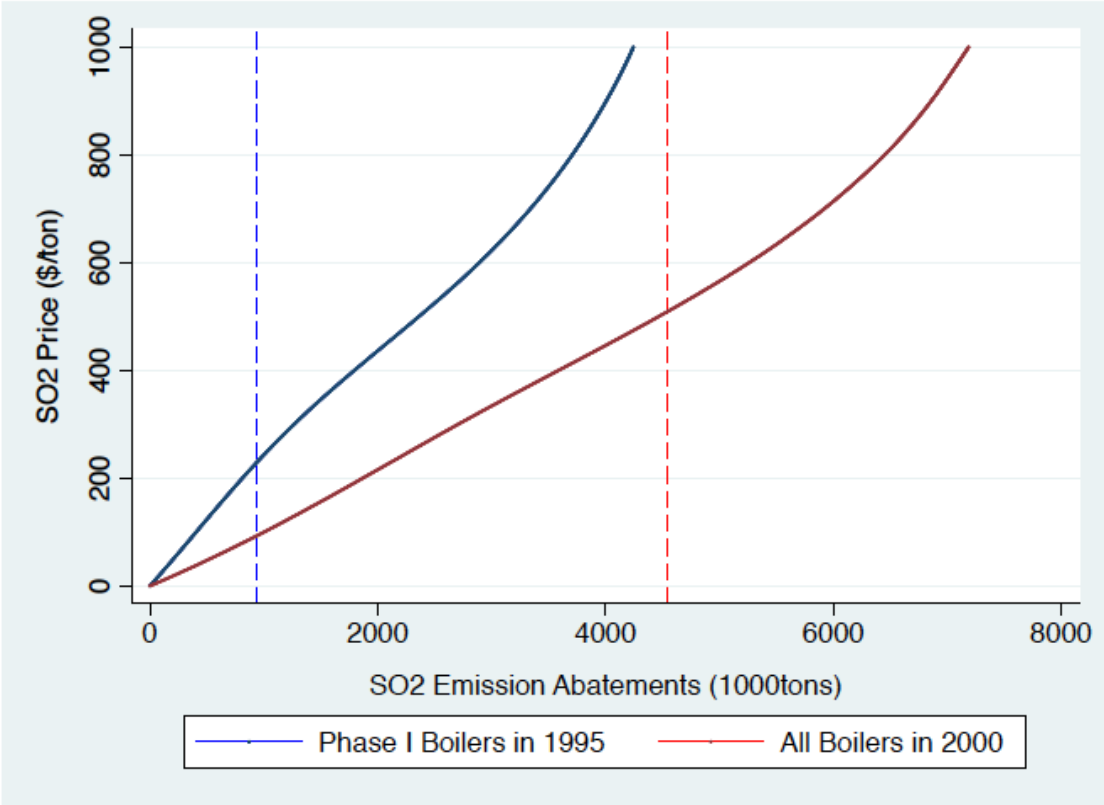

Figure 2.6: Simulated Marginal Abatement Cost Curves of Sample Boilers under the ARP

Notes: This figure shows the corresponding marginal cost curves obtained from Figure 2.5. The blue and red dashed vertical lines represent the SO_2 emissions abatements associated with the SO_2 emissions caps in 1995 and 2000, respectively. Integrating under the marginal abatement curves and left to the abatement lines gives the simulated abatement costs for sample units under the ARP.

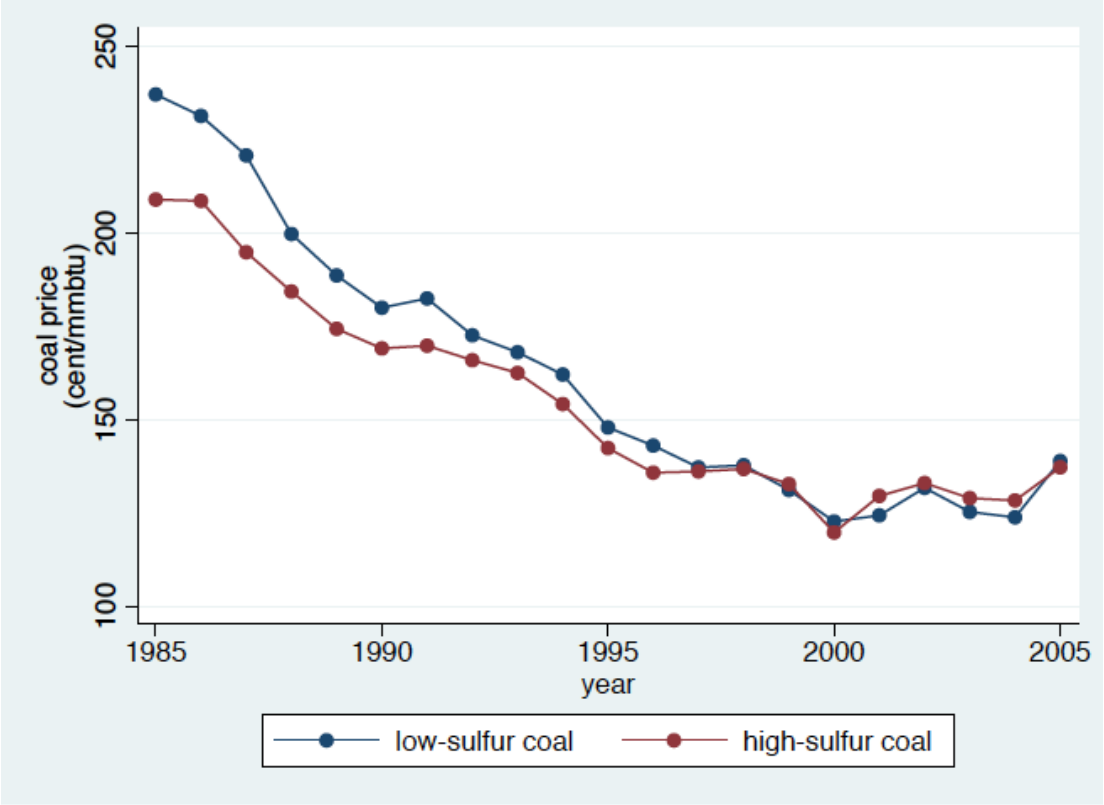

Figure 2.7: Contract Prices of Low-Sulfur Coal and High-Sulfur Coal

Notes: All dollar values are in 1996 constant dollars.

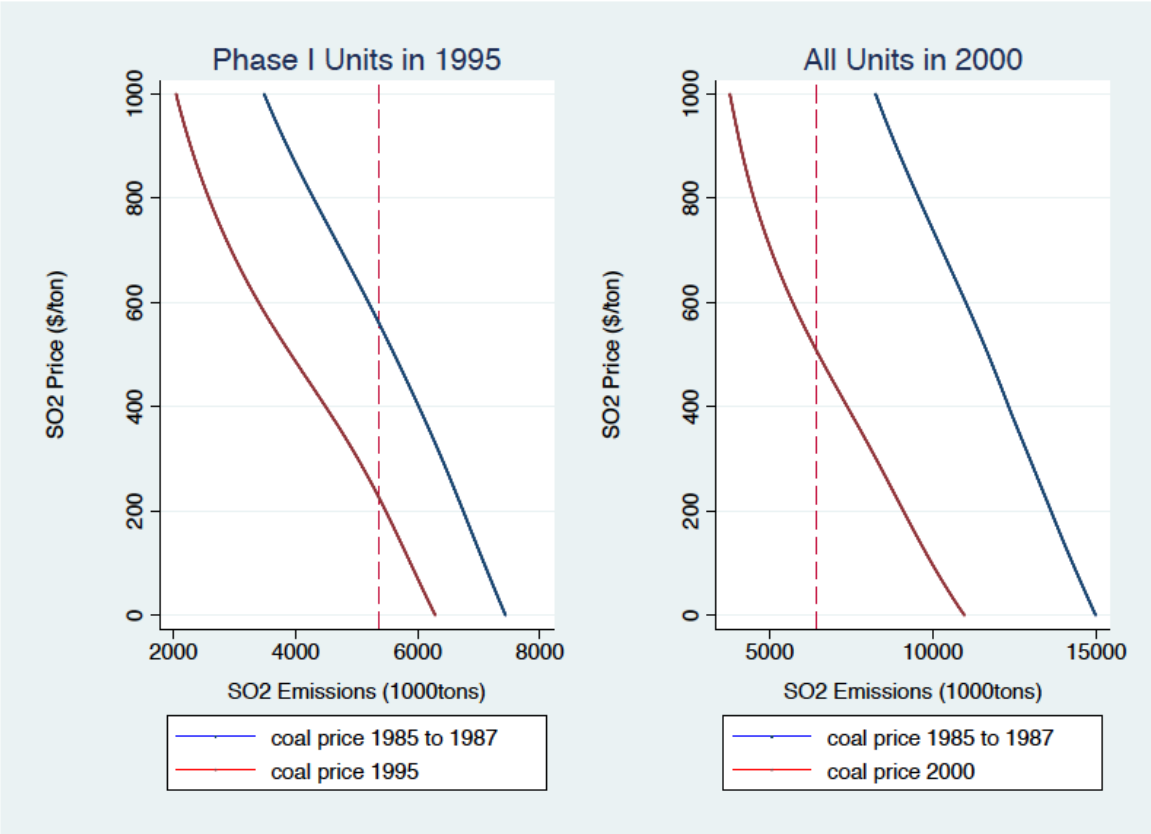

Figure 2.8: The Effect of the Coal Price Decrease on SO2 Emissions

Notes: The left panel illustrates Phase I units' compliance costs in 1995. The red solid curve is simulated using the estimates of the discrete-continuous model and average coal contract price between 1991 and 1995, the blue solid curve is simulated using the same model estimates but average coal contract price during the baseline period 1985-1987, and the dashed vertical line represents the 1995 cap. The right panel replicates the graphical exercise in the left panel for all units in 2000. The average coal contract price used to simulate the red solid curve in the right panel is over 1991-995 for Phase I units and over 1996-2000 for Phase II units, to simulate the blue solid curve is over the baseline period 1985-1987. And the dashed vertical line represents the 2000 cap.

Table 2.1: Number of Sample Units by Compliance Strategy and ARP Phase

Phase	Strategy	No. of units	No. of plants	No. of utilities
I	Non-scrubbing	341	127	63
	Scrubbing	39	19	19
II	Non-scrubbing	467	191	105
	Scrubbing	25	12	10

Table 2.2: Number of Units Starting to Scrub by Year

Year	Number of Units	
	Phase I	Phase II
1992	3	0
1993	1	1
1994	3	1
1995	23	0
1996	4	0
1997	0	1
1998	0	8
1999	3	3
2000	0	0
2001	0	4
2002	2	3
2003	0	4*
Total Scrub Units	39	25
Total Units	380	492

Notes: *These 4 scrubbers were installed no later than 1999 according to the costs reported in the EIA-767 form. 34 of 39 Phase I scrubbing units started to scrub between 1992 and 1996, and 20 of 25 Phase II scrubbing units started to scrub between 1997 and 2001.

Table 2.3: Summary Statistics of Cost Related Characteristics by Compliance Strategy and Phase

Phase	Strategy	Scrubbing Capital Cost ($/KW)	Sulfur Content (lb/mmbtu)		Price (cent/btu)	
			high sulfur coal	low sulfur coal	high sulfur coal	low sulfur coal
I	Non-scrubbing	/	1.73	0.47	159.43	164.59
			(0.69)	(0.08)	(29.86)	(36.41)
	Scrubbing	276.36	2.06	0.49	147.82	164.34
		(182.96)	(0.52)	(0.06)	(29.01)	(26.39)
II	Non-scrubbing	/	1.3	0.46	134.28	138.80
			(0.60)	(0.10)	(35.52)	(36.47)
	Scrubbing	163.14	1.16	0.45	125.57	127.47
		(126.88)	(0.64)	(0.08)	(33.35)	(27.74)

Notes: Standard deviations in parentheses. Summary statistics generated using 1990-1995 data for the 380 Phase I units and 1996 to 2000 data for the 492 Phase II units used to estimate the discrete choice model. All dollar values are in 1996 constant dollars.

Table 2.4: Unit Characteristics During the Baseline Period

Phase	Strategy	Capacity (MW)	Age by 1990	Heat Rate (btu/KWh)
I	Non-scrubbing	261.27	29.67	11889.02
		(227.32)	(9.16)	(6943.97)
	Scrubbing	494.33	22.18	10359.15
		(362.57)	(9.69)	(1181.43)
II	Non-scrubbing	261.64	27.70	11805.66
		(248.47)	(11.03)	(6490.05)
	Scrubbing	348.52	20.48	10489.92
		(272.64)	(7.89)	(2356.57)

Notes: Standard deviations in parentheses. Summary statistics generated using baseline 1985-1987 data for the 872 sample units.

Table 2.5: Summary Statistics of Variables for the Emission Rate Equation

Phase	Strategy	Emission Rate (lb SO2/mmbtu)	Price Premium (cents/mmbtu)	Electricity Generation ($\text{MWh}\times10^4$)
I	Non-scrubbing	2.01	-9.96	146.96
		(1.26)	(28.22)	(134.20)
	Scrubbing	0.83	-6.34	305.24
		(1.14)	(24.10)	(246.41)
II	Non-scrubbing	1.30	-5.99	162.44
		(0.75)	(40.77)	(164.44)
	Scrubbing	0.57	-7.68	243.71
		(0.89)	(42.79)	(195.84)

Notes: Standard deviations in parentheses. Summary statistics generated using 1995-2003 data for the 380 Phase I units and 2000 to 2003 data for the 492 Phase II units used to estimate the continuous emission rate equation. All dollar values are in 1996 constant dollars.

Table 2.6: Estimates of Scrubbing Capital Costs

Dependent Variable: $ln(K_{n1})$

Explanatory Variable	(1) OLS	(2) Heckman
$ln(capacity)$	-0.452***	-0.687***
	(0.166)	(0.191)
$ln(sulfur\ content)$	-0.168	-0.420
	(0.289)	(0.297)
$ln(heat\ rate)$	1.402**	1.330**
	(0.558)	(0.534)
period 2	-0.521	-0.930**
	(0.401)	(0.423)
extreme low	-4.981***	-5.800***
	(0.390)	(0.540)
cons	-4.516	-1.716
	(5.591)	(5.479)
Mills lambda		-0.480**
		(0.224)
R2	0.796	

Notes: The dependent variable is log of scrubbing capital cost. Results presented in column (1) are estimated using 64 scrubbing units and OLS. Results presented in column (2) are estimated using all 872 units and from the second-step OLS regression with selection bias correction of Heckman's two-step method. Standard errors are in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 2.7: Estimates of Non-Scrubbing Choice Using Predict Scrubbing Capital Costs

Choice = non-scrubbing	
$ln(K_{n1})$	0.725*** (0.164)
$ln(w_{n1}^{h}) - ln(w_{n0}^{h})$	-28.387 (64.017)
$-ln(w_{n0}^{l})$	74.915*** (25.203)
$(ln(w_{n1}^{h}))^2 - (ln(w_{n0}^{h}))^2$	5.275 (6.902)
$-(ln(w_{n0}^{l}))^2$	-7.055*** (2.391)
$-ln(w_{n0}^{h})ln(w_{n0}^{l})$	-0.515 (0.314)
$ln(y_n) \cdot (ln(w_{n1}^{h}) - ln(w_{n0}^{h}))$	-2.559 (3.182)
$-ln(y_n) \cdot ln(w_{n0}^{h})$	-0.044 (0.050)
$age\ by 90$	0.062*** (0.019)
$PROTECT_n$	0.034 (0.268)
period 2	1.702*** (0.376)
NERC region dummies	Yes
cons	179.956*** (60.663)

Notes: This table presents results of estimating the probit model of units' discrete choice between scrubbing and non-scrubbing. The dependent variable is 1 if the unit chooses non-scrubbing. Robust standard errors are in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 2.8: Predicted vs. Observed Compliance Choices

Choice	Scrubbing	Non-Scrubbing	Total
		Cutoff = 0.5	
Observed	64	808	872
Predicted	29	843	872
Correctly Predicted	24	803	872
	(37.5%)	(99.4%)	(94.8%)
		Cutoff = 0.68	
Observed	64	808	872
Predicted	64	808	872
Correctly Predicted	43	787	830
	(67.2%)	(97.4%)	(95.2%)

Notes: This table compares predicted choices from the probit model estimation with observed choices to test the performance of the model. A unit is predicted to switch to/blend with low-sulfur coal if the estimated probability of non-scrubbing is greater than 0.5 or 0.68. The latter cutoff value, 0.68, is the one that results in the same number of predicted choices as actual choices in the data.

Table 2.9: Estimates of Emission Rate with Selection Bias Correction

Dependent Variable: Emission Rate

Explanatory Variable	(1) OLS	(2) Heckman	(3) Heckman IV
ln(*price premium*)	-0.226*	-0.285*	3.241***
	(0.132)	(0.168)	(0.983)
ln(*y*)	0.099**	0.119***	0.189***
	(0.039)	(0.043)	(0.042)
ln(*capacity*)	-0.118	-0.356***	-0.585***
	(0.076)	(0.081)	(0.119)
ln(*sulfur content*)	0.840***	0.996***	1.042***
	(0.070)	(0.078)	(0.097)
ln(*heat rate*)	-0.011	0.009	-0.114
	(0.076)	(0.045)	(0.130)
ln(*age by*90)	0.016***	0.003	-0.008
	(0.005)	(0.006)	(0.009)
Year fixed effects	Y	Y	Y
Mills		0.409*	0.287
		(0.211)	(0.356)
cons	1.688*	3.067***	5.562***
	(0.952)	(0.769)	(1.704)

Notes: The dependent variable is emission rate of Phase I non-scrubbing units from 1995 to 2003, and Phase II non-scrubbing units from 2000 to 2003. Results presented in column (1) are estimated of OLS. Results presented in column (2) and (3) are estimated conditional on units' choosing non-scrubbing, and a selection bias correction term (Mills) from the probit model is included. In column (3), log of low-sulfur coal price premium for non-scrubbing units is instrumented using the distance of the plant to the Powder River Basin, and the results are estimated using 2SLS. Unit random effects are included in all three columns. Robust standard errors are in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 2.10: Simulation Results Under Implemented and Counterfactual Policies

		ARP	Counterfactual
1995	Annual Emissions (1000 tons)	5353.95	5353.95
	Allowance Price ($/ton)	228	[0, 3618]
	Compliance Costs ($10^6$$)	107.46	533.62
2000	Annual Emissions (1000 tons)	6414.17	6414.17
	Allowance Price ($/ton)	510	[0, 1276]
	Compliance Costs ($10^6$$)	1136.43	1772.54

Notes: This table summarizes the results from simulating units compliance choices and associated federal level aggregate compliance costs under the observed ARP and the counterfactual policy design for given caps in 1995 and 2000, respectively. Under the ARP, a nationwide SO2 emissions cap was imposed and allowances could be traded between any two units. Under the counterfactual policy design, state-specific caps were set to the sum of all ARP allowances allocated to the units within the state, only intrastate allowance trading was allowed, and interstate trading prohibited.

Table 2.11: Allowance Prices and Compliance Costs by State with Trade Restrictions (1995 Cap)

State	Cap (10^3 tons)	Compliance Costs (10^6 \$)	Allowance Price (\$/ton)
AL	224.86	0.00	0
FL	129.63	2.04	272
GA	526.85	0.00	0
IA	38.29	7.72	484
IL	383.88	10.50	245
IN	677.92	65.43	368
KS	4.11	0.00	0
KY	261.03	19.67	254
MD	135.87	2.90	402
MI	41.23	131.07	3618
MN	4.16	2.30	556
MO	343.70	0.78	58
MS	53.17	3.00	322
NH	31.34	7.46	751
NJ	20.23	2.02	517
NY	56.26	40.18	846
OH	922.16	6.17	163
PA	520.09	0.00	0
TN	376.26	0.00	0
WI	118.14	39.77	1405
WV	484.77	3.03	133

Notes: This table summarizes the results from simulating units compliance choices and associated compliance costs state by state under the counterfactual policy design for given state-specific caps in 1995.

Table 2.12: Allowance Prices and Compliance Costs by State with Trade Restrictions (2000 Cap)

State	Cap (10^3tons)	Compliance Costs (10^6 \$)	Allowance Price (\$/ton)	State	Cap (10^3tons)	Compliance Costs (10^6 \$)	Allowance Price (\$/ton)
AL	264.02	137.12	1251	ND	59.11	1.58	264
AR	96.59	1.75	191	NE	68.37	0.00	0
AZ	83.53	0.00	0	NH	18.14	24.08	1131
CO	78.11	0.00	0	NJ	42.58	2.51	441
DE	34.89	0.00	0	NM	40.79	36.48	1000
FL	208.88	39.87	578	NV	53.22	0.00	0
GA	374.37	168.57	1178	NY	84.55	57.84	815
IA	89.17	32.32	546	OH	603.51	152.35	527
IL	356.19	2.47	80	OK	107.70	0.00	0
IN	402.28	193.65	610	PA	436.87	270.11	813
KS	24.09	0.00	0	SC	90.68	68.66	1482
KY	240.36	67.32	405	SD	13.72	20.96	1071
LA	82.64	1.45	151	TN	266.34	146.51	721
MA	60.59	23.17	971	TX	317.94	55.97	388
MD	89.46	30.11	981	UT	18.12	0.00	0
MI	323.31	0.00	0	VA	112.46	27.18	1331
MN	49.16	9.56	343	WA	39.41	0.00	0
MO	247.24	18.86	369	WI	150.66	42.92	522
MS	48.45	5.35	420	WV	387.49	91.62	559
MT	5.06	0.09	156	WY	30.05	11.75	773
NC	314.11	30.38	927				

Notes: This table summarizes the results from simulating units compliance choices and associated compliance costs state by state under the counterfactual policy design for given state-specific caps in 2000.

Chapter 3. What Matters in Residential Location Choice in a Transition Economy: Evidence from Beijing, China[1]

[1] I would like to thank Siqi Zheng from Tsinghua University for kindly providing the housing dataset, and comments from Gordon Rausser, Meredith Fowlie, Sofia Villas-Boas, Denis Nekipelov, Jeff Perloff, Jaremy Magruder, Keneth Train, and Ben Handel. All errors are my own.

3.1 Introduction to Chapter

In China, for a long time, urban residents generally obtained welfare housing for free from "working units" they work for within the regime of the planned economy, and as expected, housing had been always a scarcity at that time. Since 1998, the Chinese central government started to abolish the long-established welfare housing system, and in Beijing, due to the resistance from lots of public sectors/enterprises, it was until 2002 that such a system had been totally abandoned. From then on, households had to resort to the housing market to get new houses or update their current houses. On the other hand, this also meant that households had more and more opportunities to sort themselves into neighborhoods within the city according to their heterogeneous preferences for neighborhood attributes. This sorting process, intertwined with the local fiscal system, tends to bring about large variation across neighborhoods within a city in terms of socio-demographic composition, income, education level, etc.

This essay applies an equilibrium sorting model to investigate the residential location choice in the brand-new housing market in Beijing. The focus will be on heterogeneous household preferences for neighborhood public goods provision, including public transportation services, public primary schools, government-owned parks and non-park green space.

The equilibrium sorting model developed by Bayer et al. (2005) is built on the random utility model of McFadden (1978) and the specification of Berry, Levinsohn, and Pakes (1995). It relies on the properties of market equilibrium and household residential location choice behavior to recover the varied household preferences for neighborhood characteristics (Kuminoff, Smith, and Timmins, 2010). Subsequent studies have applied this framework to investigate how the sorting behavior in housing markets relates to school quality (Bayer, Ferreira, and McMillan, 2007), climate amenities (Timmins, 2007), open space (Klaiber and Phaneuf, 2010), and air quality (Tra, 2010).

Although Chinese housing reforms have been widely discussed in the literature, there are only a few studies that have paid attention to individual household's location choice behavior (Zheng and Kahn, 2008; Zheng, Sun, and Wang, 2014). Wu, Zhang, and Dong (2013) investigate the determinant of residential location choice in Bejing, using a discrete choice model and a survey data set in 2009. However, housing price is assumed to be exogenous and neighborhood unobserved attributes are not controlled in their estimation. Relying on a unique, detailed data set collected from Beijing Household Transportation Survey in 2005, this essay provides the first empirical study of individual household's heterogeneous preference to neighborhood attributes using an equilibrium sorting model. Knowledge of this preference heterogeneity can be used in evaluating the effects of government policies on social welfare, and also can shed light on future policy designs.

In the model, a household's utility from choosing a particular neighborhood is decomposed into three parts. First is the mean indirect utility provided by the neighborhood, which is determined by all observed and unobserved neighborhood attributes including average housing price, and is common to all households. The second part is the portion of utility unique to each household through interacting household characteristics with neighborhood attributes, which represents individual household's heterogeneous preference. And the third part is an idiosyncratic error term. By including unobserved neighborhood attributes into the mean indirect

utility, the parameters for the interaction terms can be identified. The endogenous housing price in each neighborhood is instrumented using attributes in other neighborhoods.

My findings suggest that people do take into consideration public goods provision when they choose their residences. In general, lower housing price, better environmental amenities, and being closer to job centers will increase the choice opportunity of a neighborhood. Public transportation systems play a more important role in the neighborhoods far away from urban centers. Moreover, different households show varying preferences for these public goods. A distinct fact is that in addition to income, people's preferences vary greatly with generation (head age of households) and job type (whether there are public employees), which reveal the significant differences between generations and illustrate the welfare for public employees within the context of transitional economy.

The remainder of this chapter is organized as follows: Section 3.2 introduces background on the housing market in Beijing; Section 3.3 presents the model and the estimation procedure; Section 3.4 reports descriptive statistics of the data employed; Section 3.5 discusses the estimation results; and Section 3.6 concludes.

3.2 Background on the Housing Market in Beijing

In late 1970s, China initiated its economic transition from a centrally planned economy into a market-oriented economy. During this process, land and housing reform has had major implications for residential location choice by restructuring housing provision and reshaping urban spatial structure.

In the original planned economy, housing provision was dominated by a work unit-based system, in which housing was allocated to employees for free by the government or the work unit. The workplace-residence relationship at that time was characterized by well-balanced job-housing distribution, featuring self-contained work unit compounds. But this system (in particular, the lack of land markets) also produced enormous land-use deficiencies, represented by the low land-use densities in urban center (Bertaud and Renaud, 1997).

The introduction of land and housing markets has significantly changed the urban spatial structure. First, high-rise office buildings have gradually replaced the old housing and industry enterprises in urban centers. Second, large-scale housing production was mainly carried out far away from the traditional urban center. These two aspects have reduced land use mixing, greatly increased workplace-residence separation, and led to large-distance commuting.

More importantly, the housing policies established by the central government in 1998 greatly facilitated the creation of an urban housing market. The two major policies were to abolish the direct work-unit based welfare housing provision system, and to introduce a new housing finance system for both developers and individuals (State Council of China, 1998).

Due to the resistance from lots of public sectors/enterprises, the housing market in Beijing hadn't grown very quickly until the welfare housing system was totally abandoned in 2002. From then on, annual newly built housing floor area has been no less than 20 million square meters, which amounts to about 200,000 apartments assuming that the average floor area

per apartment is 100 square meters. Usually, the municipal government appropriates a piece of land and sells it (70-year land use permit for residential land) through auction to developers for constructing new housing projects. Apartment instead of single-family housing is the most common product in Beijing housing market. A new housing project usually consists of hundreds of apartments. A common pricing strategy adopted by a developer is to first set an average price (RMB per square meter) for a housing project according to its location, amenities, housing quality, etc., and then to adjust the price for each apartment in this project by its sunshine, number of floor, view, etc.

As to households, without the welfare housing allocation after 2002, they have to resort to the housing market to get new houses or update their current houses; on the other hand, this also means that households have more and more opportunities to relocate themselves within the city totally by their own preferences.

3.3 Model and Identification Strategies

3.3.1 Equilibrium Sorting Model

The framework of this essay follows closely the equilibrium sorting model developed by Bayer et al. (2005), which is based on discrete choice models of household residential location decisions. This model is extremely flexible in its treatment of preference heterogeneity by allowing households to differ in their relative preferences for multiple housing characteristics and amenities.[2]

A household i's utility from choosing a particular neighborhood j is

$$U_j^i = V_j^i(X_j, z^i, p_j, d_j^i, \xi_j) + \varepsilon_j^i \tag{3.1}$$

where X_j represents the observable characteristics of neighborhood j including public transportation services, public primary schools and environmental amenities in this essay, z^i represents observable household characteristics including income, job type and household structure in the following analysis, p_j is average housing price in neighborhood j, d_j^i is the average distance (in terms of all working adults) from the centroid of neighborhood j to the job location of household i, ξ_j reflects unobserved attributes of neighborhood j, and ε_j^i is an idiosyncratic error term.

The utility function is specified in a linear form as

$$U_j^i = \alpha_X^i X_j + \alpha_p^i p_j + \alpha_d^i d_j^i + \xi_j + \varepsilon_j^i \tag{3.2}$$

[2] Some unique features of this new housing market in Beijing might weaken the applicability of the model that is based on a mature housing market. In Beijing, huge housing demand was suddenly released since around 2002 due to the gradual abolishment of the welfare housing system. Consequently, housing supply has been increasing quickly since then, which makes the exogenous and fixed housing supply assumption in the model not very plausible.

where each parameter associated with observed neighborhood characteristics and distance varies with a household's own characteristics according to

$$\alpha_k^i = \alpha_{0k} + \sum_{l=1}^{L} \alpha_{lk} z_l^i \tag{3.3}$$

where $k \in \{X, p, d\}$, L is the number of household characteristics. This specification gives rise to a horizontal model of sorting that allows for heterogeneous preferences of households over each neighborhood characteristics.

Household i chooses neighborhood j if the utility from this neighborhood exceeds the utility from all other possible neighborhood choices, that is

$$U_j^i > U_{j'}^i \Rightarrow V_j^i + \varepsilon_j^i > V_{j'}^i + \varepsilon_{j'}^i \Rightarrow \varepsilon_j^i - \varepsilon_{j'}^i > V_{j'}^i - V_j^i \qquad \forall j' \neq j \tag{3.4}$$

The inequality (3.4) implies that the probability of any household i choosing a neighborhood j depends in general on the characteristics of all neighborhoods across the market and can be written as

$$Pr_j^i = f_j(z^i, \mathbf{X}, \mathbf{p}, \xi) \tag{3.5}$$

where bold letters are vectors of observed and unobserved characteristics, and average housing prices for all choice neighborhoods. Integrating the probabilities in equation (3.5) over the distribution of household characteristics yields the predicted demand for each neighborhood j, that is

$$D_j = \int f_j(z^i, \mathbf{X}, \mathbf{p}, \xi) g(z^i) dz^i \tag{3.6}$$

where $g(z^i)$ is the distribution function for household characteristics. Assume an exogenous supply of housing S_j in each neighborhood. In order for the housing market to clear, the demand for neighborhood j must equal the supply for such neighborhood, so

$$D_j = S_j \quad \forall j \quad \Rightarrow \quad \int f_j(z^i, \mathbf{X}, \mathbf{p}, \xi) g(z^i) dz^i = S_j \quad \forall j \tag{3.7}$$

Given the decentralized nature of the housing market, prices are assumed to adjust in order to clear the market. Bayer et al. (2005) prove that a unique vector of housing prices solves the system of equations given in (3.7).

3.3.2 Estimation Procedure and Identification Strategies

To identify the parameters in equation (3.2), I follow a two-step procedure discussed in Berry, Levinsohn, and Pakes (henceforth, BLP, 1995). In doing so, I rewrite the indirect utility function in equation (3.2) as

$$U_j^i = \delta_j + \gamma_j^i + \varepsilon_j^i \tag{3.8}$$

In equation (3.8) δ_j captures the portion of utility provided by neighborhood j (neighborhood fixed effects), which is determined by all observed and unobserved attributes of neighborhood j and is common to all households. When the household characteristics are demeaned, δ_j can be interpreted as the mean indirect utility provided by neighborhood j. Explicitly, δ_j is given by

$$\delta_j = \alpha_{0X} X_j + \alpha_{0p} p_j + \xi_j \tag{3.9}$$

γ_j^i captures the portion of utility unique to household i through interactions with household characteristics, which is given by

$$\gamma_j^i = (\sum_{l=1}^{L} \alpha_{lX} z_l^i) X_j + (\sum_{l=1}^{L} \alpha_{lp} z_l^i) p_j + (\sum_{l=1}^{L} \alpha_{ld} z_l^i) d_j^i + \alpha_{0d} d_j^i \tag{3.10}$$

By including unobserved neighborhood attributes ξ_j into δ_j, it is assumed that once the mean utility components are swept out, the remaining utility components (i.e. the individual specific deviations) are uncorrelated with the error term, provided that neighborhood-specific attributes do not vary systematically across individuals within a given market.

Then, assuming ε_j^i is independent and identically distributed extreme value, the conditional logit probability of household i choosing neighborhood j is given as

$$Pr_j^i = \frac{exp(\delta_j + \gamma_j^i)}{\sum_n exp(\delta_n + \gamma_n^i)} \tag{3.11}$$

and the log likelihood function is

$$ll = \sum_i \sum_j Y_j^i ln(Pr_j^i) \tag{3.12}$$

where Y_j^i equals 1 if household i chooses neighborhood j and equals 0 otherwise. By a maximum likelihood process the first step estimation recovers the heterogeneous parameters $\alpha_{lk}'s$ and the mean indirect utilities $\delta_j's$.

The second step estimation recovers the parameters $\alpha_{0k}'s$ by decomposing the mean indirect utility into its observable and unobservable components according to equation (3.9). In this step, the endogeneity of housing price and other neighborhood attributes is a big concern. First, housing price is almost certainly correlated with the unobserved neighborhood attributes. For example, better locations, better air quality and/or other neighborhood amenities captured in part by the unobserved variables are likely to command a higher price. Second, neighborhood public goods and services quality, which influences household location decisions, is also possible to be determined endogenously, especially in US cities where public goods provision is a collective choice under majority rule (Epple et al. 2001).

In order to solve price endogeneity, this essay uses government-planned residential land supply in neighborhoods within 5km-10km distance as an instrument for price of a

neighborhood. The total residential land supply of the city is determined by the government according to the anticipated growth of population and economy. These residential land supply quotas are then allocated to each neighborhood through a complicated decision-making process. However, this is more a political process during which the mayor, the planners, the district government, and even the town government all are involved, and this policy may quickly change due to the unexpected political changes.[3] Then, Developers achieve the developing rights to these residential lands through open auctions. Therefore, residential land supply in a neighborhood will directly affect housing price in that neighborhood. In addition, based on competition between developers, observed housing price in a particular neighborhood is determined not only by the attributes of the neighborhood itself but also by the prices of alternative neighborhoods in the wider region. Thus, the residential land supply of neighborhoods at a reasonable distance from a particular neighborhood *i* could serve as suitable instruments for the neighborhood *i*'s averaging housing price, since it affects the price through the housing market and is uncorrelated with the unobservable attributes in neighborhood *i* that the utility derived from living in that neighborhood.

As to other neighborhood attributes (public bus and subway services, public primary schools, government-owned parks and non-park green space) in this essay, I simply assume that they are all predetermined. This assumption is partially acceptable because the public investment decisions are mostly made at the city level government and urban residents are barely involved in this process, which implies that public services at the neighborhood level are sorting independent. However, the estimated coefficients on these public services could still be biased because of the omitted variables. For example, one big concern of the omitted variables is housing quality in each neighborhood, and it is highly likely that developers tend to provide high class housing projects in the neighborhoods with better green space. Therefore, green space ratio is probably positively correlated with the error term, which then leads to over-evaluated estimation of the WTP for green space. Consequently, it would be better to interpreter green space as a proxy of neighborhood/housing quality. Other neighborhood attributes, as public bus, subway, primary school, might also encounter similar omitted variable problems and it is even more difficult to talk about the signs of these correlations.

3.4 Data

The data mainly comes from Beijing Household Transportation Survey (BHTS) in 2005, which covers all 18 districts/counties in the whole Beijing city and uses transportation analysis zones (TAZs)[4] as the smallest geographic unit of aggregation. This essay merely focuses on the Beijing Metropolitan Area (BMA), which approximately amounts to 8 urban districts and is located inside the 5th ring road (Figure 3.1).[5] The BMA roughly represents a unified labor market of

[3] For example, in certain year the southern neighborhoods got more quotas just because the mayor wanted to balance the development across the city.

[4] A transportation analysis zone (TAZ) is the unit of geography most commonly used in conventional transportation planning models.

[5] In Beijing, people usually use ring road to measure the geographic proximity to urban center (See Figure 1). And psychologically people take the area within 3rd ring road as the central city around 2005.

central Beijing, covering 580 km^2 land (358 TAZs). The sample size in this research area is 48,705 households, 90 percent of which (43,974 households) still lives in state-allocated welfare housing units and 10 percent (4,731) has moved to commercial housing units located in 132 TAZs. Only the latter observations, 4,731 households, are used in the following model estimation, because there is no active residential location choice for those still living in welfare housing units. As mentioned before, housing market in Beijing hadn't grown very quickly until the welfare housing system was totally abandoned in 2002. So roughly, we assume that households who live in new commercial housing made their location choices between 2002 and 2005. And most of the following neighborhood variables are estimated to reflect the situation at that time.

3.4.1 Summary Statistics

This subsection presents detailed characteristics of these households (4,731 observations) and attributes of their choice sets (132 TAZs).

Households Characteristics

The survey data provides a wealth of information on the individuals and households in the sample. To permit flexibility in preferences across different types of households, the following household characteristics are included in the model: household monthly income,[6] whether there is any household member who works in a public sector, age of household head, and whether there are children under 13 years old. As shown in Table 3.1, a typical household has monthly income 400-550 US$. The average head age is around 45. The average commuting distance for all household members who work is 10.43km, which is obviously higher than the average level for all employees in Beijing. This indicates that the new commercial housing projects generally are located farther away from job centers than the housing stock. Of these households, about 26% has at least one member employed in public sector, and 22% has at least one kid under age 13.

Choice Set

Choice set in this case is neighborhood rather than housing units due to data limitation. However, in China, developers usually build large-scale housing projects that consist of hundreds of similar apartments, so the variation in housing attributes within/between a neighborhood is not as significant as that in US cities. Again, limited by available data, TAZs where households reside are taken in this essay as local neighborhoods.[7] Combined with geographic information on public facilities and transportation infrastructure obtained from Beijing City Planning Institute, location attributes of neighborhoods including density of bus-lines, density of subway stops, whether there is government-owned parks, ratio of non-park green space to the neighborhood land area, average distance to the nearest 3 key primary schools as well as distance to the two main job centers, are estimated using the GIS technique.

---

[6] It is categorical in the data with 1: <300, 2: 300-400, 3: 400-550, 4: 550-850, 5: 850-1500, 6: 1500-3000, 7: 3000-4500, 8: >4500 $/month.

[7] A possible bias is that for some large TAZs, the internal variance cannot be well presented by the overall statistics.

Housing Transaction Data

The survey doesn't provide any information on housing transaction price. A data set containing transaction prices and location information of 968 housing projects is used to estimate average housing price in each neighborhood by the inverse distance weighting (IDW) interpolation method.[8] This transaction dataset is drawn from Beijing's Housing Transaction Registration System, which keeps the records of all new housing transaction contracts. All the 968 projects in the dataset are new commercial housing projects and account for more than 80% of the total housing transactions in Beijing from 2004-2005. Within each project, there are several residential buildings and each building has many housing units, similar to a condominium building in the United States. The average project in the sample has 791 housing units. However, to what extent could the assessment price represent the actual average transaction price in each neighborhood is suspected.

The details of all the variables and their sources, and summary statistics are given in Table 3.1.

3.4.2 Move or Stay?

The characteristics of households who chose to stay and those who chose to move are presented in Table 3.2. Such comparison may roughly show what kinds of households embraced the opportunity provided by housing reform and newly-booming real estate industry to relocate themselves and who chose to stay (after an implicit choice).

As indicated in Table 3.2, there are significant differences between those who chose to move and those who chose to stay in terms of household income, job status, age of household head and children (under 13). And the results are consistent with our expectation. For example, households with higher income could afford the new, modernized but also expensive housing. And generally households with any member working in public sectors preferred the new housing than staying in old housing. Actually, this helps illustrate the effect of income in this decision process for public employees usually have relatively higher incomes at that time. As for the age of household head, it is expected that the housing reform facilitated the decrease of household size. Before the housing reform, housing shortage may force three generations (or more) crowded in a small apartment. With the development of housing industry, the younger generation then could really start their "independent" lives by buying a new commercial housing, generally aided by the older generation. This made the head of households who live in commercial housing younger than that in state-allocated housing. Similarly, households living in commercial housing were more likely to have children under age 13.

3.5 Estimation Results

[8] To use IWD method, the research area is divided into 10m*10m grids. Price in each grid is a weighted average price of all 968 housing projects, with distance as the weight. And average neighborhood housing price is a simple average of the interpolated grid price in each neighborhood.

For households who chose to buy new commercial housing, the 2-step discrete choice model is used to estimate their preferences for various neighborhood-specific public goods.

3.5.1 Preference Heterogeneity Estimation from the First Step

The first step of estimation recovers 28 parameters of interactive terms between household characteristics and neighborhood public goods provision as well as a vector of mean indirect utilities for the 132 TAZs in the sample. The model also captures the average effect of commuting distance in households' residential location choice.[9] The estimation results are reported in Table 3.3.

It turns out that households' preferences vary with income, which is consistent with that in western cities. For example, with higher monthly income, households show higher preferences for neighborhoods with more convenient public transportation services. Note that at that time the private car ownership is still low (40% in the sample). This is also consistent with the estimation result on the interactive term between income and commuting distance. It is shown that commuting distance hurts poor households more than wealthy households possibly because the latter cares other amenities, such as non-park green space in this case, or because the former prefer living close to workplace due to longer work time.[10] Moreover, with higher monthly income, households relatively prefer for high housing price. In other words, for high-income people, the utility decreasing caused by high housing price is not as severe as that for low-income people. This probably shows that housing is an important commodity in indicating one's identify/class, and high-income residents prefer living with people of similar income.

An interesting result is the heterogeneous households' preferences associated with head age. Although other studies also identify similar difference across generations, the high heterogeneity caused by age here may reveal much more differences between generations within the context of transitional economy. The quickly changing values as well as the education system that help shape households' preferences may also bring about such heterogeneity. Due to income (or saving) effects, older people/households have similar preferences for housing price and public bus services as those with higher income. Also, though private primary school began to emerge in Beijing, older people may be least willing to accept such an education alternative due to their long-established ideas. Besides non-park green space, older people even have higher valuation to government-owned parks, because in China older people are used to morning exercises and most parks are free to access in the early morning.

The heterogeneous preferences by job type well illustrate the welfare for public employees in a transitional economy. Households having members employed in public sectors have high preference for neighborhoods with high ratio of non-park green space, as it implies less housing density, less congestion, better air quality etc. However, they do not care bus services, for most of them can use firm-provided bus or car (in the sample, 8% of households which have public-employees have firm-provided cars). And longer commuting distance hurt this type of households more. This is because in some sense, the welfare associated with public

[9] The implicit assumption here is that workplaces of working adults in each household are predetermined and will not be affected by households' residential location decisions.
[10] Note that overall nobody likes longer commuting distance.

sector is also localized. Only by living close to workplace (and thus state-allocated housing for this working unit) can these households enjoy the welfare. And the negative valuation on housing price is probably due to career concerns, since buying expensive housing may sometimes be a little harmful to their reputations and thus promotions.

Households with children under age 13 have no significant preference for public transportation services. In Beijing, parents prefer picking up their children by themselves by bike or by car (actually, in the sample, about 60% of this type of households have private cars). The coefficients of the interactive terms both with the primary school and with non-park green space are significant and conform to expectation. And since the parents in such households need spend lots of time in peak hours in picking up their children, longer commuting distance hurt them most heavily.

There is no significant preference heterogeneity for subway stop density. One possible reason is that the measure adopted in this essay at a neighborhood level cannot discover the localized effect of the subway. Another reason is that at that time there were only 4 subways lines in operation which meant low accessibility and high transfer cost, thus subway service is not a big concern to household's residential location choice. These two reasons could also explain the insignificant estimation results of subway stop density in the second step.

To test the robustness of the first step estimation, I change the algorithm method from Newton-Raphson to DFP and get exactly the same results. I also estimate the model by leaving out several interaction terms such as interactions with price, or with primary school, and I get stable results. However, the preference heterogeneity is only partially captured by the interaction terms in the model, and it would be better to test whether the preference systematically changes with other unobserved household socio-demographic characteristics. I will incorporate a mixed-logit estimation in the future study.

3.5.2 Estimation Results from the Second Step

In the second step, the mean indirect utilities of 132 neighborhoods are decomposed into observed and unobserved neighborhood attributes as in the equation (3.9).

First a simple OLS regression is applied to estimate this equation, and the results are reported in the first result column of Table 3.4. The coefficient of average housing price is significantly negative, indicating that the price increase will lower households' utility. The coefficient of bus-lines density is negative and insignificant. A possible reason is that in the city center, public bus system is already well-developed and can well satisfy households' daily travel demand, so adding an extra bus-line might even be painful for urban residents because of the congestion, pollution, and noise problems caused by it. To test whether there is differentiated effect of bus-lines in the city center and in suburb, I add an interactive term that is the product of bus-line density and a dummy variable of whether a neighborhood locates beyond the 3rd ring road. The estimation results are reported in the second result column of Table 3.4. The coefficient of this new variable validates our explanation. Adding an additional bus line will increase the utility of households who live far away from urban center. In the second regression, distance to two main job centers in Beijing is also included to capture the location advantage of neighborhoods. The estimate results on these two variables conform to our expectation but are not significant.

As mentioned before, to address the price endogeneity issue, government-planned residential land supply in neighborhoods within 5km-10km distance is used as an instrument variable for the price of a specific neighborhood. The IV regression results are reported in result columns 3 to 5 of Table 3.4. Generally the price coefficient is more negative than that in the OLS regression, reflecting positive correlation between price and unobserved variables. The estimated coefficients on two bus-related variables still conform to our previous explanation as in the 4th column. The coefficients associated with non-park green space are positive and significant as expected. In addition, the coefficients of variables that aim to identify the effect of job market now become significant. Households gain more utility when approaching these two job centers, especially the high-tech job center in Haidian district.

As to other public goods (subway, primary school, and government-owned park), the coefficients are insignificant, not stable or even with unexpected signs. Besides omitted variable problems discussed before, another important reason leading to these results is that there are high correlations between these regressors, as shown in Table 3.5. Especially, the bus-line density, subway stop density and primary school distance are highly correlated to the ring road variable, because in Beijing these public facilities highly concentrate around the urban center and their densities decrease with the distance to the urban center. Therefore, when a "ring" dummy variable is included in the regression to further control district fixed effects,[11] it turns out that all public service coefficients except green space become insignificant, as shown in the result column (5) of Table 3.4. In this case, average housing price, location and housing quality (represented by the ratio of green space) almost captures all important factors influencing household's residential location choice decision.

3.5.3 Marginal Willingness to Pay

To better illustrate the estimation results, I present marginal willingness to pay (MWTP) for a typical household that combines both first and second stage estimates (IV regression results in the result column (5) of Table 3.4) in Table 3.6. Due to the insignificant estimators in the model, the standard errors of the MWTP are relatively big.

A typical household would like to pay \$0.0165 for 1% increase of bus-line density, \$0.1357 for 1% increase of subway stop density, \$-13.7736 for one kilometer closer to primary school and \$28.5937 for government-owned park. The highest MWTP is for non-park green space, which is \$36.1984, which possibly reflects severe scarcity of these environmental amenities in Beijing. As mentioned before, another possible reason for this high MWTP is that what the green space ratio measures is not only the environmental amenity but also quality of the housing projects in the neighborhood. For example, in neighborhoods with parks and/or more green space, the housing density is likely to be lower and developers tend to provide high-class housing projects, which may not be well captured by the housing price.

[11] District dummy is not included in the regression because the government at the district level usually plays a less important role in Beijing compared with the municipal government and the internal difference within a district might be bigger than that between districts due to large geographic area of a district. For example, Haidian district covers 430.8km^2 land extending from within 3rd ring road to beyond 6th ring road.

3.6 Conclusions

With the development of housing market in Beijing, China, more and more households relocate themselves within the city by their own preferences. Knowledge of the heterogeneity of household preferences could be used to evaluate government policies, especially when the neighborhood public goods are mostly provided by the government. By applying the equilibrium sorting model to this brand-new housing market, this essay tests heterogeneous household preferences for local public goods in making their residential location decisions.

In general, public bus system in suburb areas and non-park green space are desirable neighborhood attributes in household location choices. Therefore, it would be better to increase the public transportation investment in suburb areas. And the high MWTP for non-park green space may reflect the increasing household's valuation in environmental amenities as well as demands for environment-friendly policies with rapid economic growth.

The estimates also reveal that household preferences for these public goods vary greatly with observable household characteristics, including income, job type and household structure. Households with higher monthly income, older households and household having children have higher preferences for living in neighborhoods providing better public transportation services, environmental amenities and education resources. Especially, in addition to income, head age of households plays a significant role in location choice in Beijing, which might lead to homogeneity in age profile within neighborhoods and further affect the efficiency of current public goods provision. For example, periodically, there could be geographic mismatch between the demand for and the supply of education resources across neighborhoods, especially considering that the latter is still controlled by the government. Therefore, this preference heterogeneity implies future policies should be more geographically asymmetric, locally targeted and tailored based on specific socio-economic characteristics.

3.7 Tables and Figures

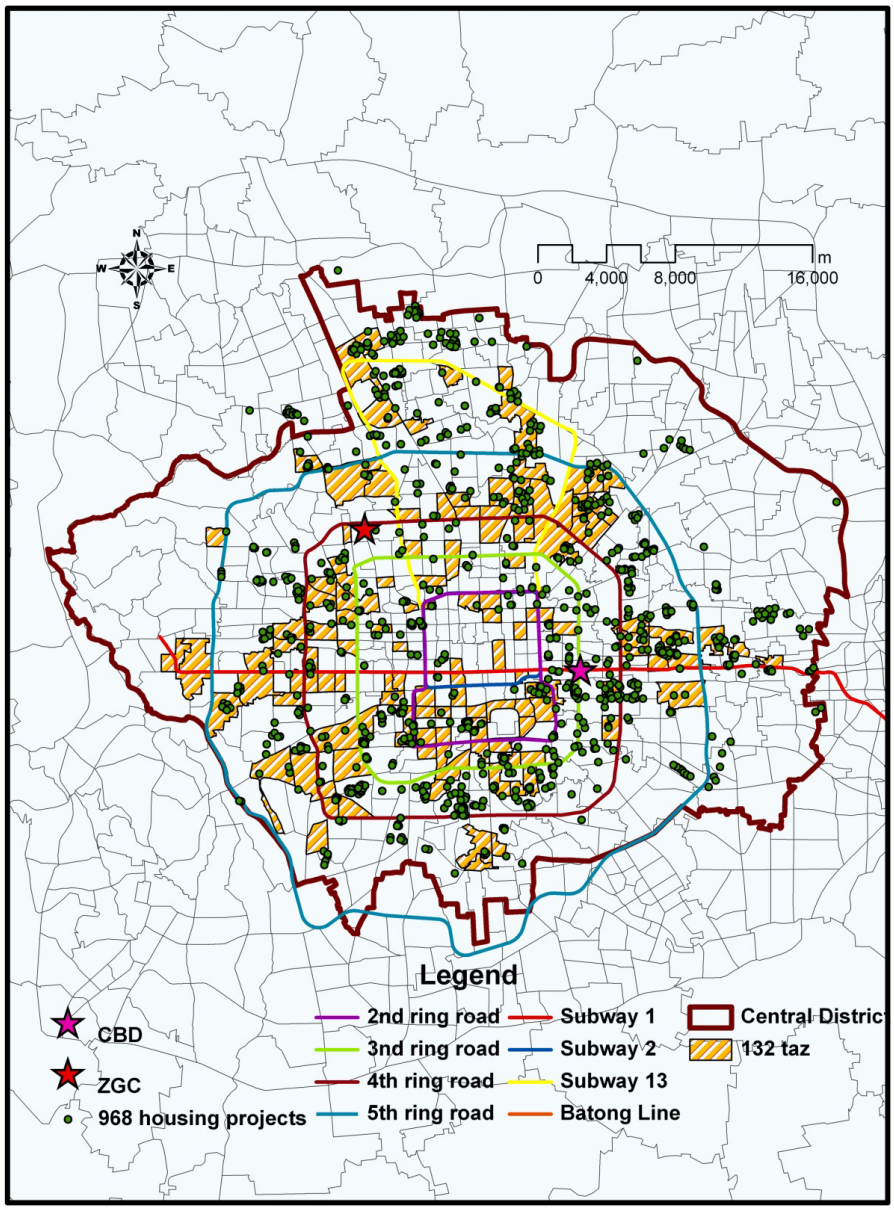

Figure 3.1: Research Area

Notes: The research area is the central district of Beijing, which approximately amounts to 8 urban districts and is located inside the 5th ring road. The yellow shaded areas (132 transportation analysis zones) are the smallest geographic unit of aggregation and are the choice sets for households. The green dots represent 986 housing projects, the housing prices of which are used to generate average housing price in each zone. CBD and ZGC represent two job centers of the city.

Table 3.1: Variable definition, source, and summary statistics

variables	Definitions and sources	Obs	Mean	S.D.	Min	Max
TAZ attributes						
Area	Land area of each TAZ, km^2. Source: GIS	132	1.46	0.93	0.33	5.71
Average Price	Estimated average housing price, 1000\$/m^2	132	1.24	0.42	0.54	3.36
Bus-line	Density of bus lines, line/km^2. Source: GIS	132	19.17	20.19	0	144.5
Subway Stop	Density of subway stops, stop/km^2. Source: GIS	132	0.30	1.03	0	8.37
Primary School	Average distance from the centroid of neighborhood to the nearest 3 key primary schools, km. Source: GIS	132	3.26	2.09	0.21	10.58
Park	Whether there are government-owned parks in each neighborhood, 1 yes, 0 no. Source: GIS	132	0.11	0.32	0	1
Non-park Green Space	Ratio of non-park green space to the neighborhood land area	132	0.04	0.06	0	0.40
CBD	Distance to the central business district, km. Source: GIS	132	10.39	5.07	1.34	22.76
ZGC	Distance to the high-tech job center, km. Source: GIS	132	10.55	4.65	0.82	24.00
Household attributes						
Monthly Income	Monthly income of each household, categorical data. Source: BHTS	4731	3.60	1.48	1	8
Public Employee	Whether there are members employed in public sectors in each household. Source: BHTS	4731	0.26	0.44	0	1
Head Age	Age of each household's head person. Source: BHTS	4731	44.84	13.35	11	88
Kids under 13	whether there are kids under 13 of each household. Source: BHTS	4731	0.22	0.41	0	1
Commuting Distance	Average commuting distance of working adults in each household. Source: BHTS & GIS	4731	10.43	7.42	0	69.55

Table 3.2: Comparison between Moving and Staying Households

	Mean of households in welfare housing (G1)	Mean of households in commercial housing (G2)	Difference (G1-G2)	t-statistic
Monthly income	3.046	3.596	-0.550	-28.002
Whether there are public employees	0.235	0.263	-0.027	-4.294
Head age	51	45	6.599	30.876
Kids under 13	0.143	0.217	-0.0738	-13.185

Notes: This table is generated using data of 43,974 households who chose to stay in state-allocated welfare housing units, and of 4,731 households who chose to move to commercial housing units.

Table 3.3: Interaction Parameter Estimates from the First Step

	Household Characteristic			
Neighborhood Attributes	Income	Head Age	Public Employee	Kids under 13
Average housing Price	0.040***	0.003***	-0.067**	-0.119***
	(0.009)	(0.001)	(0.028)	(0.031)
Bus-lines Density	0.008**	0.001***	-0.015**	-0.003
	(0.001)	(0.000)	(0.004)	(0.005)
Subway Stop Density	-0.000	-0.003	0.084	-0.084
	(0.023)	(0.003)	(0.070)	(0.085)
Distance to Primary School	0.003	-0.003***	-0.020	-0.074 ***
	(0.007)	(0.001)	(0.023)	(0.024)
Whether There are Parks	-0.022	0.008*	-0.134	-0.594***
	(0.044)	(0.005)	(0.143)	(0.164)
Ratio of Non-park Green Space	0.020	0.048*	2.007***	3.412***
	(0.247)	(0.026)	(0.754)	(0.825)
Commuting Distance	0.032 ***	0.000	-0.060***	-0.055 ***
	(0.004)	(0.000)	(0.012)	(0.013)
Commuting Distance overall		-0.321***		
		(0.005)		

Notes: This table presents results from the first step logit estimation, which recovers 28 parameters of interactive terms between household characteristics and neighborhood public goods provision as well as the average effect of commuting distance in households' residential location choice. Standard errors are in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 3.4: Estimation Results from the Second Step
Dependent Variable: Mean Indirect Utility

Variables	(1) OLS	(2) OLS	(3) 2SLS	(4) 2SLS	(5) 2SLS
Average Price	-0.174***	-0.215***	-0.570***	-0.552**	-0.276**
	(0.057)	(0.050)	(0.192)	(0.228)	(0.138)
Bus-line Density	-0.010	-0.012	0.001	-0.008	0.008
	(0.011)	(0.011)	(0.013)	(0.011)	(0.012)
Bus-line Density in Area beyond the 3rd ring		0.080***		0.069***	0.031
		(0.022)		(0.023)	(0.029)
Subway Stop Density	-0.039	0.003	-0.080	-0.037	-0.194
	(0.158)	(0.151)	(0.181)	(0.157)	(0.166)
Distance to Primary School	0.227**	0.257**	0.000	0.268**	-0.029
	(0.098)	(0.129)	(0.162)	(0.133)	(0.138)
Whether there are government-owned parks	-0.003	0.154	-0.016	0.168	0.034
	(0.471)	(0.491)	(0.496)	(0.503)	(0.478)
Ratio of Non-park Green Space	4.610	4.842*	9.184**	7.279**	6.886**
	(3.059)	(2.791)	(3.685)	(3.043)	(3.405)
Distance to CBD, km		-0.042		-0.143*	
		(0.045)		(0.080)	
Distance to ZGC, km		-0.078		-0.205**	
		(0.062)		(0.102)	
Ring fixed effect	No	No	No	No	Yes
Constant	1.395*	2.589**	4.990***	7.562**	2.137
	(0.761)	(1.222)	(1.931)	(3.606)	(1.601)

Notes: This table presents estimation results from the second step, which regresses the mean indirect utilities of 132 neighborhoods on the observed and unobserved neighborhood attributes. Results presented in column (1) and (2) are estimated using OLS. Results presented in the last three columns are estimated using an instrument variable for housing price and 2SLS. Standard errors are in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 3.5: Correlation Between Regressors

	Average Price	Bus-line	Subway Stop	Primary School	Parks	Green Space	Ring Road
Average Price	1.000						
Bus-line	0.281* (0.001)	1.000					
Subway Stop	0.118 (0.179)	0.346* (0.179)	1.0000				
Primary School	-0.517* (0.000)	-0.517* (0.000)	-0.150* (0.086)	1 .0000			
Parks	0.122 (0.163)	0.084 (0.336)	-0.049 (0.578)	-0.159* (0.069)	1.000		
Green Space	0.248* (0.004)	0.026 (0.768)	-0.116 (0.185)	-0.065 (0.459)	0.322* (0.0002)	1.000	
Ring Road	-0.462* (0.000)	-0.486* (0.000)	-0.216* (0.013)	0.685* (0.000)	-0.128 (0.145)	-0.047 (0.596)	1.000

Notes: Significance level in parentheses. Asterisks indicate statistical significance at the 10% * levels.

Table 3.6: Marginal Willingness to Pay Measures of a typical household

Public Goods Provision (Change)	Bus-line Density (0.01)	Subway Stop Density (0.01)	Distance to Schools (1 km)	Park (1)	Green space (0.01)
MWTP ($)	0.017 (0.16)	0.136 (0.30)	-13.77 (89.77)	28.59 (444.59)	36.20 (77.91)

Notes: Marginal willingness to pay is generated by combining the first step logit estimates and the second step estimates from the specification (5). It is transferred from in terms of average housing price to in terms of average housing value by multiplying the average housing size ($93m^2$) in the sample, and the average housing value in the research area is 115,320 US$. A typical household has monthly income 400-550, 26% probability to have at least one public employee, household head age around 44.84 and 22% probability to have at least one child under 13 years. Standard errors are in parentheses.

Chapter 4. Impacts of Climate Change on Crop Yields in China[1]

[1] I would like to thank Wenlai Jiang and Qiyou Luo from Chinese Academy of Agricultural Science for kindly providing the dataset, and comments from Gordon Rausser, Meredith Fowlie, Catherine Wolfram, Peter Berck, Yang Liu, Xi Lu, Huayong Zhi, and Lunyu Xie. I also thank Yang Ju for excellent GIS support. All errors are my own.

4.1 Introduction to Chapter

Over the past 100 years, the average temperature in China has risen by 0.5~0.8 ℃, and the annual precipitation has decreased significantly in most of northern China, eastern part of the northwest, and northeastern China (NDRC, 2007). By 2050, the average temperature in China is projected to increase by 2.3~3.3 ℃ as compared with that in 2000, the nationwide precipitation is projected to increase by 5~7%, and extreme weather/climate events are expected to occur more frequently (NDRC, 2007). Such climate change will bring unprecedented challenges to China's agriculture sector because of the sensitivity of agricultural production to climate variables. This essay aims to investigate the effects of climate change on the yields of major crops in China, including rice, wheat, and corn. Understanding the precise link between climate variables and crop yields could help to enhance the adaptive capacity of agriculture and thus ensure food security in the context of climate change.

A great deal of efforts has been put into analyzing the effects of climate change on agriculture in the United States. One commonly used microeconomic approach is the Ricardian model developed in an innovative paper of Mendelsohn, Nordhaus, and Shaw (henceforth, MNS, 1994). Using county-level cross-sectional variation in temperature and precipitation, MNS estimate a reduced-form hedonic equation with the net value of farmland as the dependent variable, and find that global warming may have economic benefits for agriculture. Schlenker, Hanemann, and Fisher (2005) estimate a similar hedonic farmland value equation but take into account irrigation conditions, and find a significant loss for dry-land non-urban counties due to climate change.

One of the main concerns with the Ricardian model is that the coefficient estimates might be biased due to the omitted variables that influence farmland value and also might be correlated with the climate variables. Deschênes and Greenstone (henceforth, DG, 2007) propose an alternative strategy to examine the effects of climate change on agriculture. Using a county-level panel on agricultural production features and climate factors, DG (2007) regress agricultural profits and crop yields on annual temperature and precipitation, conditional on county and state by year fixed effects. The year-to-year variation in temperature and precipitation is plausibly random and thus better solves the omitted variable bias problems. DG (2007) conclude that climate change might be slightly beneficial for agricultural profits and yields. The panel approach has been widely adopted in the literature to estimate effects of climate change (Schlenker and Roberts, 2006; McCarl, Villavicencio, and Wu. 2008; Schlenker and Roberts, 2009; Lobell, Schlenker, and Costa-Roberts, 2011; Burke and Emerick, 2015).

Most of the relevant researches in China use crop simulation models, which focus on crop physiology and assume limited if any farmer adaptation actions, and therefore tend to overstate the climate change influences (Xiong et al. 2007; Ye et al. 2013; Tao F., M. Yokozawa, and Z. Zhang. 2009; Tao F., and Z. Zhang. 2013). Two studies apply the Ricardian model to empirically estimate the effects of climate change on agricultural net revenue in China using cross-sectional variation, but draw different conclusions (Liu et al. 2004; Wang et al. 2009). Based on pooled county-level data from 1985-1991, Liu et al. (2004) find an overall positive impact of climate change with higher temperature and more precipitation, though the impacts vary seasonally and regionally. Using household level survey data in 2001, Wang et al. (2009)

find that the global warming is likely to be harmful to rainfed farms but beneficial to irrigated farms. As mentioned before, the Ricardian model is vulnerable to unmeasured and omitted time-invariant variables.

Given the importance of the question, this essay provides the first empirical estimates of climate change on major crop yields in China in a panel setting. Following Dischênes and Greenstone (2007), using a detailed county-level panel dataset in China from 2000 to 2010, I regress the annual yields of three major crops (rice, wheat, and corn) against two annual weather parameters, namely, temperature and precipitation, conditional on city fixed effects and province-by-year fixed effects. I replace county fixed effects with city fixed effects in the analysis because the research period is relatively short and the county-fixed effects absorb a great deal of temperature variation, which will significantly damage the estimation results. By including city fixed effects and province-by-year fixed effects, the weather parameters are identified from the county-specific deviations in yearly weather from the average climate within the same city after adjusting for shocks common to all counties in the same province. Since the variations in weather are plausibly exogenous and random after controlling for time-invariant indiosyncratic features of the city and time-variant shocks of the province, the effects of weather on crop yields are therefore clearly identified in the panel data setting.

The estimation results show that the relationship between all three crops yields and temperature is inverted U-shaped, and the relationship between rice and corn yields and precipitation also follows the inverted U-shape pattern, but wheat yield is decreasing with respect to precipitation during the research period. These estimates are then used to predict the impacts of future climate change on the three crops yields in China. Overall, the impacts of future climate change on rice and corn yields would be negative, but the impacts on wheat yields would be positive. The average rice yield is expected to decrease by 0.08%-1.18%, depending on different emission scenarios, the average corn yield would have a larger decrease by 1.64%-2.21%, and the average wheat yield would increase by 5.48%-6.68%. These yield effects imply an economic loss of $\$343\times 10^6$ to $\$524\times 10^6$ in the rice sector, and $\$361\times 10^6$ to $\$361\times 10^6$ in the corn sector, while an economic gain of $\$783\times 10^6$ to $\$955\times 10^6$ in the wheat sector. The adverse effects in the rice and corn sectors could be mitigated through long run adaptation actions, such as adopting stress-resistant varieties, switching to new crops well-adaptive to local climate conditions, improving field infrastructure (e.g. irrigation systems), and even exiting from farming.

The remaining of the chapter is structured as follows. Section 4.2 summarizes the data. Section 4.3 introduces the empirical methodology and reports the main estimation results. Section 4.4 provides simulation results. Section 4.5 concludes.

4.2 Data

I mainly rely on two data sets to create the variables necessary for the estimation. One is county level agricultural production data from China County Statistical Yearbook (2000-2010) compiled by the Chinese Academy of Agricultural Sciences. Another is daily weather data during the same period from China Meteorological Data Sharing Service System of the Chinese Meteorological Bureau. Two scenarios of climate change predictions are then used to simulate the effects on

crop yields in China. I describe relevant variables in detail below and some summary statistics are reported in Table 4.1.

Agriculture Data

Counties are the finest geographic unit observed in this data set and used for regression analysis in the following sections. All urban counties with no rural population are removed from the sample. The rural counties with missing crucial agricultural data, such as, total planted acreage and output of major crops, are also removed. I also removed counties with varying county codes from year to year due to changes in administrative divisions. Therefore, the final sample has an 11-year balanced panel of 1,378 rural counties in 30 provinces and autonomous regions of China.

County-level yields of rice, wheat, and corn from 2000 to 2010 are the dependent variables. The yields for each crop are calculated by dividing the output of the crop by its acreage planted, measured as tons of output per hectare planted. I also control for irrigation conditions using the ratio of cropland irrigated.

Panel 1 of Table 4.1 provides county-level summary statistics for the dependent variables. The average yields from 2000 to 2010 are 6.70 ton/ha, 3.77 ton/ha, and 5.15 ton/ha for rice, wheat, and corn, respectively. Panel 2 of Table 4.1 lists some other characteristics of the sample counties. The average rural population in a sample county is 37.57×10^4, total cultivated land is 42.49×10^3 ha, agricultural income is 692.04×10^6 yuan (about $\$112\times10^6$), average chemical fertilizer input is 0.94 ton/ha, average pesticide input is 0.24 ton/ha, and the ratio of cropland irrigated is around 0.67

Weather Data

The Chinese Meteorological Bureau provides daily temperature and precipitation data observed from 824 national benchmark and basic ground meteorological stations for the 1951-2012 period. Figure 4.1 shows locations of these weather stations along with county boundaries. The station-level daily temperature and precipitation data are first matched to the 1,378 rural counties for the regression analysis. I use thin plate smoothing spline interpolation method to generate the daily temperature and precipitation for the counties without any ground meteorological stations.

Following the standard agronomic approach, I convert daily temperature into growing season degree days (GDD) as an independent variable measuring temperature. GDD is defined as the sum of degrees above a lower threshold and below an upper threshold during the growing season. Ritchie and NeSmith (1991) suggested the lower threshold at 8 ℃ and the upper threshold at 32 ℃, which are broadly representative and commonly used in the related literature. A day with a mean temperature below 8 ℃ contributes to 0 degree days; between 8 ℃ and 32 ℃ contributes the number of degrees above 8 ℃; and above 32 ℃ contributes 24 degree days (Schlenker, Hanemann, and Fisher, 2006; Deschênes and Greenstone, 2007). The GDD is then calculated by summing the daily measures of degree days over the entire growing season, which is defined as the months of April through October for rice, March through August for spring wheat, October through June for winter wheat, April through September for spring and summer corn, and August through November for autumn corn in China. Using the same method of

calculating GDD, I construct a separate variable of growing season harmful degree-days (GHDD) to control for the effects of temperatures above 34 ℃, which may be damaging for crop growth (Guiteras, 2009).

Another weather related independent variable is total precipitation in the growing season, which is calculated as the sum of daily precipitation across the growing season months in the relevant year.

Panel 3 of Table 4.1 provides summary statistics for the county-level weather data from 2000 to 2010. The average GDD's are 2896C, 1560C, and 1820C for rice, wheat, and corn, respectively. The growing season precipitations are 907mm, 476mm, and 431mm for rice, wheat, and corn, respectively.

Residual Variations

Because I rely on various fixed effects to isolate the effects of yearly weather variation on crop yields, it is important to examine how much variation in weather variables will be left over after these fixed effects have been adjusted for. Ideally, if the remaining variation in the weather variables after accounting for these fixed effects is as large as the predicted changes in climate, the potential effects on crop yields of future climate change will be identified from the data, rather than from functional form extrapolations.

Table 4.2 summarizes regressions of crop-specific GDD and total precipitation on various sets of fixed effects and reports the R^2 of the regression, the standard deviation of the residuals, and the percentage of the county-year observations with the residuals larger than certain cutoffs.

Because the research period is short, the county-fixed effects absorb a great deal of temperature variation as shown in the second row of panel 1 of Table 4.2, which will significantly damage the estimation results. To recapture some of the variation, I replace county fixed effects with city fixed effects. On average, a city has nine counties in China. This way, I remove potential confounding factors within a city, such as soil quality, large-scale infrastructure conditions, agricultural policies and support programs, etc. Consider the fifth row of panel 1 of Table 4.2, after conditioning on city fixed effects and province-by-year fixed effects, there are still 2,315 county-year observations (15.3% of total observations) and 845 county-year observations (5.57% of total observations) with residuals larger than 1 ℃ and 2 ℃, respectively. This range overlaps predictions of changes in temperature from global climate models. Therefore, the predicted impacts of temperature will be well identified from the data, rather than by out-of-sample extrapolations.

Regarding precipitation, the residual variation remains large after accounting for the fixed effects, as shown in panel 2 of Table 4.2.

4.3 Econometric Strategy and Estimation Results

Using standard panel methods, I regress the yields of three major crops (rice, wheat, and corn) on growing season degree-days, growing season total precipitation, and growing season harmful degree-days, conditional on city and province-by-year fixed effects. Following the convention in

the literature, I include both linear and quadratic terms of the climate variables. The regression equation takes the form of

$$y_{ct}^{i} = \alpha_c + \gamma_{st} + \theta_1 GDD_{ct} + \theta_2 GDD_{ct}^2 + \theta_3 Prec_{ct} + \theta_4 Prec_{ct}^2 + \theta_5 GHDD_{ct} + \varepsilon_{ct}$$

where the dependent variable is the yield of major crops (tons per hectare) in county c and year t with $i = rice, wheat, or\ corn$, GDD_{ct} and GDD_{ct}^2 are the county-year linear and quadratic terms of growing season degree-days, respectively, $Prec_{ct}$ and $Prec_{ct}^2$ are the county-year linear and quadratic terms of total precipitation during the growing season, respectively, and $GHDD_{ct}$ is the county-year growing season harmful degree-days. A full set of city fixed effect α_c is included in the equation, which absorbs all time-invariant idiosyncratic features of the city in which the county is located. γ_{st} is the province-by-year fixed effect, which captures shocks that are common across all counties within a province. By conditioning on the city and province-by-year fixed effects, the weather parameters are identified from the county-specific deviations in yearly weather from the city average climate after adjusting for shocks common to all counties in the province.

I run separate regressions for each crop. Table 4.4-4.6 report estimation results for rice, wheat, and corn, respectively. The specification of column (1) in each table regresses the corresponding crop yield on the five weather variables including growing season degree-days (GDD) and its quadratic term, growing season total precipitation and its quadratic term, and growing season harmful degree-days (GHDD), conditional on city fixed effects. Column (2) considers the irrigation conditions for each county by incorporating the ratio of irrigated land in estimation. Column (3) adds year fixed effects, and column (4) replaces year fixed effects with province-by-year fixed effects.

As shown in Table 4.4, rice yield is increasing in the linear temperature (GDD) and precipitation terms, but decreasing in the squares in all four model specifications. Once province-by-year fixed effects are included (column (4) of Table 4.4), the coefficients of both GDD and its quadratic term are significantly different from zero at 1% level, the coefficient of precipitation is significantly different from zero at 5% level, and the coefficient of precipitation square is significantly different from zero at 1% level. Figure 4.2 graphically illustrates the estimation results of the last three model specifications for county-level rice yield. The left panel of Figure 4.2 shows the relationship between rice yield and GDD, plotting the parameter estimates on GDD and GDD square for deciles of the distribution of GDD at the midpoint of each decile's range, conditional on the mean of precipitation and the percentage of irrigated land. The effect of GDD on rice yield manifests an inverted U-shape, with the optimal GDD around 2600-2800 depending on model specifications. The right panel replicates the graphical exercise in the left one but for growing season precipitation rather than GDD, which also manifests an inverted U-shape pattern and is peaking around 400-700mm. Moreover, the growing season harmful temperature (GHDD) has significantly negative effects on rice yield, while the ratio of effective irrigated land has significantly positive effects on rice yield in all four model specifications.

As shown in Table 4.5, wheat yield is also significantly increasing with respect to linear GDD and decreasing with respect to GDD square in all four model specifications, but precipitation has a negative effect on wheat yield, and this effect turns insignificant when province-by-year fixed effects are included. Therefore, the relationship between wheat yield and

GDD is inverted U-shaped as illustrated in the left panel of Figure 4.3, with the optimal GDD around 1900-2100, depending on model specifications. The relationship between wheat yield and growing season precipitation is decreasing during the research period as illustrated in the right panel of Figure 4.3. The effect of GHDD is negative but not significant on wheat yield in all four model specifications. Irrigation also has significantly positive effects on wheat yield.

For corn, the coefficient of GDD is positive and the coefficient of GDD square is negative in all four model specifications. However, the effects of GDD and its quadratic term is only significantly different from zero at 10% level when province-by-year fixed effects are included, and is not significant if year fixed effects are included (Table 4.6). As illustrated in the left panel of Figure 4.4, the inverted U-shaped relationship between corn yield and GDD turns flatter when year or province-by-year fixed effects are accounted for. The optimal GDD for corn is around 1750-2300, depending on model specifications. The coefficients of growing season precipitation and its square are significantly positive at 1% and significantly negative at 1% level, respectively, across all four model specifications, which manifests a stable inverted U-shaped relationship between corn yield and growing season precipitation as shown in the right panel of Figure 4.4. The optimal growing season precipitation for corn is around 500-600mm. The effect of GHDD is not significant on corn yield in all four model specifications. Irrigation also has significantly positive effects on corn yield.

4.4 Impacts of Climate Change

Using the estimation results from model specification (4), I evaluate the potential effects of future climate change on the yields of three major crops, namely, rice, wheat, and corn in China.

The effects of climate change are calculated in this section as the discrete difference between the predicted yields at the projected temperature and precipitation scenarios and the predicted yields at the historical (1960-2000) climate. I consider two representative concentration pathways (RCPs) scenarios (RCP8.5 and 4.5) over the period from 2040 to 2060, both of which are used by the IPCC for its fifth Assessment Report (AR5) in 2014. RCP8.5 represents a scenario with high greenhouse gas concentration levels, and RCP4.5 represents a stabilization scenario. Growing season average temperature and precipitation for each crop are drawn from Wei et. al (2014), which are calculated by aggregating gridded data from an existing downscaling climate data set. As shown in Table 4.6, the growing season average temperature for rice, wheat, and corn increases around 3 ℃ and 2 ℃ under scenario RCP8.5 and RCP4.5, respectively. And the growing season precipitation increases around 7-8% and 4-5% under scenario RCP8.5 and RCP4.5, respectively.

Table 4.7 presents the impacts of climate change on the three crops yields in China. In general, the impacts of climate change over the short to medium term will be negative on rice and corn yields, but positive on wheat yield for the two scenarios. Specifically, the effects of precipitation increase on rice yield are mildly positive and outweighed by negative temperature effects for the two scenarios, which will lead to an aggregate 1.18% and 0.08% decrease of rice yield for the RCP8.5 and RCP4.5 scenarios, respectively. The effects of precipitation increase on wheat yield are mildly negative and outweighed by positive temperature effects for the two scenarios, which will lead to an aggregate 6.68% and 5.48% increase of wheat yield for the

RCP8.5 and RCP4.5 scenarios, respectively. The effects of both temperature and precipitation are negative on corn yields, leading to 2.21% and 1.64% decrease for the RCP8.5 and RCP4.5 scenarios, respectively. To provide some intuition of the results, I multiply the estimated yield changes in each crop by the respective average total planted acreage of all sample counties over the research period (2000-2010) to get an estimate of the changes in total production. For the RCP8.5 scenario, the total productions of rice and corn will decrease by 1.30×10^6 tons and 1.63×10^6 tons, respectively, and the total production of wheat will increase by 2.99×10^6 tons. For the RCP4.5 scenario, the total productions of rice and corn will decrease by 0.85×10^6 tons and 1.21×10^6 tons, respectively, and the total production of wheat will increase by 2.45×10^6 tons. I then multiply each crop's total production change by the respective average market price by the end of 2010 to get a rough estimate of economic impacts of climate change. For the RCP8.5 scenario, the rice sector and corn sector will suffer economic losses of around $\$524\times 10^6$ and $\$487\times 10^6$, respectively, while the wheat sector will gain $\$955\times 10^6$. For the RCP4.5 scenario, the rice sector and corn sector will suffer economic losses of $\$343\times 10^6$ and $\$361\times 10^6$, respectively, while the wheat sector will gain $\$783\times 10^6$.

One criticism to the panel data approach is that it relies on year-to-year variation in weather and cannot capture the full range of adaptations of farmers in response to long-term climate change, therefore, it may overstate the damage associated with climate change (Fisher, Hanemann, Roberts, and Shlenker, 2007). For comparison, I also explore the relationship between climate variables and crop yields using a cross sectional model and run regressions for each single year separately and for samples pooled from 2000 to 2010. Appendix C provides more details on the cross sectional estimation. The results show significant variation in the estimated impact of climate change on crop yields, depending on model specification and data sources. Moreover, the point estimates of climate change impacts on the three crop yields from panel approach are covered in the range of impacts generated from cross sectional estimations. For example, for rice, the point estimate of climate change impacts under the comparatively high emission scenario (RCP8.5) is -0.081 using the panel approach, while the cross sectional estimates range between -0.503 and 0.119. For wheat, the panel estimate is 0.236, and the cross sectional estimates range between -0.407 and 0.439. And for corn, the panel estimate is -0.115, and the cross sectional estimates range between -0.027 and 0.713. These results imply that the concern to the panel approach may not be that severe in the context of Chinese agriculture. One possible reason is that it is not easy for small-scale farmers to adapt to climate changes quickly in short time. Therefore, the panel estimation results could plausibly predict the short to medium-run impacts of climate change on crop yields in China.

4.5 Conclusion

This essay estimates the effects of random year-to-year variation in weather on three major crops (rice, wheat, and corn) yields in China, using an 11-year county-level panel data set covering more than 1,000 counties.

The estimation results are overall as expected. The relationship between rice yield and growing season degree-days (GDD) in a range of 8-32 ℃ is inverted U-shaped, with the optimal GDD around 2600-2800. The relationship between rice yield and growing season precipitation is

also inverted U-shaped peaking around 400-700mm. The relationship between wheat yield and GDD is also inverted U-shaped with the optimal GDD around 1900-2100, but wheat yield is generally decreasing with respect growing season precipitation during the research period. Corn yield is also following the inverted U-shaped with respect to both GDD and growing season precipitation, with the optimal GDD for corn around 1750-2300 and optimal precipitation around 500-600mm.

The estimation results are then used to predict the impacts of future climate change on the three crops yields in China. During short to medium term (2040-2060), the average rice yield in China is expected to decrease by 1.18% under the comparatively high emission scenario and by 0.08% under a medium-low scenario. The wheat yield is expected to increase by 5.48% and 6.68% under the two emission scenarios, respectively. And the corn yield is expected to decrease by 1.64% and 2.21% under the two emission scenarios, respectively. These yield effects imply an economic loss of $\$343\times10^6$ to $\$524\times10^6$ in the rice sector, and $\$361\times10^6$ to $\$487\times10^6$ in the corn sector, while an economic gain of $\$783\times10^6$ to $\$955\times10^6$ in the wheat sector. The adverse effects in the rice and corn sectors could be mitigated through long run adaptation actions, such as adopting stress-resistant varieties, switching to new crops well-adaptive to local climate conditions, improving field infrastructure (e.g. irrigation systems), and even exiting from farming.

4.6 Tables and Figures

Figure 4.1: Locations of the Weather Stations

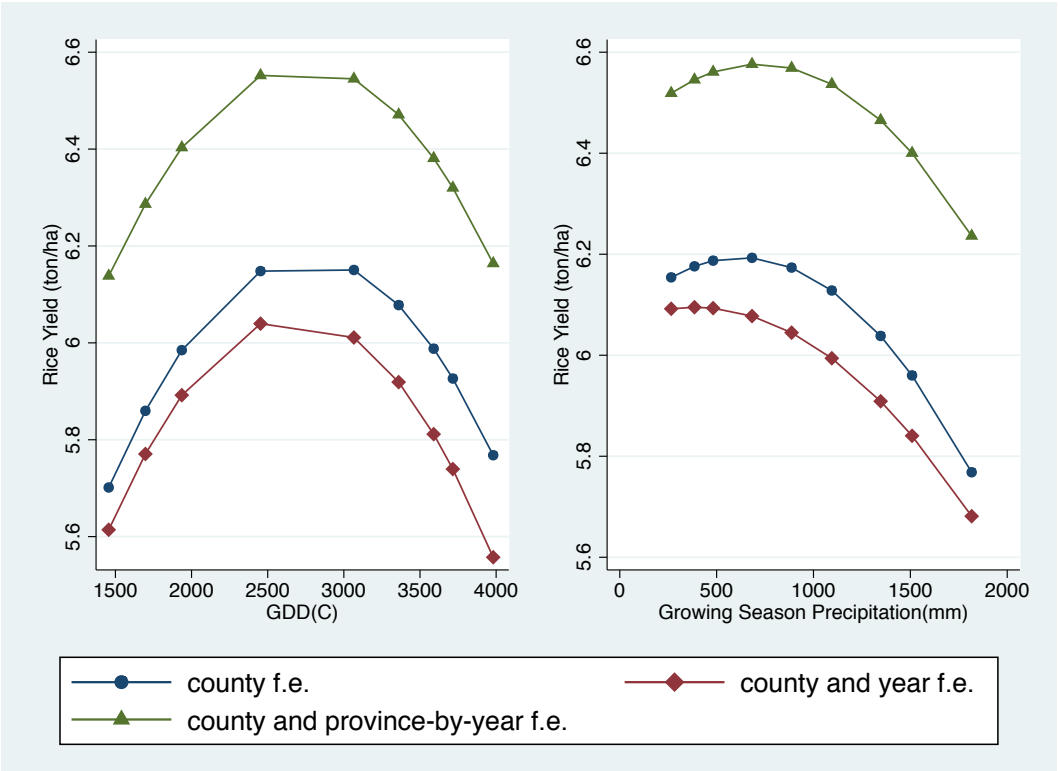

Figure 4.2: Estimated Relationship Between Rice Yield and Weather Variables

Notes: The left panel of this figure plots the parameter estimates of rice yield on GDD and GDD square for deciles of the distribution of GDD at the midpoint of each decile's range. The three lines represent the regressions that include county fixed effects, county and year fixed effects, and county and province-by-year fixed effects as labeled, respectively. The right panel replicates the graphical exercise in the left panel, except for growing season precipitation rather than GDD.

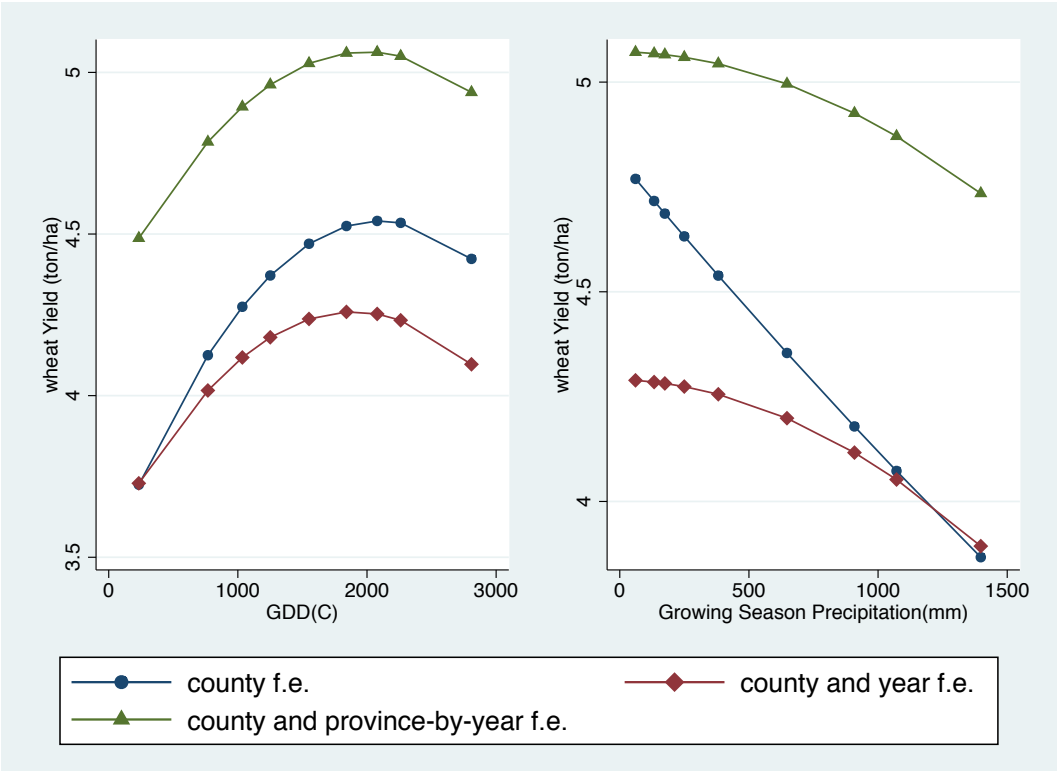

Figure 4.3: Estimated Relationship Between Wheat Yield and Weather Variables

Notes: This figure replicates the graphical exercise in Figure 4.2, except for wheat yield rather than rice yield.

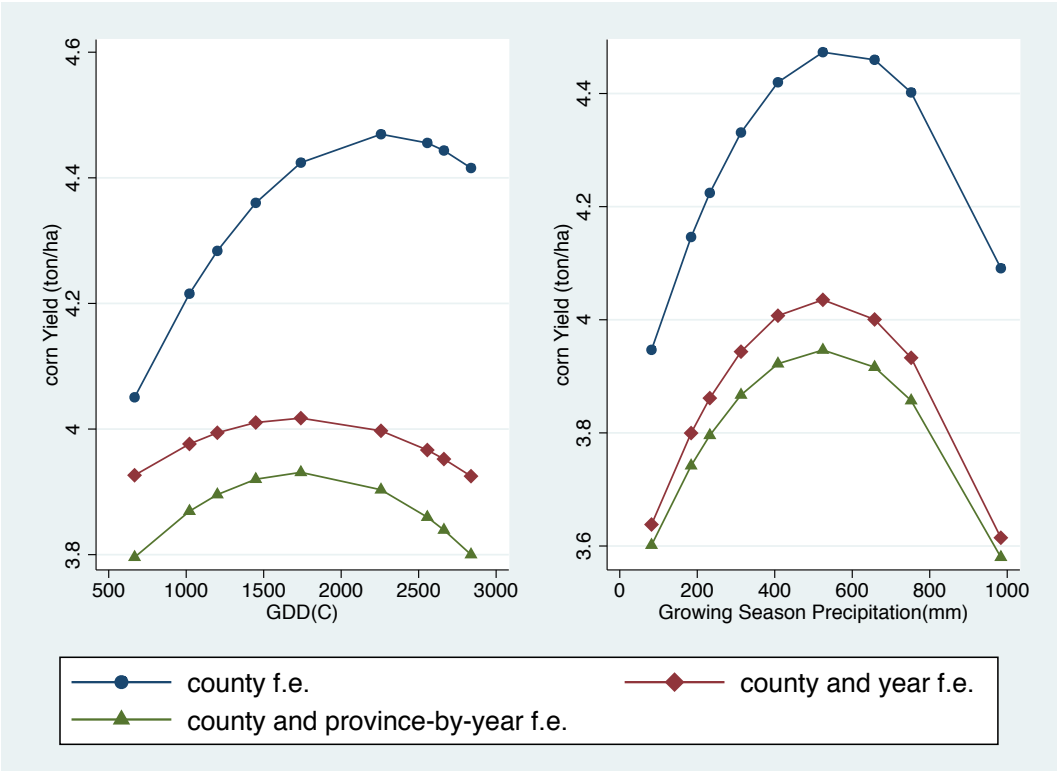

Figure 4.4: Estimated Relationship Between Corn Yield and Weather Variables

Notes: This figure replicates the graphical exercise in Figure 4.2, except for corn yield rather than rice yield.

Table 4.1: County-Level Summary Statistics (2000-2010)

Variable	Mean	St. dev.	Min	Max
Dependent Variable: Crop Yield (ton/ha)				
Rice	6.70	1.44	0.32	13
Wheat	3.77	1.63	0.013	9.80
Corn	5.15	1.76	0.013	15.92
County Characteristics				
Rural Population (10,000)	37.57	30.28	0.002	653.98
Total Cultivated Land (1,000 ha)	42.49	52.43	0.002	1478.37
Ag. Income (10^6 yuan)	692.04	823.04	0.02	28995.22
Ave. Fertilizer Input (ton/ha)	0.94	24.47	0.0001	2321.06
Ave. Pesticide Input (ton/ha)	0.24	4.12	0	244.34
Share of Cropland Irrigated	0.67	0.32	0	1
Weather Variables				
Rice				
GDD (1000C)	2.896	0.631	0.326	4.309
Growing Season Prec (1000mm)	0.907	0.338	0.013	3.401
Wheat				
GDD (1000C)	1.560	0.462	0.160	4.161
Growing Season Prec (1000mm)	0.476	0.302	0.003	2.103
Corn				
GDD (1000C)	1.820	0.521	0.089	3.025
Growing Season Prec (1000mm)	0.431	0.178	0.005	1.607

Notes: Summary statistics generated using 2000-2010 data for 1378 counties used to estimate the model.

Table 4.2: Residual Variation in Weather Variables

	GDD					
	R^2	σ_e	|e|>0.5C	|e|>1C	|e|>1.5C	|e|>2C
No F.E.		3.30C	87.8%	76.2%	65.2%	54.6%
County F.E.	0.984	0.39C	20.1%	1%	0.05%	0.01%
County and Year F.E.	0.988	0.35C	14.7%	0.76%	0.03%	0
County and Province-by-Year F.E.	0.996	0.17C	2.5%	0.09%	0.01%	0
City and Province-by-Year F.E.	0.905	1.02C	33.8%	15.3%	9.1%	5.57%
	Precipitation					
	R^2	σ_e	|e|>25mm	|e|>35mm	|e|>45mm	|e|>55mm
No F.E.		336.8mm	94.7%	92.3%	90.1%	87.9%
County F.E.	0.794	163.7mm	84.8%	79.1%	73.7%	68.5%
County and Year F.E.	0.776	159.5mm	84.5%	78.8%	73.1%	68.0%
County and Province-by-Year F.E.	0.871	121.2mm	78.5%	70.5%	63.1%	55.9%
City and Province-by-Year F.E.	0.823	141.6mm	98.9%	98.5%	98.1%	97.8%

Notes: This table summarizes regressions of growing season degree-days (GDD) and growing season total precipitation on various sets of fixed effects and reports the R^2 of the regression, the standard deviation of the residuals, and the percentage of the county-year observations with the residuals larger than some cutoffs. The regressions are based on 15,158 county-year observations.

Table 4.3: Estimated Effects of Weather Variables on Rice Yields

	Dependent Variable: Rice Yield (ton/ha)			
VARIABLES	(1)	(2)	(3)	(4)
GDD (1000C)	1.514***	1.527***	1.576***	1.451***
	(0.338)	(0.337)	(0.333)	(0.336)
GDD^2 ($1000C^2$)	-0.274***	-0.276***	-0.294***	-0.265***
	(0.0602)	(0.0600)	(0.0599)	(0.0610)
Precipitation (1000mm)	0.386*	0.378*	0.158	0.405**
	(0.202)	(0.201)	(0.195)	(0.194)
Precipitation ($1000mm^2$)	-0.305***	-0.301***	-0.203***	-0.282***
	(0.0805)	(0.0804)	(0.0779)	(0.0735)
GHDD (C)	-0.0977***	-0.0977***	-0.0930***	-0.0445***
	(0.0234)	(0.0235)	(0.0220)	(0.0114)
Irrigated Land (%)		0.129**	0.352***	0.348***
		(0.0519)	(0.0536)	(0.0589)
City F.E.	Yes	Yes	Yes	Yes
Year F.E.	No	No	Yes	No
Province-by-Year F.E.	No	No	No	Yes
Constant	4.005***	3.884***	3.733***	4.220***
	(0.495)	(0.494)	(0.492)	(0.508)

Notes: The dependent variable is rice yield measured as ton/ha. Results presented in all four columns are estimated using 858 rice production counties 2000-2010 data. Robust standard errors are in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 4.4: Estimated Effects of Weather Variables on Wheat Yields

	Dependent Variable: Wheat Yields (ton/ha)			
VARIABLES	(1)	(2)	(3)	(4)
GDD (1000C)	1.009***	0.983***	0.730***	0.744***
	(0.125)	(0.124)	(0.124)	(0.123)
GDD^2 ($1000C^2$)	-0.241***	-0.234***	-0.193***	-0.187***
	(0.0313)	(0.0313)	(0.0317)	(0.0306)
Precipitation (1000mm)	-0.530***	-0.739***	-0.0200	-0.0126
	(0.165)	(0.164)	(0.163)	(0.187)
$Precipitation^2$ ($1000mm^2$)	-0.0373	0.0440	-0.189**	-0.164
	(0.0878)	(0.0873)	(0.0871)	(0.105)
GHDD (C)	-0.0590	-0.0265	-0.111	-0.0572
	(0.184)	(0.157)	(0.127)	(0.125)
Irrigated Land (%)		0.494***	0.723***	1.009***
		(0.0365)	(0.0396)	(0.0451)
City F.E.	Yes	Yes	Yes	Yes
Year F.E.	No	No	Yes	No
Province-by-Year F.E.	No	No	No	Yes
Constant	3.881***	3.514***	3.130***	3.681***
	(0.187)	(0.183)	(0.194)	(0.419)

Notes: The dependent variable is wheat yield measured as ton/ha. Results presented in all four columns are estimated using 955 wheat production counties 2000-2010 data. Robust standard errors are in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 4.5: Estimated Effects of Weather Variables on Corn Yields

	Dependent Variable: Corn Yields (ton/ha)			
VARIABLES	(1)	(2)	(3)	(4)
GDD (1000C)	0.746***	0.742***	0.272	0.398*
	(0.213)	(0.210)	(0.206)	(0.211)
GDD^2 ($1000C^2$)	-0.167***	-0.164***	-0.0778	-0.113*
	(0.0639)	(0.0634)	(0.0621)	(0.0652)
Precipitation (1000mm)	2.499***	2.548***	2.117***	1.840***
	(0.267)	(0.271)	(0.271)	(0.288)
$Precipitation^2$ ($1000mm^2$)	-2.187***	-2.241***	-2.011***	-1.749***
	(0.239)	(0.243)	(0.245)	(0.245)
GHDD (C)	-0.0135	-0.0125	-0.0202	0.000416
	(0.0157)	(0.0154)	(0.0149)	(0.0145)
Irrigated Land (%)		0.427***	0.697***	0.750***
		(0.0471)	(0.0504)	(0.0561)
City F.E.	Yes	Yes	Yes	Yes
Year F.E.	No	No	Yes	No
Province-by-Year F.E.	No	No	No	Yes
Constant	3.016***	2.660***	2.774***	2.610***
	(0.292)	(0.293)	(0.285)	(0.573)

Notes: The dependent variable is corn yield measured as ton/ha. Results presented in all four columns are estimated using 1232 corn production counties 2000-2010 data. Robust standard errors are in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 4.6: Climate Predictions under Different Scenarios

	Temperature			Precipitation		
	1960-2000	2040-2060		1960-2000	2040-2060	
		RCP8.5	RCP4.5		RCP8.5	RCP4.5
Rice	20.9	23.4	22.9	613.7	662.2	645.2
Wheat	10.7	13.4	12.7	416.6	444.1	438.2
Corn	21.7	24.3	23.8	561.2	607.2	584.4

Notes: This table is drawn from Table 5 in Wei et. al. (2014). It shows mean of growing season temperature and precipitation during the historical period (1960-2000) and short to medium-run period (2040-2060) under two RCP scenarios (RCP8.5 and RCP4.5).

Table 4.7: Projected Impact of Climate Change on Crop Yields

	RCP8.5		
	Rice	Wheat	Corn
Temperature Effect	-0.083	0.243	-0.105
	(0.266)	(0.217)	(0.354)
Precipitation Effect	0.002	-0.004	-0.009
	(0.268)	(0.223)	(0.326)
Total Effect	-0.081	0.239	-0.115
	(0.266)	(0.217)	(0.354)
	RCP4.5		
	Rice	Wheat	Corn
Temperature Effect	-0.054	0.199	-0.081
	(0.266)	(0.218)	(0.346)
Precipitation Effect	0.002	-0.003	-0.004
	(0.268)	(0.223)	(0.326)
Total Effect	-0.053	0.196	-0.085
	(0.266)	(0.218)	(0.346)

Notes: This table is generated using estimation results from model specification (4) for each crop. The effects of climate change are calculated as the discrete difference between the predicted yields at the projected temperature and precipitation scenarios and the predicted yields at the historical climate. Standard errors are in parentheses.

Bibliography

Arimura, Toshi H. "An empirical study of the SO_2 allowance market: effects of PUC regulations." *Journal of Environmental Economics and Management* 44, no. 2 (2002): 271-289.

Banzhaf, H. Spencer, and Randall P. Walsh. "Do people vote with their feet? An empirical test of Tiebout's mechanism." *The American Economic Review* (2008): 843-863.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan. *A unified framework for measuring preferences for schools and neighborhoods*. No. w13236. National Bureau of Economic Research, 2007.

Bayer, Patrick, Robert McMillan, and Kim Rueben. *An equilibrium model of sorting in an urban housing market*. No. w10865. National Bureau of Economic Research, 2004.

Berry, Steven, James Levinsohn, and Ariel Pakes. "Automobile prices in market equilibrium." *Econometrica: Journal of the Econometric Society* (1995): 841-890.

Burke, Marshall, and Kyle Emerick. "Adaptation to climate change: Evidence from US agriculture." *Available at SSRN 2144928* (2012).

Busse, Meghan R., and Nathaniel O. Keohane. "Market effects of environmental regulation: coal, railroads, and the 1990 Clean Air Act." *The RAND Journal of Economics* 38, no. 4 (2007): 1159-1179.

Carlson, Curtis, Dallas Burtraw, Maureen Cropper, and Karen L. Palmer. "Sulfur dioxide control by electric utilities: What are the gains from trade?." *Journal of political Economy* 108, no. 6 (2000): 1292-1326.

CHINA, I. "China's National Climate Change Programme." *People's Republic of China: National Development and Reform Commission* (2007).

Deschenes, Olivier, and Michael Greenstone. "The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather." *The American Economic Review* (2007): 354-385.

Dubin, Jeffrey A., and Daniel L. McFadden. "An econometric analysis of residential electric appliance holdings and consumption." *Econometrica: Journal of the Econometric Society* (1984): 345-362.

Epple, Dennis, Thomas Romer, and Holger Sieg. "Interjurisdictional sorting and majority rule: an empirical analysis." *Econometrica* 69, no. 6 (2001): 1437-1465.

Epple, Dennis, and Holger Sieg. *Estimating equilibrium models of local jurisdictions*. No. w6822. National Bureau of Economic Research, 1998.

Farrow, R. Scott, Martin T. Schultz, Pinar Celikkol, and George L. Van Houtven. "Pollution trading in water quality limited areas: Use of benefits assessment and cost-effective trading ratios." *Land Economics* 81, no. 2 (2005): 191-205.

Ferreyra, Maria Marta. "Estimating the effects of private school vouchers in multidistrict economies." *The American Economic Review* (2007): 789-817.

Fisher, Anthony, Michael Hanemann, Michael Roberts, and Wolfram Schlenker. "Potential impacts of climate change on crop yields and land values in US agriculture: negative, significant, and robust." *University of California, Berkeley*(2007).

Fowlie, Meredith. "Emissions trading, electricity restructing, and investment in pollution abatement." *The American Economic Review* (2010): 837-869.

Fowlie, Meredith, and Nicholas Muller. *Market-based emissions regulation when damages vary across sources: What are the gains from differentiation?*. No. w18801. National Bureau of Economic Research, 2013.

Guiteras, Raymond. "The impact of climate change on Indian agriculture." Manuscript, Department of Economics, University of Maryland, College Park, Maryland (2009).

Fullerton, Don, Shaun P. McDermott, and Jonathan P. Caulkins. "Sulfur dioxide compliance of a regulated utility." *Journal of Environmental Economics and Management* 34, no. 1 (1997): 32-53.

Heckman, James J. "Sample selection bias as a specification error." *Econometrica: Journal of the econometric society* (1979): 153-161.

Joskow, Paul L., and Richard Schmalensee. "The Political Economy of Market-Based Environmental Policy: the US Acid Rain Program 1." *The journal of law and economics* 41, no. 1 (1998): 37-84.

Keohane, Nathaniel O. "Environmental policy and the choice of abatement technique: evidence from coal-fired power plants." In *2nd World Congress of Environmental and Resource Economists, Monterrey, CA*. 2002.

Klaiber, H. Allen, and Daniel J. Phaneuf. "Valuing open space in a residential sorting model of the Twin Cities." *Journal of Environmental Economics and Management* 60, no. 2 (2010): 57-77.

Krupnick, Alan J., Virginia D. McConnell, Matthew Cannon, Terrell Stoessell, and Michael Batz. *Cost-effective NOx control in the eastern United States*. Resources for the Future, 2000.

Kuminoff, Nicolai V. "Decomposing the structural identification of non-market values." *Journal of Environmental Economics and Management* 57, no. 2 (2009): 123-139.

Liu, Hui, Xiubin Li, Guenther Fischer, and Laixiang Sun. "Study on the impacts of climate change on China's agriculture." *Climatic Change* 65, no. 1-2 (2004): 125-148.

Lobell, David B., Wolfram Schlenker, and Justin Costa-Roberts. "Climate trends and global crop production since 1980." *Science* 333, no. 6042 (2011): 616-620.

McCarl, Bruce A., Xavier Villavicencio, and Ximing Wu. "Climate change and future analysis: is stationarity dying?." *American Journal of Agricultural Economics* 90, no. 5 (2008): 1241-1247.

Mendelsohn, Robert. "Regulating heterogeneous emissions." *Journal of Environmental Economics and Management* 13, no. 4 (1986): 301-312.

Mendelsohn, Robert, William D. Nordhaus, and Daigee Shaw. "The impact of global warming on agriculture: a Ricardian analysis." *The American economic review* (1994): 753-771.

Montero, Juan-Pablo. "Voluntary compliance with market-based environmental policy: Evidence from the US Acid Rain Program." *Journal of Political Economy*107, no. 5 (1999): 998-1033.

Montgomery, W. David. "Markets in licenses and efficient pollution control programs." *Journal of economic theory* 5, no. 3 (1972): 395-418.

Muller, Nicholas Z., and Robert Mendelsohn. "Efficient pollution regulation: getting the prices right." *The American Economic Review* (2009): 1714-1739.

Newell, Richard G., and William A. Pizer. "Carbon mitigation costs for the commercial building sector: Discrete–continuous choice analysis of multifuel energy demand." *Resource and Energy Economics* 30, no. 4 (2008): 527-539.

Newey, Whitney K. "Nonparametric Continuous/Discrete Choice Models." *International Economic Review* 48, no. 4 (2007): 1429-1439.

Ritchie, J. T., and D. S. NeSmith. "Temperature and crop development." *Modeling plant and soil systems. Agron. Monogr* 31 (1991): 5-29.

Schlenker, Wolfram, W. Michael Hanemann, and Anthony C. Fisher. "Will US agriculture really benefit from global warming? Accounting for irrigation in the hedonic approach." *American Economic Review* (2005): 395-406.

Schlenker, Wolfram, W. Michael Hanemann, and Anthony C. Fisher. "The impact of global warming on US agriculture: an econometric analysis of optimal growing conditions." *Review of Economics and Statistics* 88, no. 1 (2006): 113-125.

Schlenker, Wolfram, and Michael J. Roberts. "Nonlinear effects of weather on corn yields." *Applied Economic Perspectives and Policy* 28, no. 3 (2006): 391-398.

Schlenker, Wolfram, and Michael J. Roberts. "Nonlinear temperature effects indicate severe damages to US crop yields under climate change." *Proceedings of the National Academy of sciences* 106, no. 37 (2009): 15594-15598.

Schmalensee, Richard, and Robert Stavins. "The SO2 allowance trading system: The ironic history of a grand policy experiment." *Journal of Economic Perspectives* 27, no. 1 (2013): 103-22.

Tao, Fulu, Masayuki Yokozawa, and Zhao Zhang. "Modelling the impacts of weather and climate variability on crop productivity over a large area: a new process-based model development, optimization, and uncertainties analysis." *Agricultural and Forest Meteorology* 149, no. 5 (2009): 831-850.

Tao, Fulu, and Zhao Zhang. "Climate change, wheat productivity and water use in the North China Plain: A new super-ensemble-based probabilistic projection." *Agricultural and Forest Meteorology* 170 (2013): 146-165.

Tietenberg, Tom. "Tradeable permits for pollution control when emission location matters: what have we learned?." *Environmental and Resource Economics* 5, no. 2 (1995): 95-113.

Timmins, Christopher. "If You Cannot Take the Heat, Get Out of the Cerrado...Recovering the Equilibrium Amenity Cost of Nonmarginal Climate Change in Brazil." *Journal of Regional Science* 47, no. 1 (2007): 1-25.

Walsh, Randy. "Endogenous open space amenities in a locational equilibrium." *Journal of urban Economics* 61, no. 2 (2007): 319-344.

Wang, Jinxia, Robert Mendelsohn, Ariel Dinar, Jikun Huang, Scott Rozelle, and Lijuan Zhang. "The impact of climate change on China's agriculture." *Agricultural Economics* 40, no. 3 (2009): 323-337.

Wei, Taoyuan, Todd L. Cherry, Solveig Glomrød, and Tianyi Zhang. "Climate change impacts on crop yield: Evidence from China." *Science of the Total Environment* 499 (2014): 133-140.

Wu, Wenjie, Wenzhong Zhang, and Guanpeng Dong. "Determinant of residential location choice in a transitional housing market: Evidence based on micro survey from Beijing." *Habitat International* 39 (2013): 16-24.

Xiong, Wei, Erda Lin, Hui Ju, and Yinlong Xu. "Climate change and critical thresholds in China's food security." *Climatic Change* 81, no. 2 (2007): 205-221.

Ye, Liming, Wei Xiong, Zhengguo Li, Peng Yang, Wenbin Wu, Guixia Yang, Yijiang Fu et al. "Climate change impact on China food security in 2050." *Agronomy for Sustainable Development* 33, no. 2 (2013): 363-374.

Zheng, Siqi, and Matthew E. Kahn. "Land and residential property markets in a booming economy: New evidence from Beijing." *Journal of Urban Economics*63, no. 2 (2008): 743-757.

APPENDICES

APPENDIX A. Estimates of Scrubbing Capital Cost

I use Heckman's two-step method to correct for selection bias in estimating the scrubbing capital cost equation. The scrubbing capital cost $ln(K_{n1})$ is expressed as a function of L unit characteristics as

$$ln(K_{n1}) = \beta_0 + \sum_{l=1}^{L} \beta_l \cdot ln(X_{nl}) + \varepsilon_n \tag{A.1}$$

where X_{nl} represents a vector of unit characteristics, and $\varepsilon_n \sim N(0,\sigma_\varepsilon^2)$. Guided by engineering models (EPA IPM 2010; IAPCS 1999), I choose baseline unit nameplate capacity, coal sulfur content, and heat rate as X_{nl} in the equation.

A unit chooses to install a scrubber and therefore its scrubbing capital cost is observed if the anticipated total generating cost with a scrubber (C_{n1}) is smaller than the cost with switching/blending (C_{n0}) under the influences of certain regulation factors, that is,

$$C_{n0} > C_{n1} \cdot exp\{\eta \cdot R_n\} \tag{A.2}$$

where R_n represents other factors affecting unit decisions, including boiler age by 1990, and a dummy variable $PROPECT_n = 1$ if the unit (plant) is located in the states of Kentucky, Illinois, Indiana, Ohio, or Pennsylvania, where the burning of high-sulfur coal is protected (Arimura, 2002).

Taking the log of both sides of (A.2) gives

$$ln(C_{n0}) > ln(C_{n1}) + \eta \cdot R_n \tag{A.3}$$

The cost functions are specified in a translog form as follows,

$$\begin{aligned} ln(C_{n0}) = {} & \alpha_0 + \gamma_2 \cdot ln(w_{n0}^h) + \gamma_3 \cdot ln(w_{n0}^l) + \gamma_4 \cdot ln(y_n) + \frac{1}{2}\gamma_5 \cdot (ln(w_{n0}^h))^2 \\ & + \frac{1}{2}\gamma_6 \cdot (ln(w_{n0}^l))^2 + \frac{1}{2}\gamma_7 \cdot (ln(y_n))^2 + \frac{1}{2}\gamma_8 \cdot ln(w_{n0}^h) \cdot ln(w_{n0}^l) \\ & + \frac{1}{2}\gamma_9 \cdot ln(w_{n0}^h) \cdot ln(y_n) + \frac{1}{2}\gamma_{10} \cdot ln(w_{n0}^l) \cdot ln(y_n) + \xi_{n0} \end{aligned} \tag{A.4}$$

and

$$\begin{aligned} ln(C_{n1}) = {} & \alpha_1 + \gamma_1 \cdot ln(K_{n1}) + \gamma_2 \cdot ln(w_{n1}^h) + \gamma_4 \cdot ln(y_n) + \frac{1}{2}\gamma_5 \cdot (ln(w_{n1}^h))^2 \\ & + \frac{1}{2}\gamma_7 \cdot (ln(y_n))^2 + \frac{1}{2}\gamma_9 \cdot ln(w_{n1}^h) \cdot ln(y_n) + \xi_{n1} \end{aligned} \tag{A.5}$$

where w_{n0}^{h} and w_{n0}^{l} are the allowance-price-included prices for high-sulfur coal and low-sulfur coal, respectively, when unit n chooses non-scrubbing; w_{n1}^{h} is the allowance-price-included price for high-sulfur coal when unit n chooses scrubbing; y_n is the electricity output; and the error terms are assumed to be normally distributed and independent across units, that is, $\xi_{n0} \sim N(0,\sigma_0^2)$, and $\xi_{n1} \sim N(0,\sigma_1^2)$.

Therefore, the probability of unit n choosing scrubbing is

$$Pr_{n0} = Prob(ln(C_{n0}) - ln(C_{n1}) - \eta_1 \cdot PROTECT_n - \eta_2 \cdot age\_by90_n > 0\} \tag{A.6}$$

Substituting the expressions of $ln(C_{n0})$, $ln(C_{n1})$, and $ln(K_{n1})$ into A.6, and rearranging the equation, I have

$$Pr_{n1} = Prob\{\varepsilon_n - \nu_n < \mathbf{z_n}\delta\} \tag{A.7}$$

where

$$\begin{aligned}\mathbf{z_n}\delta = &(\frac{\alpha_1 - \alpha_0}{\gamma_1} - \beta_0) - \sum_{l=1}^{L}\beta_l \cdot ln(X_{nl}) - \frac{\gamma_2}{\gamma_1}\cdot \Delta ln(w_n^h) + \frac{\gamma_3}{\gamma_1}\cdot ln(w_{n0}^l) \\ &- \frac{1}{2}\frac{\gamma_5}{\gamma_1}\cdot \Delta(ln(w_n^h))^2 + \frac{1}{2}\frac{\gamma_6}{\gamma_1}\cdot (ln(w_{n0}^l))^2 + \frac{1}{2}\frac{\gamma_8}{\gamma_1}\cdot ln(w_{n0}^h)\cdot ln(w_{n0}^l) \\ &- \frac{1}{2}\frac{\gamma_9}{\gamma_1}\cdot \Delta ln(w_n^h)\cdot ln(y_n) + \frac{1}{2}\frac{\gamma_{10}}{\gamma_1}\cdot ln(w_{n0}^l)\cdot ln(y_n) - \frac{\eta_1}{\gamma_1}\cdot PROTECT_n \\ &- \frac{\eta_2}{\gamma_1}\cdot age\_by90_n + \nu_n\end{aligned} \tag{A.8}$$

with $\Delta ln(w_n^h) = ln(w_{n1}^h) - ln(w_{n0}^h)$ and $\Delta(ln(w_n^h))^2 = (ln(w_{n1}^h))^2 - (ln(w_{n0}^h))^2$.

Moreover, $\nu_n = \frac{\xi_{n0} - \xi_{n1}}{\gamma_1} \sim N(0,\sigma_\nu^2)$, and $\varepsilon_n \sim N(0,\sigma_\varepsilon^2)$, so I have $(\varepsilon_n - \nu_n) \sim N(0,\sigma^2)$ with $\sigma^2 = \sigma_\varepsilon^2 + \sigma_\nu^2 - 2\sigma_{\varepsilon\nu}$.

Define $u_n = \frac{\varepsilon_n - \nu_n}{\sigma}$ and $\delta^* = \frac{\delta}{\sigma}$, I can further rewrite equation (2.A.7) as

$$Pr_{n1} = Prob\{u_n < \mathbf{z_n}\delta^*\} = \Phi(\mathbf{z_n}\delta^*) \tag{A.9}$$

where $\Phi(\cdot)$ denotes the standard normal cumulative distribution function.

Following Heckman's two-step method (Heckman, 1979), I conduct the estimation in the following steps: 1) Estimate equation (A.9) using the probit model; 2) Generate estimates of the hazard function $\frac{\phi(\mathbf{z_n}\widehat{\delta^*})}{\Phi(\mathbf{z_n}\widehat{\delta^*})}$; 3) Run the OLS of the scrubbing cost equation (A.1) with the estimated hazard function on the right-hand side to account for the selection bias and generate an

estimate for the scrubbing capital cost $ln(\widehat{K_{n1}})$ for all units; and 4) Estimate the original non-scrubbing choice equation (2.14) using $ln(\widehat{K_{n1}})$ and other observed variables.

To identify the coefficients in the scrubbing capital cost equation (A.1) using Heckman's two-step method, a valid exclusion restriction is needed. Here, the excluded variables are the prices of high-sulfur coal and low-sulfur coal, which significantly affect the selection to scrubbing, but do not directly affect the scrubbing capital cost.

APPENDIX B. First-Stage and Reduced Form Regressions

The first-stage relationship between the distance to the Powder River Basin (PRB) and the price premium of low-sulfur coal is strongly positive, as shown in Table B.1. The coefficient of $ln(distance\ to\ PRB)$ is positive and significantly different from zero at the 1% level. Longer distances to PRB are associated with significantly higher SO_2 emission rates in the reduced-form regression, as shown in Table B.2.

Table B.1: Distance to PRB and Price Premium of Low-Sulfur Coal (First-Stage)

Explanatory Variable	OLS
$ln(distance\ to\ PRB)$	0.17***
	(0.014)
$ln(y)$	-0.005
	(0.007)
$ln(capacity)$	0.02
	(0.012)
$ln(sulfur\ content)$	-0.015*
	(0.009)
$ln(heat\ rate)$	0.028**
	(0.011)
$ln(age\ by 90)$	0.034*
	(0.018)
Year fixed effects	Y
Mills	0.117***
	(0.026)
cons	-0.995***
	(0.167)

Notes: This table presents estimation results of the first-stage regression of the 2SLS estimation for the emission rate equation. The dependent variable is the price premium of low-sulfur coal of Phase I non-scrubbing units from 1995 to 2003, and of Phase II non-scrubbing units from 2000 to 2003. The instrument variable is $ln(Distance\ to\ PRB)$. Unit random effects are include. Robust standard errors are in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table B.2: Distance to PRB and Emission Rate (Reduced-Form)

Explanatory Variable	OLS
$ln(distance\ to\ PRB)$	0.589
	(0.101)***
$ln(y)$	0.057
	(0.028)**
$ln(capacity)$	-0.37
	(0.061)***
$ln(sulfur\ content)$	0.958
	(0.076)***
$ln(heat\ rate)$	0.032
	(0.051)
$ln(age\ by 90)$	-0.081
	(0.094)
Year fixed effects	Y
Mills	0.635
	(0.212)***
cons	1.832
	(0.940)*

Notes: This table presents estimation results of the reduce-form regression of the 2SLS estimation for the emission rate equation. The dependent variable is SO_2 emission rates of low-sulfur coal of Phase I non-scrubbing units from 1995 to 2003, and of Phase II non-scrubbing units from 2000 to 2003. The instrument variable is $ln(Distance\ to\ PRB)$. Unit random effects are included. Robust standard errors are in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

APPENDIX C. Cross Sectional Estimates of the Impacts of Climate Change

This section explores the relationship between climate variables and crop yields using a cross sectional model specified as the following equation (C.1):

$$y_c^i = \alpha + \theta_1 GDD_c + \theta_2 GDD_c^2 + \theta_3 Prec_c + \theta_4 Prec_c^2 + \theta_5 GHDD_c + X_c'\beta + \varepsilon_c \qquad \text{(C.1)}$$

where the dependent variable is the yield of major crops (tons per hectare) in county c in any year with $i = rice, wheat, or\ corn$, GDD_c and GDD_c^2 are the linear and quadratic terms of growing season degree-days, respectively, $Prec_c$ and $Prec_c^2$ are the linear and quadratic terms of total precipitation during the growing season, respectively, $GHDD_c$ is growing season harmful degree-days, and X_c is a vector of county characteristics such as soil quality. The error term is composed of a time-invariant county-specific component α_c and an idiosyncratic shock u_c, specified in equation (C.2)

$$\varepsilon_c = \alpha_c + u_c \qquad \text{(C.2)}$$

To consistently estimate the coefficients of interest θ_i, it requires that

$$E[W_c \cdot \varepsilon_c \mid X_c] = 0 \tag{C.3}$$

Put in other way, conditional on observed county characteristics X_c, climate variables must be uncorrelated with the remaining unobserved determinants of crop yields. Therefore, the specification of X_c plays an important role in identifying the coefficients θ_i. Due to data availability, I incorporate the ratio of cropland irrigated to control for irrigation conditions, and I also add province level fixed-effects to control for other policy and socioeconomic characteristics.

Table A.1 report the predicted impacts of climate change under the comparatively high emission scenario (RCP8.5) on rice, wheat, and corn yields, respectively, using the cross sectional estimates. In column (1) of Table A.1, for each crop, the regressors include the five climate parameters, along with the irrigation ratio. In column (2), province fixed-effects are included.

The results show significant variation in the estimated impacts of climate change on crop yields. Let's first consider column (1) for each crop, without province fixed-effects, the estimated impacts of climate change on rice yield range between -0.062 and -0.503, on wheat yield range between -0.407 and 0.216, and on corn yield range between -0.027 and 0.713. Once province fixed effects are taken into account, the impacts on rice yield and wheat yield are drawn toward more positive, while the impacts on corn yield are drawn toward more negative. The wide variation indicates that the cross sectional approach is sensitive to model specification and data sources.

Table C.1 Cross Sectional Estimates of Impact of Climate Change on Rice Yield

	Rice		Wheat		Corn	
	(1)	(2)	(1)	(2)	(1)	(2)
2000	-0.317	-0.026	-0.407	-0.101	0.493	-0.131
	(0.152)	(0.324)	(0.132)	(0.266)	(0.217)	(0.519)
2001	-0.401	-0.184	-0.251	0.108	0.713	0.103
	(0.146)	(0.348)	(0.172)	(0.286)	(0.229)	(0.559)
2002	-0.062	0.119	0.095	0.370	-0.027	-0.301
	(0.202)	(0.379)	(0.148)	(0.272)	(0.241)	(0.543)
2003	-0.212	0.095	-0.287	0.251	0.159	-0.575
	(0.150)	(0.352)	(0.139)	(0.265)	(0.284)	(0.566)
2004	-0.387	0.029	-0.036	0.327	0.202	-0.566
	(0.135)	(0.298)	(0.152)	(0.264)	(0.222)	(0.455)
2005	-0.397	-0.016	0.230	0.439	0.224	-0.152
	(0.125)	(0.306)	(0.161)	(0.260)	(0.182)	(0.473)
2006	-0.265	-0.131	0.216	0.175	0.354	-0.369
	(0.123)	(0.319)	(0.168)	(0.274)	(0.189)	(0.437)
2007	-0.497	-0.076	-0.024	0.167	0.599	-0.141
	(0.130)	(0.321)	(0.183)	(0.272)	(0.178)	(0.484)
2008	-0.350	0.108	-0.097	0.024	0.150	-0.089
	(0.163)	(0.324)	(0.189)	(0.327)	(0.217)	(0.553)
2009	-0.503	-0.130	0.106	0.219	0.669	-0.088
	(0.136)	(0.312)	(0.178)	(0.287)	(0.204)	(0.501)
2010	-0.416	-0.174	-0.061	0.067	0.562	-0.160
	(0.180)	(0.339)	(0.142)	(0.303)	(0.229)	(0.526)
Pooled 2000-2010	-0.355	-0.053	-0.047	0.186	0.411	-0.164
	(0.077)	(0.118)	(0.073)	(0.099)	(0.090)	(0.166)
Province f.e.	No	Yes	No	Yes	No	Yes

Notes: This table summarizes the predicted impacts of climate change under RCP8.5 scenario on rice, wheat, and corn yields using cross sectional estimates for each year separately and for the pooled sample from 2000-2010. The regressors in column (1) include the five climate variables along with the ratio cropland irrigated, and province fixed effects are included in column (2). The effects of climate change are calculated as the discrete difference between the predicted yields at the projected temperature and precipitation scenarios and the predicted yields at the historical climate. Standard errors are in parentheses.