Impacts of Environmental Policy in China's Beijing-Tianjin-Hebei Region: A Case Study of the Iron and Steel Industry

by

Danielle L. Wilson

Submitted to the Institute for Data, Systems, and Society on May 26, 2017 in Partial Fulfillment of the Requirements for the degree of Master of Science in Technology and Policy

Abstract

This thesis focuses on the drivers of regional development and environmental control decisions in the iron and steel industry in the Beijing-Tianjin-Hebei region of China. The analysis presented includes (1) a county-level econometric analysis of the geographic distribution of the iron and steel industry between 2004 and 2009 in Hebei province and (2) a firm-level case study that analyzes the value of a flexible space option for a hypothetical, relocated steel plant that faces uncertainty in the timing and stringency of future environmental policy. Results from the econometric analysis of industrial distribution show a significant negative relationship between the initial rate of economic growth of a county, measured in Gross Domestic Product (GDP) change between 2004 and 2005, and growth of iron and steel shares in that county between 2005 and 2009. Results also show a significant negative relationship between population density in 2005 and both growth and concentration of iron and steel industrial shares. These results provide regional context and a descriptive baseline for a stylized case study. Inspired by the government-mandated relocation of major steel company Shougang in advance of the 2008 Beijing Olympics, the case study uses real options analysis to evaluate the value of a flexible option to install flue-gas desulfurization (FGD) technology for sulfur dioxide ($SO_2$) removal in response to uncertain developments in environmental policy. Case study results show a significant increase in the NPV of a steel facility that has the ability to employ the flexible option to construct an FGD system in the face of policy that begins in year six. This result is sensitive to policy timing, FGD space cost, and future FGD costs. Insights from this analysis are designed to help policymakers understand past patterns of industry evolution and environmental decision-making in firms during a period of very rapid economic development. This understanding can be used to identify ways to increase the effectiveness of environmental regulations in dynamic, rapidly-developing China.

Thesis Supervisor: Valerie J. Karplus
Title: Class of 1943 Career Development Professor
Assistant Professor of Global Economics and Management
Sloan School of Management
Acknowledgments

“If you aren’t in over your head, how do you know how tall you are?”
—T.S. Eliot

It takes a village. I would not be here today without the support of a truly incredible network of people. I must begin with my thesis advisor, Dr. Valerie J. Karplus, who has been the most dedicated and inspiring mentor I have ever had. Thank you for your unending patience as I worked through this unbelievably challenging academic endeavor and also for your nuggets of wisdom shared along the way. You pushed me hard, never gave up on me. For that I’m forever grateful.

Thanks to the tremendous amount of resources and support available at MIT, my dream turned into my journey. I extend extreme gratitude to the Samuel Tak Lee Real Estate Entrepreneurship Lab (STL Lab) for providing me with resources to conduct such rich and engaging research during my first year. I also want to thank TPP and ODGE for providing me with resources that allowed me to continue research in my second year. Thank you to all of the staff and faculty members in the Joint Program for providing me with all of the resources I could ever need, including a space to call my own in the office, throughout my entire MIT career.

In addition to resources, TPP offers a wonderful source of institutional support. Thank you to my entire TPP cohort of classmates for sharing your brilliant light with me. I couldn’t have survived this experience without the core group of TPP 2017 gals who endured my unique ability to have an entire conversation in GIFs, memes, and emojis. Also, I extend a special thank you to my CECP/KEG colleagues and JP cube mates for keeping me sane, dealing with my random questions, and for not judging me when I inevitably fell into my hole a couple of times each semester. Finally, I’m forever indebted to the beautiful individuals who do the hard work of running TPP every single day: Barb, Ed, Frank, and Noelle. I thank you all for your amazing support in all things from personal matters to funding for my dream summer internship in DC. TPP literally would not exist without you all.

Outside of TPP, my family and close friends held me up when MIT life proved nearly soul-crushing. My chosen family, friends from all of my past experiences, I love you all and cherish you dearly. Chelsea, David, Louis, Steph, Sam, Sanya, Elizabeth, Esha, Jillian, and many others from home, Boston, Tufts, JASC, and Yume Kata. My soul sister Tracy since fourth grade, words can’t describe how grateful I am to continue to experience your beautiful friendship. To Jacki, Titi Petra, Titi Carmen, Titi Myrna, Titi Gloria, Mimi, and my powerful tribe of cousins, aunts, uncles, and family from all over: you all can’t know how much your support and love drive me to continue to go out into this big, strange, scary world and do my thing. Without my grandmother I would not be here on this Earth or in this graduate program. I love you and miss you every single day. Thank you to Harry for always loving me. Last, but certainly not least, Mom and Nini: always know I love you both more.
## Contents

**List of Figures**  
11

**List of Tables**  
13

### 1 Introduction

1.1 **Context**  
15

1.2 **Motivation**  
17

1.3 **Scope of this project**  
18

1.4 **Structure of Thesis**  
19

### 2 Analysis of Iron and Steel Industry Growth in the Jing-Jin-Ji region, 2005-2009

2.1 **Motivation**  
22

2.1.1 Environmental goals outlined in the 11th Five Year Plan  
23

2.1.2 Green Olympics in 2008  
24

2.2 **Description of Jing-Jin-Ji Region**  
24

2.2.1 Beijing, national capital city  
25

2.2.2 Tianjin, regional port city  
25

2.2.3 Hebei, surrounding province with rural and industrial areas  
26

2.3 **Quantitative Analysis of Iron and Steel Industrial Location Patterns**  
27

2.3.1 **Methodology**  
27

2.3.2 **Description of data**  
28

2.3.3 Iron and steel shares and growth  
30

2.3.4 County characteristics of interest  
30

2.3.5 Measuring shares and growth of iron and steel industry  
32
List of Figures

1-1 Map that shows the location of Jing-Jin-Ji (created in ArcGIS) .......................... 16
1-2 Rapid increase in GDP across Jing-Jin-Ji between 2005 and 2009 (Source: CMR) .......................................................... 17
2-1 Changing intensity of the iron and steel industry, measured as a share of total main business revenue, across Jing-Jin-Ji counties in 2005 and 2009 (Source: CIC, Annual China Industrial Census Surveys) .......................................................... 27
3-1 Different forms of environmental policy uncertainty ............................................. 40
3-2 Decision tree that outlines hypothetical options for steel plant with policy uncertainty .......................................................... 51
3-3 Sensitivity analysis of ENPV rankings to changing assumptions about policy probability ......................................................... 52
3-4 Sensitivity analysis of ENPV rankings to changing assumptions about cost of extra space required for FGD construction ......................................................... 53
3-5 Sensitivity analysis of ENPV rankings to changing assumptions about the timing of policy ......................................................... 55
3-6 Sensitivity analysis of ENPV rankings to higher or lower FGD costs, for the case of policy uncertainty in year five ......................................................... 56
3-7 Cumulative distribution function for base uncertainty scenario .......................... 57
3-8 Decision tree that outlines outcomes based on enforcement uncertainty ............. 58
3-9 Sensitivity of ENPV to enforcement uncertainty .................................................. 59
3-10 Cumulative distribution function for enforcement scenario ............................. 60
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Summary statistics of county average iron and steel industry main business revenue (mill RMB)</td>
<td>29</td>
</tr>
<tr>
<td>2.2</td>
<td>Definitions of dependent variables and county characteristics explored</td>
<td>32</td>
</tr>
<tr>
<td>2.3</td>
<td>Determinants of annualized growth and share changes: OLS regression</td>
<td>34</td>
</tr>
<tr>
<td>3.1</td>
<td>Characteristics of a 1 million ton per year steel plant</td>
<td>45</td>
</tr>
<tr>
<td>3.2</td>
<td>Plan descriptions and purpose</td>
<td>46</td>
</tr>
<tr>
<td>3.3</td>
<td>Uncertainty that can impact steel plant</td>
<td>48</td>
</tr>
<tr>
<td>3.4</td>
<td>Basic data and cost estimates used for base case</td>
<td>48</td>
</tr>
<tr>
<td>3.5</td>
<td>ENPV for each plan given 75% chance of policy (million US$)</td>
<td>51</td>
</tr>
<tr>
<td>3.6</td>
<td>Values to assess VARG with chance of getting caught (US$million)</td>
<td>60</td>
</tr>
<tr>
<td>3.7</td>
<td>Sensitivity Analysis Results</td>
<td>61</td>
</tr>
<tr>
<td>B.1</td>
<td>Determinants of annualized growth, measured by main business revenue: OLS regression</td>
<td>88</td>
</tr>
<tr>
<td>B.2</td>
<td>Determinants of share changes, as a total of main business revenue: OLS regression</td>
<td>89</td>
</tr>
<tr>
<td>B.3</td>
<td>Base case with FGD and policy in year 6, example excerpt from spreadsheet of NPV calculations</td>
<td>90</td>
</tr>
<tr>
<td>B.4</td>
<td>Base case without FGD, example excerpt from spreadsheet of NPV calculations</td>
<td>91</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

China’s vast and diverse industrial sector has played an important role in the nation's economic development. Its industrial sector accounts for more than 40% of China’s gross domestic product (GDP) (The World Bank 2015). Industries such as mining, manufacturing, construction, power, and water enterprises serve as important drivers of national economic growth. As the most populous country and the number one exporter in the world, China’s industrial sector heavily impacts not only the global economy but also the local environment.

This thesis explores the iron and steel industry of the Beijing-Tianjin-Hebei region of China. At a macro level, I analyze drivers of the spatial distribution of industrial activity across regional counties between 2005 and 2009, a time of rapid economic growth. At a micro level, I then focus on a major actor with an analysis of impacts of environmental policy on decision-making in a new iron and steel facility. I present these two levels of analysis to better understand the context of an industrial region grappling with how to pursue sustainable economic development.

1.1 Context

Sustainable economic development encompasses many facets of transitions in industries, workforce, and resource efficiency. One definition of sustainable economic transition offered by the Paulson Institute emphasizes economic movement away from heavily polluting industries and towards low-carbon, energy efficient, and diversified industries (Hove et al. 2015). China’s eastern capital region of Beijing, Tianjin, and Hebei (referred to as Jing-Jin-
Ji, an abbreviated combination of the Chinese characters that stand for Beijing, Tianjin, and Hebei, respectively) represents one area in particular in which industrial activity has played a major role in regional development. Figure 1-1 shows Jing-Jin-Ji along the coast of the Bohai Sea in northern China. Jing-Jin-Ji offers a compelling case for understanding opportunities and challenges of sustainable economic transition in the Chinese context.

Figure 1-1: Map that shows the location of Jing-Jin-Ji (created in ArcGIS)

Rapid development in Jing-Jin-Ji has brought with it a rising environmental toll. Figure 1-2 demonstrates the kind of rapid economic growth that took place in the region since the early 2000s. Issues of air pollution accompanied this break-neck economic growth. China’s Ministry of Environmental Protection claims that more than half of China’s most polluted cities come from Hebei province [China Daily, 2017]. In particular, implementation and enforcement of environmental regulation in Hebei has proven challenging as the region has developed. Hebei’s environmental watchdog received the highest number of complaints about pollution [Shanghai Daily, 2013].
1.2 Motivation

Iron and steel production is an energy-intensive and heavily-polluting manufacturing process. The iron and steel industry in China accounts for nearly 30% of primary energy consumption from manufacturing industries (NBS, 2015) and contributed to approximately 20% of sulfur dioxide ($SO_2$) emissions and 27% of dust in 2013 (Wang et al., 2016). To reduce air pollution emissions, the Chinese government sets standards and control technology requirements as components of overarching environmental policies.

An important open question is to what extent environmental policy pressure, alongside other factors, shape the location decisions of industry in rapidly developing regions. Researchers have explored evidence of pollution havens across China (Shen et al., 2017; Zheng and Shi, 2017). Rapid growth after the great economic opening in 1980 was largely driven by industrial activity in eastern coastal urban centers and expansion of production for export, which accelerated after China’s entry into the World Trade Organization in 2001. Studying the Jing-Jin-Ji mega-region presents a chance to understand what drives the emergence of certain energy-intensive industries in different locations.

---

The pollution haven hypothesis refers to the concept that polluting industries relocate to jurisdictions with less stringent environmental regulations.
China’s winning bid for the 29th Olympic Games in 2001 stimulated increased attention to pollution issues in the form of stronger environmental policies issued by the Chinese government. Researchers examined this increased attention both before (Beyer, 2006) and after the games (UNEP, 2009; Chen et al., 2013). Before the games, commentary flagged plans to relocate energy-intensive industries from Beijing city center to outer cities in Hebei province as a problematic approach to actually solving regional air pollution issues (Beyer, 2006). After the games, the United Nations Environment Programme (UNEP) conducted an independent assessment of outcomes from environmental measures relative to commitments expressed in the 2000 bid process and concluded that these measures, including industrial relocations, overall resulted in significant improvements to air quality (UNEP, 2009). Findings from other studies agreed with UNEP assessment of air quality improvements (Chen et al., 2013). From an environmental policy perspective, the period around the 2008 Beijing Olympic Games offers a chance to examine impacts of policy on industrial structure at a vibrant time for the region.

With this thesis, I contribute to existing research through my study of regional trends and decision-making of firms during the years immediately prior to the 2008 Olympics in response to environmental policy. Given the central Chinese government’s recent focus on economic integration of Jing-Jin-Ji as a mega-region (Yao and Liu, 2014; Johnson, 2015), industrial location patterns present a unique opportunity to analyze drivers of industrial growth throughout the region and to observe potential impacts of environmental policies on firm location decisions.

1.3 Scope of this project

This analysis includes (1) county-level econometric analysis on geographic distribution of the iron and steel industry between 2004 and 2009 across the surrounding province of Hebei and (2) firm-level case study analysis on the value of a flexible space option for a hypothetical, relocated steel plant that faces uncertainty in future industrial SO₂ control policies.

Using national data from 147 counties, the econometric analysis involved ordinary least squares regression to assess key, pre-existing factors associated with changes in iron and steel industrial concentration and growth between 2004 and 2009. Results from the econometric analysis of industrial distribution show a significant negative relationship between economic
growth of a county, measured in GDP change between 2004 and 2005, and growth of iron and steel shares in that county. These results provide regional context and a descriptive baseline for a stylized case study.

Inspired by the mandated relocation of major steel company Shougang in advance of the 2008 Beijing Olympics, the case study analyzes the value of flexibility in plant design to explore industrial decision-making that accounts for future policy uncertainties. Specifically, I explore the value of a flexible option to install $SO_2$ control equipment at an iron and steel plan in response to uncertain developments in $SO_2$ control policies. Case study results show a significant increase in the NPV of a steel facility that employs the flexible option to set aside space for delayed construction of $SO_2$ control technology.

These two analytical pieces combined construct a larger descriptive picture of Jing-Jin-Ji as an emerging mega-region that aims to reconcile continued economic development with environmental protection goals. Insights from this thesis can support and inform current regional policymaking by evaluating whether iron and steel industrial activity tended to grow in parts of Jing-Jin-Ji with certain characteristics and also by demonstrating how uncertainty in timing, stringency, and enforcement of policy affects firm decision to invest in environmental controls. This proves especially salient now as national and local government administrators use various policy mechanisms to integrate regional cities and rebalance economic development across Jing-Jin-Ji.

1.4 Structure of Thesis

This thesis is organized as follows:

- **Chapter 2** - provides in-depth context on Beijing-Tianjin-Hebei area and summarizes county-level econometric analysis of factors associated with the geographic distribution of growth in the iron and steel industry

- **Chapter 3** - presents a case study that values flexibility in the timing of installation of environmental control technology for a hypothetical steel plant

- **Chapter 4** - discusses conclusions and suggests ideas for future work
Chapter 2

Analysis of Iron and Steel Industry

Growth in the Jing-Jin-Ji region, 2005-2009

Geographic distribution of energy-intensive, heavy industries reflects variation in the drivers of regional development patterns. Interactions between economic development objectives and environmental goals can shape the location and technology adoption decisions of local firms. However, characterizing these interactions is not a straightforward task because it is often difficult to identify the mechanism behind observed outcomes. Successful implementation of policies to promote environmental goals in rapidly developing China requires a deep understanding of the underlying drivers of geographic distribution. This question is especially important for industry that generates substantial pollution, such as the iron and steel industry. Few studies have empirically characterized drivers of industrial distribution or impacts of environmental policy on where firms locate in the Chinese context.

This chapter investigates geographic distribution of industry at the county-level in Jing-Jin-Ji. I used firm-level economic and financial data to analyze the evolution of economic activity, measured with main business revenue, of the iron and steel industry across Jing-Jin-Ji counties. I performed cross-sectional regression analysis of over 134 counties in Hebei province to assess key determinants of iron and steel industrial location. Results from this regression analysis suggest that population density and changes to GDP immediately prior to the 11th FYP are associated with observed growth rate and level of industrial share
throughout Hebei counties.

The results presented here provide a descriptive basis for further analysis of determinants of industrial growth in burgeoning regions in China. This is especially important as the country seeks to integrate regions into mega-cities and balance economic growth with environmental protection.

The descriptive analysis presented here establishes context of the Jing-Jin-Ji region around the time of the 2008 Beijing Olympics, for which China’s central government made aggressive policy moves to showcase blue skies for the major event. This context illustrates the need for further analysis into causal relationships between specific locational characteristics and industrial location patterns in China. Additionally, this context provides background for the case study examined in chapter 3 of this thesis.

The following sections are organized as follows: section 2.1 describes overarching motivation, section 2.2 describes the industrial region of Jing-Jin-Ji, section 2.3 outlines the quantitative approach and methods employed with a brief discussion, and section 2.4 summarizes concluding remarks.

### 2.1 Motivation

Motivation for this work stems from two overarching themes. The first relates to increasing concerns of environmental pollution. From here, the Olympics provide a guidepost and working indicator of how attitudes shifted and policies intensified to meet some of these concerns. The second relates to an increasing focus of the Jing-Jin-Ji region as an emerging mega-region that is expected to integrate the capital region in similar ways as occurred in the Yangtze or Pearl Delta River regions.

Heavy industries played a key role in powering China's economy through a period of rapid economic expansion. After China's modern economic reform in the late 1970s, the country needed heavy industries to provide energy, construction materials, and consumer goods for an increasingly wealthy and urbanized population.

However, China has reached a point in its economic development where issues of pollution, especially from energy-intensive industrial activities, impact economic growth. Industrial pollution impacts economic growth through additional deaths, employee retention, quality of life, productivity of the working population, and illness in vulnerable populations.
This chapter examines the relationship between location characteristics and share of iron and steel firms across Jing-Jin-Ji counties. Specifically, the time period between 2005 and 2009 presents an opportunity to observe changes to Jing-Jin-Ji industrial structure in light of the 2008 Olympics and more stringent environmental policies. Stringent policies implemented in advance of the Olympic Games included measures such as industrial plant closures, restriction on vehicle use, halting of construction activity, and expanded use of geothermal energy resources (Cheung [2010]). The analysis presented here provides a quantitative description of correlations between location characteristics and the evolution of iron and steel firms across Jing-Jin-Ji during this time.

2.1.1 Environmental goals outlined in the 11th Five Year Plan

Air pollution goals outlined in China’s Five-year Plans (FYPs) identify national priorities that regional and local governments are responsible for implementing. Policy mechanisms to achieve these priorities can include various market-based instruments, such as pollution fees and subsidies, or command and control instruments, such as technology mandates or performance standards. For example, the 10th FYP (2001-2005) outlined national goals to reduce \( SO_2 \) emissions by 10% below 2000 emission levels and included a variety of policy mechanisms. Key policy elements featured in the 10th FYP included pollution fees of 0.63 RMB per kilogram of \( SO_2 \) emitted, green pricing premiums of 0.015 RMB per kilowatt-hour of electricity generated from power plants with FGD installed, and \( SO_2 \) concentration limits (Schreifels et al. [2012]). During the 10th FYP, \( SO_2 \) emissions actually increased by 28% over 2000 levels. This resulted in increased efforts to strengthen policies for the next FYP.

For the 11th FYP (2006-2010), the government maintained the 10% reduction in \( SO_2 \) emissions goal (over 2005 levels). This time, the policies featured pollution fees of 1.26 RMB per kilogram of \( SO_2 \) emitted, green pricing premiums of 0.015 RMB per kilowatt-hour of electricity generated for power plants with satisfactory operation of FGD, and closures of small, inefficient plants. Emissions during the 10th FYP reduced by 14% below 2005 levels.

While some researchers attribute 10th FYP success to effective integration of multiple policy instruments and structures that improved accountability, incentives, and political support (Schreifels et al. [2012]), others posit that emissions reductions during this time had more to do with the “campaign style” regulation implemented in preparation for the 2008 Olympics.
The latter hypothesis highlights unique aspects of the time around the Beijing Olympics. It also raises important questions about how environmental policies and enforcement mechanisms impact regional actors.

2.1.2 Green Olympics in 2008

As in past cases when China had hosted major sporting events, the 2008 Olympics increased policymakers attention on cleaning the environment. As far as pre-Olympic policies, findings suggest that aggressive policies, such as those listed above, helped improve short-term air quality in the area. In particular, one study examined the cost-benefit ratio of environmental policies and measures implemented before the Olympics on short-run and long-run environmental improvements (Cheung, 2010). The 2008 Olympics took place in Beijing and also in various surrounding cities, which made the entire Jing-Jin-Ji region a focus of environmental cleanup efforts.

The remainder of this chapter is organized as follows: section 2 presents a brief overview of Jing-Jin-Ji; section 3 describes the quantitative analysis and discusses results; and section 4 summarizes main findings.

2.2 Description of Jing-Jin-Ji Region

Modern China has witnessed the rapid rise of megaregions, or adjacent metropolitan cities that feature a distinct and cohesive regional socio-economic identity. In the south of China, the Pearl River Delta and Yangtze River Delta regions offer two early examples of megaregions that exhibit such social and economic integration (Yusuf, 2007; World Bank, 2015). Adjacent to the Bohai Bay in northern China, Jing-Jin-Ji represents a third major eastern coastal megaregion that has been the target of integrated development plans and policies. Successful economic integration of the region requires thoughtful assessment of existing social divisions and economic trends that define the three areas that comprise Jing-Jin-Ji.

Campaign-style regulation refers to policy implementation that involves aggressive mobilization of political support and economic resources at all administrative levels to stimulate an urgent need. See Liu et al. for a detailed description of this type of regulation in China (Liu et al., 2015).
2.2.1 Beijing, national capital city

As China’s capital city, Beijing naturally occupies the political, cultural, and educational center of Jing-Jin-Ji. Important governmental and political institutions, including the National People’s Congress, reside within Beijing’s borders. Governed as a provincial-level municipality, the city encompasses 16 county-level districts that represent a mix of urban, suburban, and rural administrative divisions connected by ring roads. Combined, these counties host a total population of over 20 million residents spread across approximately 16,400 square kilometers of land area (CEIC, 2004-2015; CMR, 2005).

Culturally, the “Northern Capital” (from the Chinese characters bei for north and jing for capital) features traditional architecture from imperial China and major national historical sites such as the Forbidden City and the Temple of Heaven. More than 70 institutions of higher learning contribute to the rich educational resources possessed by this historical heartland. These educational resources contribute greatly to the types of industry that propel the provincial-level city to the tops of wealth and development rankings. Major industries in Beijing include finance, technology, healthcare, tourism, media, and other service-oriented businesses characteristic of a post-industrial economy.

Economically speaking, Beijing ranks high among the wealthiest and most developed cities of China. In 2015 the municipality generated nearly 2.3 trillion RMB in GDP, which put it at thirteenth among provincial-level administrative units (CEIC, 2004-2015). This ranks Beijing’s per capita GDP, approximately 93,000 RMB GDP per capita in 2015, at number two behind Tianjin.

2.2.2 Tianjin, regional port city

Tianjin, the second provincial-level city in Jing-Jin-Ji, produces the highest economic output in terms of GDP per capita. In 2015, the port city situated directly on the Bohai Bay generated nearly 1.6 trillion RMB in GDP (CEIC, 2004-2015). With a population of roughly 15 million people, Tianjin’s per capita GDP of approximately 108,000 RMB represents one economic measure by which the 11,760 square kilometer city outperforms Beijing and the nation (CEIC, 2004-2015; CMR, 2005). Historically, Tianjin has provided the global gateway to Beijing as a major source of marine and import activities.

Dominant industries in modern Tianjin include mostly high-end manufacturing busi-
nesses. Automobile, aerospace, petrochemical, pharmaceutical, textile, and mechanical pro-
duction industries comprise the relatively high-technology manufacturing sector in the port
area [ECN, 2007].

2.2.3 Hebei, surrounding province with rural and industrial areas

By comparison, Hebei province, with which Tianjin shares most of its borders, specializes in
more traditional manufacturing industries. The largest steel-producing province in China,
Hebei province expands both north and south of Beijing and Tianjin provincial cities. The
province covers approximately 200,000 square kilometers and includes a population of over
73 million residents. With a GDP of 2.98 trillion RMB in 2015, Hebei’s average per capita
GDP of nearly 41,000 RMB indicates the per-capita income disparity of Hebei residents
relative to Beijing and Tianjin residents [CEIC, 2004-2015]. Dominant industries in Hebei
include agriculture and heavy industry.

Important industrial cities in Hebei include the capital and largest city Shijiazhuang, cen-
tral Baoding, Handan in the south, and Qinhuangdao and Tangshan in the east. Tangshan,
the second largest city in Hebei, has a long industrial past. Located 105 kilometers north-
east of its immediate neighbor Tianjin, Tangshan historically served as a core mining city
for Hebei with its coal and iron resources. The city received much national attention after
a 7.8 magnitude earthquake devastated the population and infrastructure in 1976. Post-
earthquake development focused on innovative urbanization and industrial policies such as
the re-development of Caofeidian, a town along the coast of the Bohai Sea [Li et al. 2017].
Contemporary Tangshan specializes in steel, cement, porcelain, textile, and machinery in-
dustries as well as thermal power generation.

The analysis in this thesis began with a visualization of observed changes in industrial
shares. Figure 2-1 shows changes in the county-level distribution of iron and steel firms across
Jing-Jin-Ji between 2005 and 2009. These two maps show notable decreases in iron and steel
industry shares across Hebei cities both in northern Chengde and in southern Xingtai. This
visualization of changes in iron and steel shares over time shows how industrial concentration
for this one industry evolved in less than five years.
Figure 2-1: Changing intensity of the iron and steel industry, measured as a share of total main business revenue, across Jing-Jin-Ji counties in 2005 and 2009 (Source: CIC, Annual China Industrial Census Surveys)

2.3 Quantitative Analysis of Iron and Steel Industrial Location Patterns

This section describes the quantitative approach used to analyze determinants of iron and steel industrial location in Jing-Jin-Ji. The empirical analysis presented here, while descriptive in nature, offers a basis for further exploration on determinants of industry location. Due to the limited amount of existing quantitative research on drivers of industrial location patterns in China, this description of location patterns in Jing-Jin-Ji’s largest industrial sector serves as a novel exploration.

2.3.1 Methodology

This analysis examines 146 Jing-Jin-Ji counties, four sub-sectors of the ferrous metal smelting and rolling manufacturing sector (puddling or iron making, steel making, steel rolling, and ferroalloy smelting), and six consecutive years (2005-2009). To explain location patterns, this analysis considers the impact of five predictor variables on county-level shares of iron and steel economic output relative to county-wide economic output from all other industries such as textile manufacturing, electricity production, chemical manufacturing,
paper making, and non-metallic mineral production. The selection of determinant factors to explore came from available literature.

A large body of literature has studied the drivers of spatial location of industries. Many scholars have focused on explaining the spatial agglomeration and regional specialization of industries (Krugman 1991, Kim 1995), while others have focused on locational drivers of trade (Midelfart et al. 2000, Martinez-Galarraga 2011).

Most of the previous literature has come from, and therefore focused on, contexts in advanced industrialized countries. A growing body of literature has examined industrial location patterns in developing economies such as China. Some studies have acknowledged the impact of environmental regulation on location of China’s manufacturing firms (Lian et al. 2016). Studies have acknowledged both the power and limitations of applying theoretical models constructed from studies of western economies to other, developing economies that face unique globalization, marketization, and decentralization challenges (He et al. 2008).

Other literature has examined the relationship between environmental regulations and industrial location patterns (Jeppesen et al. 2002, Henderson 1997).

### 2.3.2 Description of data

Data used for this thesis include two proprietary datasets and one semi-public dataset. Firm-level economic data derives from the Chinese Annual Industrial Survey, which is a subset of the Chinese Industrial Census (CIC), and includes financial characteristics such as gross economic output (RMB), sales, and revenue. The National Bureau of Statistics in China (NBS) populates the industrial section of the annual China Statistical Yearbook with CIC data (Brandt et al. 2014). Equally important, county-level economic, social, and demographic data come from China Premium Database (CEIC) and include basic characteristics such as GDP, education level, and government input.

To construct a panel of data for this study, I primarily relied on firm-level location and asset information included in the CIC. The CIC contains detailed break-downs of key company characteristics such as address, founding year, firm type, industry classification, primary sector code, sales, revenues, profits, and various types of investments. China’s National Bureau of Statistics (NBS) builds on CIC to present relevant macroeconomic factors specific to the industrial sector. The CIC covers all registered firms with gross revenues above 5 million RMB (approximately US $700,000) (CIC 2004-2010). In 2005, this represented
over 90% of China’s industrial sector. This dataset represents the most comprehensive
dataset of Chinese industrial firms available [Brandt et al., 2014]. Table 2.1 shows descriptive
data of the sample data considered. This helps provide an idea of the range of values for
each variable across the Jing-Jin-Ji region.

Table 2.1: Summary statistics of county average iron and steel
industry main business revenue (mill RMB)

<table>
<thead>
<tr>
<th>Year</th>
<th>Num. of Counties</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>116</td>
<td>3.22</td>
<td>6.12</td>
<td>0.0000</td>
<td>32.3</td>
</tr>
<tr>
<td>2006</td>
<td>113</td>
<td>3.69</td>
<td>7.66</td>
<td>0.0000</td>
<td>38.1</td>
</tr>
<tr>
<td>2007</td>
<td>123</td>
<td>5.13</td>
<td>9.94</td>
<td>0.0000</td>
<td>49.3</td>
</tr>
<tr>
<td>2008</td>
<td>117</td>
<td>5.70</td>
<td>10.4</td>
<td>0.0025</td>
<td>52.9</td>
</tr>
<tr>
<td>2009</td>
<td>89</td>
<td>5.55</td>
<td>10.0</td>
<td>0.0008</td>
<td>51.5</td>
</tr>
</tbody>
</table>

Notes: Dataset covers both state owned and non-state owned firms
with sales over 5 million RMB. Industrial data includes manufacturing,
utilities, and mining industries.

Certain characteristics of the CIC present challenges and limitations for constructing a
panel dataset for the regression analysis. Other research that has used this dataset have
cited issues of data reliability for years after 2008 [Brandt et al., 2014]. For the purposes
of this study, I chose to examine through 2009 because I wanted to use the most reliable data.
The following paragraphs outline the specific challenges encountered in building the panel
data set and approaches to tackle the problem for this study. I also present a few alternative
solutions to solving these challenges.

The first problem arose with regards to inconsistently defined location names in Chinese
in the dataset. Primarily location names for county were in Chinese. This presented a
problem because names in the CEIC were in English. Additionally, the Chinese names in
CIC were inconsistently entered. For example, a county-level administrative unit could just
be the name or the name plus the character for district. To overcome this challenge, I ran
a script to consistently match county names in Chinese with corresponding English county
names (Appendix A).

The other data set I relied on to build location-specific characteristics for this study came
from the CEIC China Premium Database. CEIC maintains a comprehensive repository of
macroeconomic data from China’s NBS that expands back as early as 1949 in some cases
and records certain economic or demographic factors at the county level [CEIC (2004-2015)].
The CEIC dataset proved easier to work with than CIC because of CEIC’s online interface
and abundant granularity for many macroeconomic factors of interest to this study.

This study primarily examines the years immediately before, leading up to, and immediately after the 2008 Beijing Olympics. The sample period of 2005 through 2009 offers an opportunity to clearly define and bookend pre- and post-intervention phases, where the intervention here is policies enacted prior to the Olympics that required firms to relocate.

To conduct this analysis, I used cross-sectional data constructed from the two data sources mentioned. Panel data, also referred to as longitudinal or cross-sectional time-series data, includes a set of behaviors or observations for a large number of units over time. For this panel data, the unit was individual firms. The observations were different economic quantities observed for the firms in different years. Panel data allow for analysis of observations across different types of units, or firms in this case. It accounts for individual heterogeneity. Some drawbacks with panel data involve sample design and coverage of assembled panel (Baltagi, 2009).

2.3.3 Iron and steel shares and growth

The two models used for this study examine what factors predict changes in iron and steel industry shares. Estimates of economic output come from annual industrial census reports of main business revenue or, income a firm receives from their primary business activity. Because this study focuses on one specific industry, iron and steel, main business revenue offers an appropriate indicator of economic output for firms of interest. County shares are represented by the sum of main business revenue for iron and steel firms divided by the sum of main business revenue for all industries in each county. Similarly, annualized growth rate for the iron and steel industry is calculated using the equation below:

\[
G_{t_0,t_n}^{is} = \left( \frac{MBUS^{is}(t_n)}{MBUS^{is}(t_0)} \right)^{1/(t_n-t_0)} - 1
\]  

(2.1)

where \( MBUS^{is}(t_n) \) is main business revenue in year \( t_n \) divided by main business revenue in year \( t_0 \).

2.3.4 County characteristics of interest

This analysis considers five county characteristics that are expected to affect industrial location across Jing-Jin-Ji counties between 2005 and 2009 shown in Table 2.2.
determinant factor displayed in the table, population $POP$ (thousands of people), captures population as a driver of economic activity. Counties with larger populations are expected to have higher iron and steel shares due to having more people available for work. Here, I expect a positive relationship between a county’s population in 2005 and industrial growth and share changes. The second determinant factor shown in the table considers average wages per capita of each county, $WGPC$ (RMB/thousands of people). This variable captures the effect of prevailing wages prior to the start of the study period. For this analysis, which considers a region with heavy industrial output, average wage per capita may indicate the relative balance of supply and demand for labor, as well as the relative skill level of workers in that county. For this factor, I expect a positive relationship between a county’s average wages per capita in 2005 and industrial growth and share changes.

I then consider two factors that may have been especially important in China’s context during this period. These are the extent of regional specialization in agricultural activities and local government revenue. The variable $FFRU$, considers the number of workers employed in the agricultural sector as a percentage of total population. Agricultural sector employment normalized by population serves as a proxy for economic specialization in the agricultural industries. Given China has a deep history in agricultural industries, normalized agricultural employment also may indicate how traditional or rural a particular county’s economy might be. This in turn can provide insight on levels of industrialization observed in a county. I expect a negative relationship between a county’s normalized agricultural worker population in 2005 and industrial growth and share changes.

How much money a local government brings in can also offer key insights to geographic distribution of industry throughout counties. Total government revenue, $GOVR$, as a percentage of GDP (%) is computed by dividing government tax revenue by a county’s GDP. Since local government revenue can come from various sources, total government revenue as a percentage of GDP provides a measure of government involvement in the local economy. This variable could suggest the degree to which a local government controls a county’s resources. A county that features a relatively high ratio of revenue in GDP terms that falls closer to the OECD average range of between 30% and 35%, might be expected to have heavy government involvement in matters that impact use of local resources required for

\footnote{Data from national census sources present agricultural employment as a sub-sector of rural employment (CEIC 2004-2015). The sector typically encompasses farming, forestry, animal husbandry, and fishery sub-sectors.}
production (OECD 2017). For government revenue normalized by GDP in 2005, I expect a positive relationship with iron and steel industrial growth and share changes.

Related to government revenue, fixed asset investments \( FAIV \) as a percentage of GDP measures capital spending of a firm. FAI typically includes physical assets such as machinery, infrastructure, or other capital investments owned by a firm for long periods of time. For FAI normalized by GDP in 2005, I expect a positive relationship with iron and steel industrial growth and share changes. It makes sense that larger shares of fixed assets in GDP correlate with iron and steel industry expansion.

Population density, \( POPD \) (people per 10 square kilometers) represents number of people per square kilometer. As a determinant variable, population density can indicate characteristics about relatively urban counties compared with sparse, rural counties with far-flung populations. I might expect a negative relationship between a county’s population density in 2005 and iron and steel industry growth and share changes. The final variable, \( GDPC \), measures the percent change in GDP from 2004 to 2005, the base year considered for this chapter. This provides an indicator of economic growth of counties at a time of rapid regional development. The goal in examining these variables is to understand effects of pre-existing county characteristics on iron and steel industrial growth between 2005 and 2009.

Table 2.2: Definitions of dependent variables and county characteristics explored

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( G^*)</td>
<td>Annualized growth rate of iron and steel industry, dependent (%)</td>
</tr>
<tr>
<td>( S^*)</td>
<td>Share of iron and steel industry, dependent (%)</td>
</tr>
<tr>
<td>POP</td>
<td>Population (thousands of people) for size effects</td>
</tr>
<tr>
<td>WGPC</td>
<td>Average wages per capita (RMB/thousands of people)</td>
</tr>
<tr>
<td>FFRU</td>
<td>Agricultural workers, percentage of population (%)</td>
</tr>
<tr>
<td>GOVR</td>
<td>Government revenue as a percentage of GDP (%)</td>
</tr>
<tr>
<td>FAIV</td>
<td>Fixed asset investment as a percentage of GDP (%)</td>
</tr>
<tr>
<td>POPD</td>
<td>Population density (persons/10 sq km)</td>
</tr>
<tr>
<td>GDPC</td>
<td>Change in GDP from 2004 to 2005 (%)</td>
</tr>
</tbody>
</table>

2.3.5 Measuring shares and growth of iron and steel industry

The approach applied in this thesis represents a simplified version of spatial regression analysis described in industrial location literature (Midelfart et al. 2000; Martinez-Galarraga 2011). To analyze the relationship between location characteristics and iron and steel share
at the county level, the empirical model used for regressions is estimated with ordinary least squares (OLS) and Huber-White sandwich estimators for standard error. The model is expressed in the following equation:

\[
S_{jt} = \alpha \cdot POP_{jt} + \sum_j ((\beta_{kt})_j \cdot (X_{kt})_j) + T_t + \epsilon
\] (2.2)

where \(S_{jt}\), the dependent variable, is the share of iron and steel industrial output in county \(j\) for year \(t\). The first variable \(POP_{jt}\) represents the population of county \(j\) in year \(t\). \(POP_{jt}\) and its associated coefficient \(\alpha\) capture size effects that exist because larger counties are expected to have larger corresponding populations. The five independent variables are represented by \(X_k\) where \((X_{kt})_j\) is the county characteristic \(k\) in year \(t\) for county \(j\). Relatedly, coefficients \((\beta_{kt})_j\) associated with each characteristic \(k\) estimate the level to which these characteristics impact shares of iron and steel across Jing-Jin-Ji counties in a year \(t\). The time variable \(T_t\) represents an estimated constant that captures time-varying effects common to all counties. Model error is captured in the final term, \(\epsilon\).

To estimate impacts of key factors on iron and steel industry growth, instead of shares, I use the model above to examine a second dependent variable for annualized growth rate, \(G_{j0,t_n}^{is}\), of the iron and steel industry between a start year \(t_0\) and an end year \(t_n\) in county \(j\). To calculate iron and steel shares and growth by county, I used main business revenue reported in CIC. Main business revenue represents the income a firm receives from their primary business activity, in this case iron and steel related activities.

The CIC also reports other financial data that could be used to indicate economic output of firms such as gross output and sales revenue. However, these values are less appropriate for the purposes of this study. Values for sales revenue include other sources of business revenue, which has the potential to introduce unaccounted for bias in my analysis. I use deflated values of main business revenue to measure economic output for Jing-Jin-Ji counties.

2.3.6 Empirical results

I used OLS regression to estimate impacts of pre-existing county characteristics on iron and steel industry growth outcome across Hebei counties. To conduct this analysis, I used the statistical software package STATA for processing raw downloaded data from CEIC and CIC, merging and manipulating the datasets, and regressing county characteristics to
growth and share variables (code included in Appendix A).

Table 2.3 summarizes results from the two dependent variables explored in this regression exercise. To facilitate interpretation of results, I examined the log of 2005 population, average wages per capita, and government revenue in GDP. I based the decision to log these three variables on observation of the distribution of values for each variable. Appendix B includes additional output from these regressions (Tables B-1 and B-2).

Table 2.3: Determinants of annualized growth and share changes: OLS regression

<table>
<thead>
<tr>
<th>Log 2005 Pop (th people)</th>
<th>0.190</th>
<th>1.584</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1.73)</td>
<td>(1.37)</td>
</tr>
<tr>
<td>Log 2005 Avg Wages Per Cap (RMB/th people)</td>
<td>0.0939</td>
<td>0.392</td>
</tr>
<tr>
<td></td>
<td>(0.87)</td>
<td>(0.42)</td>
</tr>
<tr>
<td>Agr Workers in 2005, %Total Pop in 2005</td>
<td>-0.765</td>
<td>-8.619</td>
</tr>
<tr>
<td></td>
<td>(-1.60)</td>
<td>(-1.61)</td>
</tr>
<tr>
<td>Log of Govt Rev, %GDP in 2005</td>
<td>-0.165</td>
<td>-1.617</td>
</tr>
<tr>
<td></td>
<td>(-1.96)</td>
<td>(-1.77)</td>
</tr>
<tr>
<td>Fixed Asset Investment, %GDP in 2005</td>
<td>0.0180</td>
<td>0.433</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.27)</td>
</tr>
<tr>
<td>Pop Dens in 2005 (person/10 sq.km)</td>
<td>-0.0000530**</td>
<td>-0.000477*</td>
</tr>
<tr>
<td></td>
<td>(-2.73)</td>
<td>(-2.10)</td>
</tr>
<tr>
<td>Change in GDP from 2004 to 2005</td>
<td>-0.00335*</td>
<td>-0.0290</td>
</tr>
<tr>
<td></td>
<td>(-2.60)</td>
<td>(-1.94)</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.543</td>
<td>-11.66</td>
</tr>
<tr>
<td></td>
<td>(-1.58)</td>
<td>(-1.28)</td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.068</td>
<td>0.029</td>
</tr>
<tr>
<td>No. of cases</td>
<td>134</td>
<td>134</td>
</tr>
</tbody>
</table>

Notes: CEIC, CIC, CMR (sources); * $p < 0.05$, ** $p < 0.01$

2.3.7 Discussion

Results from the regression presented above suggest a significant relationship between two pre-existing county characteristics and the two dependent variables of concern. For iron and steel industrial growth, the coefficient for population density in 2005 is -0.000053 person/10 square kilometers. This coefficient indicates that for every person/10 square kilometers decrease in population density there is an increase in iron and steel industrial growth. This means that iron and steel industry located in places with lower population density. The
coefficient for change in GDP from 2004 to 2005, -0.00335, indicates that lower GDP growth correlates with an increase in industrial growth. This result suggests that the iron and steel industry grew less in places that were already experiencing rapid GDP growth.

For iron and steel share differences between 2005 and 2009, the coefficient for GDP growth (-0.00335) further supports the finding that iron and steel industry grows less in places with rapid GDP growth. Table 2.3 shows that the coefficients for population density and change in GDP growth are significant at the 1% and 5% levels respectively with respect to industrial growth. Furthermore, the coefficient for population density with respect to industrial share differences is significant at the 1% level.

Regression results also indicate nonsignificant correlations for the remaining five county characteristics in 2005 and iron and steel industry growth and share difference. As expected, the coefficients for log of 2005 population, log of 2005 average wages per capita, and 2005 fixed asset investment in GDP indicate a positive relationship with growth and share changes in iron and steel industry. These results intuitively suggest that larger populations, higher per capita wages, and larger shares of fixed asset investment represent pre-existing conditions correlated with iron and steel industry growth and share changes. The coefficients for log of government revenue in terms of GDP indicate that for every % decrease in this term, iron and steel industry growth increases by 0.165 and shares increase by 1.617. These results, although difficult to interpret, could suggest that iron and steel industry tends to locate in counties with lower government involvement. Finally, the coefficients for agricultural workers over total population in 2005 indicate that every unit of decrease in this term correlates to high increases in iron and steel industry growth and share differences. This result could demonstrate that more iron and steel industry tends to located in counties that have moved away from more traditional forms of labor featured in rural populations.

The findings presented here offer insight into industrial relocation patterns in China to create the descriptive foundation for further study on precise predictors and driving factors of industrial development. This will prove especially useful as China’s economy slows, as it has showed signs of doing since 2011 (The Economist, 2015). Findings from other studies have supported the importance of availability of labor and other NEG factors in industrial location patterns. This study builds on this overall field of work by contributing context unique and specific to China.
2.4 Summary of Findings

Analysis in this thesis focused on counties in Hebei province for the iron and steel industry between 2005 and 2009. The main findings indicate that two pre-existing characteristics, county population density and GDP growth, have a significant impact on iron and steel industry growth and shares. These findings suggest that iron and steel industry tended to locate in areas with lower population density and less rapid GDP growth. National policy-makers interested in a coordinated integration effort across Jing-Jin-Ji would benefit from a deeper understanding of how locational characteristics relate to and impact industrial concentration and growth. This type of analysis could inform focused policy-making tailored down to the county level that directly meets the needs of the region.

Recent attempts to economically and socially integrate Jing-Jin-Ji cities have raised concerns of how China will (a) accomplish this integration effort over a large area that contains counties with different locational characteristics and also (b) balance economic development with environmental protection. China’s past experience in advance of the 2008 Beijing Olympics offers a useful opportunity to examine how industrial development has permeated throughout the region and also how responses to concerns of environmental pollution can impact this classically industrial structure in a very real way. Chapter 2 examines county-level characteristics across the region at the macro level. This descriptive analysis sets the stage for a deeper dive into how a major regional actor might make decisions in light of these dynamics at the micro level.
Chapter 3

The Value of Flexibility Under Environmental Policy Uncertainty

In the years leading up the Beijing ‘Green Olympics,’ many pollution-intensive firms were relocated from Beijing to Hebei and surrounding provinces. These relocations provided an opportunity to redesign plants. Here I perform a stylized case study of design flexibility for a new steel plant given uncertainty in future environmental policies. Results suggest that incorporating flexibility such as an extra space option for sulfur dioxide control technologies can offer significant increase in the net present value (NPV) of a steel facility. This case study builds on the descriptive background of Jing-Jin-Ji’s iron and steel industry in 2005 provided in chapter 2.

Chapter 3 is organized as follows: the first section provides background on the 2008 Olympics and describes the approach taken, real options analysis; the second section defines the system under consideration and describes model setup; the third and fourth sections describe results and sensitivity analysis for two types of uncertainty examined in the case study; and the fifth section summarizes findings and thoughts for discussion.

3.1 Background

This case study stemmed from a desire to understand how policy uncertainty affects the pollution abatement decisions of emitting firms against a backdrop of rapid economic growth and externally-driven environmental pressure. A deeper understanding of these decision drivers requires an understanding of (1) the real world context that motivated this study,
(2) various forms of policy uncertainty, and (3) the theory behind real options analysis, the method employed in this case study.

3.1.1 Jing-Jin-Ji’s iron and steel industry

In their bid to host the games, the Organizing Committee outlined plans to revamp parts of the city and to improve air quality. These plans integrated and fast-tracked aspects of Beijing’s “Environmental Master Plan,” an existing 20-year initiative from the municipal government to promote environmental protection through 2015 (UNEP, 2007). Along with other initiatives, the relocation of firms comprised one of the major features incorporated into the overall preparation for the Games.

To deliver clear, blue skies for the Olympics, Beijing required nearly two hundred medium and large manufacturing plants to relocate from urban neighborhoods in Beijing to cities in the surrounding province of Hebei (UNEP, 2009; Fu et al., 2014). This policy intervention to reduce air pollution in Beijing opened a decision-making window for plant executives who needed to construct new facilities elsewhere.

One firm mandated to move, Shougang Iron and Steel, offers a unique opportunity to serve as a model for this case study because of its history, size, and the amount of attention paid to its relocation. A classic iron and steel company that had been headquartered in Beijing since its inception nearly 80 years prior, Shougang represented one of Beijing’s largest steel producers with a capacity of around 8 million tons of steel output per year (Etienne et al., 1992; Dai and Song, 2012). The central government issued a mandate in 2005 for Shougang to move its main production facility by 2010 from a Beijing neighborhood to the port city of Caofeidian, Tangshan in Hebei province some 250 kilometers south of Beijing (The People’s Daily, 2005; China Daily, 2005).

As a notorious polluter in the area, Shougang’s relocation and accompanied production cuts (Hornby, 2008) drew much attention from national and international media outlets as well as from local citizens right before the Olympics (Oster, 2007; Cha, 2008; Hornby, 2008; Watts, 2008). After the Olympics, Shougang’s relocation continued to draw attention from researchers and policymakers interested in understanding the effects of this particular hybrid of economic and industrial policy. I rely on available information about Shougang’s relocation, with respect to FGD installed (Chen et al., 2013) and planned emissions reductions (China Daily, 2007), to develop this stylized case study and provide real-world context for
the analysis. The context of mandated relocations provides a prompt for decision-makers to consider alternative facility designs in the face of uncertainty in environmental policy stringency, timing, and enforcement.

3.1.2 Relevant sources of environmental policy uncertainty

Sources of environmental policy uncertainty can compound sources of market uncertainty for manufacturing plants. Market uncertainty refers to changes in future raw materials costs or products prices that would impact a firm’s investment strategy. For the steel industry, iron ore and fuel prices represent a major raw material cost to firms on the production side. On the revenue side, steel prices heavily depend on global production levels and economic conditions. Environmental policy objectives impact costs and revenue through additional costs (implicit or explicit, direct or indirect) imposed for polluting. Uncertainty in regards to how much it costs to pollute, timing of policy introduction, specific technology requirements, and extent of enforcement all serve to encourage firms to delay action, maybe indefinitely.

As environmental issues in heavily populated, industrializing economies come to the forefront, there is a clear need for quantitative investigation into how environmental policy uncertainty may impact firm decision-making of the heaviest, most polluting industries. Some literature has examined the impact of environmental policy uncertainty on innovation and adoption of environmental technologies. Kalamova et al. examined how an increase in environmental policy uncertainty causes a decrease in environmental patent activity, which could impact adoption of environmental technologies more broadly (Kalamova et al., 2013). Figure 3-1 describes four main categories of policy uncertainty scenarios relevant to environmental policy.

---

1This refers to the Organization for Economic Co-operation and Development (OECD) definition of environmental policy stringency, which describes stringency as intensity of an environmental policy signal in terms of implicit and explicit costs from environmentally harmful behavior.
3.1.3 Project valuation with real options analysis

Corporate decision-makers rely on financial assessments of return on capital invested in a particular project over the project lifetime. These financial assessments use cashflow analysis to estimate the value of a particular project or set of projects. Dominant methods used to estimate expected cashflows include Net Present Value (NPV), which employs a chosen discount rate, and Internal Rate of Return (IRR), which is the required interest rate for a project to result in a zero NPV \cite{deNeufville2011}. Project developers can expect a positive NPV when the IRR exceeds the cost of capital.

Real options analysis accounts for uncertainty in a manner largely absent from traditional project valuation methods. To account for uncertainty, this form of analysis relies on cashflow estimates that derive from a consideration of possible outcomes whereby managerial flexibility can occur. Managerial flexibility refers to strategic consideration of available options that a decision-maker pays to possess but need not necessarily exercise. As a methodology, real options analysis enables a decision-maker to exercise flexibility in their investment choices and to avoid financial downsides while taking advantage of potential upsides \cite{deNeufville2011}. Real options analysis presents an alternative framework to deterministic NPV analysis and facilitates strategic decision-making for industrial projects.
The past few decades have witnessed an emergence of literature that applies real options analysis and related modeling methods to major engineering design projects for certain industries. A few studies have applied real options analysis to the iron and steel industry in particular. Studies have shown value of real options as a risk assessment tool especially for small and medium enterprises in the steel industry (Muharam, 2011).

Other studies use real options to assess flexibility in steel plant design. Ozorio et al. evaluated managerial flexibility in response to market uncertainty for steel product demand (2013). In their study, the authors valued a product switch option for a hypothetical steel facility equipped with lamination equipment and found that the option could generate a significant increase in project NPV.

Real options offers a method to assess potential investments that consider the costs and benefits of design flexibility in response to uncertainty. The real options approach builds on discounted cash flow analysis to estimate future valuations through results from a simulation meant to capture uncertain outcomes. Decision-makers armed with alternative valuations from simulation can make strategic decisions that optimize value and accounts for uncertainty. This technique proves especially useful for iron and steel firms that seek to value alternative options for a new plant design that account for costly and uncertain future environmental policy uncertainty.

Broadly speaking, the iron and steel industry employs two main production processes. The first process, basic oxygen steelmaking (BOF), converts liquid iron to steel by blowing oxygen through carbon-rich molten iron in a blast furnace that burns coal. The second, and more recent, process involves melting steel scraps in an electric arc furnace (EAF). BOF requires coke, iron ore, and fossil fuels as primary feed materials, whereas EAF requires electricity and steel scraps. In China, the BOF process dominates the iron and steel industry due to abundant access to iron ore mines and limited access to scrap steel. The hypothetical steel plant considered in this study most likely continued to employ the BOF process.

Technologies employed in the iron and steel industry include well-established processes that fall into two primary categories. In 2005, a relocated iron and steel facility in China likely encountered a number of energy efficient technology updates for a new facility. Such

---

2Laminators convert crude steel or steel slabs into different end products such as coils, rods, plates, bars, and other laminated profiles. Steel plants can, and frequently do, sell steel slabs without further processing. However, plants can also invest in laminators, which require significant capital cost, to further process steel slabs and produce various steel products with higher market prices.
updates might have included coke dry quenching (CDQ), Top-pressure Recovery Turbine (TR), recycling gas converters, continuous casting, adjustable product laminators, and other measures to improve energy efficiency and modernize blast furnaces used in steel production \cite{Hasanbeigi2017}. Higher pressure CDQ is the largest contributor to overall energy savings of steelmaking in China between 2005 and 2010 \cite{Wang2017}.

### 3.1.4 A stylized case study

Inspired by Shougang’s pre-Olympic relocation, I analyzed tradeoffs decision-makers of an iron and steel firm might face in the presence of significant environmental uncertainty. For my analysis, I developed a modified cash flow model to estimate expected net present value of three potential design plans, including one flexible plan, for a hypothetical iron and steel facility given future uncertainty in sulfur dioxide ($\text{SO}_2$) control policies. The remainder of chapter 3 describes this case study with the purpose of exploring how a fairly aggressive environmental policy, mandated relocations, impacts other decisions firms encounter as they seek to maximize the value of future engineering design projects under significant policy uncertainty.

Flexibility in the form of air pollution mitigation options in a steel plant could generate great value to the firm given future environmental policy uncertainty. Additionally, flexible options could also allow a firm to take advantage of market upsides while protecting from extreme downsides. This chapter explores how decision-makers in a hypothetical steel plant might value flexible design options given uncertainty in environmental policies, specifically in future $\text{SO}_2$ emissions regulations that require flue gas desulfurization (FGD) technology. Results suggest that air pollution mitigation options can offer a significant increase in the NPV of steel facility.

To accommodate uncertainty in design plans, firms typically explore potential valuations of various plans or build options meant to account for sources of economic uncertainty. For heavy industries such as iron and steel, such sources of economic uncertainty include future fluctuations in prices of production input materials, such as iron ore, future global steel prices, and future demand for various steel products. Firms try to capture these future circumstances through financial analysis of design plans based on either trend predictions or average estimations of prices and demand. Some firms consider build options that allow for future flexibility in response to economic uncertainty. For example, flexibility in the
The steel industry might manifest as switch options or assets that enable production lines to change and meet new mixes of steel product demand. Firms typically acknowledge the existence of these sources of economic uncertainty and find quantitative ways to evaluate future investment options.

One important source of political uncertainty includes future environmental regulations and policies that can result in significant financial cost for firms through noncompliance fees, additional costs for technology upgrades, or even through a total shutdown of business. Manufacturing and heavily-polluting industries are particularly vulnerable to these financial costs from environmental policies. One way decision-makers might account for uncertainty in future environmental policies in particular includes real options analysis of flexible design options.

To explore how an industrial firm might value flexible options in response to uncertain future policy outcomes, I develop multiple scenarios of environmental policy uncertainty that may impact a hypothetical iron and steel plant. The exploration of flexible design options serves to simulate the kinds of analysis and decisions executives of the Chinese firm might have faced as they sought to consider the uncertain economic and political future. Since the relocation occurred nearly a decade ago, this case offers a unique opportunity to retrospectively consider the types of decisions the firm actually faced at the time of relocation.

Methodology

This case study values the additional benefit of air pollution mitigation options in a newly constructed steel plant and a real options approach to achieve this. The case study is organized as follows. First, I consider a hypothetical integrated steel plant and FGD technology. I use basic cost estimations and data to construct a deterministic base case of plant valuation for three potential plans. I then consider the valuation of the plans under future policy scenarios in which the probability and timing of policy that requires FGD operation varies. Finally I consider the valuation of the plans under a scenario in which chance of getting caught operating or not operating FGD varies.

Installation and operation of FGD technologies at a steel plant requires a nontrivial capital cost. While costs vary widely, FGD installation and operation costs in the Chinese power sector, for which the most data exists, researchers have estimated cost increases of 3.8% and
2.4% respectively for a 600MW plant to construct a facility with FGD (Cao et al., 2013). Relatively low noncompliance costs, insufficient compliance monitoring, and lax enforcement policies can encourage firms to delay, perhaps indefinitely, FGD investments. Firms that do not swiftly adapt to this level of policy uncertainty, which exists on top of obvious variabilities in steel demand and price, stand at risk of prohibitively high noncompliance penalties, forced reductions in production, or even closure. For this reason, flexible strategies offer an attractive way of mitigating uncertainty in these respects.

3.2 Define the System

To understand how this model uses various cost assumptions and accounts for uncertainty requires an understanding of the system considered in my model.

3.2.1 Overview

This system involves a 1 million ton/year iron and steel plant in Jing-Jin-Ji prior to the 2008 Beijing Olympics. As mentioned in the previous chapter, this period represented an era of explosive economic growth for Jing-Jin-Ji as urban populations and infrastructure grew. To accommodate this regional economic growth, as well as domestic growth across China, the iron and steel production industry in Jing-Jin-Ji plays an important role. Policy mechanisms meant to target increasing concerns of industrial air pollution must consider the important role that steel and other industrial firms play in Jing-Jin-Ji.

FGD technologies scrub \( \text{SO}_2 \) from industrial effluent streams and represent a fairly common air pollution control technology employed by Chinese steel plants. The energy-intensive reduction and refining processes in steel plants result in high emissions of particles and compounds that contribute to local air pollution. As such, the Chinese government has prioritized limits on \( \text{SO}_2 \) emissions that require FGD installation. These requirements have primarily focused on the thermal power sector. Schreifels et al. describe the evolution of \( \text{SO}_2 \) policies and also summarize \( \text{SO}_2 \) reduction goals, results, and the role of FGD technology in both the 10th and 11th Five-year Plans (Schreifels et al., 2012). Lessons learned from \( \text{SO}_2 \) policies implemented in the power sector offer a path forward for similar policies for the iron and steel industry.

Cost data for construction, operation, and installed pollution technologies in Chinese
steel plants proves hard to come by. Because of this, many of the cost estimates used for this analysis represent best estimations extracted from literature. Cost estimates for basic steel facility operation costs primarily came from Ozorio et al. (2013). Table 3.1 summarizes capital and operational costs used for this analysis.

Table 3.1: Characteristics of a 1 million ton per year steel plant

<table>
<thead>
<tr>
<th>Cost Type</th>
<th>Without FGD</th>
<th>With FGD</th>
<th>Cost Increase With FGD versus Without FGD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capital Expenditure ($ million)</td>
<td>500</td>
<td>700</td>
<td>+40%</td>
</tr>
<tr>
<td>Operational Expenditure ($ per ton)</td>
<td>150</td>
<td>170</td>
<td>+13%</td>
</tr>
</tbody>
</table>

Table 3.1 shows the assumed capital and operational expenditures. Adding FGD results in more than 50% higher cost to the firm, with increased capital costs of 40% and increased operation costs of 13%.

This case study primarily considers three possible plans for a new facility with or without FGD over a lifetime of twenty years. The three plans include:

- **Plan A**: Install FGD at time of plant construction. Operate technology whenever future policy gets enacted to avoid environmental compliance penalties.

- **Plan B**: Pay an incremental cost to set aside extra space for FGD installation at plant construction time. Install FGD at a later time to comply with future policy. Pay environmental compliance penalties for one year during FGD installation.

- **Plan C**: Do not install FGD at time of plant construction, do not set aside extra space for later installation, and pay future penalties for not having FGD indefinitely.

These three plans have various advantages and disadvantages. Table 3.2 summarizes the three plans and key reasons a decision-maker would select each plan. Simply put, while plan A requires high upfront capital costs to install FGD technology right away, it completely avoids future penalties from not having FGD. By comparison, plan C requires no additional upfront capital costs for FGD technology. However, a facility that implements plan C faces the risk of costly penalties in the future that could shut down a business entirely. As a
flexible option, plan B incorporates elements from both plans A and C. Importantly, the flexible plan allows the firm to choose to install FGD technology with one year of lead time should policy require FGD installation. A firm would only choose to install FGD if policy required it to. While a plan B facility would pay some upfront cost to reserve space for FGD technology, and potentially forgo profit from the unused space, it does avoid FGD installation costs if future policies turn out to not require FGD installation. The remainder of this case study explores the relative attractiveness of these three possible design plans under different types of policy uncertainty.

Table 3.2: Plan descriptions and purpose

<table>
<thead>
<tr>
<th>Plan</th>
<th>Description</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| A    | Install FGD | ● Equip plant with FGD technology from build time to prevent interruptions to business and costly updates later  
● Continue operation without falling into noncompliance of environmental laws and policies  
● Avoid environmental penalties from not having FGD |
| B    | Set Aside Space | ● Defer decision to install FGD technology until gain more clarity about future environmental policies  
● Pay to set aside extra space for FGD technology that may or may not be installed in the future  
● Allows for flexible planning in event of future environmental policies  
● Enables staged FGD installation for quick build and minimal disruption to regular operation in event of future requirement |
| C    | Nothing Extra | ● Avoid additional cost of FGD installation if future policies or laws do not require technology  
● Allows for consistent planning that incorporates environmental penalties into company financials |

3.2.2 Key policy probability, timing, and enforcement uncertainties

As mentioned in the background section, environmental policies evolve over time and political administrations. This constant policy evolution in the relative short-term creates significant uncertainty for decision-makers making long-term investment decisions. This case study examines three scenarios of uncertainty in timing, stringency, and enforcement of future SO$_2$ control policies, specifically for future FGD installation requirements.
The first uncertainty focuses on the probability of FGD installation policy. To capture this, I assume a future year in which policy is introduced and analyze the value of the three plans when chance of policy ranges from 0% to 100%. This probability of policy reflects the degree of certainty that a policy will be officially announced. Once announced, firms are assumed to be required to comply. For example, a scenario with 75% policy in year 5 represents a world where policy is officially announced 75 times out of 100. This exercise enables an understanding of how certainty of future policy impacts the relative value of each plan.

Relatedly, the second scenario of policy uncertainty deals with timing of FGD installation policy. I assumed a 100% chance of FGD installation policy in various future years. To study the impact of changing the start time of a policy, I consider degree of certainty that the policy will start in year five versus year eight. This policy uncertainty scenario demonstrates how changes in policy timing triggered by administration changes, redirected priorities, or some other external event might make a certain plan look better than another.

The third policy scenario this case study examines uncertainty in policy enforcement, which represents an important challenge in China. As mentioned in chapter 2, China’s central government typically relies on regional and local governments to implement and enforce policies. With environmental policies, local and regional governments have unique dynamics that result in lax enforcement of issues environmental policies. Research has shown that failures in enforcement of environmental regulations come from weak local legitimacy caused by stakeholders with different political interests (van Rooij, 2006). For FGD installation requirements in China, other researchers have examined how lack of political accountability, political and financial incentives, and political support (Schreifels et al., 2012; Cao et al., 2009) contributed to failed emissions targets outlined in the 10th Five-year Plan (2001-2005). For this analysis, relative attractiveness of the plans in this scenario depend on the value of the plans weighted by the probability that officials will check for policy enforcement.

This model assumes constant annual increases in steel slab and iron ore prices for simplicity. Table 3.3 summarizes the main sources of policy uncertainty this case study examines.

### 3.2.3 Steel plant specifications and cost assumptions

This section describes the cost assumptions and various specifications required to define the model. The model uses a discount rate of 10%, which represents a typical estimate for the
Table 3.3: Uncertainty that can impact steel plant

<table>
<thead>
<tr>
<th>Source of Uncertainty</th>
<th>Probability or Timing</th>
<th>Plans Examined</th>
</tr>
</thead>
<tbody>
<tr>
<td>Policy probability in year 5</td>
<td>75% yes policy/ 25% no policy</td>
<td>A, B, C</td>
</tr>
<tr>
<td>Policy timing</td>
<td>Year 5 vs. year 8</td>
<td>A, B, C</td>
</tr>
<tr>
<td>Policy enforcement</td>
<td>75% get caught/ 25% do not get caught</td>
<td>A and C</td>
</tr>
</tbody>
</table>

an industrial steel plant, and examines annual cash flows for twenty years (Ozorio et al., 2013). Table 3.4 summarizes the basic assumptions applied here. Estimates for these cost assumptions are based on data specific to the iron and steel industry (Ozorio et al., 2013) and data specific to FGD technology installed on thermal power plants in China (Xu, 2011).

Table 3.4: Basic data and cost estimates used for base case

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Basic Data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discount rate</td>
<td>10</td>
<td>%</td>
</tr>
<tr>
<td>Time horizon</td>
<td>20</td>
<td>years</td>
</tr>
<tr>
<td>Capacity</td>
<td>1,000,000</td>
<td>tons/year</td>
</tr>
<tr>
<td><strong>Price</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steel price</td>
<td>400</td>
<td>dollar/ton produced</td>
</tr>
<tr>
<td>Annual steel price increase</td>
<td>2</td>
<td>%</td>
</tr>
<tr>
<td><strong>Variable Costs</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iron ore cost</td>
<td>20</td>
<td>dollar/ton of capacity</td>
</tr>
<tr>
<td>Annual ore cost increase</td>
<td>2</td>
<td>%</td>
</tr>
<tr>
<td>Fuel and other production costs</td>
<td>100</td>
<td>dollar/ton</td>
</tr>
<tr>
<td>Plant O&amp;M</td>
<td>30</td>
<td>dollar/ton</td>
</tr>
<tr>
<td>FGD O&amp;M</td>
<td>20</td>
<td>dollar/ton</td>
</tr>
<tr>
<td>Environmental fines</td>
<td>100</td>
<td>dollar/ton</td>
</tr>
<tr>
<td><strong>Fixed Costs</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plant capital</td>
<td>500,000,000</td>
<td>dollars</td>
</tr>
<tr>
<td>FGD capital</td>
<td>200</td>
<td>dollars/ton</td>
</tr>
<tr>
<td>Extra space capital</td>
<td>100,000</td>
<td>dollars</td>
</tr>
</tbody>
</table>

On the revenue side, I made certain assumptions to build the model. Steel slab prices used for the model are set to equal world prices faced by firms in 2005 and begin at US$400. Historical prices for various steel products between 2000 and 2015 increase on average be-
tween 2% and 5% per year (London Metal Exchange, 2017). This model incorporates constant 2% annual increase in steel prices to simulate realistic growth.

On the cost side, iron ore prices represent a major cost to steel production. The model uses an iron ore price of US$20 per ton of capacity based on average prices in 2005 and assumes a 2% annual increase in prices. Energy or fuel costs represent another primary cost to steel production. This model considers cost of fuel and other production materials, coke and coal, together at US$100 per ton of capacity. Operation and management (O&M) costs for the plant total US$30 per ton of capacity. Other than iron ore, ballpark cost estimations for appropriate steel facility variable costs came from Ozorio et al. (2013). Table 3.4 summarizes the various assumptions and cost estimates the model employs.

I use discount cash flow (DCF) analysis to estimate net present value of the new plant. To estimate annual cash flows of the facility, the model relies on a basic equation of costs minus revenue

\[ CF_t = Cap \times (P_t - CI_t - COP - COM) \]  

(3.1)

where \( CF_t \) is cash flow at time \( t \), \( Cap \) is plant capacity, \( P_t \) is the price of steel at time \( t \), \( CI_t \) is the price of iron ore at time \( t \), \( COP \) is the cost of fuel and other production costs, and \( COM \) is the cost of general operation and maintenance for the plant.

The next step involves discounting, on an annual basis, those cash flows to the fixed discount rate specified above of 10%. The equation for discounting is

\[ DCF_t = CF_t / ((1 + r)^T) \]  

(3.2)

where \( DCF_t \) is discounted cash flow at time \( t \), \( r \) is discount rate, and \( T \) is the time horizon. Annual discounted cash flows summed over the entire time horizon of consideration result in a present value of cashflow. Net present value represents the final number that results from a subtraction of capital costs incurred in year zero from the present value of cashflow.

\[ NPV = \sum_{n=0}^{T} DCF_t - FGDCAP - PLANTCAP \]  

(3.3)

As a note, FGD capital costs, \( FGDCAP \), depend on the year of FGD installation. Costs
for FGD technology installed in any year after year zero are discounted in the model.

### 3.3 Analysis of Flexible Options with Uncertainty in the Probability and Timing of Policy

I analyze the attractiveness of a flexible option for steel firms that face environmental policy uncertainty. This option involves paying for and reserving extra space in the new facility to accommodate FGD technology without installing the technology at build time. With this option, firms have the capability of adding and operating FGD technology later if stringent policies come into play. If future policies do not require FGD technology to meet future emissions limits, firms only face the cost of extra reserved space upfront. What this means is that (a) firms can acquire the option to exercise flexibility for a fairly cheap upfront cost and (b) firms can build FGD technology at a later date when the time value of money to build the FGD systems is lower.

#### 3.3.1 Decision analysis

I assumed that decision-makers want to construct a new steel production facility. They know that future policies might require installation and operation of FGD to meet $SO_2$ emissions limits. Decision-makers must decide whether or not to install FGD technology when the plant is initially constructed. Additionally, decision-makers can decide to pay an incremental cost to set purchase and reserve extra space that would allow for future installation of FGD.

The separate build plans are characterized by different cost and time tradeoffs. For example, while the Install FGD option described in plan A starts off with costly capital investments in year zero, estimated at around $200 million for the 1 million ton per year facility, this plan ultimately allows the firm to avoid environmental compliance penalties should policy get enacted. With plan A, firms can choose to operate FGD technology immediately when it proves cost-effective for them to do so given policy changes.

By comparison, plan B incorporates an incremental cost for extra FGD space upfront. This reserved space allows the firm the option to build FGD at a later date should policy get enacted. I assume that the FGD system could be constructed and operated within one year of policy implementation. A plan B facility could face costly FGD capital investments at a later time and would need to pay noncompliance fees for a year while the FGDs are built.
This facility could also not face future FGD capital investments if it turns out that policy will not get enacted. The ability to add FGD at a later date in response to policy is what makes this plan the flexible option. From a managerial perspective, plan B makes sense if the ability to exercise the option in the future offers significant value for firm stakeholders.

Under plan C, decision-makers choose to not install FGD at build time. However, the firm that implements plan C faces costly compliance penalties indefinitely if policy is introduced because they did not install FGD or leave space for FGD technology, which have fairly large space requirements. For simplicity, I assume that the cost of failing to comply with the policy is some costly unknown value that remains constant over the years. Figure 3-2 shows a conceptual model that outlines the three options for the base case considering one flexible option.

![Decision Tree](image)

**Figure 3-2: Decision tree that outlines hypothetical options for steel plant with policy uncertainty**

### 3.3.2 Valuation results

Table 3.5 shows results from NPV analysis. This results demonstrate Expected Net Present Values (ENPV) for the deterministic base case of a facility under each plan.

<table>
<thead>
<tr>
<th>Plan</th>
<th>NPV without policy</th>
<th>NPV with policy</th>
<th>ENPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1894</td>
<td>1800</td>
<td>1823</td>
</tr>
<tr>
<td>B</td>
<td>2094</td>
<td>1813</td>
<td>1883</td>
</tr>
<tr>
<td>C</td>
<td>2094</td>
<td>1622</td>
<td>1740</td>
</tr>
</tbody>
</table>

*Notes: Where ENPV equals NPV with policy times the probability of policy times NPV without policy times the probability of no policy.*
3.3.3 Sensitivity analysis

To explore sensitivity of valuations, I conducted four analyses of how variations of different parameters impacted ENPV for the three plans. The checks included (1) the probability of policy versus no policy, (2) the cost of extra space for FGDs, (3) the probability of policy occurring in two comparative years, and (4) variations in future cost of FGD technology. For this analysis, I use an equation for ENPV that weights NPV for two policy cases as probability for policy on changes.

**Sensitivity test 1: no policy versus policy**

Figure 3-3 shows sensitivity of ENPV to the probability of policy in year 5. For this analysis, I sought to examine the impact of higher or lower policy probabilities on ENPVs. Specifically, I want to know if there is a probability that causes the flexible option, plan B, to not be the best option.

![Figure 3-3: Sensitivity analysis of ENPV rankings to changing assumptions about policy probability](image)

Figure 3-3: Sensitivity analysis of ENPV rankings to changing assumptions about policy probability

Figure 3-3 highlights a few notable results from this analysis. As I would expect, plans B and C start off with the highest NPV values (plan C slightly higher) when there’s no chance of policy occurring since those plans include very little or no upfront investment in FGD at
build time. At 0% policy chance, plan A comes out significantly lower since it includes full FGD install at the time of facility construction. Plan A outperforms plan C at around 50% policy probability. This makes sense because plan C, as the plan with no FGD technology, includes no mitigating financial opportunity to alleviate penalties from future environmental fines, which serves to significantly decline ENPV for plan C as policy probability moves from 0% to 100%.

At 100% policy probability, ENPVs for plans A and B are approximately equal (with ENPV for plan B slightly higher). This makes sense because at 100%, both plans A and B include FGD build. The delayed cost of installation for plan B results in a slightly higher ENPV at 100%. While plan A might look attractive since it results in same ENPV at 100% policy probability, plan B still comes out on top because (1) there are financial benefits from waiting that offset costs of paying one year of fines while building FGD for remaining time and (2) plan B outperforms all other plans at any policy chance less than 100%.

**Sensitivity test 2: FGD space cost**

For this analysis, I examine the impact of higher FGD space costs on ENPVs. To examine this sensitivity, I kept the probability of policy fixed at 75%.

![Figure 3-4: Sensitivity analysis of ENPV rankings to changing assumptions about cost of extra space required for FGD construction](image-url)
I want to know at what cost does plan B, the flexible option that pays the incremental space cost, not appear attractive. Figure 3-4 shows sensitivity of ENPV to the incremental cost of reserving FGD space. Plan A outperforms plan B when FGD space cost is around $70 million or 700 times the assumed cost for the space in this model. This means that FGD space costs would need to approach 15% of total plant capital costs or 35% of FGD capital costs for this form of flexibility to not offer a significant increase in value. By design, the flexibility of plan B is meant to offer major benefits (significant increases to ENPV) for a minor cost (cost of space). At an assumed cost of $100,000, the cost extra space falls within the realm of an error term for overall NPV of the steel plant. As a decision-maker, a marked increase in NPV for such a low cost makes the flexible option appear extremely attractive.

**Sensitivity test 3: policy timing**

The third sensitivity assesses timing of policy in future years. For this analysis, I examine ENPV of the three plans for policy that arrives in year 5 versus in year 8. This sensitivity demonstrates that while there might be some benefit to a firm implementing plan A if future policy occurs later, in year 8, ultimately plan B continues to outperform the other two options.

Figure 3-5 shows changing magnitudes of the gaps between year 5 ENPV and year 8 ENPV as a function of probability for all three plans. Plan B features the widest magnitude increase between the two lines as policy probability increases to 100%. The increasing ENPV of plan B year 8 as policy probability increases, which drives the wide magnitude increase for this plan, suggests that plan B provides significant value for later policy arrival due to deferred costs of installation. The magnitude between the two lines for plan A have the smallest magnitude difference as policy probability increases. This suggests that policy timing does not hugely impact ENPV for a plant that installs the technology at time of construction. The intersections between plans A and C in the two timing scenarios indicate that plan C outperforms plan A at a higher probability when policy arrives later.

From the perspective of a policy-maker, results from this sensitivity analysis are most interesting when policy probability is closer to 100%. At lower policy probabilities, less than 50% chance, plan A clearly offers much lower ENPV when compared with plans B and C. Plan A becomes more attractive at higher probability because it represents the more cautious plan.
Figure 3-5: Sensitivity analysis of ENPV rankings to changing assumptions about the timing of policy

**Sensitivity test 4: future FGD cost**

The fourth sensitivity analysis focuses on increased or decreased cost of FGD technology in future years. Intuitively, it makes most sense that direct capital costs of FGD technology would decrease in the future as learning effects accumulate for the new technology. Innovation and market incentives to manufacture FGDs would result in lowered future costs as companies learn to make the technology. In China, the cost of FGD technology decreased by more than 50% in 20 years as Chinese manufacturers learned to produce FGD systems primarily for the coal-fired power sector (Cao et al., 2009).

Alternatively, direct capital costs of FGD could increase if the technological market stagnates or if future policies abandon the technology in favor of other air pollution control mechanisms. This sensitivity tests how NPV changes when FGD costs increase or decrease by 50% in year five. Fifty percent represents a middle-of-the-road cost increase that might represent either a stagnated technology market for FGDs or accelerated innovation with scale that reduces cost per unit of $SO_2$ removed.

Figure 3-6 shows that at a probability of policy at around 80%, plan A outperforms plan B in the scenario where future FGD costs are +50% more expensive than present FGD costs.
Higher cost assumptions for FGD change the ranking of plans at higher policy probabilities because the attractiveness of reserving space for something that is getting more expensive diminishes the value of holding out. A prudent company decision-maker would likely benefit from selecting plan A with policy probability above 80%.

Figure 3-6: Sensitivity analysis of ENPV rankings to higher or lower FGD costs, for the case of policy uncertainty in year five

This sensitivity analysis suggests that a firm with information that future FGD costs might be more expensive and that policy is almost certain in year 5 should choose plan A. This plan offers higher ENPV at the higher percentage range. Otherwise, at probabilities lower than 80% of future policy, firms would still experience higher ENPVs with plan B, the flexible option.

3.3.4 Value-at-risk and gain

Value at risk and value at gain (VARG) represent a commonly-used financial concept that probabilistically illustrates downside and upside potential of a major project. Conceptually, value at risk represents the maximum loss a firm can expect due to unavoidable market and political uncertainties. It allows a decision-maker to assess probabilistic downside losses and project risk. Conversely, value at gain represents potential upside and project gains
available if the flexible option is exercised. Flexible engineering design provides an approach that decreases downside risk and also increases upside potential. The cumulative distribution function (CDF) illustrates VARG.

I calculated NPVs from the base case. However, this assumed deterministic probabilities of future policy in year 5. This assumption oversimplifies what would happen in the real world since this probability will vary according to political climate, economic growth, external factors, etc. Due to this uncertainty, I used Monte Carlo simulation to generate a range of potential NPVs. Such a range allows me to conduct statistical analysis (e.g. minimum, maximum, distributions, mean values) helpful to decision-making.

For the base case, I examine how the various plans differ in their best and worst performance. Figure 3-7 represents probabilities that realized ENPV will be less than some target value. For plan A, there’s a 10% chance of ENPV below $1,809 million and a 10% chance of ENPV higher than $1,885 million. For plan B, there’s a 10% chance of ENPV below $1,840 million and a 10% chance of ENPV higher than $2,067 million. For plan C, there’s a 10% chance of ENPV below $1,665 million and a 10% chance of ENPV higher than $2,047 million. When compared with the results outlined in Figure 3-2, plan B continues to appear attractive because the worst case simulated scenario for plan B outperforms ENPV values estimated with decision analysis for plans A and C.

![Figure 3-7: Cumulative distribution function for base uncertainty scenario](image-url)
3.4 Analysis of Options for FGD Installation Uncertainty in Policy Enforcement

3.4.1 Decision analysis

The second scenario tests the probability of getting caught either operating or not operating FGD in future years with 100% chance of policy that requires FGD. This analysis examines how ENPV responds to probability of a plant being checked for operational FGD in two extreme cases where the firm either operates or does not operate FGD. The cost of failing to comply with the policy is constant at $100,000 ($100 per ton of capacity) for each year of noncompliance. For this test, I looked only at plans A (install) and C (don’t install). I test three scenarios: plan A operate, plane A don’t operate, and plan C don’t operate. I don’t include plan B (flexible) here because, as outlined earlier in this chapter, plans A and B are similar at 100% future policy. The probability of a firm getting caught not operating their FGD scrubbers depends on effective monitoring and enforcement efforts.

This final sensitivity represents, for example, a case where the central government issues policy directives but leaves enforcement up to regional or local governments. Technology checks for appropriate emissions control systems represent a real-world challenge of this gap between policy issuance and enforcement.

Figure 3-8 shows a conceptual model that outlines options for a policy scenario with enforcement uncertainty.

Figure 3-8: Decision tree that outlines outcomes based on enforcement uncertainty
3.4.2 Valuation Results

Figure 3-9 shows that for policy enforcement probabilities around 40%, the relative value of operate versus don’t operate intersect. This scenario assumes that local compliance teams check plants and facilities once per year for compliance enforcement purposes. This suggests that firms operating in plan A, where a firm has already installed FGD, are better off operating the technology they have installed. Specifically, as the probability of being checked increases, which it’s likely to do for a firm operating in plan A world that has already acknowledged the need for FGD technology in the first place, and policy enforcement improves, firms should operate installed FGD. The precipitous drop in ENPV for plan C highlights the great risk of plan C.

![Figure 3-9: Sensitivity of ENPV to enforcement uncertainty](image)

3.4.3 Value-at-Risk and Gain

For the enforcement uncertainty scenario, I examine how the three plans differ in their best and worst performance. Figure 3-10 shows the distribution of probabilities that realized ENPV will be less than or greater than some target value. ENPV calculations for plan A, operate, include the cost of FGD O&M. Plan A, no operate, and plan C ENPV calculations are based on a 30% chance of getting caught not operating installed FGD. Table 3.6...
summarizes these results.

![Cumulative distribution function for enforcement scenario.](image)

**Figure 3-10:** Cumulative distribution function for enforcement scenario.

**Table 3.6:** Values to assess VARG with chance of getting caught (US$million)

<table>
<thead>
<tr>
<th>Value</th>
<th>Plan A-Operate</th>
<th>Install Plan A-Don’t Operate</th>
<th>Don’t Install Plan C</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENPV</td>
<td>1724</td>
<td>1767</td>
<td>1876</td>
</tr>
<tr>
<td>Std Dev ENPV</td>
<td>0</td>
<td>245</td>
<td>246</td>
</tr>
<tr>
<td>VAR 10%</td>
<td>1724</td>
<td>1221</td>
<td>1331</td>
</tr>
<tr>
<td>VAG 10%</td>
<td>1724</td>
<td>1910</td>
<td>2010</td>
</tr>
<tr>
<td>CAPEX</td>
<td>700</td>
<td>700</td>
<td>500</td>
</tr>
</tbody>
</table>

These results Presenting firms with a table of these probabilities and ranges can help executives decide the plan that offers them a value they can live with.

### 3.5 Discussion

This work employed NPV analysis to value flexibility in the installation of \( S O_2 \) controls in plant design for a new steel facility given uncertainty in future sulfur dioxide emissions policies and enforcement mechanisms. Using financial data assumptions for a 1 million ton/year new steel facility, I constructed a model based on discounted cash flow values to
estimate ENPV of three different potential build plans under different scenarios of policy uncertainty.

### 3.5.1 Value of Flexibility in Steel Industry

Main findings from this exercise demonstrate that flexibility in constructing a scrubber ready facility generates significant value for this hypothetical steel plant. Sensitivity analysis shows that while most of the time the flexible option makes sense, certain scenarios of policy probability result in other options making more sense. Table [3.7] summarizes conditions where other options outperform the flexible one. However, costly $SO_2$ equipment and incomplete policy enforcement could undermine firm responses to control.

**Table 3.7: Sensitivity Analysis Results**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Outcome</th>
<th>Implication</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No Policy vs. Policy</strong></td>
<td>Plan B outperforms A and C</td>
<td>At 100% policy, very little difference between Plans A and B</td>
</tr>
<tr>
<td><strong>FGD Space Cost</strong></td>
<td>Intersection between plans A and B at space cost of $70mill; Space would need to be 700x for the cost of FGD capital to not be worth it</td>
<td>Space cost assumptions make sense for flexibility estimation</td>
</tr>
<tr>
<td><strong>Policy Timing</strong></td>
<td>Plan A, slightly outperforms plan B with 100% probability of policy in year 8</td>
<td>Plan B, later policy at 100% chance, significantly outperforms other options</td>
</tr>
<tr>
<td><strong>Future FGD Cost</strong></td>
<td>At 100% policy probability in year 5, Plan A slightly outperforms flexible option where FGD more expensive in the future</td>
<td>If have info about significantly cheaper future FGD costs, should do plan B</td>
</tr>
</tbody>
</table>

### 3.5.2 Directions for Future Work

This case study only scratches the surface of how flexibility can be valued for a major industrial firm that faces environmental policy uncertain. A variety of increasingly complex and robust tools to value flexibility with the kinds of policy uncertainty exist for explo-
ration. Such tools can offer benefit for firm decision-makers who seek to nimbly adapt to rapid changes in government administrations and policy directions with regards to outlined environmental goals.

This chapter details how I used NPV analysis and real options to assess the value of flexibility in air pollution mitigation options for an iron and steel facility. I find the option of flexibility in this case to generate significant value for decision-makers of the hypothetical steel plant.

In the case with policy uncertainty, plan B, the flexible option outperforms the other two plans most of the time. The required conditions to not make plan B the best option include: when policy is 100% going to happen; when space is greater than 100x the cost of FGD capital; when policy is guaranteed to arrive soon at 100%; when future FGD costs are going to be more 50% more expensive in the future.

Future work might explore alternative flexible designs such as partial build options for required technology, installation of different technology options based on available FGD (e.g. installation of inefficient but cheap FGDs over others), or an option that accounts for intentional production reductions as a result of market behavior or other external events. Future work might test for these alternative flexible designs. The space flexibility option explored here provides a simple yet meaningful exercise for this thesis because it uses real world policy context and plant parameters from China to examine the value of flexibility in decision related to environmental control in an iron and steel firm.

This type of data-driven, technical analysis that incorporates real policy uncertainty poses real value to firms in rapidly developing countries like China.
Chapter 4

Conclusions

This thesis examines county-level location and firm-level decision-making in Jing-Jin-Ji's iron and steel industry between 2005 and 2009, a time of rapid economic growth for China's largest steel producing region, and assesses the value of flexibility in FGD installation for a steel plant that faces environmental policy uncertainty. Insights from these two analytical pieces have important implications for future economic, industrial, and environmental policy for the area, as well as across China.

Chapter 2 examines county-level characteristics across the region at the macro level. I used national data from 147 Jing-Jin-Ji counties to assess key locational characteristics of iron and steel industry growth and concentration between 2005 and 2009. For this assessment, I used financial data from annual national industrial census surveys and socio-economic data from national population census surveys to conduct ordinary least squares regression analysis of six county characteristics. This econometric analysis examined any corresponding relationships between key county factors and iron and steel geographic distribution.

This quantitative analysis confirmed the relationship between key economic factors such as GDP growth and industrial location patterns. Results from the econometric analysis show a significant negative relationship between economic growth of a county, measured in GDP change between 2004 and 2005, and growth of iron and steel shares in that county. These results provide regional context and a descriptive baseline for a stylized case study that zooms into a main character of the Jing-Jin-Ji story, a Shougang-like iron and steel plant.
Chapter 3 addresses firm-level dynamics at the micro level. The case study analyzes the value of flexibility in plant design to explore industrial decision-making in stringency, timing, and enforcement uncertainty in future environmental policies. Case study results show a significant increase in the NPV of a steel facility that employs the flexible option to set aside space for delayed construction of sulfur dioxide scrubbing technology. Implications from the case study include a need for strong policy signals and mechanisms of policy enforcement.

This thesis presents analysis that draws broader lessons for the industrial Jing-Jin-Ji region as it evolves from an economic perspective. One lesson comes from the iron and steel industry, the largest industrial sector in the region. Case study results suggest that weak enforcement of environmental policy diminishes the value of installing $SO_2$ emissions control technology in a steel plant, even with existence of a flexible option. Given the locational patterns of the iron and steel industry before and after the 2008 Beijing Olympics, these insights fit into each other and offer starting points for policy-makers who seek to promote effective economic development plans that incorporate environmental policy.

Insights from the results of technically-focused policy analysis such as outlined in this thesis can support and inform current regional policymaking. A quantitative understanding of drivers of industrial location patterns and of flexible options available to firms can promote focused policies appropriate for different areas. This type of focused policy-making based on county characteristics would complement policies directed to industry. Additionally, this type of analysis can provide a common platform for all stakeholders to discuss and work through their various interests. Stakeholders in this context include Jing-Jin-Ji residents, local businesses, municipal-level policy-makers, Jing-Jin-Ji regional coordination entities. This type of integrated and focused approach proves especially salient now as national and local government administrators use various policy mechanisms to integrate regional cities and balance sustainable economic development across Jing-Jin-Ji in the future.
Appendix A

STATA Code
**Table of Contents for Code**

1.0 CODE to compile CIC data for analysis
   1.1 PREPARE CIC data set
   1.2 SUM main business revenue, total and industry, by county
   1.3 CALCULATE industry shares, by county
   1.4 REPEAT steps 1.1 through 1.3 for years of concern
   1.5 APPEND the CIC year datasets
   1.6 REPLACE: Add county codes to CIC dataset for ease of merging later
   1.7 ADD province and city names

2.0 CODE to compile CEIC data for analysis
   2.1 PROCESS data downloaded from CEIC for POPULATION
   2.2 ADD county codes to facilitate merge
   2.3 REPEAT steps 2.1 through 2.2 for variables of concern
   2.4 MERGE the CEIC datasets for each unique county_code and year variable
   2.5 PROCESS CMR datasets for Hebei in 2005
   2.6 PROCESS CMR datasets for Beijing in 2005
   2.7 PROCESS CMR datasets for Tianjin in 2005
   2.8 APPEND three CMR files
   2.9 REPEAT step 2.2 to add county codes to CMR dataset
   2.10 MERGE CMR data with CEIC data
   2.11 GENERATE new column with 2005 area populated for counties
   2.12 ADD independent variables

3.0 CODE to merge the CEIC-CMR data with the CIC data

4.0 CODE to regress county characteristics to growth and share variables
   4.1 PREPARE data for regression
   4.2 CREATE independent variable columns for earliest and latest years
   4.3 CREATE new independent variable for share difference
   4.4 CREATE new independent variable for main business revenue growth
   4.5 CONVERT some variables to log based on visual distribution
   4.6 DEFINE labels
   4.7 CHECK correlation of independent variables to dependent variables
   4.8 REGRESS county variables and annualized growth with OLS
   4.9 REGRESS county variables and shares with OLS

**Program Listings**

```stata
// Open the "CIC_FINAL_FULL_REAL.dta" panel dataset provided by DZ in 2016.
use "CIC_FINAL_FULL_REAL.dta"
```
// Save data file to maintain original file integrity.
save newcicreal.dta, replace

// Keep variables required for analysis.
keep year city_code county_name_cn province_code firm_id four_digit_sector_code
gross_output main_bus_rev sector_code location_code_6digit

// Drop all years except for one, manually change for additional years.
drop if year!=2005

// Drop all provinces that aren't 1, 2, 3 (JJJ).
drop if province_code!=1 & province_code!=2 & province_code!=3

// Drop counties with no Chinese county name.
gen county_name_has_value = cond(county_name_cn != ".", 1, 0)
bysort location_code_6digit:
gen location_code_identifiable = sum(county_name_has_value)
drop if location_code_identifiable == 0

// Fix inconsistent Chinese county names to facilitate analysis.
replace county_name_cn = "三河市" if county_name_cn == "三河"
replace county_name_cn = "下花园" if county_name_cn == "下花园区"
replace county_name_cn = "东丽区" if county_name_cn == "东丽"
replace county_name_cn = "丰南区" if county_name_cn == "丰南"
replace county_name_cn = "丰台区" if county_name_cn == "丰台"
replace county_name_cn = "丰润区" if county_name_cn == "丰润"
replace county_name_cn = "井陉区" if county_name_cn == "井陉"
replace county_name_cn = "内丘县" if county_name_cn == "内丘"
replace county_name_cn = "内丘县" if county_name_cn == "内邱县"
replace county_name_cn = "北辰区" if county_name_cn == "北辰"
replace county_name_cn = "吴桥县" if county_name_cn == "吴桥"
replace county_name_cn = "和平区" if county_name_cn == "和平"
replace county_name_cn = "复兴区" if county_name_cn == "复兴"
replace county_name_cn = "滨海新区" if county_name_cn == "大港"
replace county_name_cn = "孟村回族自治县" if county_name_cn == "孟村"
replace county_name_cn = "宁河县" if county_name_cn == "宁河"
replace county_name_cn = "安次区" if county_name_cn == "安次"
replace county_name_cn = "宝坻区" if county_name_cn == "宝坻"
replace county_name_cn = "宽城满族自治县" if county_name_cn == "宽城县"
replace county_name_cn = "平山县" if county_name_cn == "平山"
replace county_name_cn = "平泉县" if county_name_cn == "平泉"
replace county_name_cn = "平谷区" if county_name_cn == "平谷"
replace county_name_cn = "开平区" if county_name_cn == "开平"
replace county_name_cn = "徐水县" if county_name_cn == "徐水"
replace county_name_cn = "房山区" if county_name_cn == "房山"
replace county_name_cn = "抚宁县" if county_name_cn == "抚宁"
replace county_name_cn = "文安县" if county_name_cn == "文安"
replace county_name_cn = "新乐市" if county_name_cn == "新乐"
replace county_name_cn = "昌平区" if county_name_cn == "昌平"
replace county_name_cn = "昌黎县" if county_name_cn == "昌黎"
replace county_name_cn = "景县" if county_name_cn == "景 县"
replace county_name_cn = "曲阳县" if county_name_cn == "曲阳"
replace county_name_cn = "滦南县" if county_name_cn == "滦南"
replace county_name_cn = "武安市" if county_name_cn == "武安"
replace county_name_cn = "武清区" if county_name_cn == "武清"
replace county_name_cn = "滨海新区" if county_name_cn == "滨海新" 
replace county_name_cn = "河东区" if county_name_cn == "河东" 
replace county_name_cn = "河北区" if county_name_cn == "河北" 
replace county_name_cn = "河西区" if county_name_cn == "河西" 
replace county_name_cn = "津南区" if county_name_cn == "津南" 
replace county_name_cn = "海港区" if county_name_cn == "海港开" 
replace county_name_cn = "玉田县" if county_name_cn == "玉田" 
replace county_name_cn = "井陉矿区" if county_name_cn == "矿区" 
replace county_name_cn = "肃州区" if county_name_cn == "肃州" 
replace county_name_cn = "蓟县" if county_name_cn == "蓟 县"
replace county_name_cn = "行唐县" if county_name_cn == "行唐"
replace county_name_cn = "西青区" if county_name_cn == "西青"
replace county_name_cn = "路南区" if county_name_cn == "路南"
replace county_name_cn = "迁安市" if county_name_cn == "迁安"
replace county_name_cn = "迁西县" if county_name_cn == "迁西"
replace county_name_cn = "运河区" if county_name_cn == "运河"
replace county_name_cn = "遵化市" if county_name_cn == "遵化"
replace county_name_cn = "邯山区" if county_name_cn == "邯郸"
replace county_name_cn = "邯山区" if county_name_cn == "邯山区"
replace county_name_cn = "长安区" if county_name_cn == "长安"
replace county_name_cn = "青龙满族自治县" if county_name_cn == "青龙县"
replace county_name_cn = "静海县" if county_name_cn == "静海"
replace county_name_cn = "鸡泽县" if county_name_cn == "鸡泽"
replace county_name_cn = "鹿泉区" if county_name_cn == "鹿泉"
replace county_name_cn = "黄骅市" if county_name_cn == "黄骅"
replace county_name_cn = "南开区" if county_name_cn == "南开"
replace county_name_cn = "北辰区" if county_name_cn == "双街"
replace county_name_cn = "滨海新区" if county_name_cn == "滨海新区"
replace county_name_cn = "滨海新区" if county_name_cn == "大港区"
replace county_name_cn = "红桥区" if county_name_cn == "西青区"
replace county_name_cn = "静海县" if county_name_cn == "静海"
replace county_name_cn = "乐亭县" if county_name_cn == "南开"
replace county_name_cn = "曹妃甸区" if county_name_cn == "曹妃甸区"
replace county_name_cn = "滦南县" if county_name_cn == "滦南"
replace county_name_cn = "山海关区" if county_name_cn == "山海关"
replace county_name_cn = "开发区" if county_name_cn == "开发区"
replace county_name_cn = "邢台县" if county_name_cn == "邢台"
replace county_name_cn = "莲池区" if county_name_cn == "莲池区"
replace county_name_cn = "任丘市" if county_name_cn == "任丘市"
replace county_name_cn = "任丘市" if county_name_cn == "任丘市"
replace county_name_cn = "河间市" if county_name_cn == "河间市"
replace county_name_cn = "海兴县" if county_name_cn == "海兴县"
replace county_name_cn = "大厂回族自治县" if county_name_cn == "大厂回族自治县"
replace county_name_cn = "永清县" if county_name_cn == "永清县"
replace county_name_cn = "霸州市" if county_name_cn == "霸州市"
replace county_name_cn = "饶阳县" if county_name_cn == "饶阳县"
replace county_name_cn = "滦南县" if county_name_cn == "滦南县"
replace county_name_cn = "任丘市" if county_name_cn == "任丘市"
replace county_name_cn = "抚宁县" if county_name_cn == "抚宁县"
replace county_name_cn = "清苑区" if county_name_cn == "清苑区"
replace county_name_cn = "南和县" if county_name_cn == "南和县"
replace county_name_cn = "藁城市" if county_name_cn == "藁城市"
// Dropped unknown/ambiguous county names because they mess up data merge.
drop if county_name_cn == "天津"
drop if county_name_cn == "市辖区"
drop if county_name_cn == "保税"
drop if county_name_cn == "军粮城"
drop if county_name_cn == "大寺"
drop if county_name_cn == "开发区"

drop if county_name_cn == "开发区"

drop if county_name_cn == "天津" 

********************************************************************************
* 1.2 SUM main business revenue, total and industry, by county                *
********************************************************************************
// Sum values of all observations of a variable using egen.
sort county_name_cn
by county_name_cn: egen county_bus_rev = sum(main_bus_rev)

// Sum main business revenue by county for puddling industry.
sort county_name_cn sector_code
by county_name_cn: egen isp_county_bus_rev = sum(main_bus_rev) ///
    if sector_code==3210

// Sum main business revenue by county for steel smelting industry.
sort county_name_cn sector_code
by county_name_cn: egen iss_county_bus_rev = sum(main_bus_rev) ///
    if sector_code==3220

// Sum main business revenue by county for steel rolling industry.
sort county_name_cn sector_code
by county_name_cn: egen isr_county_bus_rev = sum(main_bus_rev) ///
    if sector_code==3230

// Sum main business revenue by county for ferroalloy industry.
sort county_name_cn sector_code
by county_name_cn: egen isf_county_bus_rev = sum(main_bus_rev) ///
    if sector_code==3240

// Sum business revenue for 4 sector codes to get total iron and steel industry.
egen istot_county_bus_rev = rsum(isp_county_bus_rev iss_county_bus_rev ///
    isr_county_bus_rev isf_county_bus_rev)

********************************************************************************
* 1.3 CALCULATE industry shares, by county                                     *
********************************************************************************
// Keep only iron and steel industry (Ferrous_metal_smelting_and_rolling).
drop if sector_code!=3210 & sector_code!=3220 & sector_code!=3230 & ///
    sector_code!=3240

// Share of iron and steel industry.
sort county_name_cn
by county_name_cn: ///
    generate istot_shr_county = istot_county_bus_rev/county_bus_rev

// Contract data set to include iron and steel industry shares by county.
contract year istot_shr_county city_code province_code county_name_cn ///
    county_bus_rev

// Drop the variable that encodes for number of county observations.
drop _freq

// Replace missing values with zeros.
replace istot_shr_county = 0 if istot_shr_county==.
// Sum firm shares to get county-level shares.
sort county_name_cn
by county_name_cn: egen istot_shr_temp = sum(istot_shr_county)

// Sum firm business revenue to get county-level shares.
sort county_name_cn
by county_name_cn: egen istot_county_bus_rev_temp = sum(istot_county_bus_rev)

// Contract so that each row includes all share information.
contract year istot_shr_temp city_code province_code county_name_cn istot_county_bus_rev_temp

// Drop the variable that encodes for number of observations.
drop _freq

// Save as a .dta file for merging with format year+is+real.
// Manually change for additional years.
save 2005isreal.dta, replace

********************************************************************************
* 1.4 REPEAT steps 1.1 through 1.3 for years of concern
********************************************************************************
* 2005isreal.dta
* 2006isreal.dta
* 2007isreal.dta
* 2008isreal.dta
* 2009isreal.dta

********************************************************************************
* 1.5 APPEND the CIC year datasets
********************************************************************************
// Open first observed year .dta file.
clear
use 2005isreal.dta
// Merge data sets with the same variables and different cases.
append using 2006isreal.dta
append using 2007isreal.dta
append using 2008isreal.dta
append using 2009isreal.dta

// Sort for ease
sort year
sort county_name_cn

********************************************************************************
* 1.6 REPLACE: Add county codes to CIC dataset for ease of merging later
********************************************************************************
// Create a new column for county code.
gen county_code=0
tostring county_code, replace

// Replace Chinese county name with county code. (Source: GIS, Google Translate)
replace county_code = "130729" if county_name_cn == "万全县"
replace county_code = "130822" if county_name_cn == "三河市"
replace county_code = "130706" if county_name_cn == "下花园区"
replace county_code = "120110" if county_name_cn == "东丽区"
replace county_code = "130207" if county_name_cn == "丰南区"
replace county_code = "110106" if county_name_cn == "丰台区"
replace county_code = "130207" if county_name_cn == "丰南区"
replace county_code = "130208" if county_name_cn == "丰南区"
replace county_code = "130523" if county_name_cn == "内丘县"
replace county_code = "120113" if county_name_cn == "北辰区"
replace county_code = "130928" if county_name_cn == "吴桥县"
replace county_code = "120101" if county_name_cn == "和平区"
replace county_code = "130404" if county_name_cn == "复兴区"
replace county_code = "130930" if county_name_cn == "孟村回族自治县"
replace county_code = "120221" if county_name_cn == "宁河区"
replace county_code = "131002" if county_name_cn == "安次区"
replace county_code = "130827" if county_name_cn == "宝坻区"
replace county_code = "130131" if county_name_cn == "平山县"
replace county_code = "130823" if county_name_cn == "开平区"
replace county_code = "130115" if county_name_cn == "宁河区"
replace county_code = "130827" if county_name_cn == "宝坻区"
replace county_code = "130184" if county_name_cn == "新乐市"
replace county_code = "130114" if county_name_cn == "昌平区"
replace county_code = "130322" if county_name_cn == "昌黎县"
replace county_code = "131112" if county_name_cn == "玉田县"
replace county_code = "130225" if county_name_cn == "乐亭县"
replace county_code = "130229" if county_name_cn == "玉田县"
replace county_code = "130107" if county_name_cn == "井陉矿区"
replace county_code = "130125" if county_name_cn == "行唐县"
replace county_code = "120111" if county_name_cn == "西青区"
replace county_code = "130202" if county_name_cn == "路南区"
replace county_code = "130283" if county_name_cn == "迁安市"
replace county_code = "130227" if county_name_cn == "迁西县"
replace county_code = "130903" if county_name_cn == "迁安市"
replace county_code = "130281" if county_name_cn == "迁西县"
replace county_code = "130481" if county_name_cn == "迁安市"
replace county_code = "130114" if county_name_cn == "昌平区"
replace county_code = "130625" if county_name_cn == "徐水区"
replace county_code = "130111" if county_name_cn == "徐水区"
replace county_code = "130205" if county_name_cn == "徐水区"
replace county_code = "130185" if county_name_cn == "晋州市"
replace county_code = "130225" if county_name_cn == "乐亭县"
replace county_code = "130209" if county_name_cn == "曹妃甸区"
replace county_code = "130303" if county_name_cn == "山海关区"
replace county_code = "130607" if county_name_cn == "遵化市"
replace county_code = "130223" if county_name_cn == "遵化市"
replace county_code = "130109" if county_name_cn == "遵化市"
replace county_code = "130102" if county_name_cn == "遵化市"
replace county_code = "130281" if county_name_cn == "遵化市"
replace county_code = "130227" if county_name_cn == "遵化市"
replace county_code = "130274" if county_name_cn == "遵化市"
replace county_code = "130402" if county_name_cn == "遵化市"
replace county_code = "130302" if county_name_cn == "遵化市"
replace county_code = "130321" if county_name_cn == "青龙满族自治县"
replace county_code = "130982" if county_name_cn == '任丘市'
replace county_code = "130305" if county_name_cn == '任丘市'
replace county_code = "130608" if county_name_cn == '任丘市'
replace county_code = "130527" if county_name_cn == '任丘市'
replace county_code = "110101" if county_name_cn == '任丘市'
replace county_code = "130826" if county_name_cn == '任丘市'
replace county_code = "130522" if county_name_cn == '任丘市'
replace county_code = "130132" if county_name_cn == '任丘市'
replace county_code = "130822" if county_name_cn == '任丘市'
replace county_code = "131181" if county_name_cn == '任丘市'
replace county_code = "130581" if county_name_cn == '任丘市'
replace county_code = "130604" if county_name_cn == '任丘市'
replace county_code = "130324" if county_name_cn == '任丘市'
replace county_code = "130803" if county_name_cn == '任丘市'
replace county_code = "130204" if county_name_cn == '任丘市'
replace county_code = "110115" if county_name_cn == '任丘市'
replace county_code = "130705" if county_name_cn == '任丘市'
replace county_code = "130532" if county_name_cn == '任丘市'
replace county_code = "131003" if county_name_cn == '任丘市'
replace county_code = "130730" if county_name_cn == '任丘市'
replace county_code = "130424" if county_name_cn == '任丘市'
replace county_code = "130721" if county_name_cn == '任丘市'
replace county_code = "130821" if county_name_cn == '任丘市'
replace county_code = "130435" if county_name_cn == '任丘市'
replace county_code = "130602" if county_name_cn == '任丘市'
replace county_code = "130435" if county_name_cn == '任丘市'
replace county_code = "110105" if county_name_cn == '任丘市'
replace county_code = "130124" if county_name_cn == '任丘市'
replace county_code = "130124" if county_name_cn == '任丘市'
replace county_code = "131102" if county_name_cn == '任丘市'
replace county_code = "130702" if county_name_cn == '任丘市'
replace county_code = "130503" if county_name_cn == '任丘市'
replace county_code = "130703" if county_name_cn == '任丘市'
replace county_code = "130103" if county_name_cn == '任丘市'
replace county_code = "130702" if county_name_cn == '任丘市'
replace county_code = "130502" if county_name_cn == '任丘市'
replace county_code = "130104" if county_name_cn == '任丘市'
replace county_code = "130703" if county_name_cn == '任丘市'
replace county_code = "130503" if county_name_cn == '任丘市'
replace county_code = "130123" if county_name_cn == '任丘市'
replace county_code = "131122" if county_name_cn == '任丘市'
replace county_code = "130429" if county_name_cn == '任丘市'
replace county_code = "130582" if county_name_cn == '任丘市'
replace county_code = "130921" if county_name_cn == '任丘市'
replace county_code = "110108" if county_name_cn == '任丘市'
replace county_code = "130426" if county_name_cn == '任丘市'
replace county_code = "130623" if county_name_cn == '任丘市'
replace county_code = "130630" if county_name_cn == '任丘市'
replace county_code = "130681" if county_name_cn == '任丘市'
replace county_code = "130534" if county_name_cn == '任丘市'
replace county_code = "130621" if county_name_cn == '任丘市'
replace county_code = "130223" if county_name_cn == '任丘市'
replace county_code = "130824" if county_name_cn == '任丘市'
replace county_code = "130126" if county_name_cn == '任丘市'
replace county_code = "110107" if county_name_cn == '任丘市'
replace county_code = "130427" if county_name_cn == '任丘市'
replace county_code = "130428" if county_name_cn == '任丘市'
replace county_code = "130129" if county_name_cn == '任丘市'
replace county_code = "130133" if county_name_cn == '赵县'
replace county_code = "130203" if county_name_cn == '路北区'
replace county_code = "130181" if county_name_cn == '辛集市'
replace county_code = "130112" if county_name_cn == '通州区'
replace county_code = "110109" if county_name_cn == '门头沟区'
replace county_code = "130624" if county_name_cn == '阜平县'
replace county_code = "130727" if county_name_cn == '阜城县'
replace county_code = "130825" if county_name_cn == '阳原县'
replace county_code = "130922" if county_name_cn == '隆化县'
replace county_code = "110113" if county_name_cn == '顺义区'
replace county_code = "130104" if county_name_cn == '香河县'
replace county_code = "130804" if county_name_cn == '高邑县'
replace county_code = "130182" if county_name_cn == '鹰手营子矿区'

// Drop unknown/mistaken counties.
drop if county_name_cn=="马头生态工业城"
drop if county_name_cn=="丰南区" & province_code==1

********************************************************************************
* 1.7 ADD province and city names          *
********************************************************************************
// Generate a city column.
gen city = 0
tostring city, replace
// Make city_code a string.
tostring city_code, replace
// Add city names as per city codes.
replace city = "Beijing" if city_code == "1"
replace city = "Tianjin" if city_code == "2"
replace city = "Shijiazhuang" if city_code == "3"
replace city = "Tangshan" if city_code == "4"
replace city = "Qinhuangdao" if city_code == "5"
replace city = "Handan" if city_code == "6"
replace city = "Xingtai" if city_code == "7"
replace city = "Baoding" if city_code == "8"
replace city = "Zhangjiakou" if city_code == "9"
replace city = "Chengde" if city_code == "10"
replace city = "Cangzhou" if city_code == "11"
replace city = "Langfang" if city_code == "12"
replace city = "Hengshui" if city_code == "13"
// Drop city_code.
drop city_code

// Generate a province column.
gen province = 0
tostring province, replace
// Make province_code a string.
tostring province_code, replace
// Add province names as per province codes.
replace province = "Beijing" if province_code == "1"
replace province = "Tianjin" if province_code == "2"
replace province = "Hebei" if province_code == "3"
// Drop province_code.
drop province_code
// Save file.
save jjj-isreal-codes.dta, replace

*=============================================================================*
* 2.0 CODE to compile CEIC data for analysis
* Dataset used: JJJ-County-pop-CEIC-05142017.xlsx
* Output: jjj-char.dta
*=============================================================================*
capture log close
set more off
clear all
set matsize 10000
macro drop _all
set linesize 80

*********************************************************************************
* 2.1 PROCESS data downloaded from CEIC for POPULATION *
*********************************************************************************
// Manually add format the year variable name as "data####" in spreadsheet.
// Import manually edited CEIC output spreadsheet.
import excel using "JJJ-County-pop-CEIC-05142017.xlsx", firstrow

// Split column one by indicated separator.
split SelectthislinkandclickRefre, p(" ")
// Rename province, city, and county (la var = label variable).
rename SelectthislinkandclickRefre3 province
rename SelectthislinkandclickRefre4 city
rename SelectthislinkandclickRefre5 county

// Drop columns don't need.
drop SelectthislinkandclickRefre Country Frequency Unit Source Status ///
   SeriesCode FunctionInformation FirstObsDate LastObsDate LastUpdateTime ///
   Remarks Mean Variance StandardDeviation Skewness Kurtosis ///
   CoefficientVariation Min Max Median NoofObs SelectthislinkandclickRefre1 ///
   SelectthislinkandclickRefre2

// Use reshape command to put into panel format.
reshape long data, i(province city county) j(time)

// Move city names for Beijing and Tianjin to the county column.
replace county = city if province == "Beijing"
replace county = city if province == "Tianjin"

// Replace city names for Beijing and Tianjin with province names.
replace city = "Beijing" if province == "Beijing"
replace city = "Tianjin" if province == "Tianjin"

// Rename data column to variable of interest and add label.
rename data pop
la var pop "Population (Person th)"

// Rename time column.
rename time year

*********************************************************************************
2.2 ADD county codes to facilitate merge

// Create a new column for county code.
gen county_code=0
tostring county_code, replace

// Replace English county name with county code. (Source: GIS, Google Translate)
replace county_code = "130683" if county == "Anguo"
replace county_code = "131125" if county == "Anping"
replace county_code = "130632" if county == "Anxin"
replace county_code = "130524" if county == "Baixiang"
replace county_code = "130981" if county == "Botou"
replace county_code = "130637" if county == "Boye"
replace county_code = "130732" if county == "Chicheng"
replace county_code = "130733" if county == "Chongli"
replace county_code = "131025" if county == "Dacheng"
replace county_code = "130425" if county == "Daming"
replace county_code = "130626" if county == "Dingxing"
replace county_code = "130682" if county == "Dingzhou"
replace county_code = "130628" if county == "Gaoyang"
replace county_code = "131022" if county == "Gu'an"
replace county_code = "130432" if county == "Guangping"
replace county_code = "130531" if county == "Guangzhou"
replace county_code = "130635" if county == "Guangzong"
replace county_code = "130433" if county == "Guangzhou"
replace county_code = "131126" if county == "Guangzhou"
replace county_code = "130724" if county == "Guangzhou"
replace county_code = "130728" if county == "Guangzhou"
replace county_code = "110116" if county == "Huairou"
replace county_code = "130121" if county == "Jingxing"
replace county_code = "130723" if county == "Kangbao"
replace county_code = "130635" if county == "Liu"
replace county_code = "130535" if county == "Linxi"
replace county_code = "130423" if county == "Linzhang"
replace county_code = "130525" if county == "Longyao"
replace county_code = "130525" if county == "Longzao"
replace county_code = "110228" if county == "Miyun"
replace county_code = "13052" if county == "Nanping"
replace county_code = "130528" if county == "Nieguangping"
replace county_code = "130622" if county == "Qingyuan"
replace county_code = "130430" if county == "Qiu"
replace county_code = "130526" if county == "Ren"
replace county_code = "130629" if county == "Rongcheng"
replace county_code = "130725" if county == "Shangyi"
replace county_code = "130128" if county == "Shanxian"
replace county_code = "131182" if county == "Shenzhen"
replace county_code = "130636" if county == "Shunping"
replace county_code = "130926" if county == "Suning"
replace county_code = "130627" if county == "Tangshan"
replace county_code = "130230" if county == "Tangshan"
replace county_code = "130631" if county == "Wangdu"
replace county_code = "130533" if county == "Wei & city == "Xingtai"
replace county_code = "130434" if county == "Wei & city == "Handan"
replace county_code = "130828" if county == "Weichang"
replace county_code = "130130" if county == "Wuji"
replace county_code = "131123" if county == "Wuqiang"
replace county_code = "130530" if county == "Xianhe"
replace county_code = "130638" if county == "Xiong"
replace county_code = "131121" if county == "Zaoqiang" 
replace county_code = "130722" if county == "Zhangbei" 
replace county_code = "130731" if county == "Zhuofu" 
replace county_code = "130729" if county == "Wanquan" 
replace county_code = "131082" if county == "Sanhe" 
replace county_code = "130207" if county == "Fengnan" 
replace county_code = "130208" if county == "Fengrun" 
replace county_code = "130523" if county == "Neiqiu" 
replace county_code = "130928" if county == "Wuqiao" 
replace county_code = "130930" if county == "Mengcun" 
replace county_code = "120221" if county == "Ninghe" 
replace county_code = "120115" if county == "Baodi" 
replace county_code = "130827" if county == "Kuancheng" 
replace county_code = "130131" if county == "Pingshan" 
replace county_code = "130823" if county == "Pingquan" 
replace county_code = "110117" if county == "Pinggu" 
replace county_code = "130625" if county == "Xushui" 
replace county_code = "130323" if county == "Funing" 
replace county_code = "131026" if county == "Wen'an" 
replace county_code = "130324" if county == "Xinle" 
replace county_code = "130322" if county == "Changli" 
replace county_code = "131127" if county == "Jing" 
replace county_code = "130634" if county == "Quyang" 
replace county_code = "130224" if county == "Luanan" 
replace county_code = "130481" if county == "Wu'an" 
replace county_code = "130429" if county == "Yutian" 
replace county_code = "130107" if county == "Jingxingkuangqu" 
replace county_code = "120225" if county == "Jixian" 
replace county_code = "130127" if county == "Qian'an" 
replace county_code = "130126" if county == "Qianxi" 
replace county_code = "130291" if county == "Zunhua" 
replace county_code = "130421" if county == "Handan" 
replace county_code = "130321" if county == "Qinglong" 
replace county_code = "120223" if county == "Jinghai" 
replace county_code = "130431" if county == "Ji" 
replace county_code = "130185" if county == "Luzhou" 
replace county_code = "130983" if county == "Huanghua" 
replace county_code = "130183" if county == "Jinzhou" 
replace county_code = "130225" if county == "Laoting" 
replace county_code = "130225" if county == "Luyi" 
replace county_code = "130982" if county == "Rengiu" 
replace county_code = "130984" if county == "Heijian" 
replace county_code = "130924" if county == "Haixing" 
replace county_code = "131028" if county == "Dachang" 
replace county_code = "131023" if county == "Yongqian" 
replace county_code = "131081" if county == "Bazhou" 
replace county_code = "131124" if county == "Raoyang" 
replace county_code = "130527" if county == "Nanhe" 
replace county_code = "130826" if county == "Fengning" 
replace county_code = "130522" if county == "Licheng" 
replace county_code = "130132" if county == "Yuanshi" 
replace county_code = "130822" if county == "Xinglong" 
replace county_code = "131181" if county == "Jizhou" 
replace county_code = "130581" if county == "Nangong" 
replace county_code = "130324" if county == "Luoliang" 
replace county_code = "110115" if county == "Daxing" 
replace county_code = "130705" if county == "Xuanhua" 
replace county_code = "130529" if county == "Julu" 
replace county_code = "130532" if county == "Pingxiang" 
replace county_code = "130730" if county == "Huaibei" 
replace county_code = "130424" if county == "Cheng'an"
replace county_code = "130821" if county == "Chengde"
replace county_code = "130435" if county == "Quzhou"
replace county_code = "130124" if county == "Luancheng"
replace county_code = "130123" if county == "Zhengding"
replace county_code = "131122" if county == "Wuyi"
replace county_code = "130429" if county == "Yongnian"
replace county_code = "130582" if county == "Shahe"
replace county_code = "130521" if county == "Cang"
replace county_code = "130426" if county == "She"
replace county_code = "130623" if county == "Lai shui"
replace county_code = "130630" if county == "Laiyuan"
replace county_code = "130681" if county == "Zhuozhou"
replace county_code = "130534" if county == "Qinghe"
replace county_code = "130621" if county == "Mancheng"
replace county_code = "130223" if county == "Luan"
replace county_code = "130824" if county == "Luanping"
replace county_code = "130126" if county == "Lingshou"
replace county_code = "130427" if county == "Ci"
replace county_code = "130428" if county == "Feixiang"
replace county_code = "130129" if county == "Zanhuang"
replace county_code = "130133" if county == "Zhao"
replace county_code = "130181" if county == "Xinji"
replace county_code = "131128" if county == "Fucheng"
replace county_code = "130624" if county == "Fuping"
replace county_code = "130727" if county == "Yangyuan"
replace county_code = "130825" if county == "Longhua"
replace county_code = "130922" if county == "Qing"
replace county_code = "1301024" if county == "Xianghe"
replace county_code = "130684" if county == "Gaobelian"
replace county_code = "130127" if county == "Gaoyi"
replace county_code = "130182" if county == "Gaocheng"

// Save data set for merging later.
save "ceic-pop.dta", replace

******************************************************************************
* 2.3 REPEAT steps 2.1 through 2.2 for variables of concern
******************************************************************************
* ceic-gdp.dta
* ceic-emprur
* ceic-emprurff.dta
* ceic-govrev.dta
* ceic-fai.dta
* ceic-wages.dta
******************************************************************************

******************************************************************************
* 2.4 MERGE the CEIC datasets for each unique county code and year variable
******************************************************************************

// Open first dataset.
clear
use "ceic-pop.dta"

// Merge 1:1 since each county and each year has only one value.
merge 1:1 year county_code using "ceic-gdp.dta"

// Drop _merge column so can merge more datasets.
drop _merge

// Repeat above for each additional dataset.
merge 1:1 year county_code using "ceic-emprur.dta"
drop _merge
merge 1:1 year county_code using "ceic-emprurff.dta"
drop _merge
merge 1:1 year county_code using "ceic-govrev.dta"
drop _merge
merge 1:1 year county_code using "ceic-fai.dta"
drop _merge
merge 1:1 year county_code using "ceic-wages.dta"
drop _merge

// Sort for easy viewing of final dataset.
sort province
sort city
sort year
sort county

// Save for merging later.
save ceic-fullpanel.dta, replace

********************************************************************************
* 2.5 PROCESS CMR datasets for Hebei in 2005 *
********************************************************************************
// Import CEIC output spreadsheet.
clear
import excel using "Hebei-Land-2005-CMR-05142017.xlsx", ///
cellrange(A4:F142) firstrow

// Split the district column and rename county column.
split District
rename District1 county

// Drop columns that do not need.
drop District District2 District3 District4 District5 District6 NumberofTownsunit NumberofVillagersCommittees TotalAgriculturalMachineryPow

// Rename two other variables.
rename AreaofAdministrativeRegion10 area
rename NumberofLocalTelephoneSubscr telesub

// Fix triple Wei county problem.
gen city = ""
replace city="Xingtai" if county=="Wei" & telesub==4.5
replace city="Handan" if county=="Wei" & telesub==8.3
replace county="Yu" if county=="Wei" & telesub==6.5

// Add province column for number matching.
gen province = "Hebei"

// Add year column.
gen year = 2005

// Save for merge
save "cmr-hebei-land-2005.dta", replace

********************************************************************************
* 2.6 PROCESS CMR datasets for Beijing in 2005 *
********************************************************************************
// Import CEIC output spreadsheet.
clear
import excel using "Beijing-Land-2005-CMR-05142017.xlsx", ///
cellrange(A4:F9) firstrow

// Split the district column and rename county column.
split District
rename District1 county

// Drop columns that do not need.
drop District District2 NumberofTownsunit NumberofVillagersCommittees //
TotalAgriculturalMachineryPow

// Rename two other variables.
rename AreaofAdministrativeRegion10 area
rename NumberofLocalTelephoneSubscr telesub

// Add province column for number matching.
gen province = "Beijing"

// Add year column.
gen year = 2005

// Save for merge
save "cmr-beijing-land-2005.dta", replace

********************************************************************************
* 2.7 PROCESS CMR datasets for Tianjin in 2005
********************************************************************************
// Import CEIC output spreadsheet.
clear
import excel using "Tianjin-Land-2005-CMR-05142017.xlsx", //
cellrange(A4:F8) firstrow
// Split the district column and rename county column.
split District
rename District1 county
// Drop columns that do not need.
drop District District2 NumberofTownsunit NumberofVillagersCommittees ///
TotalAgriculturalMachineryPow

// Rename two other variables.
rename AreaofAdministrativeRegion10 area
rename NumberofLocalTelephoneSubscr telesub

// Fix Ji/Jixian/Jizhou Tianjin naming problem.
replace county="Jixian" if county="Ji"
// Add province column for number matching.
gen province = "Tianjin"

// Add year column.
gen year = 2005

// Save for merge.
save "cmr-tianjin-land-2005.dta", replace

********************************************************************************
* 2.8 APPEND three CMR files
********************************************************************************
// Open first Hebei .dta file.
clear
use cmr-hebei-land-2005.dta
// Merge data sets with the same variables and different cases.
append using cmr-beijing-land-2005.dta
append using cmr-tianjin-land-2005.dta
sort county

* 2.9 REPEAT step 2.2 to add county codes to CMR dataset
********************************************************************************
* 2.10 MERGE CMR data with CEIC data

// Use 1:1 merge command.
merge 1:1 year county_code using "ceic-fullpanel.dta"
drop _merge

* 2.11 GENERATE new column with 2005 area populated for counties

// Generate a new variable for area and populate with zeros.
// Replace zeros in new variable column with variables from the original variable column for year of interest by county.
// Generate a second new variable column with values from original variable column for year of interest by county.
// Rename second new variable column with old variable name since has numeric values desired.
sort county year
gen area2005 = 0
replace area2005 = area if year==2005
by county: egen area2005ref = max(area2005)
drop area2005
rename area2005ref area2005
// Drop original area column.
drop area

* 2.12 ADD independent variables

// Calculate average wages per capita.
gen WGPC = wage / pop

// Calculate ratio of agricultural employment to total population.
gen FFRU = emprurff / pop

// Calculate ratio of government revenue to GDP.
gen GOVR = govrev / gdp

// Calculate ratio of FAI to GDP.
gen FAIV = fai / gdp

// Calculate population density.
gen POPD = pop / area2005

// Create variable of % GDP Change between 2004 and 2005.
// Generate a new variable for 2004 GDP and populate with zeros.
// Generate a second new variable column with values from the original variable column for year of interest.
// Drop first new variable county.
// Rename second new variable column with old variable name since has numeric values desired.
sort county year
gen gdp2004 = 0
replace gdp2004 = gdp if year==2004
by county: egen gdp2004ref = max(gdp2004)
drop gdp2004
rename gdp2004ref gdp2004

// Generate a new variable for 2005 GDP and populate with zeros.
// Replace zeros in new variable column with variables from the original
// variable column for year of interest.
// Generate a second new variable column with values from original variable
// column for year of interest by county.
// Drop the first new variable county.
// Rename second new column with old variable name since has numeric
// values desired.
sort county year
gen gdp2005 = 0
replace gdp2005 = gdp if year==2005
by county: egen gdp2005ref = max(gdp2005)
drop gdp2005
rename gdp2005ref gdp2005

// Calculate percent change.
gen GDPC = ((gdp2005/gdp2004) - 1)*100

// Save as final dataset.
save "jjj-char.dta", replace

* ==============================================================================
* 3.0 CODE to merge the CEIC-CMR data with the CIC data
*
* Dataset used:      j jj-char.dta
*
* Output:           fullpanel-fix.dta
*
* ==============================================================================
capture log close
set more off

clear all
set matsize 10000
macro drop _all
set linesize 80

// Open full panel dta file.
use "jjj-char.dta"

// Merge with 1:1, should work because each county and year has one value.
merge 1:1 year county_code using "jjj-isreal-codes.dta"
sort year
sort county
sort province
sort city

// Save for regression.
save fullpanel-fix.dta, replace

* ==============================================================================
* 4.0 CODE to regress county characteristics to growth and share variables
* Dataset used:      fullpanel-fix.dta
* Output:           fullpanel-fix.dta
* ==============================================================================
capture log close
set more off
clear all
set matsize 10000
macro drop _all
set linesize 80
// Open full panel dta file
use "fullpanel-fix.dta"

********************************************************************************
* 4.1 PREPARE data for regression
********************************************************************************
// Drop all counties/years without CEIC observations, replace I_S share to 0
// in all counties where activity not observed.
drop if _merge==2
replace istot_shr_temp = 0 if _merge==1
replace istot_county_bus_rev_temp = 0 if _merge==1
// Convert county to usable format for regression.
encode county, gen(county_e)
drop county
rename county_e county
// Rename independent variable.
rename pop POP
// Drop Beijing and Tianjin.
drop if province=="Beijing"
drop if province=="Tianjin"

********************************************************************************
* 4.2 CREATE independent variable columns for earliest and latest years *
********************************************************************************
// Create a new set of independent variables based on original variable names
// of interest.
// Define a new list of new variables; e.g. global [newlistname] "[element1]
// [element2] [element3]"
global ivars "POP WGPC FFRU GOVR FAIV POPD"
sort county year
// Generate new independent variables of 2005 values for each observation/row.
// Use foreach loop to execute set of commands for each element (iv) of
// specified list (ivars).
// Start with first variable listed in above variable list, ivars; single quote
// indicates replaceable text.
// Generate a new variable for a certain year based on what's listed in ivars;
// populate column w/ zeros.
// Replace zeros in the new variable column with values from the original
// variable column for year of interest.
// Generate a second new variable that populates a new column with the max
// value for each county (use egen b/c want to apply to all observations
// for certain county).
// Drop the first new variable column, the one that has combo of values for
// specific and zeros otherwise.
// Rename second new column with old variable column name since has numeric
// values desired.
// Repeat for each independent variable, back to the top.
foreach iv of global ivars{
gen "iv"_2005 = 0
replace "iv"_2005 = `iv' if year==2005
by county: egen "iv"_2005_ref = max("iv"_2005)
drop "iv"_2005
rename `iv'\textunderscore 2005\_ref `iv'\textunderscore 2005

// Sort.
sort county year
// Repeat for 2009.
foreach iv of global ivars{
    gen `iv'\textunderscore 2009 = 0
    replace `iv'\textunderscore 2009 = `iv' if year == 2009
    by county: egen `iv'\textunderscore 2009\_ref = max(`iv'\textunderscore 2009)
drop `iv'\textunderscore 2009
    rename `iv'\textunderscore 2009\_ref `iv'\textunderscore 2009
}

// Drop years that lack observations of is\textunderscore shares.
drop if year > 2009

********************************************************************************
* 4.3 CREATE new independent variable for share difference
********************************************************************************
// Create a column of iron and steel shares from earliest year for each county.
// Generate a new variable for a certain year; populate column w/ zeros.
// Replace zeros in the new variable column with values from the original variable column for year of interest.
// Generate a second new variable that populates a new column with the max value for each county (use egen b/c want max to apply to all observations for certain county).
// Drop the first new variable column, the one that has combo of values for specific and zeros otherwise.
gen istot\_shr\textunderscore 2005 = 0
replace istot\_shr\textunderscore 2005 = istot\_shr\_temp if year == 2005
by county: egen istot\_shr\_2005\_ref = max(istot\_shr\_2005)
drop istot\_shr\_2005
rename istot\_shr\_2005\_ref istot\_shr\_2005

// Repeat for last year.
gen istot\_shr\textunderscore 2009 = 0
replace istot\_shr\textunderscore 2009 = istot\_shr\_temp if year == 2009
by county: egen istot\_shr\_2009\_ref = max(istot\_shr\_2009)
drop istot\_shr\_2009
rename istot\_shr\_2009\_ref istot\_shr\_2009

// Subtract earliest year from latest year share and create new column.
gen istot\_shr\_diff = (istot\_shr\_2009/istot\_shr\_2005) - 1
replace istot\_shr\_diff = 0 if istot\_shr\_2009 == 0
replace istot\_shr\_diff = 0 if istot\_shr\_2005 == 0

********************************************************************************
* 4.4 CREATE new independent variable for main business revenue growth
********************************************************************************
// Create a column of main bus rev from earliest year for each county.
// Generate a new variable for a certain year; populate column w/ zeros.
// Replace zeros in the new variable column with values from the original

// A variable column for year of interest.
// Generate a second new variable that populates a new column with the max
// value for each county (use egen b/c want max to apply to all observations
// for certain county).
// Drop the first new variable column, the one that has combo of values for
// specific and zeros otherwise.

gen istot_county_bus_rev_2005 = 0
replace istot_county_bus_rev_2005 = istot_county_bus_rev_temp if year==2005
drop istot_county_bus_rev_2005
rename istot_county_bus_rev_2005_ref istot_county_bus_rev_2005

// Repeat for last year.
gen istot_county_bus_rev_2009 = 0
replace istot_county_bus_rev_2009 = istot_county_bus_rev_temp if year==2009
by county: egen istot_county_bus_rev_2009_ref = max(istot_county_bus_rev_2009)
drop istot_county_bus_rev_2009
rename istot_county_bus_rev_2009_ref istot_county_bus_rev_2009

// Calculate annualized growth rate of IS business revenue.
gen istot_county_bus_rev_growth = 0 if istot_county_bus_rev_2009==. 
replace istot_county_bus_rev_growth=0 if istot_county_bus_rev_2009==0
replace istot_county_bus_rev_growth=0 if istot_county_bus_rev_2005==0

// Drop years that are not 2005 to not double count.
drop if year~2005

// Get rid of unnecessary columns.
drop city county_name cn _merge

********************************************************************************
* 4.5 CONVERT some variables to log based on visual distribution        *
********************************************************************************
gen LPOP_2005 = log(POP_2005)
gen LWGPC_2005 = log(WGPC_2005)
gen LGOVRE_2005 = log(GOVR_2005)

********************************************************************************
* 4.6 DEFINE labels *
********************************************************************************
label variable istot_shr_temp "Share of Iron and Steel Revenue"
label variable year "Year"
label variable province "Province"
label variable area "Area"
label variable POP "Population (th of people)"
label variable gdp "Gross Domestic Product (RMB mn)"
label variable emprur "Num of Rural Employees (th of people)"
label variable govr "Govt Rev (RMB mn)"
label variable fai "Fixed Asset Investment (RMN mn)"
label variable county "County"
label variable area2005 "Area in 2005 (10000sq. km)"
label variable istot_shr_diff "Share Diff"
label variable istot_county_bus_rev_growth "Growth"
label variable WGPC "Avg Wages per Capitol (RMB/th of people)"
label variable FFRU "Agricultural Workers, \%Total Population"
label variable GOVR "Government revenue, \%GDP"
label variable FAIV "Fixed Asset Investment, \%GDP"
label variable POPD "Population Density (th of people/10000 sq.km)"
label variable GDP "Change in GDP from 2004 to 2005"
label variable POP_2005 "Population in 2005 (th people)"
label variable WGPC_2005 "Avg Wages Per Cap in 2005 (RMB/th people)"
label variable FFRU_2005 "Agr Workers in 2005, \%Total Pop in 2005"
label variable GOVR_2005 "Govt Rev, \%GDP in 2005"
label variable FAIV_2005 "Fixed Asset Investment, \%GDP in 2005"
label variable POPD_2005 "Pop Dens in 2005 (person/10 sq.km)"
label variable LPOP_2005 "Log 2005 Pop (th people)"
label variable LWGPC_2005 "Log 2005 Avg Wages Per Cap (RMB/th people)"
label variable LGOVR_2005 "Log of Govt Rev, \%GDP in 2005"

* 4.7 CHECK correlation of independent variables to dependent variables *

// Check for missing values by looking at descriptive statistics.

// Check correlation of independent variables.

* 4.8 REGRESS county variables and annualized growth with OLS *

eststo m1, title(Annualized Growth)

* 4.9 REGRESS county variables and shares with OLS *

eststo m2, title(Changes in Shares)
Appendix B

Tables
Table B.1: Determinants of annualized growth, measured by main business revenue: OLS regression

<table>
<thead>
<tr>
<th></th>
<th>(1) Growth</th>
<th>(2) Growth</th>
<th>(3) Growth</th>
<th>(4) Growth</th>
<th>(5) Growth</th>
<th>(6) Growth</th>
<th>(7) Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log 2005 Pop (th people)</td>
<td>0.0150</td>
<td>-0.0765</td>
<td>-0.0792</td>
<td>-0.0444</td>
<td>-0.0343</td>
<td>0.0548</td>
<td>0.190</td>
</tr>
<tr>
<td></td>
<td>(0.34)</td>
<td>(-0.58)</td>
<td>(-0.60)</td>
<td>(-0.40)</td>
<td>(-0.34)</td>
<td>(0.52)</td>
<td>(1.73)</td>
</tr>
<tr>
<td>Log 2005 Avg Wages Per Cap (RMB/th people)</td>
<td>-0.110</td>
<td>-0.111</td>
<td>-0.0697</td>
<td>-0.0635</td>
<td>-0.0427</td>
<td>0.0939</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.79)</td>
<td>(-0.79)</td>
<td>(-0.61)</td>
<td>(-0.58)</td>
<td>(-0.39)</td>
<td>(0.87)</td>
<td></td>
</tr>
<tr>
<td>Agr Workers in 2005, %Total Pop in 2005</td>
<td>-0.0291</td>
<td>-0.0392</td>
<td>-0.0520</td>
<td>-0.588</td>
<td>-0.765</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.10)</td>
<td>(-0.13)</td>
<td>(-0.17)</td>
<td>(-1.27)</td>
<td>(-1.60)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log of Govt Rev, %GDP in 2005</td>
<td>-0.0538</td>
<td>-0.0540</td>
<td>-0.151</td>
<td>-0.165</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.92)</td>
<td>(-0.91)</td>
<td>(-1.84)</td>
<td>(-1.96)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixed Asset Investment, %GDP in 2005</td>
<td>0.0470</td>
<td>0.115</td>
<td>0.0180</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.32)</td>
<td>(0.71)</td>
<td>(0.11)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pop Dens in 2005 (person/10 sq.km)</td>
<td>-0.0000437*</td>
<td>-0.0000530**</td>
<td>(-2.40)</td>
<td>(-2.73)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change in GDP from 2004 to 2005</td>
<td>-0.0294</td>
<td>0.888</td>
<td>0.917</td>
<td>0.360</td>
<td>0.260</td>
<td>-0.401</td>
<td>-1.543</td>
</tr>
<tr>
<td></td>
<td>(-0.11)</td>
<td>(0.71)</td>
<td>(0.73)</td>
<td>(0.39)</td>
<td>(0.30)</td>
<td>(-0.43)</td>
<td>(-1.58)</td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>-0.007</td>
<td>-0.012</td>
<td>-0.019</td>
<td>-0.023</td>
<td>-0.030</td>
<td>0.052</td>
<td>0.068</td>
</tr>
<tr>
<td>No. of cases</td>
<td>138</td>
<td>136</td>
<td>136</td>
<td>136</td>
<td>136</td>
<td>136</td>
<td>134</td>
</tr>
</tbody>
</table>

Notes: CEIC, CIC, CMR (sources); * $p < 0.05$, ** $p < 0.01$
Table B.2: Determinants of share changes, as a total of main business revenue: OLS regression

<table>
<thead>
<tr>
<th></th>
<th>(1) Share Diff</th>
<th>(2) Share Diff</th>
<th>(3) Share Diff</th>
<th>(4) Share Diff</th>
<th>(5) Share Diff</th>
<th>(6) Share Diff</th>
<th>(7) Share Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log 2005 Pop (th people)</td>
<td>0.424 (-1.10)</td>
<td>-0.762 (-0.52)</td>
<td>-0.932 (-0.66)</td>
<td>-0.539 (-0.49)</td>
<td>-0.402 (-0.41)</td>
<td>0.406 (0.37)</td>
<td>1.584 (1.37)</td>
</tr>
<tr>
<td>Log 2005 Avg Wages Per Cap (RMB/th people)</td>
<td>-1.429 (-0.96)</td>
<td>-1.532 (-1.03)</td>
<td>-1.063 (-0.96)</td>
<td>-0.978 (-0.96)</td>
<td>-0.789 (-0.73)</td>
<td>0.392 (0.42)</td>
<td></td>
</tr>
<tr>
<td>Agr Workers in 2005, %Total Pop in 2005</td>
<td>-1.867 (-0.55)</td>
<td>-1.980 (-0.58)</td>
<td>-2.157 (-0.67)</td>
<td>-7.016 (-1.36)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log of Govt Rev, %GDP in 2005</td>
<td></td>
<td></td>
<td></td>
<td>-0.606 (-0.94)</td>
<td>-0.608 (-0.94)</td>
<td>-1.491 (-1.60)</td>
<td>-1.617 (-1.77)</td>
</tr>
<tr>
<td>Fixed Asset Investment, %GDP in 2005</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.647 (0.42)</td>
<td>1.268 (0.75)</td>
<td>0.433 (0.27)</td>
</tr>
<tr>
<td>Pop Dens in 2005 (person/10 sq.km)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.000396 (-1.87)</td>
<td>-0.000477* (-2.10)</td>
</tr>
<tr>
<td>Change in GDP from 2004 to 2005</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-1.885 (-0.86)</td>
<td>10.04 (0.73)</td>
<td>11.89 (0.89)</td>
<td>5.604 (0.65)</td>
<td>4.228 (0.57)</td>
<td>-1.762 (-0.20)</td>
<td>-11.66 (-1.28)</td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.005 (-0.06)</td>
<td>-0.008 (-0.015)</td>
<td>-0.015 (-0.018)</td>
<td>-0.026 (-0.026)</td>
<td>0.023 (0.029)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases</td>
<td>138</td>
<td>136</td>
<td>136</td>
<td>136</td>
<td>136</td>
<td>136</td>
<td>134</td>
</tr>
</tbody>
</table>

Notes: CEIC, CIC, CMR (sources); * $p < 0.05$
Table B.3: Base case with FGD and policy in year 6, example excerpt from spreadsheet of NPV calculations

<table>
<thead>
<tr>
<th>Year</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Policy</td>
<td>YES</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steel Price ($)</td>
<td>400</td>
<td>408</td>
<td>416</td>
<td>424</td>
<td>433</td>
<td>442</td>
<td></td>
</tr>
<tr>
<td>Capacity (tons)</td>
<td>1,000,000</td>
<td>1,000,000</td>
<td>1,000,000</td>
<td>1,000,000</td>
<td>1,000,000</td>
<td>1,000,000</td>
<td>1,000,000</td>
</tr>
<tr>
<td>Revenue ($)</td>
<td>0</td>
<td>400,000,000</td>
<td>408,000,000</td>
<td>416,160,000</td>
<td>424,483,200</td>
<td>432,972,864</td>
<td>441,632,321</td>
</tr>
<tr>
<td>Iron ore cost ($)</td>
<td>-20,000,000</td>
<td>-20,400,000</td>
<td>-20,808,000</td>
<td>-21,224,160</td>
<td>-21,648,643</td>
<td>-22,081,616</td>
<td></td>
</tr>
<tr>
<td>Fuel and other production costs ($)</td>
<td>-100,000,000</td>
<td>-100,000,000</td>
<td>-100,000,000</td>
<td>-100,000,000</td>
<td>-100,000,000</td>
<td>-100,000,000</td>
<td>-100,000,000</td>
</tr>
<tr>
<td>Plant O&amp;M ($)</td>
<td>-30,000,000</td>
<td>-30,000,000</td>
<td>-30,000,000</td>
<td>-30,000,000</td>
<td>-30,000,000</td>
<td>-30,000,000</td>
<td>-30,000,000</td>
</tr>
<tr>
<td>FGD O&amp;M ($)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-20,000,000</td>
</tr>
<tr>
<td>FGD capital cost ($)</td>
<td>-200,000,000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Facility capital cost ($)</td>
<td>-500,000,000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cashflow ($)</td>
<td>0</td>
<td>250,000,000</td>
<td>257,600,000</td>
<td>265,352,000</td>
<td>273,259,040</td>
<td>281,324,221</td>
<td>269,550,705</td>
</tr>
<tr>
<td>DCF ($)</td>
<td>0</td>
<td>227,272,727</td>
<td>212,892,562</td>
<td>199,362,885</td>
<td>186,639,601</td>
<td>174,680,207</td>
<td>152,154,346</td>
</tr>
<tr>
<td>Present value of cashflow ($)</td>
<td>2,499,617,270</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Net Present Value ($)</td>
<td>1,799,617,270</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table B.4: Base case without FGD, example excerpt from spreadsheet of NPV calculations

<table>
<thead>
<tr>
<th>Year</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Policy</td>
<td>YES</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steel Price ($)</td>
<td>400</td>
<td>408</td>
<td>416</td>
<td>424</td>
<td>433</td>
<td>442</td>
<td></td>
</tr>
<tr>
<td>Capacity (tons)</td>
<td>1,000,000</td>
<td>1,000,000</td>
<td>1,000,000</td>
<td>1,000,000</td>
<td>1,000,000</td>
<td>1,000,000</td>
<td>1,000,000</td>
</tr>
<tr>
<td>Revenue ($)</td>
<td>0</td>
<td>400,000,000</td>
<td>408,000,000</td>
<td>416,000,000</td>
<td>424,000,000</td>
<td>432,000,000</td>
<td>441,632,321</td>
</tr>
<tr>
<td>Iron ore cost ($)</td>
<td>-20,000,000</td>
<td>-20,400,000</td>
<td>-20,808,000</td>
<td>-21,224,160</td>
<td>-21,648,643</td>
<td>-22,081,616</td>
<td></td>
</tr>
<tr>
<td>Fuel and other production costs ($)</td>
<td>-100,000,000</td>
<td>-100,000,000</td>
<td>-100,000,000</td>
<td>-100,000,000</td>
<td>-100,000,000</td>
<td>-100,000,000</td>
<td>-100,000,000</td>
</tr>
<tr>
<td>Plant O&amp;M ($)</td>
<td>-30,000,000</td>
<td>-30,000,000</td>
<td>-30,000,000</td>
<td>-30,000,000</td>
<td>-30,000,000</td>
<td>-30,000,000</td>
<td>-30,000,000</td>
</tr>
<tr>
<td>FGD O&amp;M ($)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>FGD capital cost ($)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Facility capital cost ($)</td>
<td>-500,000,000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cashflow ($)</td>
<td>0</td>
<td>250,000,000</td>
<td>257,600,000</td>
<td>265,352,000</td>
<td>273,259,040</td>
<td>281,324,221</td>
<td>289,550,705</td>
</tr>
<tr>
<td>DCF ($)</td>
<td>0</td>
<td>227,272,727</td>
<td>212,892,562</td>
<td>199,362,885</td>
<td>186,639,601</td>
<td>174,680,207</td>
<td>163,443,825</td>
</tr>
</tbody>
</table>

Present value of cashflow ($) | 2,594,072,809 |
Net Present Value ($) | 2,094,072,809 |
Appendix C

Bibliography


CIC (2004-2010), China Industrial Census. NBS, Beijing, China.
CMR (2005), China County Statistics. China Data Online from China Marketing Research Co.


