Applications of Applied Econometrics in the Food and Health Economic and Agribusiness Topics

Ruoding Shi

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Economics

Wen You, Chair
George C. Davis, Co-Chair
Olga Isengildina Massa
Susan L. Meacham
Kimberly L. Morgan

October 2, 2019
Blacksburg, VA

Keywords: Health disparity, Hedging Cost, Health Insurance

Applications of Applied Econometrics in the Food and Health Economic and Agribusiness Topics

Ruoding Shi

ABSTRACT

This dissertation consists of three essays in Applied Econometrics that seek a better understanding of different aspects of risk and risk management tools. The first essay is about mortality risk in Virginia coal regions. With a focus on the mortality of non-malignant respiratory diseases (NMRD), I estimate the impact of living in a coal county and find that coal-mining county residency significantly increases the probability of dying from NMRD. This statistical association is accentuated by surface coal mining, high smoking rates, lower health insurance coverage, and a shortage of doctors. The second essay evaluate the cost of a price risk management tool called futures hedging. A variety of measures illustrate considerable changes in hedging costs over time. Quantile regression results show that substantial price volatility and high margin requirements are the main factors driving high hedging costs from 2007 to 2013. The third paper investigates a health risk management tool, a public health insurance program in China called New Cooperative Medical Scheme (NCMS). I apply contract theory to characterize local governments' selective incentives in NCMS benefit designs. Empirical analysis of China Health and Nutrition Survey data indicate challenges of financial sustainability of this scheme in poor regions. The NCMS plan tends to under-cover the services that are moderately predictable and negatively correlated with plan profits, such as outpatient treatments. Preventive services are generally over-provided, perhaps due to the incentive to attract healthy participants.

Applications of Applied Econometrics in the Food and Health Economic and Agribusiness Topics

Ruoding Shi

GENERAL AUDIENCE ABSTRACT

This dissertation uses quantitative analysis to investigate three economic problems related to different aspects of risk. The first question is, what affects the respiratory health of Virginia coal mining counties' residents? Using respiratory mortality as the variable of interest, this paper finds that surface coal mining, high smoking rates, and lack of health access jointly contribute to the elevated risk of dying from respiratory diseases in our study area. The second research problem is about a price risk management tool called "hedging": purchasing contracts in the futures market to offset price movements in the cash markets. Based on historical data of corn and soybeans, I simulate the cost of hedging and find this risk management tool is not cheap, especially in 2007 to 2013. The high cost is mainly due to substantial price fluctuations in the recent decade. As a health risk management tool, health insurance is the focus of my third study. In China rural areas, a public health scheme aimed to reduce a resident's risk of suffering medical impoverishment by spreading the risk over residents in a county. County governments were relatively free to design the implementation and benefit plans. This study reveals that most New Cooperative Medical Scheme (NCMS) benefit plans are not efficient to achieve the scheme's objective. Facing high risk of fund deficits, local insurance programs in poor regions are likely to under-cover health services, such as outpatient treatments. If this scheme were allowed to charge higher prices from high-risk enrollees instead of a flat-rate premium, its efficiency might be improved.

DEDICATION

To my Dad, Mom, and professor Lanlan Wang

ACKNOWLEDGMENTS

As I write this section, my heart is overwhelmed with gratitude and inspiration. Over the past five years, I have been fortunate to have the opportunity to work with several advisors closely and to learn from them. I am deeply indebted to my committee chairs, Dr. George Davis and Dr. Wen You, for their outstanding mentorship, guidance, and support throughout the whole journey. I would not have been able to make much professional progress without the rigorous training and wisdom of Dr. Davis, who has had a fundamental influence on my understanding of research and life. Dr. You's extreme patience and insights have guided me step-by-step to become an independent researcher, and her kindness warms my heart.

I wish to thank all the members of my committee for their advice and help. Special thanks to Dr. Olga Isengildina Massa, who was dedicated to offering me a challenging and exciting learning experience in Agribusiness. My sincere gratitude goes to Dr. Susan Meacham for the opportunity to work on coal-mining health issues and her invaluable remarks on my first paper. I want to thank Dr. Kimberly Morgan because I have benefited tremendously from her advice on literature review, grant writing, and professional development. Also, I thank Dr. Elena Serrano, Dr. Klaus Moeltner and Dr. Carrie Kroehler for their help in the whole process.

Many thanks to the faculty and staff in the Department of Agricultural and Applied Economics, the Graduate Student Organization and my fellow graduate students, Yu Sun, Yanliang Yang, Weizhe Weng, Xin Ning, Ling Yu, Kate Vaiknoras, Zhen Cheng, Xiang Cao, Yuelu Xu, Jinyang Yang, Deeksha Gupta, Yuanyuan Wen, William Ferris, Mina Hejazi, Yang Song, Xinde Ji and Chaoping Xie for their friendship.

Particular thanks to Catherine Copeland, Yang Song, and Dan Yu for their emotional support in my difficult time.

I am deeply thankful to Professor Lanlan Wang, who inspired me to start this journey overseas and encouraged me to pursue the truth with a whole heart. I want to extend gratitude to my previous advisors, Dr. Hai Zhong, Dr. Yongsheng Xu and Dr. Zhifeng Gao for their mentorship and continuous support. Finally, I would like to thank my parents for tolerating my long absence. The unconditional love and encouragement from my family are the greatest gifts I have.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Table page

Introduction

According to Asteriou and Hall (2015), applied econometrics means "quantifying economic relationships using actual data (on page xxvii)." Unlike statistics, applied econometrics is often driven by economic theories and tests hypothesis using existing data. As George Box on page 61, "All models are wrong, but some are useful.(Box, Luceno, & del Carmen Paniagua-Quinones, 2011)" Econometricians have developed many techniques to reveal relationships between variables and shed light on a variety of social problems. This dissertation answers three research problems in the fields of food and health economics and agribusiness using different applied econometrics methods and provides several policy recommendations.

Previous studies have associated elevated mortality risk in central Appalachia with coal-mining activities, but few have explored how different non-coal factors influence the association within each county. My first chapter seeks to understand what factors may accentuate or attenuate mortality from non-malignant respiratory diseases (NMRD) for people living in Virginia coal mining counties. An empirical reduced-form, quasi-experimental study was designed to observe the characteristics of three populations and contrasts their differences. Records for seven coal-mining counties (n=19,692) were obtained with approval from the Virginia Department of Health for the time period between 2005 to 2012. Also requested were records from three adjacent coal counties (n=10,425) to provide a geographic comparison. For a baseline comparison, records were requested for eleven tobacco-producing counties (n=27,800). We analyzed the association of 57,917 individual mortality records in Virginia with coal-mining county residency, or for ease in communication, the "coal county effect." The development of a two-level hierarchical model allowed the coal-county effect to vary by county-level socioeconomic status, health access, behavioral risk factors, and coal production. Wald tests detected sets of significant factors explaining the variation of coal health impacts across counties.

Furthermore, to illustrate how the models help explain health disparities, two coal-mining county case studies were presented. The main results show that the coal-county effect was accentuated by surface coal mining, high smoking rates, decreasing health insurance coverage, and a shortage of doctors. In Virginia coal-mining regions, the average coal-county effect increased by 147% (p-value<0.01) when the ratio of doctor to 1000 population decreased by one, and the effect increased by 68% (p-value<0.01) with a 1% reduction of health insurance coverage, holding other factors fixed. A revised version of the first chapter was submitted to the journal of BMC Public Health.

Increased price volatility in commodity markets makes risk management more important than ever. Futures markets have traditionally been used to hedge commodity price risk. However, hedging with futures involves costs that may lead to liquidity problems, even bankruptcy. The second chapter measured the cost of futures hedging and applied quantile regression techniques to examine the underlying driving factors. Based on historical prices and margin requirements of corn and soybean futures, I simulated the borrowing costs of maintaining a margin account and the risk of premature termination of a hedge over time. Quantile regressions revealed the effects of hedging cost determinants at different parts of the cost distribution. Results suggested that the costs of hedging increased from 2007 to 2013 and decreased afterward. Meanwhile, increasing margin requirements and price volatility were two important contributing factors.

The third study focused on a nationwide public health insurance program in China called the New Cooperative Medical Scheme (NCMS). Previous literature had found that this program failed to provide sufficient financial protection from medical impoverishment. One possible reason was that some county governments had financial incentives to under-cover the health

services that were widely used by their residents. A theoretical principal-agent model was employed to evaluate service-level coverages of the NCMS and measure the degree of distortion in its benefit plans. Based on the China Health and Nutrition Survey data and two-part model predictions, we divided the whole sample into four geographical regions and assessed each region's population health status, distribution of health spending, and shadow prices of different health services. To inform the NCMS benefit modification, we empirically investigated the distortion under different risk-adjusted premiums. The results revealed challenges of financial sustainability faced by the NCMS program, especially in the Western region. Local governments in poor and sick areas tended to under-cover health services. In general, outpatient reimbursements were most vulnerable to under-provision. This study also found that the efficiency of the NCMS could be improved if its risk premiums were adjusted based on individual demographic characteristics and disease history.

Chapter 1: Factors influencing high respiratory mortality in coal-mining counties: a repeated cross-sectional study

1.1 Background

Health disparities have persisted in central Appalachia for decades (Blackley, Behringer, & Zheng, 2012; Krometis et al., 2017; Meacham, Sukpraput, Taber, & Metzger, 2013; Woolley, Meacham, Balmert, Talbott, & Buchanich, 2015). Virginia mines, in the heart of central Appalachia, in the rugged mountains of the southwestern part of the state, produce high-quality coal. Coal is the heart of the economy and a cultural icon in a region that reveres "coal as king." While rates of mortality have improved in the region, they have persisted at rates higher than regional and national averages, particularly non-malignant respiratory diseases (NMRD) (Woolley et al., 2015). Studies attribute the elevated mortality risk to environmental exposure to coal extraction, processing, and transportation activities (Cortes-Ramirez, Naish, Sly, & Jagals, 2018; Hendryx & Ahern, 2008; Hendryx, Fedorko, & Anesetti-Rothermel, 2010; Hendryx, O'Donnell, & Horn, 2008). Mining releases a large amount of coal dust and methane into the environment and results in higher concentrations of particulate matter and sulfate, impairing coal miner's respiratory system, a condition known as coal workers' pneumoconiosis (CWP) (Finkelman et al., 2002). Another coal-related lung disease is silicosis caused by inhalation of crystalline silica dust. However, the potential health effects of environmental contaminants produced by coal mining on community residents are the subject of ongoing investigations (Boyles et al., 2017).

The health effects of coal mining are likely to be back in the spotlight of health policymakers as the U.S. government is attempting to revive the coal industry. Coal production in the U.S. has risen 6% from 2016 to 2017 (Mine Safety and Health Administration, 2016-

2018). Some are concerned that the reemergence of the coal industry may have negative impacts on the health of those living in these areas, retarding or reversing the progress made to improve health metrics for those living in these areas over the past few decades (Morrice & Colagiuri, 2013).

Other factors, such as access to healthcare, could also accentuate or attenuate the adverse effects of coal mining on health. For example, an accentuating factor is noticed when Kentucky lawmakers passed a House bill (HB2-18RS) that permitted fewer doctors to read chest X-ray for miners' health claims (Kentucky Legislative Research Commission, 2018). An example of an attenuating factor was seen when Congress required governments and coal companies to pay out healthcare and guarantee benefits to retired coal workers even as coal companies faced bankruptcy (Perri, 2017). In these scenarios, the legislative actions are potentially influencing the health of coal community residents.

1.1.1 Previous literature

Following Meacham et al. (2013), we classified studies on health disparities in Appalachia into two groups: those focusing on coal mining and those not focusing on coal-related factors. In the second group, the authors have identified several determinants predominantly associated with health disparities in coal communities, such as low staffing levels in hospitals and Appalachian cultural beliefs (Blackley et al., 2012; Denham, Wood, & Remsberg, 2010; Krometis et al., 2017; McGarvey, Leon-Verdin, Killos, Guterbock, & Cohn, 2011). Based on a survey on healthcare providers, Denham et al. (2010) found that insufficient health staffing and facilities, and lack of diabetes education explained high diabetes prevalence in Appalachia. This research group also proposed that cultural and ethnic components of communities contributed to poor health outcomes as well. McGarvey et al. (2011) suggested a cultural component and revealed that Appalachian residents in Virginia were more likely to report their health status as “poor” compared to non-Appalachian residents even though there was no difference in chronic diseases reported by Appalachian and non-Appalachian groups.

Several studies have focused specifically on coal mining and poor health outcomes in central Appalachia. These poor health outcomes include high mortality rates of cancer (Hendryx, Fedorko, & Anesetti-Rothermel, 2010), cardiovascular diseases (Esch & Hendryx, 2011) and kidney diseases (Hendryx, 2009), and increased risk of hospitalization for hypertension and COPD (Hendryx, Ahern, & Nurkiewicz, 2007). For instance, Hendryx et al. (2008) examined county mortality rates and found that living in a heavy coal-mining county was a risk factor for lung cancer. Based on a telephone survey on the self-reported presence of specific chronic diseases, Hendryx and Ahern (2008) tested whether coal production had adverse effects on local residents’ health after controlling demographic characteristics and county-level covariates (smoking rate, obesity rate, poverty rate, and social capital). They found higher risks of cardiopulmonary diseases, chronic lung diseases, hypertension, and kidney diseases were associated with residents living in counties with high-level coal production, compared to residents in non-coal counties.

To identify the health effect of coal mining, most studies have attempted to handle several confounding factors in central Appalachia (Esch & Hendryx, 2011; Hendryx et al., 2008; Woolley et al., 2015). However, these health effects have often been assumed constant between coal-mining counties even after controlling socioeconomic and behavioral factors, such as poverty rates and smoking rates (Esch & Hendryx, 2011; Hendryx et al., 2008). None of the previous studies have considered if the health effects may differ by county and what factors influenced those differences.

1.1.2 Current approach

For the purposes of this study, the term "*coal-county effect*" has been adopted to refer to the health effect of living in a coal-mining county on mortality[1]. Using an ecological epidemiology protocol, we estimated the associations between the mortality risk of NMRD and coal-mining county residency and what non-coal factors affect the associations. The non-coal factors of interest represented the geography, temporal trends, and socioeconomic demographics of our study population groups.

Our study objective was twofold, prompting the following research questions:

1. Are the coal-county effects constant across counties?
2. What factors lead to non-constant coal-county effects?

With the first question, we hypothesize that the coal-county effect may depend on a county's health access, economic condition, coal production, and other health behavioral risk factors[2]. For example, limited access to health care services could accentuate coal-county health effects, because some coal-related lung diseases (e.g., CWP and silicosis) are often symptomless in the early stages but develop into severe conditions without access to screening services and treatments (Lockwood, 2012). By addressing the second question, we plan to identify and estimate the impact of selected factors contributing to the existing poor health measures in coal counties. Our study methodology is unique. The development of a novel, two-level hierarchical model allows the estimated coal-county effect to vary depending on the county's socioeconomic status, health access, health behavioral risk factors, and coal production. Following the insights from Hendryx et al. (2008) and Hendryx, Fedorko, and Halverson (2010), we consider coal

---

[1] Several previous studies identified elevated mortality rates in coal-mining areas but did not name their findings as "coal-county effects.

[2] Health behavioral risk factors refer to risk behaviors that lead to poor health outcomes.

production from both surface mining (i.e., strip mining, open-pit mining, and mountaintop removal mining) and underground mining. Surface mining practice is more likely to affect neighboring communities by air and water pollution (Boyles et al., 2017), while underground coal mining is often associated with miners' lung diseases, an occupational hazard (Lockwood, 2012).

1.2 Methods

1.2.1 Study design

Individual death records (n=57,917) were merged with county-level covariates based on their counties of residence and years of death to capture the dynamic changes from 2005 to 2012. Ethical approvals of individual mortality data were obtained from the Internal Review Boards of the Edward Via College of Osteopathic Medicine and the Virginia Department of Health Office of Vital Statistics. County-level covariates were selected to capture the multi-dimensional concepts of socioeconomic status, health access, and health behavioral risk factors in three population subgroups. Our model design allowed the coal-county effect to vary as a function of selected county-level covariates. This model framework enabled us to test the assumption of non-constant coal-county health effects. It also identified factors that explain variations across coal counties.

The majority of previous research publications simply used non-mining counties as reference groups, which ignored the potential spillover effect of coal production across county borders. The spatial analysis by Hitt and Hendryx (2010) showed that cancer mortality rates were autocorrelated between adjacent counties. Although our analysis was not of the typical spatial approach, we did analyze the counties adjacent to coal-mining counties to test the spillover effect. We considered both coal-mining counties and counties adjacent to coal-mining counties as "treated" groups. Since Virginia tobacco counties share similar economic characteristics with coal-mining counties, such as "low economic diversification, low employment in professional services, and low educational attainment rates (Hendryx & Ahern, 2009)", these tobacco counties served as a control group or "untreated" baseline counties. Then, we identified the coal-county effect by comparing the average likelihood of dying from NMRD among residents in treated groups with that in baseline counties. The choice of an "untreated" baseline aimed to reduce selection bias because of the similarity between coal-mining counties and tobacco counties.

1.2.2 Study area

With places of residence recorded, the mortality data were collected from three rural, underserved health disparity areas in Virginia: coal-mining counties, adjacent coal counties, and tobacco counties. The adjacent coal counties served as a geographic comparison group with residents living in small communities in mountainous southwest Virginia. The tobacco counties were an economic comparison group located in south central Virginia and experienced financial trends over several decades that were similar to those for coal-dependent counties.

Figure 1.1 shows the three county groups in Virginia. Seven counties in southwest Virginia were considered as coal-mining counties (Buchanan, Dickenson, Lee, Russell, Scott, Tazewell, and Wise, n= 19,692 records). Although Scott County stopped producing coal after 1995, it was classified as a coal-mining county because coal mining may have a long-run impact on the local environment and human health, particularly chronic conditions (Petsonk, Rose, & Cohen, 2013). Statistical analyses will determine the sensitivity of the results to this classification. Three Virginia counties share the county border with coal-mining counties (Bland, Smyth and Washington, n= 10,425 records). The 11 tobacco counties are located in the region historically known for tobacco production (Amelia, Brunswick, Buckingham, Charlotte, Cumberland, Halifax, Lunenburg, Mecklenburg, Nottoway, Pittsylvania, and Prince Edward, n= 27,800 records). These counties dependent on the tobacco industry as a primary source of the local economy and are economically comparable to coal-mining counties (Meacham et al., 2013). Therefore, we used them as baseline counties.

[Figure 1.1 to be here]

1.2.3 Data sources and variables

1) Individual-level data

Death records were collected from the Virginia Department of Health Office of Vital Statistics (VDH, 2013), including the primary cause of death, age, gender, place of residence, marital status, and years of education. Our outcome variable was death caused by non-malignant diseases of the respiratory system with the International Codes for Diseases (ICD) 10th revision codes J00 – J99. NMRD includes but is not limited to asthma, chronic obstructive pulmonary disease (COPD), and the pneumoconiosis. NMRD was chosen as the dependent variable of concern because this group of diseases was commonly considered as a high-risk health problem in coal-mining regions (Hendryx, Fedorko, & Halverson, 2010; Woolley et al., 2015). For example, counties in central Appalachia had the highest mortality rates of pneumoconiosis and COPD (Dwyer-Lindgren et al., 2017).

2) County-level covariates

Publicly available county annual coal production was obtained from the U.S. Energy Information Administration (Energy Information Agency, 2005-2012). Other county-level covariates were collected from multiple sources and classified into three categories: socioeconomic characteristics, accessibility of health care services, and health behavioral risk factors. Most county-level covariates were obtained from the Area Health Resources File (AHRF) (Health resources and services administration, 2015-2016). AHRF is a health resource information system maintained by the Health Resources and Services Administration. County health behavioral risk factors were obtained from the Behavioral Risk Factor Surveillance System (BRFSS) data (Centers for Disease Control Prevention, 2005-2012). Finally, additional data sources included the Census Bureau's Small Area Health Insurance Estimates (SAHIE) for health insurance rates and the U.S. Census Bureau and Rural-Urban continuum code from the United States Department of Agriculture Economic Resource Service.

Selected covariates included county unemployment rates, median household income, and rural-urban status to measure SES, which played a vital role in individuals' health outcomes and likelihoods of dying. The first SES variable was the unemployment rate at the county level as unemployment increased mortality risk by keeping jobless people from investing in health (Granados, House, Ionides, Burgard, & Schoeni, 2014). However, employment alone was insufficient to measure available resources since a majority of individuals in the sample were retired. We also considered county median household income and unobserved differences between rural and urban residents. Based on the nearest observed Rural-Urban Continuum Codes in 2003 and 2013, we constructed indicators to classify counties into rural counties, non-metropolitan urban counties, and counties in the metropolitan area[3]. The 2003 Rural-Urban Continuum Codes were used to construct indicators starting in 2005 due to a closer time reference, and then we switched to 2013 Rural-Urban Continuum Codes to classify counties after 2008.

---

[3] 1) Counties in the metropolitan area: Rural-Urban Continuum Codes 1–3 with the description of "Metro - Counties in metro areas";
2) Non-metropolitan urban counties: Rural-Urban Continuum Codes 4–7 with the description of "Nonmetro - Urban population of 2,500 or more";
3) Rural counties: Rural-Urban Continuum Codes 8–9 with the description of "Nonmetro county completely rural or less than 2,500 urban population.

To represent health access, county health insurance rates were collected from SAHIE and three county-level health access measurements from AHRF, including numbers of doctors (sum of active medical doctors and osteopathic doctors), hospital beds and health centers per 1000 population. Finally, we collected smoking rates at the county level from the study of Dwyer-Lindgren et al. (2014) and age-adjusted obesity rates and physical inactivity prevalence rates from the BRFSS.

1.2.4 Empirical model

A two-level latent index model was used to estimate the coal-county effect and adjacent-coal-county effect (Wooldridge, 2010). Multilevel modeling technique is one type of regression analyses that handles micro-level individuals and macro-level counties simultaneously in one model (Duncan, Jones, & Moon, 1998). In the context of this study, traditional regression approaches do not consider between-county heterogeneity and assume the coal-county effect is constant across all counties. A less restricted assumption is that the statistical association between coal mining and health outcome follows a distribution, and it can be different across counties and over time due to other covariates, such as SES or health behavioral risk factors. Instead of fitting a different model based on each county's individual-level data, we used a two-level model with individuals (*level 1*) nested within counties (*level 2*) and allowed key model parameters vary across counties and over time in association with other covariates. A detailed description of each level's model specification was provided below.

The *Level-1 Model* assumes that for an individual i in county j deceased in year t, the probability of dying from a certain disease y_{ijt} could be estimated through a latent index y_{ijt}^*. Intuitively, the latent index y^* reflects the severity of a disease: the individual will die when the latent index reaches a threshold ($y^* > 0$). We model the latent index as a linear combination of county-specific intercept (β_{0jt}), county group indicator (coal-mining, adjacent coal or tobacco county), individual i's demographic characteristics ($\boldsymbol{X_{ijt}}$) and year-specific effects ($\boldsymbol{d_t}$) as follows:

$$y^*_{ijt} = \beta_{0jt} + c_{1jt}d_{incoal} + c_{2jt}d_{adjcoal} + \boldsymbol{X}'_{\boldsymbol{ijt}}\boldsymbol{\beta_1} + \boldsymbol{d}'_{\boldsymbol{t}}\,\boldsymbol{\sigma} + \varepsilon_{ijt} \tag{1.1}$$

To estimate the coal-county effect and adjacent-coal-county effect, we use two binary variables indicating county groups: $d_{incoal} = 1$ if the deceased lived in a coal-mining county, and $d_{adjcoal} = 1$ if the deceased lived in a county adjacent to coal-mining counties. The baseline group consists of those residing in tobacco-producing counties due to the similarity between coal-mining counties ("treated" group) and tobacco-producing counties ("untreated" group) and their non-adjacency. Additionally, individual-level demographic variables ($\boldsymbol{X_{ijt}}$), such as age, race and gender are included. A set of year dummies ($\boldsymbol{d_t}$) is added to control unobserved time effects. We adjust errors (ε_{ijt}) for correlations between individuals in the same county.

The *Level-1 Model* (1.1) allows three parameters to vary by county *j* and year t: β_{0jt}, c_{1jt}, and c_{2jt}. The parameter β_{0jt} is the tobacco-county-specific intercept, reflecting county heterogeneity in the mean latent index at the baseline when $d_{incoal} = d_{adjcoal} = 0$. We call β_{0jt} the "*county baseline"* as a short term in the following discussion since tobacco counties are chosen as baseline counties. As the probability of dying from a specific disease is an increasing function of the latent index, a lower county baseline suggests a lower mean county probability of dying. We expect the county baseline (i.e., free of coal mining effect) to be lower if that county's residents have a higher socio-economic status (SES), better health access (HA) and lower health behavioral risk (HR) at year t. Suppose there are two counties, and county A provides better health access than county B. This expectation can be explained in two scenarios. 1) When residents in both counties are the same in SES and HR aspects, residents in county A would be less likely to die from a particular disease; 2) When residents in both counties also differ in some of SES and HR aspects, for instance, if residents in county A face a higher unemployment rate and smoking rate compared to residents in county B, this may offset their advantage with health access and result in a higher likelihood of dying. Therefore, the county baseline is determined by the specific combination of SES, HA, and HR. We further specify β_{0jt} in the *Level 2 Model* with the signs indicating prior expectations.

$$\underset{(?)}{\beta_{0jt}} = \underset{(?)}{\beta_0} + \underset{(-)}{\eta_{01}} SES_{jt} + \underset{(-)}{\eta_{02}} HA_{jt} + \underset{(+)}{\eta_{03}} HR_{jt} \tag{1.2}$$

The level-2 predictors *SES*, *HA,* and *HR* are a set of county characteristics that could affect the intercept, as introduced in the section of County-level covariates.

The parameter c_{1jt} measures the average coal-county effect by comparing mean latent indices between a coal-mining county with a tobacco county, holding other factors fixed. We expect $c_{1jt} > 0$ if living in a coal-mining county contributed to the mortality risk. Like β_{0jt}, we hypothesize the coal-county effect to be different among coal-mining counties. In addition to SES, health access and health behavioral risk factors, total coal production ($Prod$) and the percent of production from surface coal mining ($Surface\%$), may have also affected the link between coal production and mortality risks. Consequently, similar to β_{0jt}, the coefficient c_{1jt} is allowed to vary by county characteristics.

$$\underset{(+)}{c_{1jt}} = \underset{(?)}{c_1} + \underset{(-)}{\eta_{11}} SES_{jt} + \underset{(-)}{\eta_{12}} HA_{jt} + \underset{(+)}{\eta_{13}} HR_{jt} + \eta_{14} \underset{(+)}{Prod_{jt}} + \eta_{15} \underset{(+)}{Surface\%_{jt}} \tag{1.3}$$

The magnitude of coal-county effect (c_{1jt}) depends on estimated parameters and historical values of county characteristics, which change over year (t) and differ for each county (j). Therefore, c_{1jt} is heterogeneous both within and between counties. To explain this intuitively, we expect coal-mining county j's adverse health effect could be reduced over time if county j improves the economic status of residents, increases the accessibility of health care services or decreases risk factors and coal production (within-county heterogeneity). Also, a coal-county effect is expected to be smaller for a coal-mining county with higher SES, better HA, lower HR and coal production and less surface mining activities, compared with other coal-mining counties during the same year t (between-county heterogeneity). Note, although η_{11} and η_{12} are expected to be negative, c_{1jt} could still be positive if the effects of the health behavioral risk factors (*HR*), coal production ($Prod$) and surface coal percentage ($Surface\%$) offset the socioeconomic status (SES) and health access (*HA*) effects.

Similar logic applies to adjacent coal counties, so the adjacent-coal-county effect is specified as:

$$\underset{(+)}{c_{2jt}} = \underset{(?)}{c_2} + \underset{(-)}{\eta_{21}}SES_{jt} + \underset{(-)}{\eta_{22}}HA_{jt} + \underset{(+)}{\eta_{23}}HR_{jt} \tag{1.4}$$

If some coal mines are located near the county boundaries, c_{2jt} is expected to be positive. Again, η_{21} and η_{22} are expected to have negative signs, indicating higher SES and better health access reducing the adjacent-coal-county effect on mortality. Since health behavioral risk factors increase the county effect (Graber, Stayner, Cohen, Conroy, & Attfield, 2014), η_{23} is expected to be positive.

Substituting equations (1.2) to (1.4) into equation (1.1), yields:

$$\begin{gathered} y_{ijt}^* = (\beta_0 + \eta_{01}SES_{jt} + \eta_{02}HA_{jt} + \eta_{03}HR_{jt}) + (c_1 \\ +\eta_{11}SES_{jt} + \eta_{12}HA_{jt} + \eta_{13}HR_{jt} + \eta_{14}Prod_{jt} + \eta_{15}Surface\%_{jt})d_{incoal} + (c_2 \\ +\eta_{21}SES_{jt} + \eta_{22}HA_{jt} + \eta_{23}HR_{jt})d_{adjcoal} + \boldsymbol{X}_{\boldsymbol{ijt}}^{\prime}\boldsymbol{\beta_1} + \boldsymbol{d}_{\boldsymbol{t}}^{\prime}\,\boldsymbol{\sigma} + \varepsilon_{ijt} \end{gathered} \tag{1.5}$$

To answer the research questions, we test the following two hypotheses:

1) Parameters β_{0jt}, c_{1jt} and c_{2jt} vary between counties and over time. This means η_{01}, η_{02} and η_{03} are not jointly equal to zero in the intercept equation. The same logic is applied to c_{1jt} equation ($\eta_{11} \neq 0$ or $\eta_{12} \neq 0$ or $\eta_{13} \neq 0$ or $\eta_{14} \neq 0$ or $\eta_{15} \neq 0$) and c_{2jt} equation ($\eta_{21} \neq 0$ or $\eta_{22} \neq 0$ or $\eta_{23} \neq 0$);
2) The coal-county effect is affected by socioeconomic status, health access, high-risk behavioral factors, and coal production. This means the coefficients $\eta_{11} \neq 0,\ \eta_{12} \neq 0,\ \eta_{13} \neq 0, \eta_{14} \neq 0$ and $\eta_{15} \neq 0$ in equation (1.3).

1.2.5 Statistical analyses

Our statistical analyses began with a descriptive summary of all variables in the model. In order to test for the first hypothesis, we estimated the general model specified by equation (1.5) with all explanatory variables. Wald tests were conducted to test the joint significance of all county-level covariates in the β_{0jt}, c_{1jt}, and c_{2jt} equations. The model assumed that individuals were correlated within the same counties or cities. According to Cameron and Miller (2015), ordinary Wald tests often over-reject when there is a small number of counties (M = 24 clusters in our case[4]), meaning that the p-values from ordinary Wald tests are underestimated. We followed their suggestion and conducted adjusted Wald tests, which were based on a t-distribution with M-1 degrees of freedom. All statistical analyses were conducted using Stata 14 software (StataCorp LP, 2015).

For the second hypothesis, particular interest centered on the coal-county effect c_{1jt} in equation (1.3). According to the Wald tests, we adjusted the general model by excluding insignificant vectors of variables and checked sensitivities of the results to different

[4] In addition to 21 counties, there are three independent cities: Bristol City and Norton City in the coal region and Danville City in the tobacco region. So we have a total of 24 clusters.

specifications. Variance inflation factor (VIF) was used to test potential collinearity between socioeconomic and health access covariates. Next, the coal-county effects (c_{1jt}) of three coal-mining counties were predicted based on these counties' historical characteristics and the estimated parameters. The case study of two Virginia coal-mining counties (Russell County and Lee County) illustrated how our finding could be meaningful in the real world. Specifically, it explained what happened to the coal-county effect when some non-coal factor changed over time.

Although our analyses were not able to identify coal miners from the death records, we expected that male and working-age residents in our sample would have a higher mortality risk associated with coal mining, because this population would be more likely to be working in coal mines. To explore this, we ran the regressions and predicted the coal-county effects for male and female subgroups separately. Similar analyses were also conducted on working age (15-64) and retirement age (>64) subgroups.

1.3 Results

1.3.1 Descriptive statistics

Table 1.1 provides descriptive statistics for all variables at the individual level (n=57,917). From 2005 to 2012, an average of 11 out of 100 people died from NMRD. Residents in the death records obtained an average of 10 years of schooling (standard deviation (SD) = 3.56) and their average age was 72 years (SD = 17.55). The majority of deceased individuals were white (83%), and one half of the sample was female. About 39% of the deceased were married. Consistent with previous literature, SES in this region was relatively low. The average county unemployment rate was 7% (SD = 2%), and the mean of median household income was $35,880 (SD = 4,120). About 39% of residents lived in rural areas where the population was less than 2,500. On average, the age-adjusted physical inactivity prevalence rate was 28% (SD = 3%), and the age-adjusted obesity rate was 30% (SD = 3%). The average smoking rate of 28% (SD=2%) was above the national average of around 24% calculated by Dwyer-Lindgren et al. (2014). Regarding health access variables, the mean values of hospital beds, federal qualified health centers and doctors were 3.08, 0.06 and 1.11 per 1000 population, respectively. The average county health insurance rate showed that 84% of individuals had some sorts of health insurance. Among the study area, the mean county annual coal production was 1.23 million tons with a large standard deviation of 2.87, which indicated heterogeneity in coal production between counties. Except for Scott County, all coal-mining counties in Virginia were involved in surface mining, and the mean surface coal production was 0.52 million tons (SD = 1.37). Finally, of the 57,917 residents in the death records, 19,692 residents (34%) were living in seven coal-mining counties and 10,425 residents (18%) in three adjacent counties.

[Table 1.1 to be here]

1.3.2 Wald test results

Table 1.2 reports p-values from adjusted Wald tests (p-values from ordinary Wald tests are reported in the parentheses). The first row suggests that varying specifications of β_{0jt}, c_{1jt}, and c_{2jt} were preferred. For example, in the c_{1jt} column of row (1), we tested the null hypothesis $H_0: \eta_{11} = \eta_{12} = \eta_{13} = \eta_{14} = \eta_{15} = 0$ in equation (1.3) and obtained a p-value less than 0.01 from adjusted Wald test, so we rejected the null hypothesis that the coal-county effect was a constant and independent of county-level covariates. Likewise, Wald tests also rejected the null hypothesis that β_{0jt} (p-value<0.01) and c_{2jt} (p-value<0.01) were constants.

Furthermore, we tested the joint significance of socioeconomic status, health access and health behavioral risk vectors of variables separately in each level-2 equation. Column (2) and (3) in Table 1.2 show that coal-county effect c_{1jt} and adjacent-coal-county effect c_{2jt} were significantly affected by health access (HA) with p-values less than 0.01, and county SES also explained the variations in adjacent-coal-county effects (p-value=0.01). The county baseline β_{0jt} appeared to depend on health access (HA) and health behavioral risk factors (HR) with p-values less than 10%.

1.3.3 Model results of coal-county effects

Results of collinearity test are provided in Appendix AA. The maximum VIF value was less than 3, which indicated that the collinearity was not a big concern. Average marginal effects of all variables are reported in Appendix AB. The average marginal effect of the coal-county indicator was significantly positive across models.

Table 1.3 reports estimated coefficients in the equation of c_{1jt}, utilizing different model specifications. The magnitude and significance of estimated coefficients were quite robust. Results show that the coal-county effect was higher in rural and metropolitan urban areas compared to non-metropolitan urban areas. Significant coefficients were found for the number of hospital beds, doctors per 1000 population, and health insurance rates. For example, one additional doctor per 1000 population significantly reduced the coal-county effect by 0.119 to 0.147 across models, and a 1% increase in health insurance coverage rates significantly reduced the health effect by 0.065 to 0.070 across models. However, the coefficient of hospital beds per 1000 population is significantly positive. Regarding health behavioral risk factors, a 1% increase in the smoking rate at the county level significantly increased the coal-county health effect by 0.026 to 0.035 across models. Finally, the coal-county effect went up by 0.02 to 0.04 with a 10% increase in surface coal proportion. The coefficients of total coal production were insignificant, so this variable was excluded from the final estimation due to high collinearity with surface coal percentage.

[Table 1.3 to be here]

1.3.4 Case studies

Figure 1.2 (a) plots annual surface-mining coal production of three counties in Virginia. Buchanan County had produced the most coal in Virginia in the past decades, and its production started to decline after 2007. Surface coal production in Russell County and Lee County had been much lower and less than 1 million tons. The coal-county effect (c_{1jt}) was predicted using the estimated parameters from model 2 preferred by the adjusted Wald tests. Figure 1.2 (b) shows the predicted coal-county effects for these three counties: Buchanan County ($\hat{c}_{1jt}$: 0.18 to 0.40), Russell County ($\hat{c}_{1jt}$: 0.02 to 0.23) and Lee County ($\hat{c}_{1jt}$: 0.06 to 0.2). A 95% confidence interval was drawn around Buchanan County's $\hat{c}_{1jt}$ to indicate the precision of predicted values. The overall average coal-county effects in the Virginia coal region was 0.1 from 2005 to 2012. Highest coal-county effects were observed in Buchanan County because of its heavy coal production. However, the coal-county effects increased rapidly in Russell County and Lee County, although their surface coal production had been flat or decreasing.

[Figure 1.2 to be here]

A closer look at figure 1.3 provides an intuitive answer to this puzzle. Russell County's health insurance rates were declining and much lower than other coal-mining counties (Figure 1.3 (a)). By plotting the increments of Russell County's coal effects from 2007 and the fraction of increments explained by health insurance rate (shadow area). Figure 1.3 (b) reveals that Russell County's declining health insurance rates mainly drove the increasing coal-county effect. Given an average of population of 28,834, our model predicted that a 1% decrease in the health insurance rate would lead to 403 residents dying from NMRD in Russell County, and increase the average coal-county effect by 68%.

How about Lee County? Figure 1.3 (c) shows that doctors were leaving Lee County from 2006, and the decreasing number of doctors explained more than two-thirds of the increments of coal-county effects in Figure 1.3 (d). The model result suggested that the average coal-county effect increased by 147% (=0.147/0.1*100%) with one additional doctor per 1000 population leaving.

[Figure 1.3 to be here]

1.3.5 Subgroup analyses

Figure 1.4 (a) and (b) show the predicted coal-county effects from the female-only model and male-only model under the specification of model 2. The predicted coal-county effects on females were quite small, ranging between 0 to 0.1 since 2007, and the marginal effect of coal-county indicator was insignificant, suggesting that, for female residents, living in a coal-mining county was not associated with a higher likelihood of dying from NMRD. However, for males, we found that the coal-county effects were much higher, and coal-mining county residency significantly increased the probability of dying from NMRD.

Next, we estimated the coal-county effects for working age (15-64) and retirement age (>64) separately and plotted the predicted coal-county effects in Figure 1.4 (c) and (d). With an average of 0.18, the coal-county effects were stronger for the working-age population, while the average coal-county effect on the retirement-age population was 0.10. The 95% confidence interval from the working-age model was wider because its sample size was much smaller compared to other subgroups. For Russell County, a dramatic increasing coal-county effect was observed for the working population in Figure 1.4 (c), but not for the retirement-age population in Figure 1.4 (d).

[Figure 1.4 to be here]

1.4 Discussion

The positive marginal effect of the coal county indicator indicated that, compared to a tobacco county, living in a coal county increased the probability of dying from NMRD. Although residents in adjacent coal counties were exposed to similar pollutants from coal production, we did not find higher mortality risk associated with residence in an adjacent-coal county. Additionally, several non-coal factors (i.e., health insurance coverage rates, numbers of doctors and hospital beds and smoking rates) significantly affected the coal-county effect.

Our main results suggested that a decline in health insurance coverage significantly accentuated the coal county effect. County health insurance rates captured the degree of health care coverage. Without any health insurance, patients might not be able to afford medical care, which may result in higher risks of dying from several chronic diseases (Franks, Clancy, & Gold, 1993). In many coal-mining counties, the declining health insurance rate was a common problem, which reduced the affordability of health care services (Huttlinger, Schaller ‐ Ayers, & Lawson, 2004). Since the demand for coal decreased in the United States, several coal companies declared bankruptcy and stopped contributing to the healthcare benefits for their retirees (Perri, 2017). The uninsured can be expected to be more vulnerable to coal-related diseases that needed long-term medical care.

Coal-mining counties are often located in mountain areas and have limited access to health services such as fewer hospitals and physicians than the national average (Denham et al., 2010; Meacham et al., 2013, p. p. 65). The number of doctors reflected the community's ability to detect diseases and provide long-term medical services. A shortage of physicians in Appalachian counties is associated with fewer appointment times (Huttlinger et al., 2004). For example, Wellmont Health system closed the only hospital in Lee County in 2013. After that closure, Lee County's residents have to visit a hospital in a neighboring county for quick lab work or X-rays. In emergencies, travel time to hospitals can be a crucial factor. Like Lee County, some poor Appalachian rural counties faced the problem of doctors leaving (Stensland, Mueller, & Sutton, 2002). A survey by Huttlinger et al. (2004) showed that many respondents in Appalachia had to wait up to three months for a doctor's appointment due to the lack of specialty care providers. A longer waiting time may impede rural residents from seeking early treatment on their coal-related diseases and therefore increase the coal health effect. As several respiratory diseases related to coal exposure are often symptomless, regular screening tests by doctors can result in detection of these diseases at earlier stages when the treatment is more effective to prevent death. Without easy access to healthcare professionals, a patient has a lower chance of surviving as his or her disease progresses to a complicated form (Finkelman et al., 2002; Ward, Ward, & Leach, 2010).

The significantly positive coefficients of the variable, number of hospital beds were unexpected, which might be explained by reverse causality (Table 1.3) (Grafova, Freedman, Lurie, Kumar, & Rogowski, 2014). As the number of hospital beds represents the capacity of healthcare facilities (Kirby & Kjesbo, 2003), a county with a large number of hospital beds often has a big hospital. If so, the county mortality rates can be higher than neighboring counties, because patients from outside the county may go to that hospital for treatments of respiratory diseases such as mechanical ventilation and oxygen therapy (Zilberberg & Shorr, 2008).

Smoking and surface coal mining also contributed to the coal-county effect. Researchers observed much higher smoking rates (Meyer, Toborg, Denham, & Mande, 2008) in central Appalachia than the national average (Gregg et al., 2009). Similar findings from previous literature also suggested that living in a county with surface coal mining was associated with more hospitalizations for asthma (Fitzpatrick, 2018) and high mortality rates of chronic heart diseases (Esch & Hendryx, 2011).

As subgroup analyses revealed higher coal-county effects among male than female residents, we suspected that occupational health hazard from coal miners might partly drive the estimated coal-county effects. Similar findings were reported by Hendryx and Ahern (2008), who found coal effect was higher for male than female residents and interpreted this phenomenon as a miner's effect. A few previous studies found that female residents in coal-mining areas had a higher mortality risk than females in non-coal areas (Hendryx, 2009; Woolley et al., 2015). Our study did not find a significantly positive coal county effect among the female subgroup due to two potential reasons. First, the ICD diagnosis codes used in our study were J00-J99 for NMRD. Previous studies focused on other health outcomes. Second, due to some unobserved factors, less healthy people may self-select to live in economically distressed counties, and the observed health disparity has nothing to do with coal mining. Using tobacco counties as the comparison group provided a better way to separate income effect from the health effect. However, previous researchers used non-coal-mining counties as the comparison group, which did not consider the issue of selection bias and income effects.

In the second subgroup analysis, health effect of coal mining on the working-age residents was higher than that on retirement-age residents. Driven by the decline in health insurance coverage rate, an increase in coal county effect was observed for Russell County's working population, but not for the retirement-age population in the same county, which reflected the crucial role of health insurance on the working population to reduce adverse health impact from coal production.

1.4.1 Policy suggestions

Our findings assist health policymakers in identifying and choosing between alternative strategies when attempting to reduce elevated mortality rates in coal communities. First, affordability of health insurance challenges these coal communities due to declines in the coal industry during the past two decades (Huttlinger et al., 2004), and thus, loss of jobs leads to loss of health benefits. Policy makers may consider expanding health insurance coverage by introducing low-cost health insurance plans and increasing diverse job opportunities. According to Perri (2017), Congress reached a deal to provide a permanent $1.3 billion benefit for over 22,600 retired coal miners and their families, which may be helpful to increase health insurance coverage. Second, to address the shortage of doctors, healthcare facilities in coal-mining counties may consider collaborations with other healthcare facilities and increase incentives to recruit more healthcare professionals. Some rural counties may use telehealth (Singh, Mathiassen, Stachura, & Astapova, 2010), which allows patients to see a remote specialist by using video conferencing.

1.4.2 Limitations

Common to previous studies, this study has several limitations. Although our analyses were based on individual-level data, the risk of ecological bias still existed. According to Greenland and Morgenstern (1989), ecological bias means "the failure of ecological- (aggregate-) level associations to properly reflect individual-level associations." This problem happens when an inference is made for individuals based on aggregate data due to "loss of intergroup variation in the distribution of other risk factors and effect modifiers." Since our regression analysis used individual death records and controlled individual-level covariates, the risk of ecological bias was lower than studies using county-level mortality rates. However, the coal-county effect was an average health impact of living in a coal-mining county. Within each county, the coal health impact on each individual can be different. To detect ecological bias, future studies may consider analyzing the association at individual and different aggregation level to see if there is a significant difference. If yes, appropriate control of individual-level covariates can reduce ecological bias (Greenland & Morgenstern, 1989). One potential limitation of the statistical model is that it did not assess the spatial autocorrelation among counties. Future studies may incorporate spatial analyses to better understand the health effect of living in an adjacent coal county. Finally, the model revealed the statistical association between coal mining and likelihoods of dying from NMRD, but not the causal relationship. To establish a causal link, researchers need more sophisticated identification strategies, such as natural experiments, longitudinal data on both health, environment and coal mining.

Other important limitations are mainly associated with data availability. First, we used county of residence in the death records as a rough measurement of exposure to coal production, which did not capture the length of exposure. Second, separating coal miners' occupational hazard from the community health effect is another common challenge in this field. The lack of separation may overestimate coal health effects on the general population. Since almost all coal miners are male, we estimated the coal-county effects based on female subgroup as a "second-best" strategy to exclude occupational health exposure. Third, the specific mechanism through which coal affects population health is not in the scope of this study. As previous studies suggested coal mining was a significant source of air pollutants (Ghose, 2007; Palmer et al., 2010; Petsonk et al., 2013), future studies may examine environmental factors such as particulate matter distribution and concentration near Appalachian coal-mining region to investigate the mechanism and associate relevant disease incidence data.

1.5 Conclusions

This study is a step forward in understanding the underlying factors that may be associated with a "coal-county effect" and helps identify factors that can be targeted to improve health in coal-mining counties. Based on individual mortality data, we found a higher risk of dying from NMRD associated with living in a coal-mining county, but not with living in an adjacent county. This association was further accentuated by limited accessibility of health services--low health insurance coverage rates and lack of doctors.

Although positive coal-county effects may include occupational hazard, this study still contributes to the literature by showing the critical role of health access in reducing health disparities related to coal exposure, especially for the working population. To move forward, future research needs the occupation information to test whether or not living in a coal-mining county contributes to non-miners' respiratory mortality. Depending on data availability, future research may also consider better measures of coal exposure such as distance from residence to the nearest coal mine site (Saha, Pattanayak, Sills, & Singha, 2011) and occupational histories (Liu et al., 2009).

Appendix AA. VIF test results of collinearity between socioeconomic and health access covariates

	VIF	SQRT VIF	Tolerance
SES			
$R_{unemploy}$	1.47	1.21	0.6786
$Income$	2.16	1.47	0.4628
I_{metro}	1.80	1.34	0.5550
I_{rural}	1.93	1.39	0.5191
Health Access			
$Bed_{per1000}$	1.54	1.24	0.6507
$Hcenter_{per1000}$	1.22	1.10	0.8206
$Doctor_{per1000}$	1.39	1.18	0.7172
R_{insur}	1.89	1.37	0.5291

Appendix AB. Average Marginal Effects on Probability of Dying from NMRD

	(1) General Model	(2) Model 1	(3) Model 2	(4) City Adjusted Model	(5) Scott Check Model
SES					
$R_{unemploy}$	0.009**	0.004			
	(2.02)	(1.24)			
$Income$	0.003**	0.002*			
	(2.46)	(1.77)			
I_{metro}	0.021**	0.019*			
	(2.08)	(1.88)			
I_{rural}	0.014**	0.017**			
	(2.22)	(2.21)			
Health Access					
$Bed_{per1000}$	0.007***	0.007***	0.005***	0.005***	0.005***
	(8.72)	(8.26)	(6.98)	(7.47)	(8.08)
$Hcenter_{per1000}$	-0.114***	-0.130***	-0.132***	-0.132***	-0.126***
	(-9.01)	(-8.90)	(-7.40)	(-7.40)	(-9.68)
$Doctor_{per1000}$	-0.005	-0.009**	-0.008***	-0.008***	-0.008***
	(-1.15)	(-2.01)	(-4.14)	(-4.14)	(-4.45)
R_{insur}	-0.016***	-0.015***	-0.014***	-0.014***	-0.014***
	(-5.16)	(-4.85)	(-4.41)	(-4.78)	(-3.88)
Risk Factor					
$R_{obesity}$	-0.002	-0.002	-0.0004	-0.0004	-0.002
	(-0.71)	(-0.80)	(-0.21)	(-0.21)	(-1.14)
$R_{inactivity}$	-0.002**	-0.002**	-0.002*	-0.002*	-0.001
	(-2.14)	(-2.01)	(-1.75)	(-1.81)	(-1.55)
$R_{smoking}$	0.005**	0.003	0.004***	0.004***	0.004***
	(2.53)	(1.44)	(2.87)	(2.84)	(2.75)
Coal-Related Variables					
$Surface\%$	0.001***	0.001***	0.0005***	0.0005***	0.0005***
	(9.11)	(9.68)	(4.70)	(4.71)	(4.59)
d_{incoal}	0.029***	0.018**	0.014**	0.014**	0.018**
	(3.12)	(2.53)	(2.07)	(2.03)	(2.60)
$d_{adjcoal}$	0.002	-0.004	-0.008	-0.008	-0.001
	(0.19)	(-0.58)	(-1.04)	(-1.15)	(-0.14)
Control Variables					
$Race_{black}$	-0.044***	-0.044***	-0.044***	-0.044***	-0.044***
	(-11.67)	(-11.69)	(-11.63)	(-12.15)	(-11.72)
$Race_{other}$	-0.062	-0.063	-0.062	-0.062	-0.062
	(-1.32)	(-1.32)	(-1.32)	(-1.29)	(-1.31)
Sex_{female}	-0.015***	-0.015***	-0.015***	-0.015***	-0.015***
	(-3.29)	(-3.30)	(-3.31)	(-3.39)	(-3.31)

	(1) General Model	(2) Model 1	(3) Model 2	(4) City Adjusted Model	(5) Scott Check Model
Single	0.003	0.003	0.003	0.003	0.003
	(0.72)	(0.72)	(0.72)	(0.78)	(0.70)
Married	-0.008*	-0.008*	-0.008*	-0.008**	-0.008*
	(-1.92)	(-1.93)	(-1.91)	(-2.17)	(-1.92)
Divorced	-0.001	-0.001	-0.0005	-0.0005	-0.001
	(-0.09)	(-0.10)	(-0.07)	(-0.07)	(-0.08)
Age	0.001***	0.001***	0.001***	0.001***	0.001***
	(7.15)	(7.21)	(7.22)	(7.14)	(7.24)
Education year	-0.005***	-0.005***	-0.005***	-0.005***	-0.005***
	(-11.96)	(-11.96)	(-11.95)	(-11.76)	(-11.98)

Note: z test statistic in parentheses * $p < .1$, ** $p < .05$, *** $p < .01$.

[a] Year dummies are controlled, but their marginal effects are not reported in A2. County-level variables' marginal effects are the average marginal effects of coal-mining residents. Individual control variables' marginal effects are the average marginal effects of the whole sample.

References

Asteriou, D., & Hall, S. G. (2015). *Applied econometrics*: Macmillan International Higher Education.

Blackley, D., Behringer, B., & Zheng, S. (2012). Cancer mortality rates in Appalachia: descriptive epidemiology and an approach to explaining differences in outcomes. *J Community Health, 37*(4), 804-813. doi:10.1007/s10900-011-9514-z

Box, G. E., Luceno, A., & del Carmen Paniagua-Quinones, M. (2011). *Statistical control by monitoring and adjustment* (Vol. 700): John Wiley & Sons.

Boyles, A. L., Blain, R. B., Rochester, J. R., Avanasi, R., Goldhaber, S. B., McComb, S., . . . Thayer, K. A. (2017). Systematic review of community health impacts of mountaintop removal mining. *Environment international, 107*, 163-172.

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources, 50*(2), 317-372.

Centers for Disease Control Prevention. (2005-2012). Behavioral risk factor surveillance system survey data. Retrieved 24 Nov 2018 http://www.cdc.gov/brfss

Cortes-Ramirez, J., Naish, S., Sly, P. D., & Jagals, P. (2018). Mortality and morbidity in populations in the vicinity of coal mining: a systematic review. *BMC public health, 18*(1), 721.

Denham, S. A., Wood, L. E., & Remsberg, K. (2010). Diabetes care: provider disparities in the US Appalachian region. *Rural and Remote health, 10*(2), 1320.

Duncan, C., Jones, K., & Moon, G. (1998). Context, composition and heterogeneity: using multilevel models in health research. *Social Science & Medicine, 46*(1), 97-117.

Dwyer-Lindgren, L., Bertozzi-Villa, A., Stubbs, R. W., Morozoff, C., Shirude, S., Naghavi, M., . . . Murray, C. J. (2017). Trends and patterns of differences in chronic respiratory disease mortality among US counties, 1980-2014. *Jama, 318*(12), 1136-1149.

Dwyer-Lindgren, L., Mokdad, A. H., Srebotnjak, T., Flaxman, A. D., Hansen, G. M., & Murray, C. J. (2014). Cigarette smoking prevalence in US counties: 1996-2012. *Population health metrics, 12*(1), 1.

Energy Information Agency. (2005-2012). Retrieved from https://www.eia.gov/coal/annual/

Esch, L., & Hendryx, M. (2011). Chronic Cardiovascular Disease Mortality in Mountaintop Mining Areas of Central Appalachian States. *Journal of Rural Health, 27*(4), 350-357. doi:http://dx.doi.org/10.1605/01.301-0018315682.2012

Finkelman, R. B., Orem, W., Castranova, V., Tatu, C. A., Belkin, H. E., Zheng, B., . . . Bates, A. L. (2002). Health impacts of coal and coal use: possible solutions. *International Journal of Coal Geology, 50*(1), 425-443.

Fitzpatrick, L. G. (2018). Surface Coal Mining and Human Health: Evidence from West Virginia. *Southern Economic Journal, 84*(4), 1109-1128.

Franks, P., Clancy, C. M., & Gold, M. R. (1993). Health insurance and mortality: evidence from a national cohort. *Jama, 270*(6), 737-741.

Ghose, M. K. (2007). Generation and quantification of hazardous dusts from coal mining in the Indian context. *Environmental monitoring and assessment, 130*(1-3), 35-45.

Graber, J. M., Stayner, L. T., Cohen, R. A., Conroy, L. M., & Attfield, M. D. (2014). Respiratory disease mortality among US coal miners; results after 37 years of follow-up. *Occup Environ Med, 71*(1), 30-39.

Grafova, I. B., Freedman, V. A., Lurie, N., Kumar, R., & Rogowski, J. (2014). The difference-in-difference method: Assessing the selection bias in the effects of neighborhood environment on health. *Economics & Human Biology, 13*, 20-33.
Granados, J. A. T., House, J. S., Ionides, E. L., Burgard, S., & Schoeni, R. S. (2014). Individual joblessness, contextual unemployment, and mortality risk. *American journal of epidemiology, 180*(3), 280-287.
Greenland, S., & Morgenstern, H. (1989). Ecological bias, confounding, and effect modification. *International journal of epidemiology, 18*(1), 269-274.
Gregg, E., Kirtland, K., Cadwell, B., Burrows, N. R., Barker, L., Thompson, T., . . . Pan, L. (2009). Estimated county-level prevalence of diabetes and obesity-United States, 2007. *Morbidity and Mortality Weekly Report, 58*(45), 1259-1263.
Health resources and services administration. (2015-2016). Area Resource File, . Retrieved 24 Nov 2018 https://datawarehouse.hrsa.gov/topics/ahrf.aspx
Hendryx, M. (2009). Mortality from heart, respiratory, and kidney disease in coal mining areas of Appalachia. *Int Arch Occup Environ Health, 82*(2), 243-249. doi:10.1007/s00420-008-0328-y
Hendryx, M., & Ahern, M. M. (2008). Relations between health indicators and residential proximity to coal mining in West Virginia. *Am J Public Health, 98*(4), 669-671. doi:10.2105/AJPH.2007.113472
Hendryx, M., & Ahern, M. M. (2009). Mortality in Appalachian Coal Mining Regions: The Value of Statistical Life Lost. *Public Health Reports, 124*(4), 541-550.
Hendryx, M., Ahern, M. M., & Nurkiewicz, T. R. (2007). Hospitalization Patterns Associated with Appalachian Coal Mining. *Journal of Toxicology and Environmental Health, Part A: Current Issues, 70*(23), 2064.
Hendryx, M., Fedorko, E., & Anesetti-Rothermel, A. (2010). A geographical information system-based analysis of cancer mortality and population exposure to coal mining activities in West Virginia, United States of America. *Geospatial health, 4*(2), 243-256.
Hendryx, M., Fedorko, E., & Halverson, J. (2010). Pollution sources and mortality rates across rural-urban areas in the United States. *The Journal of rural health, 26*(4), 383-391.
Hendryx, M., O'Donnell, K., & Horn, K. (2008). Lung cancer mortality is elevated in coal-mining areas of Appalachia. *Lung Cancer, 62*(1), 1-7. doi:10.1016/j.lungcan.2008.02.004
Hitt, N. P., & Hendryx, M. (2010). Ecological integrity of streams related to human cancer mortality rates. *EcoHealth, 7*(1), 91-104.
Huttlinger, K., Schaller-Ayers, J., & Lawson, T. (2004). Health Care in Appalachia: A Population-Based Approach. *Public Health Nursing, 21*(2), 103-110.
Kentucky Legislative Research Commission. (2018). Retrieved from https://apps.legislature.ky.gov/record/18rs/hb2.html
Kirby, A., & Kjesbo, A. (2003). Tapping into hidden hospital bed capacity. *Healthcare financial management, 57*(11), 38-42.
Krometis, L.-A., Gohlke, J., Kolivras, K., Satterwhite, E., Marmagas, S. W., & Marr, L. C. (2017). Environmental health disparities in the Central Appalachian region of the United States. *Reviews on Environmental Health, 32*(3), 253-266.
Liu, H., Tang, Z., Yang, Y., Weng, D., Sun, G., Duan, Z., & Chen, J. (2009). Identification and classification of high risk groups for Coal Workers' Pneumoconiosis using an artificial neural network based on occupational histories: a retrospective cohort study. *BMC public health, 9*(1), 366.

Lockwood, A. H. (2012). *The silent epidemic: coal and the hidden threat to health*: MIT Press.
McGarvey, E. L., Leon-Verdin, M., Killos, L. F., Guterbock, T., & Cohn, W. F. (2011). Health disparities between Appalachian and non-Appalachian counties in Virginia USA. *J Community Health, 36*(3), 348-356. doi:10.1007/s10900-010-9315-9
Meacham, S. L., Sukpraput, S., Taber, O., & Metzger, D. (2013). Publications on chronic disease in coal dependent communities in Central Appalachia. *Journal of Health Disparities Research and Practice, 5*(3), 62-79.
Meyer, M. G., Toborg, M. A., Denham, S. A., & Mande, M. J. (2008). Cultural perspectives concerning adolescent use of tobacco and alcohol in the Appalachian mountain region. *The Journal of rural health, 24*(1), 67-74.
Mine Safety and Health Administration. (2016-2018). Retrieved from https://arlweb.msha.gov/ACCINJ/accinj.htm
Morrice, E., & Colagiuri, R. (2013). Coal mining, social injustice and health: A universal conflict of power and priorities. *Health & Place, 19*, 74-79.
Palmer, M. A., Bernhardt, E. S., Schlesinger, W. H., Eshleman, K. N., Foufoula-Georgiou, E., Hendryx, M. S., . . . Wilcock, P. R. (2010). Mountaintop Mining Consequences. *SCIENCE, 327*(5962), 148-149. doi:10.1126/science.1180543
Perri, J. (2017). Retrieved from https://www.change.org/l/us/congress-extends-healthcare-for-retired-coal-miners-and-their-families
Petsonk, E. L., Rose, C., & Cohen, R. (2013). Coal mine dust lung disease. New lessons from an old exposure. *American journal of respiratory and critical care medicine, 187*(11), 1178-1185.
Saha, S., Pattanayak, S. K., Sills, E. O., & Singha, A. K. (2011). Under-mining health: environmental justice and mining in India. *Health & Place, 17*(1), 140-148.
Singh, R., Mathiassen, L., Stachura, M. E., & Astapova, E. V. (2010). Sustainable rural telehealth innovation: a public health case study. *Health services research, 45*(4), 985-1004.
StataCorp LP. (2015). Stata data analysis and statistical Software 14.0.
Stensland, J., Mueller, C., & Sutton, J. (2002). *An analysis of the financial conditions of health care institutions in the Appalachian region and their economic impacts*: Project HOPE, Center for Health Affairs.
VDH. (2013). *Office of Vital Records and Health Statistics, Virginia Department of Health*.
Ward, J. P., Ward, J., & Leach, R. M. (2010). *The respiratory system at a glance* (Vol. 19): John Wiley & Sons.
Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*: MIT press.
Woolley, S. M., Meacham, S. L., Balmert, L. C., Talbott, E. O., & Buchanich, J. M. (2015). Comparison of Mortality Disparities in Central Appalachian Coal-and Non—Coal-Mining Counties. *Journal of Occupational and Environmental Medicine, 57*(6), 687-694.
Zilberberg, M. D., & Shorr, A. F. (2008). Prolonged acute mechanical ventilation and hospital bed utilization in 2020 in the United States: implications for budgets, plant and personnel planning. *BMC health services research, 8*(1), 242.

Figures and Tables

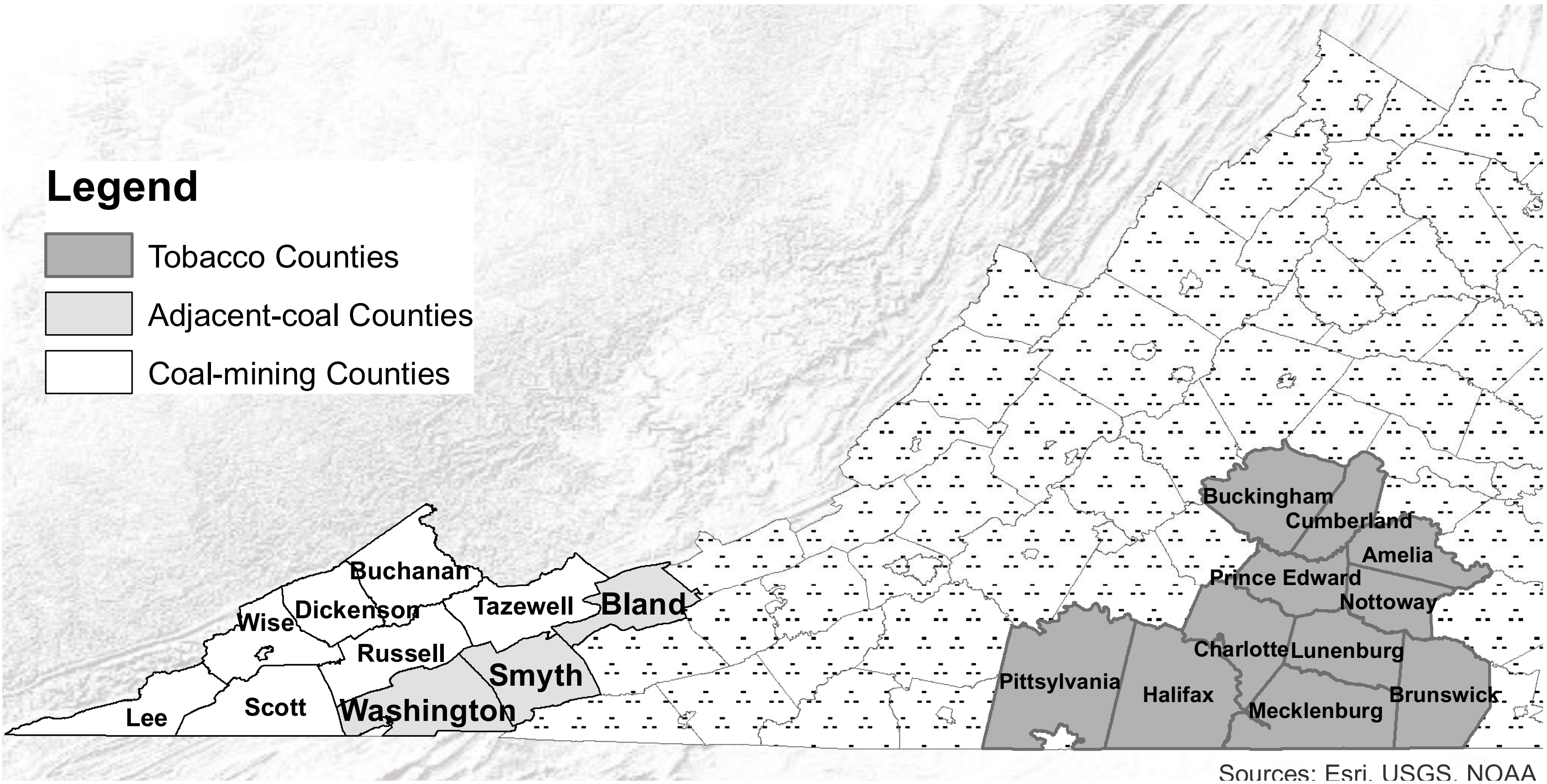

Figure 1.1 Study area by three county groups in Virginia

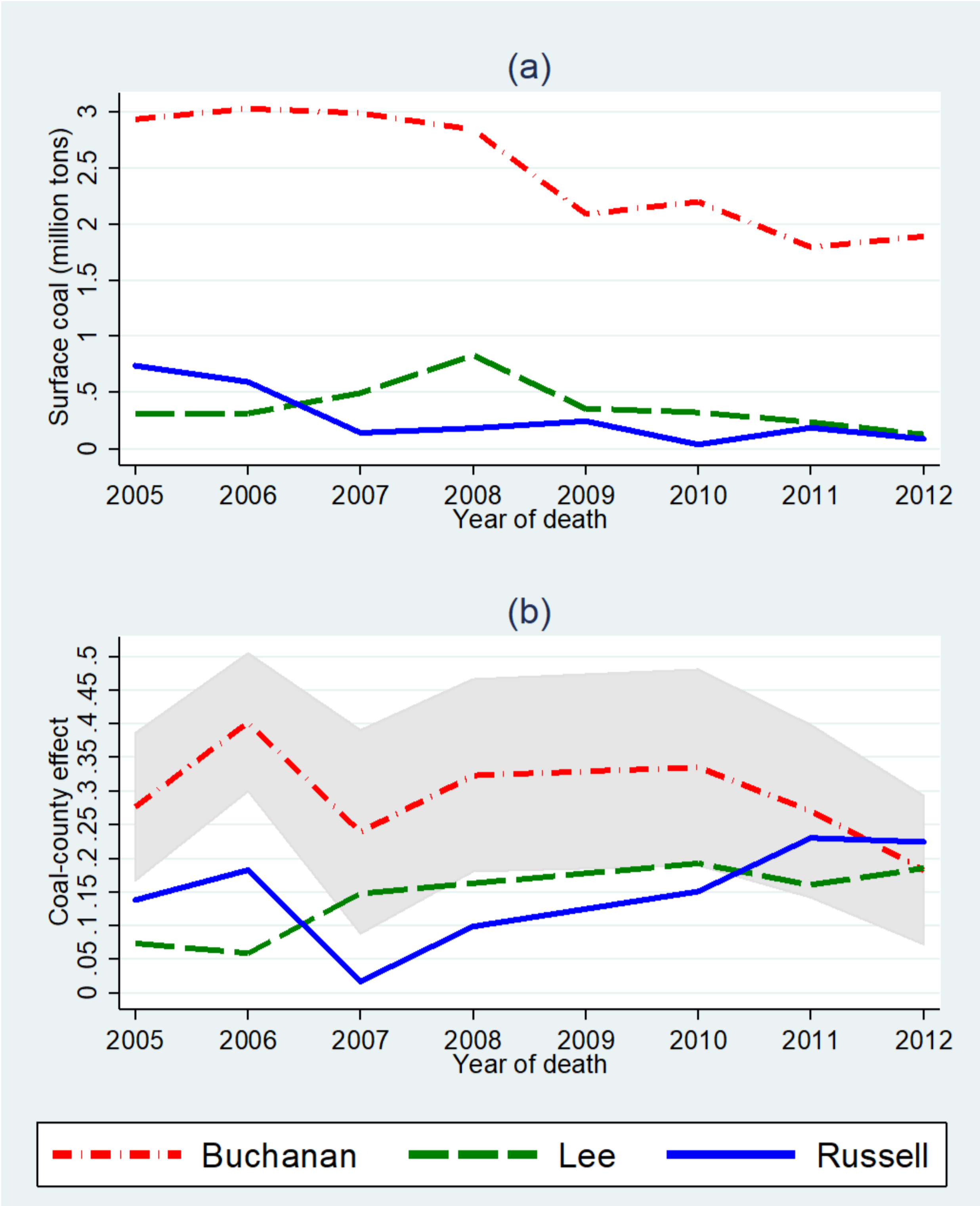

Figure 1.2 (a) Annual surface coal production and (b) Predicted coal-county effects of three Virginia coal-mining counties

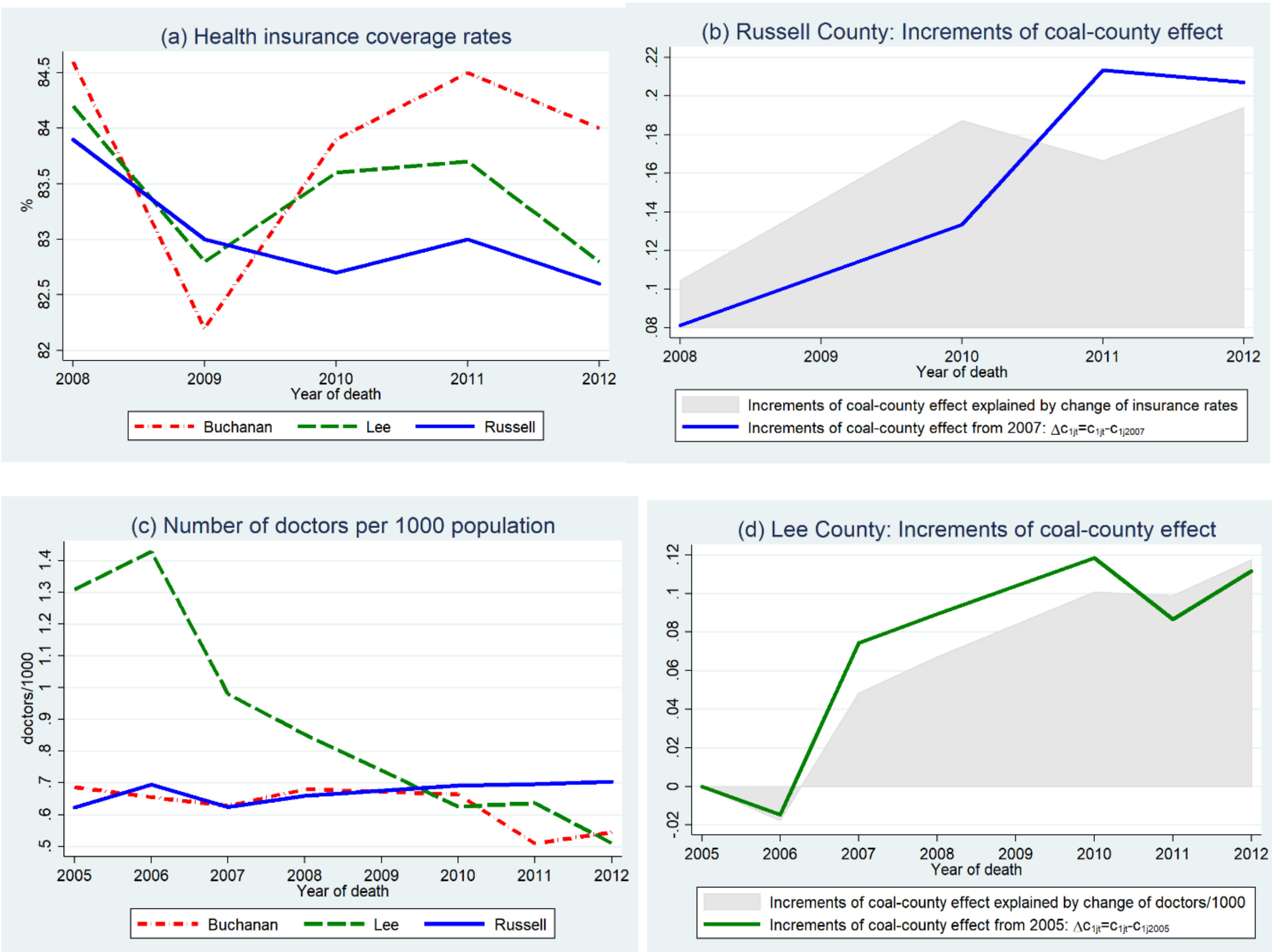

Figure 1.3 Increasing coal-county effects in two counties caused by deterioration in access to healthcare.

Note: Year-to-year comparisons of insurance rates are only appropriate after 2007 because the SAHIE program switched the data source in 2008.

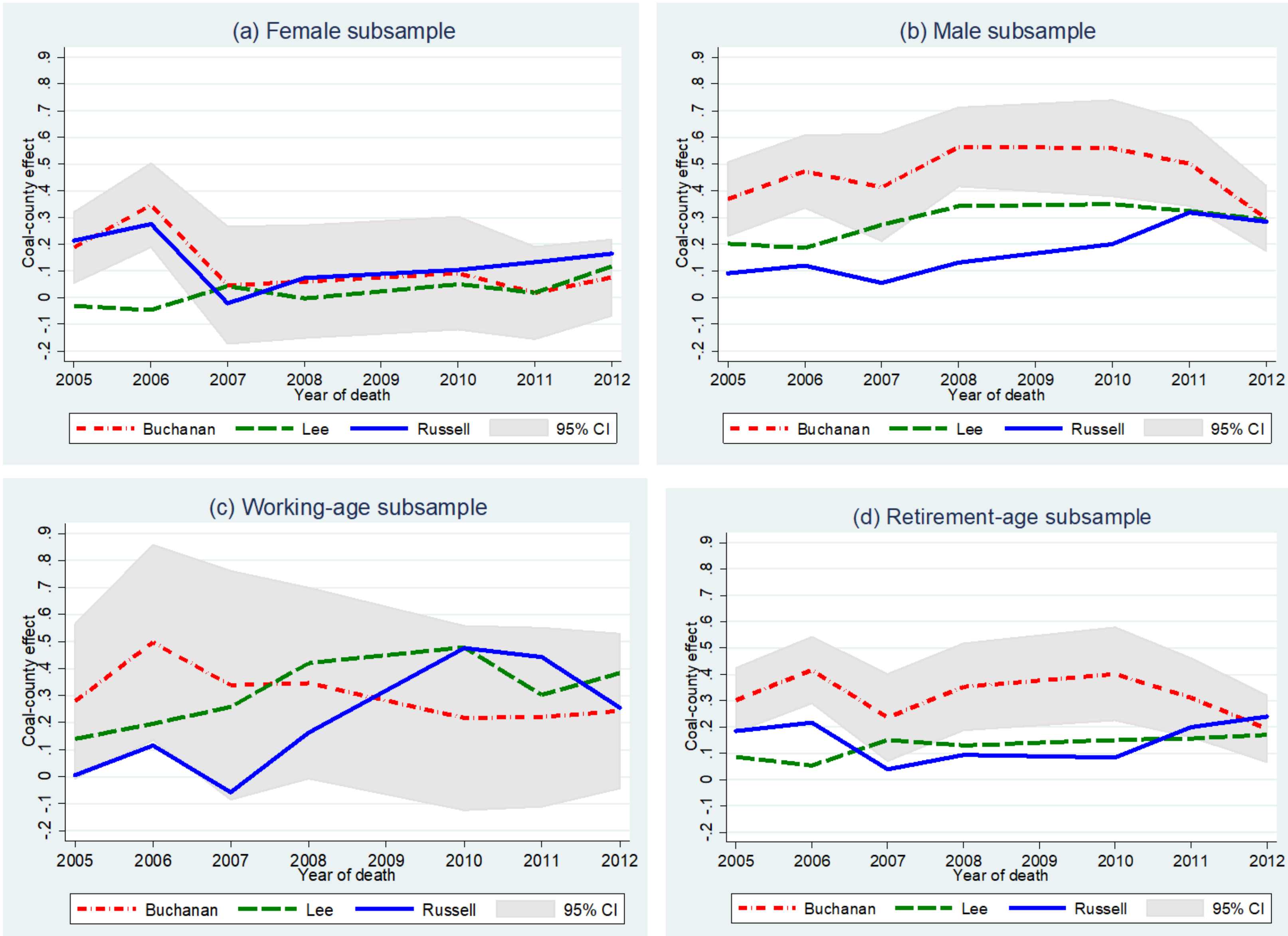

Figure 1.4 Subsample predicted coal-county effects (a) Female, (b) Male, (c) Working-age and (d) Retirement-

Table 1.1 Summary of individual and county-level characteristics from years 2005 to 2012.

Variable	Definition and Label	Mean	SD [a]	Min [b]	Max [c]
Dependent variable					
Y_{NMRD}	Death indicator: 1= Death due to Non-Malignant Respiratory Disease	0.11	0.32	0	17
Demographics					
edu	Years of education	10.08	3.56	0	17
age	Age in years	72.24	17.55	0	109
I_{white}	Race indicator: 1=white	0.83	0.38	0	1
I_{black}	Race indicator: 1=black	0.17	0.38	0	1
I_{other}	Race indicator: 1=other race except for white and black	0.002	0.05	0	1
female	Gender indicator: 1= female	0.50	0.50	0	1
I_{single}	Marital status indicator: 1= single	0.11	0.31	0	1
$I_{married}$	Marital status indicator: 1= married	0.39	0.49	0	1
$I_{widowed}$	Marital status indicator: 1= widowed	0.38	0.48	0	1
I_{divor}	Marital status indicator: 1= divorced	0.13	0.33	0	1
SES					
$R_{unemploy}$	Unemployment rate	0.07	0.02	0.03	0.1
$Income$	Median household income in 1000 dollars	35.88	4.12	25.2	52.4

Variable	Definition and Label	Mean	SD [a]	Min [b]	Max [c]
I_{metro}	Rural-urban indicator: 1= county in a metro area	0.28	0.45	0	1
$I_{nonmetro}$	Rural-urban indicator: 1= nonmetropolitan county with the urban population more than 2500	0.33	0.47	0	1
I_{rural}	Rural-urban indicator: 1= nonmetropolitan county completely rural with less than 2500	0.39	0.49	0	1
Risk factors					
$R_{inactivity}$	Age-adjusted leisure-time physical inactivity prevalence percent	0.28	0.03	0.2	0.4
$R_{obesity}$	Age-adjusted obesity rate	0.30	0.03	0.2	0.4
$R_{smoking}$	Age-standardized total cigarette smoking prevalence rate	0.28	0.02	0.2	0.3
Health access					
$bed_{per1000}$	Hospital beds per 1000 population	3.08	3.85	0	15.1
$hcenter_{per1000}$	Federal qualified health centers per 1000 population	0.06	0.07	0	0.3
$doctor_{per1000}$	M.D. and D.O. total active non-Fed & fed per 1000 population	1.11	0.84	0.1	2.9
R_{insur}	Percent insured under 65 years (%)	83.70	1.79	78.9	88.2

Variable	Definition and Label	Mean	SD [a]	Min [b]	Max [c]
$d_{>2007}{}^{d}$	Switch indicator: 1=year after 2007	0.62	0.49	0	1
Coal-related					
$Prod$	County coal production (million tons)	1.23	2.87	0	11.8
$Surface$	County surface coal production (million tons)	0.52	1.37	0	6.7
$Surface\%$	Percent of surface mining coal (%)	11.35	20.66	0	71.7
d_{incoal}	Coal indicator: 1=live in a coal-mining county	0.34	0.47	0	1
$d_{adjcoal}$	Coal indicator: 1=living in an adjacent county of coal-mining counties	0.18	0.38	0	1

Note: [a] the SD denotes the standard deviation.
[b] Min denotes the minimum values of each variable.
[c] Max denotes the maximum values of each variable.
[d] The SAHIE program calculates county-level health insurance based on national survey data. In 2008, the SAHIE program switched from using Current Population Survey (CPS) as the basis of estimation to American Community Survey (ACS). Therefore, to capture the structural change of this variable in the model, we add a product of the insurance rate with a switch indicator $d_{>2007}$, which is one after 2007.

Table 1.2 Wald test of varying parameters

Vector of Variables to Test for Joint Insignificance	**p-value:** adjusted (unadjusted)	β_{0jt} (1)	c_{1jt} (2)	c_{2jt} (3)	$\beta_{0jt} + c_{1jt}$ (4)	$\beta_{0jt} + c_{2jt}$ (5)	$c_{1jt} + c_{2jt}$ (6)	$\beta_{0jt} + c_{1jt} + c_{2jt}$ (7)
(1) All variables except for the intercept		<0.01 (<0.01)	<0.01 (<0.01)	<0.01 (<0.01)	0.02 (<0.01)	0.02 (<0.01)	0.14 (<0.01)	0.29 (<0.01)
(2) SES**:** $R_{unemploy}, Income, I_{rural}, I_{metro}$		0.31 (0.21)	0.11 (0.04)	0.01 (<0.01)	0.01 (<0.01)	0.04 (<0.01)	0.02 (<0.01)	0.06 (<0.01)
(3) HA**:** $bed_{per1000}, hcenter_{per1000}, doctor_{per1000}, R_{insur}, R_{insur} * d_{>2007}$		0.08 (0.01)	<0.01 (<0.01)	<0.01 (0.02)	<0.01 (<0.01)	<0.01 (<0.01)	<0.01 (<0.01)	<0.01 (<0.01)
(4) HR**:** $R_{obesity}, R_{inactivity}, R_{smoking}$		0.06 (0.02)	0.15 (0.09)	0.61 (0.56)	0.01 (<0.01)	0.30 (0.12)	0.30 (0.12)	0.08 (<0.01)

Table 1.3 Estimated coefficients of varying coal-county effects

	(1) General Model[a]	(2) Model 1[b]	(3) Model 2[c]	(4) City Adjusted Model[d]	(5) Scott Check Model[e]
Intercept (c_1)	4.134**	4.969***	5.028***	5.028***	5.088***
	(2.39)	(3.13)	(3.59)	(3.67)	(3.60)
SES					
$R_{unemploy}$	0.020	0.020			
	(1.09)	(1.24)			
$Income$	0.011	0.011*			
	(1.16)	(1.78)			
I_{metro}	0.146**	0.092*			
	(2.43)	(1.88)			
I_{rural}	0.107*	0.079**			
	(1.80)	(2.22)			
Health Access					
$bed_{per1000}$	0.0412***	0.041***	0.034***	0.034***	0.031***
	(8.57)	(8.71)	(7.76)	(8.38)	(8.09)
$hcenter_{per1000}$	-0.131	-0.157	-0.140	-0.140	-0.138
	(-0.52)	(-0.75)	(-0.64)	(-0.87)	(-0.67)
$doctor_{per1000}$	-0.119***	-0.143***	-0.147***	-0.147***	-0.141***
	(-3.02)	(-4.25)	(-5.92)	(-8.04)	(-5.79)
R_{insur}	-0.065***	-0.070**	-0.068***	-0.068***	-0.065***
	(-3.84)	(-4.12)	(-4.08)	(-4.15)	(-3.73)
$R_{insur} * d_{>2007}$	-0.001*	-0.002*	-0.001*	-0.001*	-0.001
	(-1.88)	(-2.76)	(-1.84)	(-1.78)	(-1.16)
Risk Factor					
$R_{obesity}$	0.005	-0.002	0.010	0.010	-0.03
	(0.38)	(-0.19)	(1.13)	(1.09)	(-0.37)
$R_{inactivity}$	-0.008	-0.006	-0.005	-0.005	-0.003
	(-0.88)	(-0.73)	(-0.71)	(-0.72)	(-0.48)
$R_{smoking}$	0.035**	0.029**	0.026***	0.026**	0.027***
	(2.52)	(1.97)	(2.64)	(2.45)	(2.71)
Coal Production					
Surface%	0.004***	0.003***	0.002***	0.002***	0.002***
	(9.14)	(9.72)	(4.68)	(4.71)	(4.55)
Pseudo R^2	0.0297	0.0296	0.0294	0.0294	0.0293
Log likelihood	-16917	-16918	-16921	-16921	-16922
BIC	34103.3	34085.3	34100.8	34068.4	34104.1
Number of observations	49437	49437	49437	49437	49437

Note: z test statistic in parentheses * $p < .1$, ** $p < .05$, *** $p < .01$

[a] General model: kept all vectors of SES, HA and HR variables in β_{0jt}, c_{1jt} and c_{2jt} in equation (1.5) as the preliminary model.

[b] Model 1: removed all SES variables in the β_{0jt} equation and all HR variables in the c_{2jt} equation from the General model

[c] Model 2: removed all SES variables in both β_{0jt} and c_{1jt} equations and all HR variables in the c_{2jt} equation from the General model

[d] City adjusted model: since there are independent cities that nest into counties in Virginia, we collapsed these cities into their belonging counties and adjusted the clustered structure of error terms in model 2 accordingly.

[e] Scott check model: provided that Scott County stop producing coal in 1996, we treated Scott County as an adjacent coal county instead of a coal-mining county to check sensitivity using model 2's specification.

Chapter 2: Double-Edged Sword: Liquidity Implications of Futures Hedging in Corn and Soybean Markets

2.1 Introduction

As global commodity markets have experienced substantial volatility growth in the recent decade, price risk management has become increasingly important. Commodity futures are widely used as a risk management tool to offset price movements in the spot markets. Previous studies discussed various costs of hedging with futures including margin requirements (Hardouvelis & Kim, 1995; Riley & Anderson, 2009), commission fees (Alexander, Prokopczuk, & Sumawong, 2013) and costs of borrowing funds (Arias, Brorsen, & Harri, 2000) to meet margin calls. However, the amounts of funds needed to sustain a futures position are often difficult to estimate and anticipate because of changing prices. Maintaining an account to meet margin calls is consistently quoted as the main impediment to using futures markets by agricultural producers (FCSA, 2017). While some argue that if the hedge is implemented correctly, the losses in futures positions should be directly offset by the gains in cash markets, large margin calls may lead to significant liquidity problems that often result in premature termination of a hedge or even bankruptcy. Some of the most notorious examples include the Metallgesellschaft Debacle (Mello & Parsons, 1995) and bankruptcies of several cotton merchant firms in 2008 (Carter & Janzen, 2009).

Despite this evidence, only a few studies have attempted to evaluate the costs of maintaining a futures hedge. Riley and Anderson (2009) estimated that the average margin requirement for corn was 13 cents/bushel in 2007, which was much higher than 4 cents/bushel in previous years. However, the margin requirement alone does not reflect the full costs of futures hedging. Alexander et al. (2013) considered transaction cost and margin cost. The transaction cost consisted of the commission fee and bid-ask spread, and the margin cost was measured as the borrowing cost of financing the initial margin plus interest losses and gains of daily cash flow from margin accounts. Other studies examined the potential liquidity problems associated with futures hedging by assessing the cash flow risk (Dahlgran and Liu (2011), or including financial constraints (e.g., Deep, 2002; Lien, 2003). However, no consensus exists in the literature on how to measure the liquidity implications of the costs of hedging.

To fill these gaps in knowledge, the goals of this study are threefold. First, a novel conceptual framework is proposed to measure the cost of maintaining a margin account for a futures hedge. We incorporate direct (e.g., margin liability and borrowing costs) and indirect cost (e.g., probability of hedging failure) into one framework and make it possible to compare the results between studies of different commodities. These costs are calculated based on corn and soybean futures prices from the Chicago Board of Trade and historical margin requirements from the Chicago Mercantile Exchange (CME). Changes in margin requirements, direct and indirect costs of hedging are measured across three sub-periods: 2004-2006, 2007-2013, and 2014-2018[5]. Second, this study examines empirical distributions of the hedging costs and estimate the impacts of hedging cost determinants at different parts of the cost distribution.

Understanding the costs of hedging, their changes over time, and driving factors is essential for a successful risk management program implementation. Our findings will also be helpful for policy analysis associated with various risk management tools. While hedging with futures and options continues to be the most common tools of price risk management, alternative instruments, such as swaps, are growing in popularity partly because of the costs associated with futures hedging. The results of this study may be helpful for designing new hedging strategies that minimize costs while offering a similar level of protection.

5 This strategy allows for comparing hedging costs with forward contracting costs calculated by Etienne, Mallory, and Irwin (2017) during similar time periods.

2.2 Conceptual Framework

Hedging with futures involves taking a position in a futures market opposite to the underlying position in the cash market. In theory, losses in the futures markets should be offset by the gains in the spot markets. However, as Mello and Parsons (1995) pointed out, this requires the maturity of future positions always to match the spot positions remaining to be traded, as well as the convergence of futures and spot prices. If there is mismatched maturity in the hedge or non-convergence problems (Garcia, Irwin, & Smith, 2015), unfavorable movements of futures prices make hedgers very vulnerable to liquidity crises. Therefore, it is essential to understand the costs associated with maintaining a futures hedge and their implications on hedger's liquidity.

The costs of hedging included in this study are associated with the amount and the cost of capital invested in the margin account [6]. Specifically, the amount of capital illustrates the access to capital required to maintain the margin account, and the cost of capital reflects borrowing costs. Liquidity problems may arise when the margin costs exceed hedger's borrowing constraints concerning either borrowing costs or credit limits. This section describes how to measure the costs of hedging from different aspects in a comprehensive framework.

2.2.1 Margin liability

First, in order to open a futures position, a trader would have to post an initial margin required by the exchange. Futures hedging offers a great deal of liquidity since margin requirements represent only a small percentage of the position value (about 2 to 10%). The initial margin requirements are stipulated by the exchange on a per-contract basis but can be

[6] Other studies (e.g., Alexander et al., 2013) also included transaction costs, which we are not taken into account in our study as they are known in advance and do not affect liquidity risk.

transferred to a per-unit level by dividing over the contact size: $M^0 \equiv \frac{\$\, initial\ margin}{size\ of\ a\ contract}$. Thus, the balance in the margin account (B) starts with an initial margin: $B_0 = M^0$. The value of the margin account balance changes every day as prices change. Using the futures settlement price p, we compute the daily changes in margin account balance before marking to market as

$$B_t = B_{t-1} + \Delta p_t, \qquad t = 1, 2, \dots, T \tag{2.1}$$

where T is a hedging horizon and $\Delta p_t = p_{t-1} - p_t$ for a short position and $\Delta p_t = p_t - p_{t-1}$ for a long position. The margin account balance is then compared to the maintenance margin requirement[7] $M^m \equiv \frac{\$\, \text{maintenance}\ margin}{size\ of\ a\ contract}$ to determine whether a margin call is required:

1) if $B_t < M^m$, a cash deposit (also called a margin call or the variation margin) is required to bring the margin account level to the initial margin generating negative cash flow for the hedger, equivalent to $M^0 - B_t$.
2) if $B_t > M^0$, the hedger is allowed to withdraw the extra margin money in excess of the initial margin, equivalent to $B_t - M^0$.

If the maintenance margin is equal to the initial margin and these requirements do not change over the hedging horizon, the size of these cash flows per unit is equivalent to Δp_t, and the balance in the account may be returned to the initial margin level after the deposits and withdrawals are made at the end of every day. Following the idea by Dahlgran and Liu (2011), if the initial margin requirement changes at day t by $\Delta M_t = M^0_{t-1} - M^0_t$, the cash flow is

$$CF_t\ (cents/bushel) = \underbrace{\Delta p_t}_{\Delta\ of\ futures\ price} + \underbrace{\Delta M_t}_{\Delta\ of\ maint.\ margin}$$

[7] Initial margin is usually set as 110% of maintenance margin for speculative participants, and 100% of maintenance margin for hedgers.

Therefore, at the end of day t, the cumulative gain (or loss) generated by the margin account (π) is the sum of previous daily cash flows minus the initial margin.

$$\pi_t = -M^0 + \sum_{j=1}^{t} \mathrm{CF_j} \ , for\ t = 1, 2, \dots, T \tag{2.2}$$

Assuming margin requirements do not change over the hedging horizon ($\Delta M_t = 0$), we illustrate cash flow ($CF_t = \Delta p_t$) and π_t in two hypothetical three-month short hedges for corn in Figure 2.1. The blue line represents π_t, the cumulative margin gain (or loss) associated with the margin account, which is affected by the initial margin requirement and daily cash flow (orange bars). In this example, the initial margin is assumed to be \$850 per contract, or 17 cents/bushel. For hedge A, $\pi_1 = -17$ cents/bushel indicates the minimum money needed to open a hedge position. Despite positive cash flows in several days, the blue line of cumulative margin gain/loss remains negative illustrating cumulative losses for the entire hedging horizon. These cumulative losses are averaged 35.78 cents/bushel over the life of the hedge and reached 48 cents/bushel on the day 22 of the hedge. For hedge B, due to mostly positive cash flows from the hedge, the cost of hedging is much less dramatic with cumulative losses averaging just 3.49 cents/bushel. This value represents the amount of capital required to maintain this hedge.

Dahlgran and Liu (2011) used the variance of cash flow as a measure of hedging risk. This assumption is not very realistic for most hedgers' concerns because cash flow is usually considered as a risk only if the margin account balance falls below the maintenance level. Instead, the cumulative losses generated by the margin account provide a better measure of hedging costs than daily cash flow as funds have to be deposited in the margin account in order to maintain a hedge (see the description of margin calls above). Thus, we concentrate on cumulative losses in this study and developed a new measure called margin liability. At day t, the margin liability is:

$$L_t = -\pi_t \cdot D(\pi_t < 0), t = 1, \ldots, T. \tag{2.3}$$

where π_t is defined in equation (2.2) and the indicator $D(.)$ returns to 1 if $\pi_t < 0$ and 0 otherwise. Then, the average daily margin liability over a hedging horizon

$$\bar{L} = \frac{1}{T}\sum_{t=1}^{T} L_t = \frac{1}{T}\sum_{t=1}^{T} -\pi_t \cdot D(\pi_t < 0) \tag{2.4}$$

is a convenient measure to compare the costs of hedging across various hedges as it adjusts the costs of hedging by the number of days in the hedging horizon. For example, in Figure 2.1, $\bar{L}$ is 35.78 cents/bushel for hedge A and 3.49 for hedge B reflecting substantially higher costs associated with hedge A. This measure reflects the average amount of money that has to be deposited in the margin account over the hedging horizon in order to maintain a hedge, and it serves as our primary measure of hedging costs.

[Figure 2.1 to be here]

2.2.2 Borrowing costs

After defining average margin liability, we can calculate a borrowing cost of hedging, which is used in direct cost studies (Alexander et al., 2013; Arias et al., 2000). We assume that the hedger has to borrow the margin liability, L, defined above. The borrowing costs are then calculated as:

$$BC = \sum_{t=1}^{T} L_t \cdot r = \bar{L} \cdot T \cdot r \tag{2.5}$$

where r is the daily interest rate. Therefore, margin liability and interest rates will directly affect the borrowing costs. BC is a more aggressive measure of borrowing costs as it does not include interest gain on excess margin. Since this appears to be the case in many margin accounts, we use this measure in our main set of results.

Alternatively, if the excess margin earns interest, the borrowing cost should include both positive and negative π_t and is defined as the sum of daily interest paid on negative π_t and interest earned on positive π_t:

$$BC^{+} = \sum_{t=1}^{T} -\pi_t * r \tag{2.6}$$

The basic idea behind BC^{+} is the same with the margin cost developed by Alexander et al. (2013), assuming that the interest earned is equal to the interest paid, but can be easily modified to reflect different rates of interest earned. BC^{+} is better than the margin cost, because it is based on the losses and gains on a daily basis. Margin cost is calculated weekly using a linear approximation of the daily changes in the margin account because of Alexander et al.'s data limitation. Note that BC^{+} can be negative if a margin account generates more positive cash flows than the funds required to maintain it, so BC^{+} would be a more conservative measure of borrowing costs associated with hedging.

2.2.3 Probability of hedging failure

This framework enables us to develop measures of hedging costs that are comparable to previous studies focusing on hedging failure. A hedging failure occurs when a hedger cannot generate enough funds to deposit into the margin account when margin losses occur, in which case a hedge would be terminated (at least partially according to the shortfall). Failure to generate funds to sustain a hedge may result from excessive borrowing costs described above as well as credit limits. Maximum credit required to maintain a hedge is best illustrated with a novel term called maximum margin liability, L_{max}, which reflects the largest cumulative loss in the margin account:

$$L_{max} = \max(L_1, L_2, \dots L_T), \text{ t=1, …, T.} \tag{2.7}$$

For instance, Figure 2.1 suggests that one has to finance at least 48 cents/bushel to avoid premature termination of hedge A. Following the idea by Deep (2002) and Lien (2003), we assume a capital constraint, C. If L_{max} exceeds this constraint, a hedger will be forced to abandon the futures position before the expected ending day. Thus, the probability of hedging failure is approximated empirically as a proportion of simulated hedges for which L_{max} exceeds C:

$$Prob(L_{max} > C) = \frac{1}{N}\sum_{i=1}^{N} D(L_{max} > C) \times 100\%, \;\; i = 1, 2, \dots, N \tag{2.8}$$

This equation has not been used in previous studies. For hedge i, the indicator function $D(L_{max} > C)$ returns to 1 if maximum daily liability exceeds the capital constraint, indicating this hedge has to be abandoned. This measure suggests that a more restrictive borrowing constraint increases the probability of premature termination of a hedge. This risk may be a crucial concern for many small producers who have insufficient credit lines. In this framework, the probability of hedging failure enables us to compare hedging costs between commodities and is applicable to empirical analysis, while previous studies' borrowing constraint is often used as a theoretical parameter.

2.3 Simulation of the Costs of Hedging

To calculate the costs of hedging, we simulated weekly short and long hedges for a single contract using historical futures prices and margin requirements from 2004 to 2018. Each week[8], a short (long) position was opened with a target ending day in one, three, and six months later (1, 3, 6-month hedging horizon). The hedge was simulated based on a target futures contract specified as the one with the nearest active delivery month to the hedge ending day[9]. Daily settlement prices of the target futures contract were used for these hedge simulations. We assumed the initial margin was 100% of the maintenance margin. Some previous study used margin requirements as the cost of hedging (Hardouvelis & Kim, 1995; Riley & Anderson, 2009). The following analyses reveal that margin requirements are inadequate to reflect the liquidity risk of hedging.

2.3.1 Changes in margin requirements

Figure 2.2 shows changes in initial margins over time. The CME margin requirements were not available from 9/20/2007 to 1/2/2009. Therefore, we estimated margin requirements in this period based on available information, as described in Appendix BA. Robustness of the results to this imputation will be examined. For corn futures, the margin requirements were quite stable at around 10 cents/bushel before 2007. From 2007 to 2013, the margin requirements had increased by 30 cents/bushel. This observation is consistent with Riley and Anderson (2009), who found the average corn margin requirement in 2007 is much higher than that of previous

[8] The hedges were initiated on Tuesdays because Disaggregated Commitments of Traders Report provides each Tuesday's open interest for futures. To understand the implication, we calculate hedging costs on the same day when market participation is reported.

[9] . If hedge i's ending day reached the delivery month of the target futures contract used by prior week's hedge, the next futures contract would be used as the target futures contract for hedge i.

years. After 2013, the exchange gradually reduced the margin requirements but still kept the values above the level in 2004-2006. In the soybean futures market, the margin requirements increased from 20 to 75 cents/bushel in 2007-2013 and decreased to 40 cents/bushel by the end of 2018.

[Figure 2.2 to be here]

2.3.2 Costs of hedging

Next, we calculated the average daily margin liability ($\bar{L}$) and maximum margin liability (L_{max}) associated with 1, 3, and 6-month hedges. Figure 2.3 shows $\bar{L}$ and L_{max} calculated for three-month hedges of corn futures, along with nearby futures prices and margin requirements. $\bar{L}$ and L_{max} are better measures of the hedging costs than margin requirements because they also capture price movements in the hedge period. The black line $\bar{L}$ measures the average amount that has to be borrowed for the duration of the hedge (three months in this example) in order to maintain a margin account. This figure demonstrates a positive relationship between short margin liability and negative relationship between long margin liability with an increase of futures price during the hedge horizon. For example, when corn prices increased to 700 cents/bushel in 2012, the short $\bar{L}$ increased dramatically and exceeded 200 cents/bushel or \$10,000 ($= 200 * 5000/100$) per contract and long $\bar{L}$ was the lowest. Meanwhile, the increase in margin requirement was less than 10 cents/bushel. Mid-2008 was the riskiest period for buying corn futures ($\bar{L} > 120$ cents/bushel), as this period coincided with a substantial drop in futures prices. The red line L_{max} illustrates the required access to capital (i.e., the highest margin liability during a hedge horizon). Figure 2.3 demonstrates that L_{max} was higher than $\bar{L}$ but followed the same patterns with maximum credit requirements reaching 371 cents/bushel on short hedges in June 2012. In 2008, the maximum margin liability on long hedges exceeded 300 cents/bushel or $15000 ($= 300 * 5000/100$) per contract.

[Figure 2.3 to be here]

Figure 2.4 shows the results for three-month soybean hedges. Again, the largest $\bar{L}$s were observed for short hedges in 2012, which exceeded 300 cents/bushel or $15,000 (=300*5000/100) per contract. Long hedgers experienced high costs of maintaining their positions in 2008 when average margin liability exceeded 200 cents/bushel at the most time. Although margin requirements fluctuated within a range of 30 cents/bushel from 2009 to 2017, the amount of funds for sustaining a position varied substantially. This observation suggests that the margin requirement is not adequate to represent the full costs of hedging. The maximum margin liability was roughly twice as large as average margin liability. Intuitively, it means the hedger's credit line had to be two times greater than the average fund invested in a position to avoid premature termination of a hedge. For example, the maximum margin liability reached 650 cents/bushel in Mid 2008, meaning the hedger's maximum credit had to be above $32,500 ($= 650 * 5000/100$) to finance the most substantial loss in a soybean contract.

[Figure 2.4 to be here]

Additional details on simulated hedging costs are provided in Table 2.1. Our sample was divided into three subsamples: 2004-2006, 2007-2013, and 2014-2018 that allowed us to take into account different market condition and compare our results to previous studies[10]. For corn futures, 2004-2006 sub-period reflected traditionally low nearby futures prices with a mean of 240 cents/bushel and low market volatility (SD=46). 2007-2013 sub-period was characterized by much higher prices that averaged 522 cents/bushel and increased market volatility (SD=149),

[10] Etienne et al. (2017) estimated the costs of forward contracting for similar sub-periods.

which were likely associated with ethanol policies and financial crisis of 2009-2010. In the last 2014-2018 sub-period, prices subsided to 376 cents/bushel, and its standard deviation reduced to 37 cents/bushel. The soybean futures market had remained at a high price level since 2011 but experienced high volatility only in 2007-2013 subperiod.

Table 2.1 also reports sub-period means of average daily margin liability ($\bar{L}$), measured in cents/bushel and percent of the average price over a hedge horizon. First, holding the same hedge horizon, there were substantial variations of $\bar{L}$ over time. Average margin liability in 2004-2006 was higher than the hedging cost of 4 cents/bushel for corn and 12 cents/bushel for soybeans reported by Riley and Anderson (2009), because their study only considered margin requirement as the hedging cost and ignored the additional funds to maintain a position. Corn and soybean hedgers experienced high average margin liabilities (in level) in 2007-2013 sub-period. We focus on the $\bar{L}$ of one-month long hedges in the following example. For corn futures, average daily margin liability increased by 165% (about 20 cents higher) in 2007-2013 compared to the previous sub-period. After 2013, $\bar{L}$ slightly dropped by 28%. Similar patterns of soybean futures are shown in the lower-left panel of Table 2.1. the sub-period mean of $\bar{L}$ (56.68 cents/bushel) was 79% higher than that in 2004-2006 (31.70 cents/bushel). The high cost of hedging corn and soybeans in 2007-2003 was mainly due to high prices, which might be caused by drought conditions in Midwestern states. When evaluating $\bar{L}$ measured in percentage of price over time, we observe a slightly different pattern for soybean futures: the most recent sub-period exhibited high average margin liability (in % of price) for long hedgers due to a decreasing trend of price movements after 2013.

Second, we find average margin liability goes up as hedge horizon extends in each sub-period. It means that more funds are needed to maintain a position with a longer horizon. Take

2007-2013 long hedges for corn as an example. The one-month $\bar{L}$ increased by 26% (or 8.5 cents/bushel) if we extended the hedge to a three-month horizon and by 51% (or 16.7 cents/bushel) if the hedge was extended to a six-month horizon.

According to equations (2.5) and (2.6), we calculated daily borrowing costs of maintaining corn and soybean positions[11] and multiplied the daily costs by seven to develop weekly borrowing costs. The results of BC and BC^+ were very similar, so we only reported BC measured in cents/bushel and percent of the average price over a hedge horizon. Etienne et al. (2017) calculated weekly costs of forward contracting[12] from February to August corresponding to December corn and November soybean contracts. The weekly cost of forward contracting in late May (week 20) was used to compare with our weekly borrowing costs of six-month long hedges because of similar hedge length. In most cases, we found borrowing costs were slightly higher than the costs of forward contracting. For corn forward contracts, the average weekly cost was 3.62 cents/bushel in 2002-2006 and 16.29 cents/bushel in 2007-2013. For corn futures, the average weekly borrowing cost of six-month corn hedges was 4.29 cents/bushel in 2004-2006 and 10.64 cents/bushel in 2007-2013. Regarding soybean futures, our calculated weekly borrowing cost was 10.35 cents/bushel in 2004-2006 and much higher than the cost of forward contracting (-0.26 cents/bushel) reported by Etienne et al. (2017). In 2007-2013 sub-period, weekly borrowing cost was 14.83 cents/bushel and was higher than weekly forward contracting cost of 11.35 cents/bushel.

[11] According to the website https://investorjunkie.com/12389/best-margin-rates/, we select 8.5% as the average annual margin cost rate with a debt balance 0-$9999. Then, we divide it by the number of days in a year to get a daily interest rate $r = 8.5\%/360$. The brokerage industry typically uses 360 days - not 365 (https://www.investopedia.com/ask/answers/07/margin_interest.asp)

[12] Their cost of forward contracting is defined as the difference between the average spot basis in October with each Thursday's forward basis.

[Table 2.1 to be here]

Table 2.2 presents means of maximum margin liability (L_{max}) in three sub-periods, as well as intervals defined by the maximum and minimum L_{max} in each period. When measured in level, L_{max} of one-month hedges increased more than 100% from 2004-2006 to 2007-2013 mainly due to an increase in the price level in corn and soybean markets. Also, maximum margin ability was higher for longer hedging horizon in each sub-period. In 2004-2006 on average, long L_{max} of three-month corn hedges was about 12 cents/bushel higher than that of one-month hedges, and the maximum value of three-month long L_{max} reached 115 cents/bushel.

Based on maximum margin liability and capital constraints (C), we calculated the probability of hedging failure as the percent of hedges with $L_{max} > C$ in a year using equation (2.8). The capital constraints were set at 20% and 30% of the average price over a hedge horizon to be comparable between commodities. Table 2.2 reports the mean capital constraints and probabilities of hedging failure. First, given a specific hedge horizon, we evaluated the probability of hedging failure over time and find 2007-2013 is the riskiest sub-period. For example, when the constraint was set as 20% of the average price, about one half of 6-month corn hedges had to be terminated before maturity in 2007-2013, while the probability of hedging failure was only around 35% in 2004-2006. Moreover, the probability of failure increased as hedge length increased. Finally, Table 2.2 suggests that hedging corn is riskier than hedging soybeans, holding the subperiod and hedging length the same.

[Table 2.2 to be here]

2.4 What Explains Changes of Hedging Costs?

A variety of measures illustrated considerable changes in hedging costs over time. However, simple descriptive statistics of historical data are limited to inform risk management practices.

This section presents a quantile regression method to detect patterns in the entire cost distribution and identify driving factors. Understanding the empirical distribution, especially the tail area, is of particular importance in effective risk management (Taylor, 1999). That is because probabilities of hedging failure are likely to be high at the tail areas, and the resulting liquidity problem impedes many small producers from using the futures market to manage price risks.

2.4.1 Unconditional quantiles

Developed initially by Koenker and Bassett (1978), quantile regression is a non-parametric approach used in estimating the quantile of a variable. Let $Q(\theta) = q$ be the θth quantile of a continuous variable y, it satisfies $Prob(y < q) = \theta$. We take average margin liability as an example and introduce unconditional sample quantile. Consider a sample of weekly average margin liability $[\bar{L}_w: w = 1, \dots, W]$, the θth sample quantile ($0 < \theta < 1$) can be obtained as the solution to

$$\min_{q \in R}[\sum_{w \in [w:\, \bar{L}_w \geq q]} \theta|\bar{L}_w - q| + \sum_{w \in [w:\, \bar{L}_w < q]} (1-\theta)|\bar{L}_w - q|] \tag{2.9}$$

One typical sample quantile is the median when $\theta = 0.5$. From a sorted sample, the unconditional sample quantile can be either identified as an order statistic or a closed interval between two adjacent order statistics (Koenker & Bassett, 1978). Next, illustrations of three-month hedges' average margin liability provide more intuitions.

Figure 2.5 shows the frequency distribution of simulated corn $\bar{L}_w$ from 2004 to 2018. The average margin liability was measured as the percentage of the opening price. $Q(.5) = 4\%$ for short hedges, meaning that in 50% of the time, the average margin liability did not exceed 4% of the opening price. However, in 10% of simulated hedges, $\bar{L}$ could be as high as 15% of the opening price ($Q(.9) = 15\%$), reaching 49% at the worst time. The distribution of long $\bar{L}$ had similar values of sample quantiles but was more concentrated on the left side. The 0.75 quantile of long $\bar{L}$ indicated there was a 25% chance that the cost exceeds 11% of the opening price. Histograms in Figure 2.6 display the distributions for three-month soybean hedges. Compared with corn cost distribution, soybean sample quantiles exhibited more symmetry between long and short sides. The average cost of hedging one bushel of soybean for three months was not more than 5% of the opening price in half of the sample. However, in 10% of the time, short $\bar{L}$ went beyond 15% of the opening price and long $\bar{L}$ exceeded 13%.

[Figure 2.5 and Figure 2.6 to be here]

2.4.2 Conditional quantiles

After providing the basic idea of quantiles, it is easy to extend the concept to conditional quantiles. Replacing q with a linear function of explanatory variables, we can obtain the θth conditional quantile by solving

$$\min_{\beta \in R}[\sum_{w \in [w:\, \bar{L}_w \geq x_w \beta]} \theta|\bar{L}_w - \boldsymbol{x_w\beta}| + \sum_{w \in [w:\, \bar{L}_w < x_w \beta]} (1-\theta)|\bar{L}_w - \boldsymbol{x_w\beta}|] \tag{2.10}$$

Then, the conditional quantile function becomes $Q(\theta|\boldsymbol{x_w})$=$\boldsymbol{x_w\beta}(\theta)$. The set of $\boldsymbol{\beta}(\theta)$ can be solved through a linear programming method, and its variance-covariance matrix is often estimated using bootstrapping method (Lien, Shrestha, & Wu, 2016). The parameter $b(\theta) \in \boldsymbol{\beta}(\theta)$ reflects the margin effect of an explanatory variable on the θth quantile of the $\bar{L}$ distribution and therefore, is allowed to vary at different part of the distribution (Bassett & Chen, 2002). Particular interest centers in the inference at the tail areas. For example, the impact of a factor may be small at low quantiles but large at high quantiles and has important liquidity implications. As the cost distribution is highly skewed (Figure 2.5 and 2.6), another advantage of quantile regressions is that the estimates are more robust against non-normal distribution and outliers (Koenker & Bassett, 1978).

The next question is: what is the main driving factor of the cost distributions? Answers to this question also depend on the measures of hedging costs. According to Riley and Anderson (2009), the benchmark margin (cost of hedging in their study) was a function of the previous day's closing prices, and mean and standard deviation of the futures returns in past 90 days. Alexander et al. (2013) showed that the margin costs were positively associated with the initial margin, interest rates of debt and net cash flow losses from margin accounts, but negatively associated with the risk-free rate of return. For those studies using indirect measures, Dahlgran and Liu (2011) suggested that the cash flow risk was affected by futures price movement, and initial and maintenance margin requirements. In Lien (2003), the decision for the premature liquidation of a hedge depended on losses in the futures position, as well as the capital constraint. Additionally, Deep (2002) added time to maturity and stated that as the hedging horizon extends, the probability of exceeding a specific capital constraint should increase.

In this study, we carefully select factors from the theoretical framework: per-bushel initial margin requirement (M_w^0), a price level change measure (DP_w), and a measure of price volatility (S_w). Given a hedge opened in week w, the dependent variable is average margin liability, measured in % of the opening price. The conditional quantile model is specified as

$$Q(\theta|\boldsymbol{x_w}) = \alpha_0(\theta) + \alpha_1(\theta) \cdot M_w^0 + \alpha_2(\theta) \cdot DP_w + \alpha_3(\theta) \cdot S_w \tag{2.11}$$

where $Q(\theta|\boldsymbol{x_w})$ is the θth conditional quantile of $\bar{L}_w$. Initial margin requirement and price level change are also measured in % of the opening price. M_w^0 determines the frequency and amount of margin calls. Holding other factors fixed, an increase in margin requirement is expected to increase average margin liability, so $\alpha_1(\theta)$ is expected to be positive. For a simulated hedge, the price level change is calculated by subtracting the opening price from the average futures price over a hedge horizon. A positive DP_w indicates that the futures price rises during a hedge period, which is favorable for long positions but not short positions. Therefore, we expect that $\alpha_2(\theta) < 0$ for long hedgers and $\alpha_2(\theta) > 0$ for short hedgers. For a hedge opened at week w, the price volatility S_w is constructed as the standard deviation of futures returns (R_t) during the whole hedge horizon, where $R_t = \ln\left(\frac{p_t}{p_{t-1}}\right) \times 100$ is the futures return at day t. The price volatility is expected to increase the conditional quantile.

Equation (2.11) can also be used to estimate maximum margin liability. Since maximum margin liability captures the largest cumulative loss in a hedge, price level change becomes the difference between opening price with maximum (for short L_{max} model) or minimum (for long L_{max} model) futures price over a hedge's horizon. Other explanatory variables are the same as specified in $\bar{L}$ models.

2.4.3 Empirical results

Empirical analyses were conducted using STATA 14.0 (StataCorp LP, 2015). Our model results revealed the sensitivity of hedging costs to three driving factors at different parts of the cost distribution (i.e., θ=0.1, 0.3, 0.5, 0.7, and 0.9). Except for price volatility, estimates of other factors were quite similar between $\bar{L}$ and L_{max} models. Therefore, we focus on average margin liability to interpret the main findings, and briefly discuss empirical results of maximum margin liability.

In Table 2.3, columns (1) and (2) show the quantile estimates for corn hedges, separated by long and short sides. The coefficients of M_w^0 were around 0.7 and statistically significant, meaning a 1% increase in margin requirements roughly raised the average margin liability by

0.7%. Alexander et al. (2013) also showed that the margin costs were positively associated with the initial margin. As we expected, estimated coefficients of price level change were symmetrically negative for long hedges and positive for short hedges. The impact was slightly smaller in the tail area. One percent higher of the average price above the opening price of a position reduced long $\bar{L}$ by 0.82% at the 0.1 quantile, and by 0.53% at the 0.9 quantile. Higher positive impacts of price volatility were observed in the upper tail of the distribution than in the lower tail, suggesting that high $\bar{L}$ tended to be more associated with price volatility than other factors. For example, at $\theta = 0.1$, one standard deviation rise of the futures returns increased average margin liability by 1%. While at $\theta = 0.9$, this led to a 4% increase in $\bar{L}$ for long positions, and 5% for short positions.

Columns (3) and (4) in Table 2.3 shows the results of equation (2.11) for soybeans. Compared with corn models, the effects of margin requirement were smaller and not significant in one model when θ=0.9. Price level change had a slightly decreasing impact on average margin liability from lower quantiles to upper quantiles, and the signs were consistent with our prior expectation. Estimates of price volatility exhibited substantial differences over the entire distribution. For instance, an additional standard deviation of soybean futures returns increased the long $\bar{L}$ by 1.35% at 0.1th quantile, and this effect was more than twice larger at the 0.9 quantile (3.5%). The substantial effect of price volatility at higher quantiles of $\bar{L}$ was also observed for long soybean hedgers. At the 0.9 quantile, one-unit rise of S_w resulted in a 4.5% rise in average margin liability.

[Table 2.3 to be here]

Next, focusing on corn three-month short hedges, we visualized quantile regression and ordinary least squares (OLS) estimates in Figure 2.7. The solid horizontal line indicates the effect of each factor on the conditional mean of $\bar{L}$, and two dashed lines indicate the 95% confidence interval of OLS estimates. The coefficients of quantile regressions were plotted at five different quantiles as the solid curve, surrounded by 95% confidence intervals. Estimates of OLS and quantile regressions differed in a few ways. First, the conditional mean was highly sensitive to price volatility (α_3=2.73) and moderately sensitive to price level change ($\alpha_2 = 0.73$), and the OLS coefficient of initial margin was only 0.56. Instead of reaching a simple conclusion, quantile regression was more informative as its estimates indicated heterogeneous effects at different parts of the cost distribution. At the left tail when $\bar{L}$ was low, the impacts of three factors were almost the same. As θ increased, the impact of price level change dropped, and price volatility became more influential than other factors. Etienne et al. (2017) stated that hedging costs mainly depend on price volatility because high price volatility "increase the opportunity cost of funds associated with maintaining a margin account." Compared with constant OLS confidence intervals, the confidence interval for quantile regression broadened towards higher quantiles because considerable uncertainty is often associated with high costs of hedging.

[Figure 2.7 to be here]

Table 2.4 reports the results of maximum margin liability models and allows for comparison with the results of average margin liability models in Table 2.3. Over the entire distribution, price level change had one-to-one effects on L_{max}, which were larger than its impacts on $\bar{L}$. We use the short side as an example to explain why. The difference between opening price with the maximum price in a hedge adequately captured the biggest cumulative loss in a short position. Our model suggested that one additional 1% higher of the maximum price over the opening price of a hedge was associated with 1% increase of L_{max}. While price volatility drove the average margin liability, this factor played a much less important role in maximum margin liability models because only extreme price movements contributed to L_{max}.

[Table 2.4 to be here]

2.4.4 Sensitivity analyses

From 9/20/2007 to 1/2/2009, historical margin requirements were not disclosed by the CME, so we had to use model-based estimates of margin requirements (see Appendix 1). To check the sensitivity of the results to this imputation, we constructed a dummy indicator ($D = 1$) for the imputed period, and estimated equation (2.11) again with the interaction $D \cdot \Delta M_w^0$. The underlying logic is: if the coefficients of this interaction term are significant, our results are sensitive to the imputation (i.e., the impact of imputed margin requirement is different from that of historical margin requirement).

As shown in Appendix BB, most estimated coefficients of $D \cdot \Delta M_w^0$ in corn models were significantly positive, indicating an underestimation of the impact of margin requirement due to imputation. Despite this, estimates of other factors changed little from the main set of results in terms of signs and magnitudes. Soybean results were insensitive to the imputation of margin requirements since most coefficients of $D \cdot \Delta M_w^0$ were not statistically significant.

2.5 Conclusion

Hedging with futures is a popular price risk management tool. However, it may incur substantial liquidity risk due to margin calls. While several previous studies have looked at various aspects of hedging costs, there is no consensus on how to measure these costs and anticipate them empirically. This study developed a novel comprehensive framework to measure costs of hedging. We started with defining a new term called margin liability to capture the cumulative loss in a margin account and created a variety of measures incorporating ideas from existing literature on direct and indirect costs. Our theoretical framework shows that the liability generated by maintaining a margin account (i.e., direct cost) is jointly determined by margin requirements and price movements over the hedge horizon. The probability of hedging failure (i.e., indirect cost) depends on biggest loss in a hedge horizon as well as the borrowing constraint.

Based on historical futures prices and margin requirements of corn and soybeans, we simulated hedging costs over time. The simulation results suggested that the costs started to increase from 2007 and then declined after 2013. This trend was consistent among different commodities, corn and soybeans. Long hedgers faced the highest liquidity risk during the periods of price declines, such as the one due to financial crisis of 2008. Short hedgers suffered the most liquidity risk during the period of rapid price appreciation such as the one during 2012 due to drought conditions. In these high-risk years, the average margin liability for corn futures exceeded $2.00/bushel, and a hedger needed as much as $15000 to sustain a three-month hedge for a single contract.

Quantile regressions provide additional insights into the driving factors and their impacts at different parts of the cost distribution. We found a large impact of price volatility on the hedging costs at higher quantiles of the cost distribution. For example, at $\theta = 0.9$, one standard deviation increase in the price volatility led to a 5% increase in the short average margin liability. Price level changes also affected the average margin liability, but its effect declined as θ increased. Finally, our findings revealed that the margin requirement was insufficient to capture the cost of maintaining a margin account, but it is still an important driving factor.

Our findings may also be used to help understand the costs of forward contracting. The companies that extend forward contracts to producers have to hedge these positions in the futures markets and incur hedging costs. They try to pass on these costs of hedging within the forward contract prices. However, our findings demonstrate that these costs may change dramatically from one year to the next and are difficult to anticipate. That is why we typically see an increase in the costs of forward contracting following a year of high hedging costs.

In the presence of liquidity risk associated with hedging, it is very important to educate financial institutions that loan funds to finance futures positions. While this study demonstrates that the costs of hedging may be very high, they tend to be temporary. If the hedge is placed correctly, the gains in the cash market tend to offset the losses in the futures market. Therefore, it is important that the hedging loans are set up separately from other operational loans and take into account the size of the cash market position of a business. A better understanding of the costs of hedging by the financial institutions may improve access to credit for hedgers and decrease the risk of hedging failure due to credit constraints.

Appendix BA. Predict margin requirements when historical values are not available

Margin requirements are implemented by the exchange to reduce the risk of default on futures contracts and consist of an initial margin (funds required to open a position) and a maintenance margin (a balance required to sustain a position). Since the CME data are missing from 9/20/2007 to 1/2/2009, we need to approximate margin requirements in this period based on available information.

Although the exchanges usually do not disclose how margin requirements are determined, according to Lam, Yu, and Lee (2010), the margin-setting committee calculates some formula-based benchmark margins as the reference level. These benchmark margins are derived as the minimum level to cover a certain probability of loss in the futures market, which is usually set at 95% or 98% by some clearinghouses (Lam, Sin, & Leung, 2004). Lam et al. (2004) provided a formula for calculating a benchmark margin (BM) at day t:

$$BM_t = p_{t-1}|\lambda_t + k\sigma_t| \tag{B1}$$

$$\lambda_t = \frac{1}{Z}\sum_{z=1}^{Z} R_{t-z}, \quad \sigma_t^2 = \frac{1}{Z-1}\sum_{z=1}^{Z}(R_{t-z} - \lambda_t)^{2}, R_t = \ln\left(\frac{p_t}{p_{t-1}}\right) \times 100$$

where p_{t-1} is the previous day's nearby futures price, λ_t and σ_t are historical mean and standard deviation of nearby futures returns over past Z days. According to Riley and Anderson (2009), the CME commonly sets $Z = 90$ and $k = 1.96$. Equation (B1) suggests that the margin requirements are affected by price level as well as the mean and standard deviation of futures returns over a relevant historical period, (e.g., the last 90 days).

Following the theoretical framework, we begin with examining margin requirements for corn and soybean futures. Figure B1 compares observed initial margins published by the CME to the reference margin calculated using equation (B1), measured in cents/bushel. The figure shows that benchmark margins of both corn and soybean futures dramatically exceed actual margin requirements in 2009, 2011 and 2013. This phenomenon suggests that the exchange was careful not to raise margins too much above its usual level, even the benchmark margin responses to the shock. After 2013, benchmark margins fell below the actual margin requirements reflecting that the clearinghouse was reluctant to reduce margin levels too fast after the shock.

This observation also suggests that the effect of past futures returns on futures margin levels may be asymmetric. In the example above, the clearinghouse appears more likely to adjust margin requirements upwards rather than downwards. Although benchmark margins are not very close to the actual margins, we can try to predict the actual initial margin requirement (M^0) based on previous day's nearby futures price p_{t-1}, historical mean return λ_t and standard deviation σ_t over the past 90 days. Additionally, two indicators $D_{\lambda<0}$ and $D_{\Delta\sigma>0}$ are added to capture the asymmetric effects of the mean and standard deviation of futures returns:

$$D_{\lambda<0} = \begin{cases} 1 & if\ \lambda_t < 0 \\ 0 & otherwise \end{cases} \text{ and } D_{\Delta\sigma>0} = \begin{cases} 1 & if\ \Delta\sigma_t = \sigma_t - \sigma_{t-90} > 0 \\ 0 & otherwise \end{cases}.$$

Margin requirement model is specified as:

$$M_t^0 = \alpha_0 + \alpha_1 p_{t-1} + \alpha_2 \lambda_t + \alpha_3 \lambda_t \cdot D_{\lambda<0} + a_4 \sigma_t + a_5 \sigma_t \cdot D_{\Delta\sigma>0} + \varepsilon_t, \quad \text{(B2)}$$

where $\varepsilon_t = \rho\varepsilon_{t-1} + v_t$, $v_t \sim_{iid} \mathrm{N}(0, \delta^2)$

The fitted models are[13]

Corn: $\widehat{M}_t^0 = \underset{(0.00)}{12.5} + \underset{(0.00)}{0.005} p_{t-1} - \underset{(0.00)}{1.08} \lambda_t + \underset{(0.59)}{0.29} \lambda_t \cdot D_{\lambda<0} + \underset{(0.00)}{2.12} \sigma_t - \underset{(0.20)}{0.05} \sigma_t \cdot D_{\Delta\sigma>0}$

Soybean: $\widehat{M}_t^0 = \underset{(0.00)}{28.01} + \underset{(0.03)}{0.002} p_{t-1} - \underset{(0.00)}{5.00} \lambda_t + \underset{(0.00)}{6.98} \lambda_t \cdot D_{\lambda<0} + \underset{(0.00)}{7.69} \sigma_t - \underset{(0.00)}{0.50} \sigma_t \cdot D_{\Delta\sigma>0}$

Results suggest that the margin requirements are positively associated with price level and volatility. For example, one unit increase in the standard deviation of futures returns over the

[13] P-values in parentheses

past 90 days increase the soybean margin requirement by 7.69 cents/bushel when $\Delta\sigma < 0$. When the volatility is increasing over time ($\Delta\sigma > 0$), one additional standard deviation of past futures returns increases the margin requirement by 7.19 cents/bushel. The margin requirement reduces with the mean of futures returns in the past 90 days. It implies that higher future returns reduce the margin requirement when the historical mean return is positive. Previous day's price level only has minimal effect on the margin requirements. As plotted in Figure B1, the fitted margin requirements follow the movements of actual margin requirements of corn and soybean futures very well. One-sample t-test suggests that the in-sample prediction errors are unbiased, so predicted margin requirements are used in the simulation when the actual margin requirements are not available

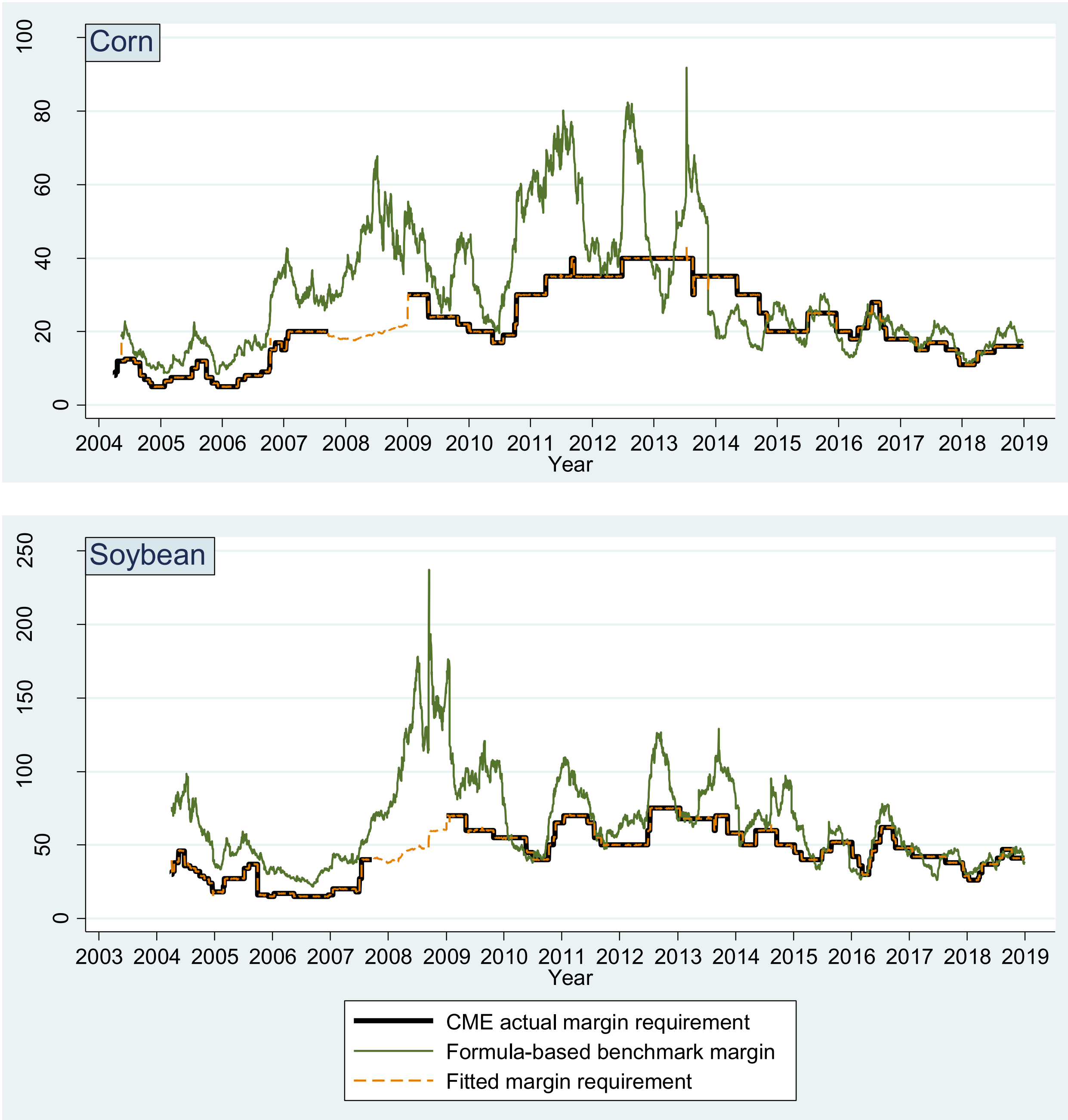

Figure B1 Historical margin requirements from CME, benchmark margins, and fitted margin requirements, 2004-2018

Appendix BB. Sensitivity analyses results

Table BB-1 Sensitivity analyses results of average margin liability for three-month hedges

		Corn		Soybeans	
Parameter		Long	Short	Long	Short
M_w^0	Q(.1)	0.668***	0.663***	0.556***	0.689***
	Q(.3)	0.713***	0.686***	0.510***	0.652***
	Q(.5)	0.755***	0.724***	0.429***	0.589***
	Q(.7)	0.854***	0.605***	0.283***	0.472***
	Q(.9)	1.011***	0.652***	0.109	0.303**
$D \cdot M_w^0$ [a]	Q(.1)	0.169**	0.0962*	-0.030	-0.041
	Q(.3)	0.216**	0.109**	0.000	-0.007
	Q(.5)	0.496***	0.161	0.012	0.068
	Q(.7)	0.432***	0.364**	0.041	0.169**
	Q(.9)	-0.063	0.189	0.097	0.103
DP_w [b]	Q(.1)	-0.807***	0.847***	-0.740***	0.901***
	Q(.3)	-0.816***	0.822***	-0.742***	0.880***
	Q(.5)	-0.749***	0.793***	-0.688***	0.843***
	Q(.7)	-0.702***	0.743***	-0.625***	0.777***
	Q(.9)	-0.533***	0.723***	-0.588***	0.734***
S_w	Q(.1)	0.886***	0.992***	1.402***	1.127***
	Q(.3)	1.064***	1.313***	1.746***	1.510***
	Q(.5)	1.340***	1.714***	2.148***	1.960***
	Q(.7)	1.931***	2.615***	2.797***	2.926***
	Q(.9)	4.097***	4.721***	3.478***	4.370***

Note: Asterisks indicate statistical significance: * p < .1, ** p < .05, *** p < .01

[a] D=1 for hedges with an opening day between 9/20/2007 to 1/2/2009, and D=0 otherwise

[b] For a hedge opened in week w, the price level change is defined as $DP_w = average\ price - opening\ price$.

Table BB-2 Sensitivity analyses results of maximum margin liability for three-month hedges

		Corn		Soybeans	
Parameter		Long	Short	Long	Short
M_w^0	Q(.1)	0.838***	0.601***	0.740***	0.739***
	Q(.3)	0.887***	0.890***	0.905***	0.882***
	Q(.5)	1.000***	0.929***	1.000***	1.000***
	Q(.7)	0.976***	0.864***	0.954***	0.900***
	Q(.9)	0.873***	0.871***	0.745***	0.801***
$D \cdot M_w^0$ [a]	Q(.1)	0.033	-0.213***	0.0554***	-0.042
	Q(.3)	-0.0291***	-0.0638**	0.0274**	-0.0379**
	Q(.5)	0.0177**	-0.0862***	0.012	0.035
	Q(.7)	0.005	-0.126	-0.009	0.093
	Q(.9)	-0.055	0.252***	-0.091	0.0555***
DP_w [b]	Q(.1)	-0.968***	1.024***	-0.982***	1.007***
	Q(.3)	-0.985***	1.008***	-0.999***	1.000***
	Q(.5)	-1.000***	1.016***	-1.000***	1.000***
	Q(.7)	-1.001***	1.038***	-1.009***	1.027***
	Q(.9)	-0.998***	1.059***	-0.995***	1.045***
S_w	Q(.1)	0.674***	0.448***	0.314***	0.250***
	Q(.3)	0.473***	0.035	0.151***	0.176**
	Q(.5)	0.000	0.0698***	0.000	0.000
	Q(.7)	0.036	0.120**	0.209***	0.159**
	Q(.9)	0.093	0.045	0.864***	0.326***

Note: Asterisks indicate statistical significance: * p < .1, ** p < .05, *** p < .01

[a] D=1 for hedges with an opening day between 9/20/2007 to 1/2/2009, and D=0 otherwise

[b] For a long hedge opened in week w, the price level change is defined as $DP_w = minimum\ price - opening\ price$. For a short hedge, $DP_w = maximum\ price - opening\ price$.

References

Alexander, C., Prokopczuk, M., & Sumawong, A. (2013). The (de) merits of minimum-variance hedging: Application to the crack spread. *Energy Economics, 36*, 698-707.

Arias, J., Brorsen, B. W., & Harri, A. (2000). Optimal hedging under nonlinear borrowing cost, progressive tax rates, and liquidity constraints. *Journal of Futures Markets, 20*(4), 375-396.

Bassett, G. W., & Chen, H.-L. (2002). Portfolio style: Return-based attribution using quantile regression *Economic applications of quantile regression* (pp. 293-305): Springer.

Carter, C. A., & Janzen, J. P. (2009). The 2008 cotton price spike and extraordinary hedging costs. *New York Times, 1600*, 70.

Dahlgran, R., & Liu, J. (2011). *Hedging and cash flow risk in ethanol refining.* Paper presented at the Proceedings of the NCCC-134's 2011 Annual Meeting. St. Louis, MO. April.

Deep, A. (2002). Optimal dynamic hedging using futures under a borrowing constraint. *BIS Working Paper, No. 109*.

Etienne, X. L., Mallory, M. L., & Irwin, S. H. (2017). Estimating the cost of pre-harvest forward contracting corn and soybeans in Illinois before and after 2007. *Agribusiness, 33*(3), 358-377.

FCSA. (2017). *Grain Marketing Survey: Producer Practices and Attitudes.* Retrieved from Farm Credit Services of America: https://www.fcsamerica.com/grain-marketing-survey

Garcia, P., Irwin, S. H., & Smith, A. (2015). Futures market failure? *American Journal of Agricultural Economics, 97*(1), 40-64.

Hardouvelis, G. A., & Kim, D. (1995). Margin requirements, price fluctuations, and market participation in metal futures. *Journal of Money, Credit and Banking, 27*(3), 659-671.

Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica: journal of the Econometric Society*, 33-50.

Lam, K., Sin, C. Y., & Leung, R. (2004). A theoretical framework to evaluate different margin-setting methodologies. *Journal of Futures Markets, 24*(2), 117-145.

Lam, K., Yu, P., & Lee, P. H. (2010). A margin scheme that advises on when to change required margin. *European Journal of Operational Research, 207*(1), 524-530.

Lien, D. (2003). The effect of liquidity constraints on futures hedging. *Journal of Futures Markets, 23*(6), 603-613.

Lien, D., Shrestha, K., & Wu, J. (2016). Quantile estimation of optimal hedge ratio. *Journal of Futures Markets, 36*(2), 194-214.

Mello, A. S., & Parsons, J. E. (1995). Maturity structure of a hedge matters: lessons from the Metallgesellschaft debacle. *Journal of Applied Corporate Finance, 8*(1), 106-121.

Riley, J. M., & Anderson, J. D. (2009). *Comparison of Hedging Cost with Other Variable Input Costs.* Paper presented at the St. Louis, MO, NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management.

StataCorp LP. (2015). Stata data analysis and statistical Software 14.0.

Taylor, J. W. (1999). A quantile regression approach to estimating the distribution of multiperiod returns. *The Journal of Derivatives, 7*(1), 64-78.

Figures and Tables

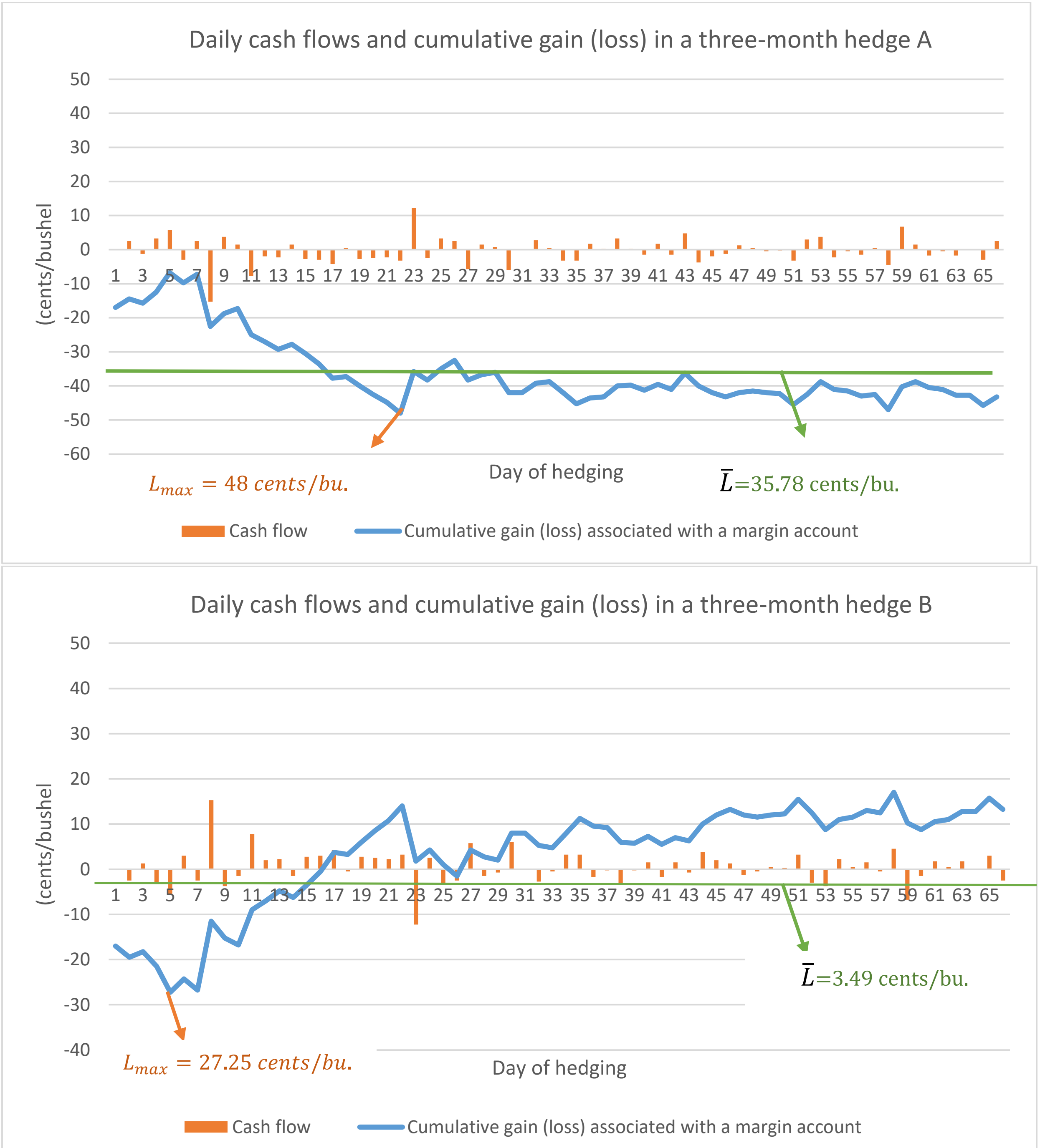

Figure 2.1 Daily cash flows, cumulative gains (losses), and margin liabilities in two hedges as an illustration.

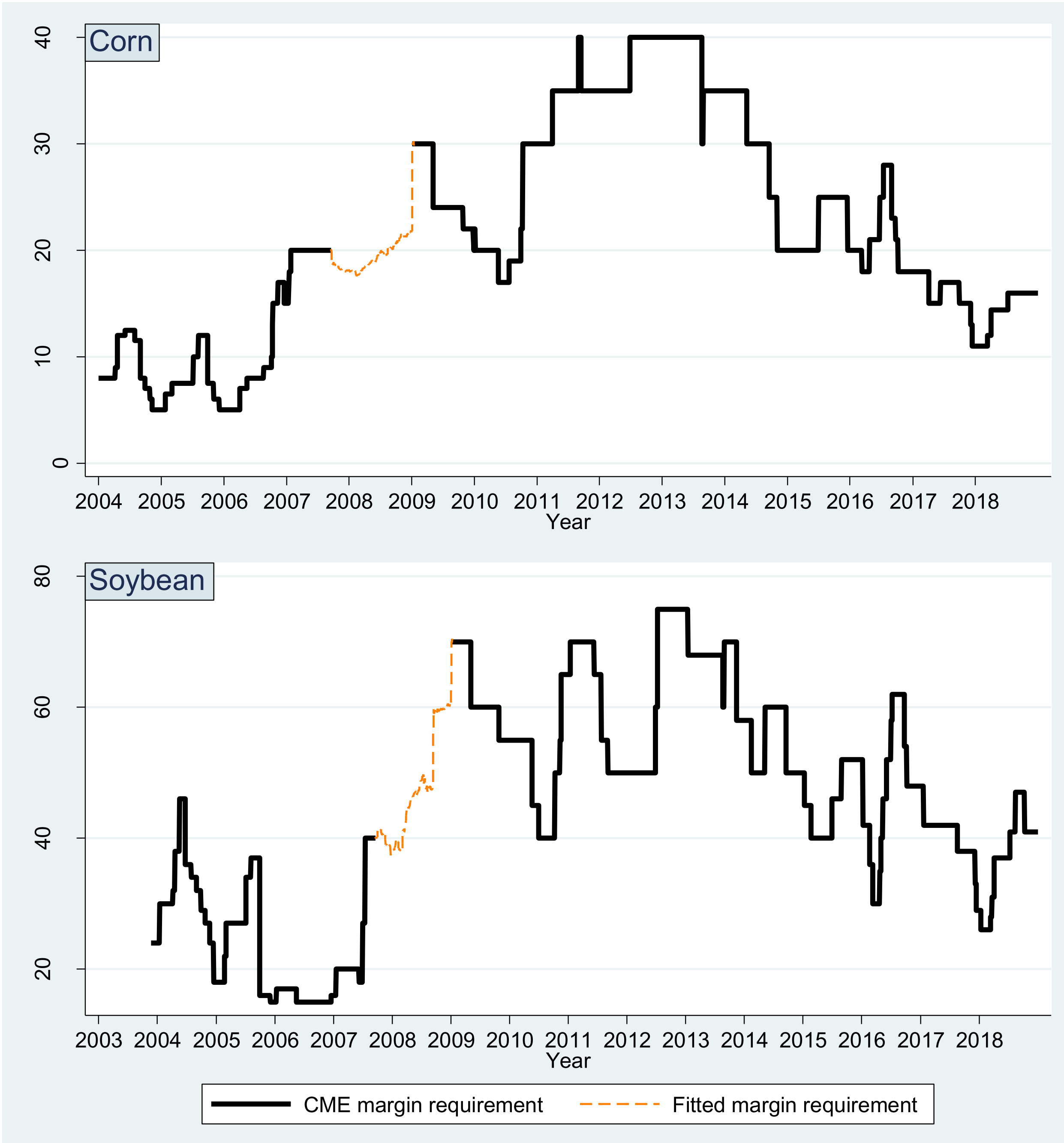

Figure 2.2 Historical initial margin requirements for corn and soybean futures, 2004-2018

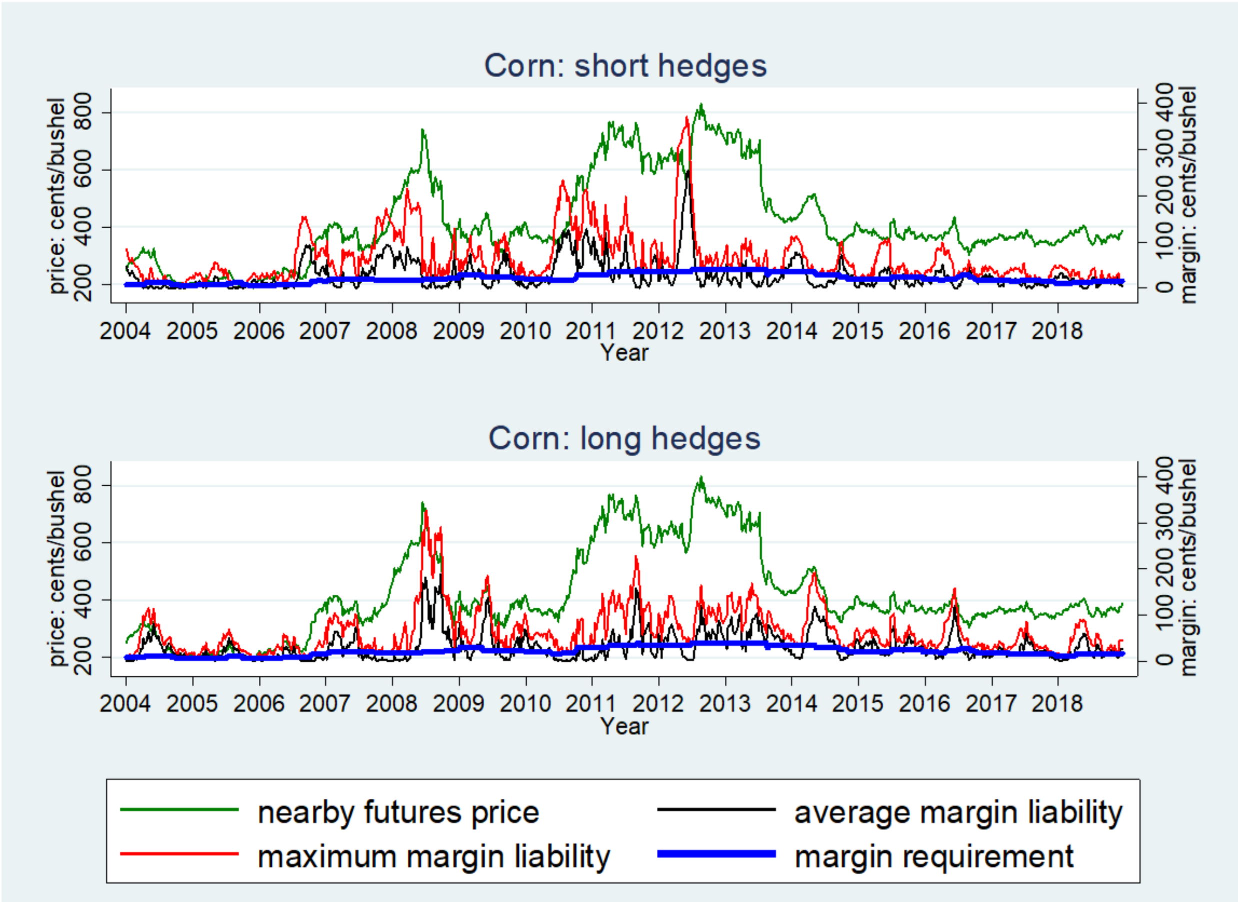

Figure 2.3 Prices of nearby corn futures, margin requirements, and simulated margin liabilities for three-month hedges, 2004-2018

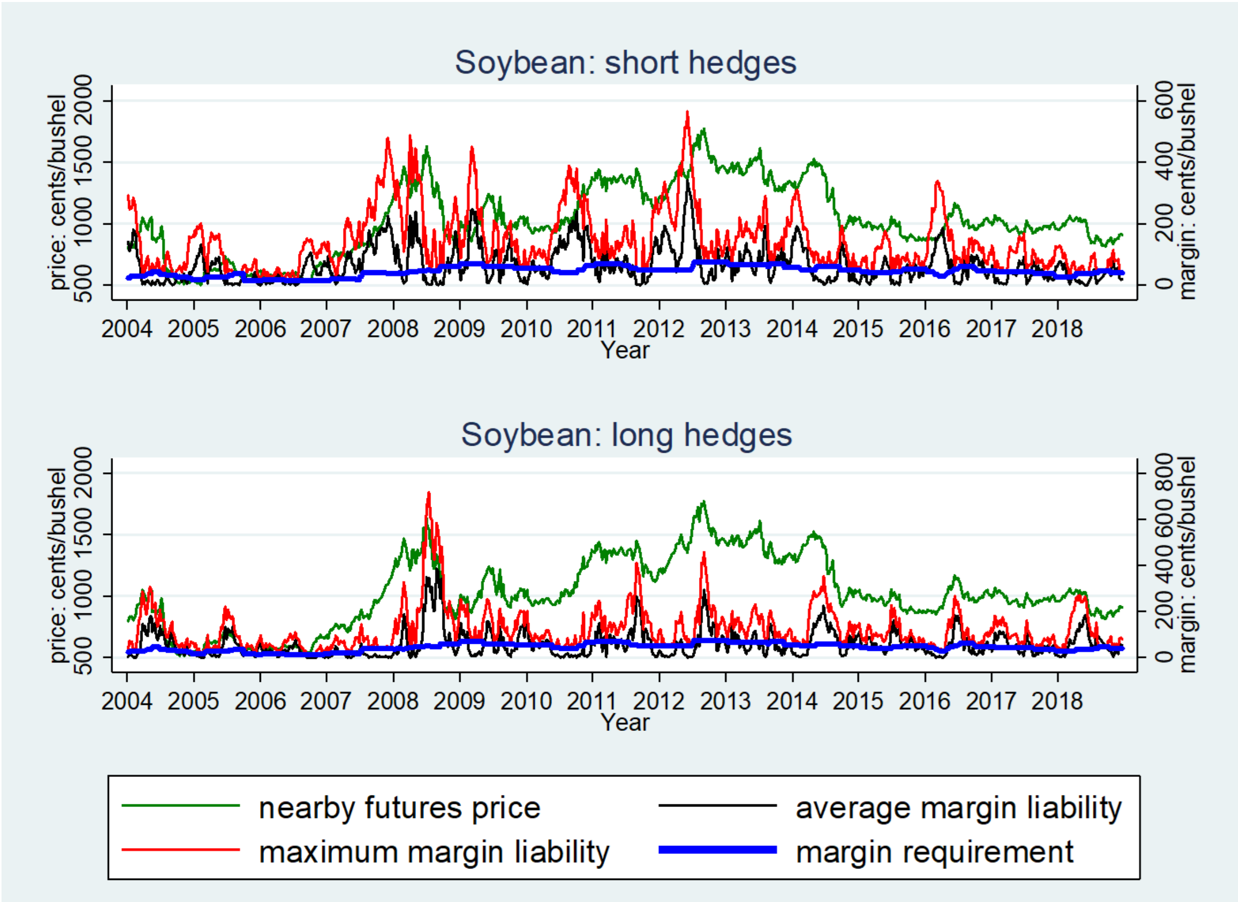

Figure 2.4 Prices of nearby soybean futures, margin requirements, and simulated margin liabilities for three-month hedges, 2004-2018

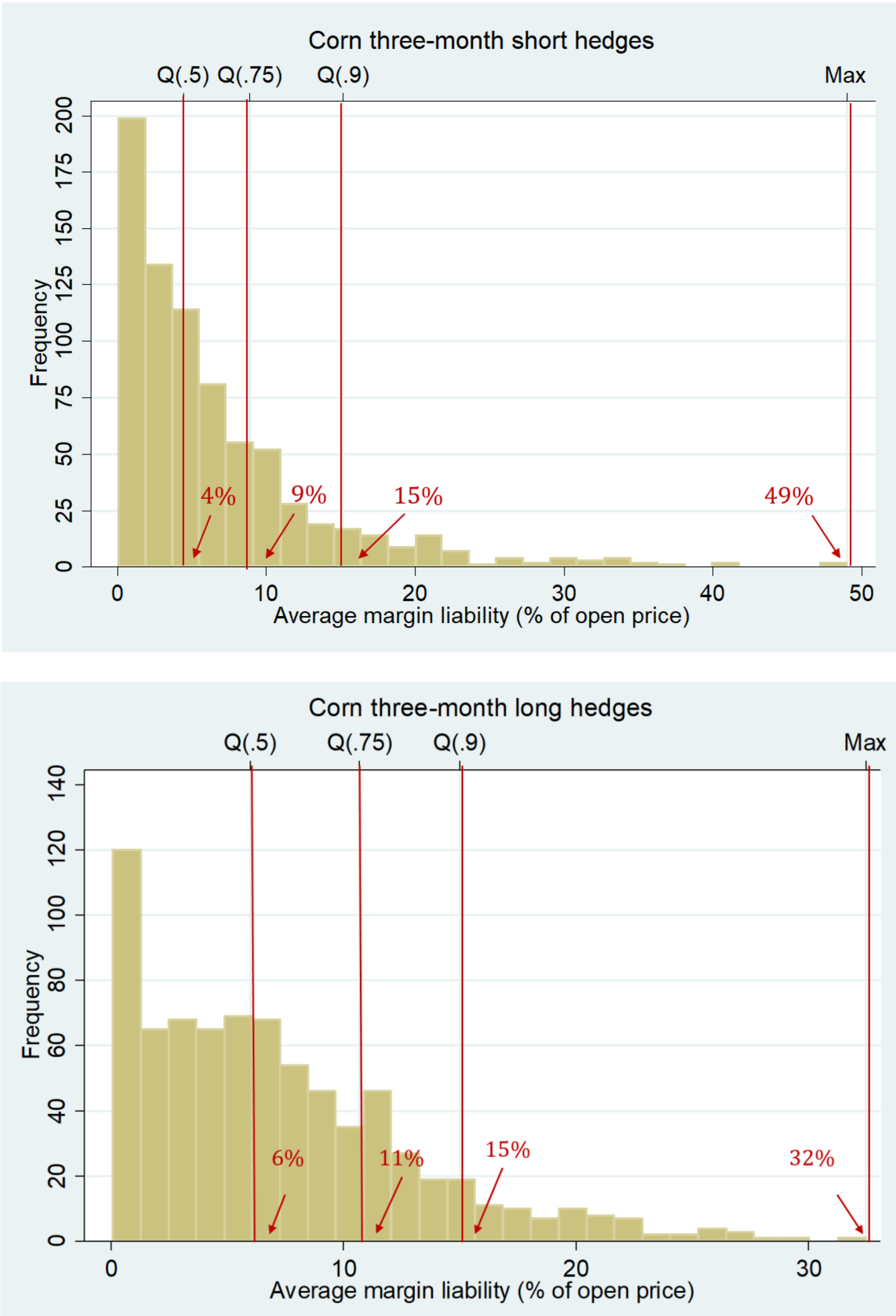

Figure 2.5 Histogram of average margin liability for three-month corn hedges with sample quantiles (θ=0.5, 0.75 and 0.9)

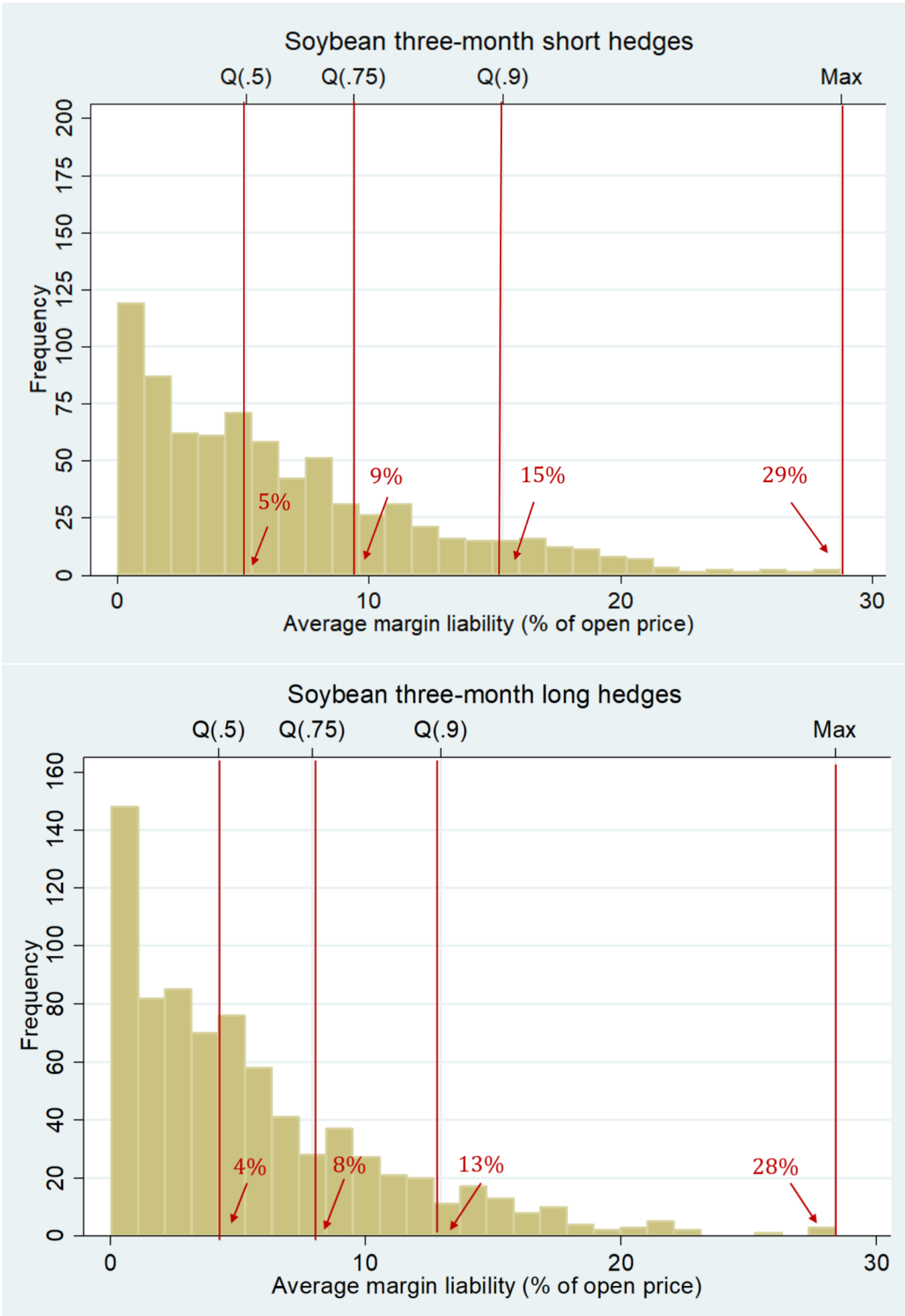

Figure 2.6 Histogram of average margin liability for three-month soybean hedges with sample quantiles (θ=0.5, 0.75 and 0.9)

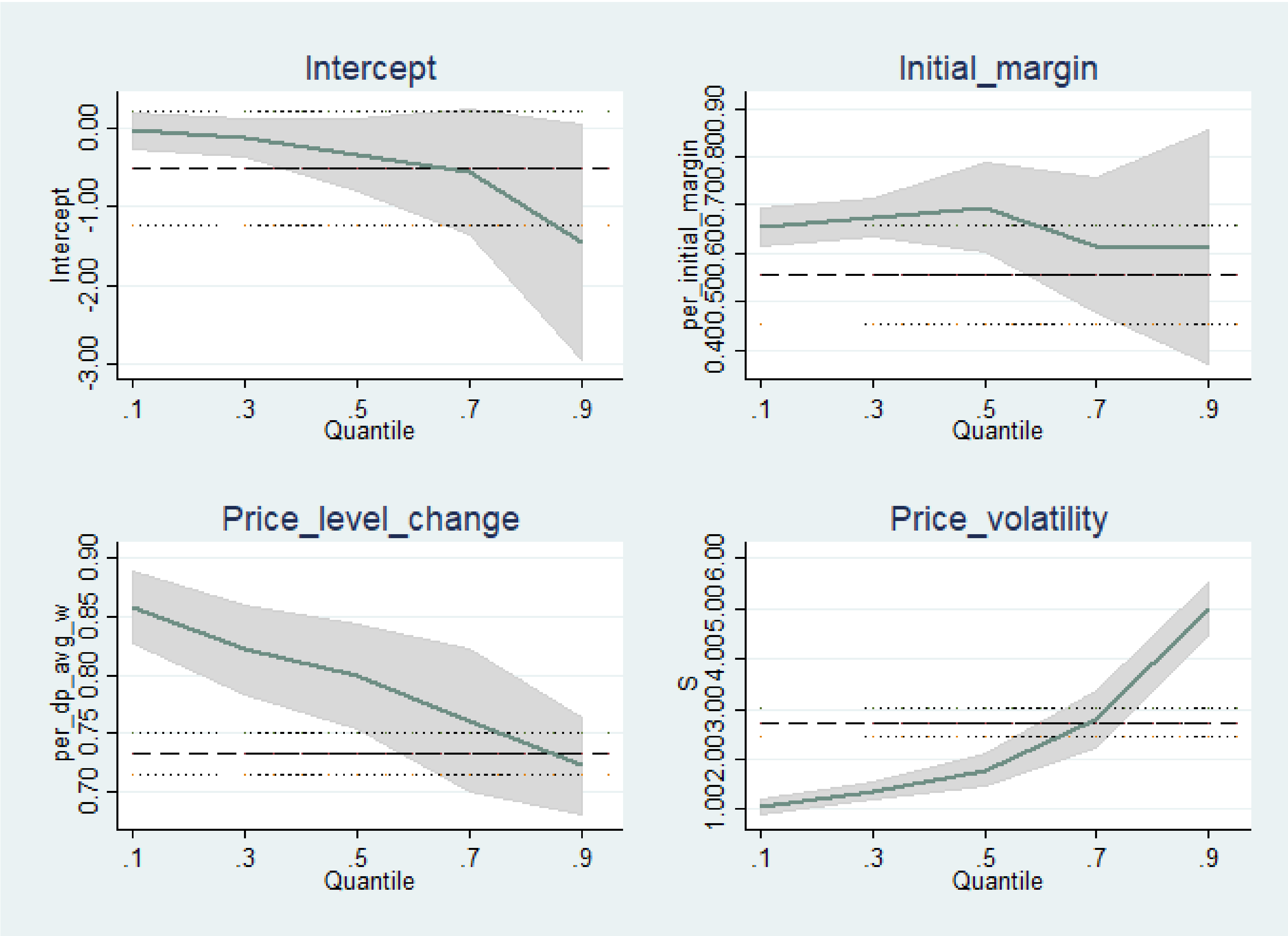

Figure 2.7 OLS and quantile regression estimates for corn short $\bar{L}$ model

Table 2.1 Simulated average daily margin liability and borrowing costs assuming no interest gain on excess margin.

Year period		Nearby futures price		Average daily margin liability ($\bar{L}$)[a]				Weekly borrowing cost ($\boldsymbol{BC}$)[a,d]			
Unit		**cents/bu.**		**cents/bu.**		**% of $\bar{p}$**[c]		**cents/bu.**		**% of $\bar{p}$**[c]	
Position	N[b]	Mean	SD	Long	Short	Long	Short	Long	Short	Long	Short
Corn Futures											
One-month hedge											
2004-2006	155	240	46	11.27	10.07	4.68	3.84	0.41	0.37	0.17	0.14
2007-2013	359	522	149	32.79	32.40	6.44	6.25	1.19	1.18	0.23	0.23
2014-2018	258	376	37	23.30	19.61	6.15	5.13	0.85	0.71	0.22	0.19
3-month hedge											
2004-2006	155	240	46	15.49	15.10	6.51	5.18	1.69	1.65	0.71	0.57
2007-2013	359	522	149	41.32	40.16	8.29	7.52	4.51	4.38	0.90	0.82
2014-2018	254	376	37	28.92	18.82	7.62	4.85	3.15	2.05	0.83	0.53
6-month hedge											
2004-2006	155	240	46	19.96	20.08	8.28	6.48	4.29	4.31	1.78	1.39
2007-2013	359	522	149	49.53	52.37	10.12	9.45	10.64	11.25	2.17	2.03
2014-2018	231	376	37	36.71	16.51	9.58	4.28	7.89	3.55	2.06	0.92
Soybean Futures											
One-month hedge											
2004-2006	155	657	136	31.70	26.81	4.93	4.05	1.15	0.97	0.18	0.15
2007-2013	361	1191	247	56.68	67.77	4.83	5.72	2.06	2.46	0.18	0.21
2014-2018	258	1018	157	49.67	45.39	5.01	4.49	1.81	1.65	0.18	0.16
3-month hedge											
2004-2006	155	657	136	40.79	35.85	6.50	5.26	4.45	3.91	0.71	0.57
2007-2013	361	1191	247	64.48	89.13	5.51	7.48	7.03	9.72	0.60	0.82
2014-2018	250	1018	157	60.98	48.82	6.20	4.74	6.65	5.33	0.68	0.52
6-month hedge											
2004-2006	155	657	136	47.42	40.52	7.76	6.17	10.35	8.84	1.69	1.35
2007-2013	361	1191	247	67.99	117.25	5.90	9.76	14.83	25.58	1.29	2.13
2014-2018	227	1018	157	71.14	48.21	7.19	4.80	15.52	10.52	1.57	1.05

Note: [a] All numbers reported are sub-period means
[b] N denotes the number of hedges simulated for each sub-period.
[c] Percent of $\bar{p}$ is calculated by dividing $\bar{L}$ over the average futures price ($\bar{p}$) during a hedge period.
[d] The daily interest rate is assumed to be 0.0236%. Weekly borrowing costs are calculated as the sum of daily borrowing costs in seven days.

Table 2.2 Simulated maximum margin liability and probabilities of hedging failure due to different liability constraints.

	Maximum margin liability (L_{max})				**Probabilities of hedging failure[b]**					
Mean [Min, Max]	**cents/bu.**		**% of $\bar{p}$[a]**		C=0.2$\bar{p}$ (cents/bu.)			C=0.3$\bar{p}$ (cents/bu.)		
Year period	Long	Short	Long	Short	C[c]	Long	Short	C[c]	Long	Short
Corn Futures										
One-month hedge										
2004-2006	21	22	9	8	49	3.9%	3.9%	74	0	1.3%
	[5, 70]	[5, 97]	[2, 28]	[3, 32]						
2007-2013	62	59	12	12	104	11.7%	7.0%	157	3.3%	1.4%
	[15, 216]	[20, 280]	[3, 43]	[3, 42]						
2014-2018	39	35	10	9	76	4.7%	1.9%	113	0.4%	0
	[12, 125]	[15, 106]	[3, 33]	[4, 26]						
3-month hedge										
2004-2006	33	38	13	14	51	21.9%	16.1%	76	9.0%	7.7%
	[6, 115]	[6, 155]	[3, 46]	[3, 52]						
2007-2013	89	92	18	18	105	30.9%	31.5%	157	9.5%	11.7%
	[17, 326]	[20, 371]	[3, 71]	[4, 53]						
2014-2018	56	46	15	12	76	17.8%	11.2%	114	7.4%	0
	[12, 191]	[15, 113]	[3, 45]	[4, 27]						
6-month hedge										
2004-2006	44	56	18	20	52	32.3%	37.4%	78	12.9%	17.4%
	[6, 137]	[8, 207]	[2, 57]	[4, 59]						
2007-2013	110	132	22	25	105	46.2%	50.4%	157	20.6%	32.9%
	[17, 502]	[20, 364]	[3, 95]	[4, 59]						
2014-2018	68	48	19	14	76	29.8%	17.1%	115	10.5%	0
	[14, 217]	[14, 109]	[4, 55]	[5, 27]						
Soybean futures										
One-month hedge										
2004-2006	58	57	9	9	129	5.8%	2.6%	193	0	0
	[15, 262]	[15, 183]	[2, 30]	[2, 28]						
2007-2013	109	121	9	10	237	5.5%	4.2%	356	0.8%	0.3%
	[16, 399]	[20, 368]	[2, 39]	[3, 31]						
2014-2018	82	80	8	8	202	1.6%	0.0%	303	0	0
	[26, 236]	[26, 191]	[2, 22]	[3, 19]						
3-month hedge										
2004-2006	82	96	13	14	128	21.3%	23.2%	192	7.7%	4.5%
	[15, 307]	[15, 343]	[2, 44]	[3, 39]						
2007-2013	148	193	13	16	238	15.0%	31.0%	357	5.0%	8.0%
	[16, 718]	[34, 566]	[2, 59]	[4, 43]						
2014-2018	114	105	12	11	201	14.0%	6.6%	301	0.4%	1.6%
	[26, 352]	[26, 340]	[3, 33]	[3, 34]						
6-month hedge										
2004-2006	100	124	16	19	126	27.7%	40.0%	189	14.2%	14.8%
	[15, 299]	[17, 374]	[2, 48]	[3, 43]						
2007-2013	171	278	15	23	239	19.9%	53.7%	359	7.5%	31.0%
	[17, 901]	[47, 685]	[2, 81]	[4, 60]						
2014-2018	136	121	15	13	199	18.2%	10.1%	298	6.2%	6.6%
	[30, 388]	[37, 354]	[3, 36]	[4, 36]						

Note: [a] Percent of $\bar{p}$ is calculated by dividing L_{max} over the average futures price ($\bar{p}$) during a hedge period.

[b] Probability of hedging failure is calculated as % of weeks when L_{max} exceeds a capital constraint.

[c] The capital constraint C is calculated as 20% or 30% of the average futures price ($\bar{p}$) during a hedge horizon.

Table 2.3 Estimation results of average margin liability for three-month hedges

Parameter		Corn		Soybeans	
		Long	Short	Long	Short
M_w^0	Q(.1)	0.667***	0.656***	0.547***	0.692***
	Q(.3)	0.704***	0.674***	0.510***	0.648***
	Q(.5)	0.729***	0.695***	0.434***	0.598***
	Q(.7)	0.778***	0.617***	0.295***	0.501***
	Q(.9)	0.943***	0.613***	0.106	0.308**
DP_w [a]	Q(.1)	-0.816***	0.858***	-0.735***	0.900***
	Q(.3)	-0.817***	0.822***	-0.742***	0.876***
	Q(.5)	-0.757***	0.799***	-0.688***	0.847***
	Q(.7)	-0.689***	0.761***	-0.626***	0.774***
	Q(.9)	-0.530***	0.723***	-0.591***	0.733***
S_w	Q(.1)	1.081***	1.069***	1.349***	1.103***
	Q(.3)	1.201***	1.370***	1.746***	1.487***
	Q(.5)	1.660***	1.778***	2.150***	1.959***
	Q(.7)	2.339***	2.789***	2.871***	3.072***
	Q(.9)	3.960***	4.992***	3.553***	4.465***

Note: Asterisks indicate statistical significance: * p < .1, ** p < .05, *** p < .01

[a] For a hedge opened in week w, the price level change is defined as $DP_w = average\ price - opening\ price$.

Table 2.4 Estimation results of maximum margin liability for three-month hedges

Parameter		Corn		Soybeans	
		Long	Short	Long	Short
M_w^0	Q(.1)	0.826***	0.646***	0.733***	0.750***
	Q(.3)	0.893***	0.971***	0.902***	0.878***
	Q(.5)	1.000***	0.953***	0.996***	1.000***
	Q(.7)	0.973***	0.876***	0.957***	0.902***
	Q(.9)	0.872***	0.869***	0.731***	0.825***
DP_w [a]	Q(.1)	-0.971***	1.013***	-0.985***	1.006***
	Q(.3)	-0.984***	1.003***	-1.000***	0.998***
	Q(.5)	-1.000***	1.014***	-1.000***	1.000***
	Q(.7)	-1.001***	1.041***	-1.009***	1.028***
	Q(.9)	-0.999***	1.057***	-0.997***	1.045***
S_w	Q(.1)	0.686***	0.154***	0.424***	0.210***
	Q(.3)	0.438***	-0.001	0.189***	0.151**
	Q(.5)	0.000	0.000	0.016	0.000
	Q(.7)	0.0420***	0.029	0.193***	0.161**
	Q(.9)	0.063	0.153	0.793***	0.524***

Note: Asterisks indicate statistical significance: * p < .1, ** p < .05, *** p < .01

[a] For a long hedge opened in week w, the price level change is defined as $DP_w = minimum\ price - opening\ price$. For a short hedge, $DP_w = maximum\ price - opening\ price$.

Chapter 3: Investigation of Service Distortion in China's New Cooperative Medical Scheme

3.1 Introduction

For decades, China has been tackling a significant challenge to provide affordable and accessible healthcare services to all the residents, especially those living in less-developed rural areas. From the 1950s to mid-1970s, most rural villages were covered by a nationwide health insurance program, called the *Cooperative Medical System* (CMS). As an integrated part of China's collective system, the CMS was a community risk-sharing program that was primarily financed by the welfare fund of agricultural communes[14] and pre-payments from individual members (N. Zhu, Ling, Shen, Lane, & Hu, 1989). The pre-payments were determined based on the commune's medical expenses in the previous year. Financial deficits, if occurred, would be covered by provincial and local governments (Wong & Chiu, 1997). This program paid primary health care and prescription drugs for rural residents and played an essential role in improving rural health (Sidel, 1993). However, when the central government transferred agricultural communes to household production units after 1978, the CMS program could not sustain itself financially. As a result, over 90% of the rural population lost their health insurance, except for those living in wealthiest rural areas (Yip & Hsiao, 2009). Additionally, the central government reduced its subsidy on total medical expenditure from 32% in 1978 to 15% in 1999 (CHEI, 2009), and many public health facilities became profit-driven and relied on the profits of drugs and services as a primary source of income (Barber & Yao, 2011; Wagstaff, Lindelow, Jun,

[14] Agricultural communes were one type of collective farms in which farmers jointly engaged in farming activities. One commune often consisted of 4000 to 5000 households, and the communes controlled and managed labor and resources in rural China, such as land and food.

Ling, & Juncheng, 2009). Decreasing subsidies led to rising healthcare costs and unaffordable healthcare for many rural uninsured households. Given the sizable rural population (around 70% of the total population) in China from the 1990s to 2000s, lack of health insurance resulted in significant medical impoverishments.

In response to this threat, the central government attempted to re-establish the original form of CMS in March 1994 for 14 pilot counties (Carrin et al., 1999), but most of them failed due to lack of funding and conflicting political interests of local governments (Y. Liu, 2004). In 2003, a redesigned health insurance scheme, called the *New Cooperative Medical Scheme* (NCMS), was launched in 257 pilot counties across 29 provinces, and it grew to about 40% rural counties in China in 2006 (CHSI, 2006). By 2008, the NCMS had expanded to all counties nationwide. The average province-level participation rates increased from 74.0% in 2003 to 98.8% in 2015, covering 670 million rural residents (NHFPC, 2015).

However, the NCMS effectiveness came under question mainly due to the evidence that NCMS participants did not obtain sufficient financial protections as compared to their urban counterpart (P. Liu, Guo, Liu, Hua, & Xiong, 2018; K. Zhu, Zhang, Yuan, Zhang, & Zhang, 2017). The literature has shown that NCMS participants had higher out-of-pocket health spending when compared to similar non-participants (Wagstaff et al., 2009) and NCMS compensation levels were much lower than the *Urban Resident Basic Medical Insurance* (URBMI) (NCMS' urban counterpart) (M. Su, Si, & Zhou, 2017). Therefore, the State Council of China initiated integration of the NCMS with the URBMI in January 2016 in order to improve the equity, sustainability, and efficiency of both programs (D. Su et al., 2019). The integration required pooling the insurance fund, unifying benefit coverages, and rural and urban area accessibility (NHFPC, 2016). By 2018, this integrated program had been established in 80%

rural areas and named as the *Urban and Rural Resident Medical Insurance* (URRMI). The integration of NCMS and URBMI followed these principles: unify coverages of drugs and services to the scheme with broader coverages; upgrade risk-pooling to municipal-level rather than county-level; allow benefit claims by rural-to-urban migrants in urban areas, and keep individual contributions to the scheme with lower individual financing levels (W. Xu et al., 2016).

Although URRMI has increased rural residents' average frequency of outpatient service utilization from 0.21 to 0.28 and that of inpatient service utilization from 0.06 to 0.11 (D. Su et al., 2019), this integrated program still faces similar challenges found in NCMS, such as insufficient funds and low utilization (K. Zhu et al., 2017). The merge of NCMS with URBMI does not significantly change actual compensation levels and service utilization rates (D. Su et al., 2019). One underlying reason is that URBMI shared similar funding sources and reimbursement designs with NCMS (Barber & Yao, 2011), and both schemes adopted the household-level-voluntary enrollment with a similar pooling fund per capita. Additionally, the current unifying process is slow, and many provinces stick to the previous separated NCMS and URBMI financing mechanism, due to the lack of policy guidance and role identification of governments (K. Zhu et al., 2017). To shed light on program effectiveness and improvement, our study seeks to understand the fundamental problems in the NCMS design that are still relevant to the new URRMI program.

One lingering problem is: why the NCMS fell short of achieving its primary goal of lifting the rural population out of the medical impoverishment despite the high participation rate? In fact, participating in the NCMS does not mean the household will actually utilize the NCMS program benefits. It has been documented that over 96% of those farmers who enrolled in the

NCMS program did not file any reimbursement claims mainly because of large deductibles and high co-payments (You & Kobayashi, 2009). Another potential reason for low utilization rates is that the NCMS's benefit plans do not cover several commonly demanded health services. Instead, the NCMS emphasized inpatient care for catastrophic illnesses (Meng & Xu, 2014), which had very low incident rates (i.e., not the common public healthcare needs). In its earlier years, many counties' programs did not cover outpatient care, which was the one most demanded by low-income people (Yip & Hsiao, 2009). Our study tackles this source of low utilization problem by investigating the existence of service distortion and local governments' incentives misalignment, which leads to inefficiency of NCMS in meeting consumers' medical needs.

Under the process that the NCMS has gradually been integrated into the URRMI, it is crucial to understand how to realign the incentives to overcome program ineffectiveness due to service distortions. Previous researchers have attempted to examine the program benefit designs by comparing NCMS reimbursement policies across regions from the supply side (P. H. Brown, De Brauw, & Du, 2009; You & Kobayashi, 2009; B. Yu et al., 2010), but only a few have conducted demand-side analyses due to lack of NCMS participants' micro-level data (Lei & Lin, 2009; Yip & Hsiao, 2009). Using China Health and Nutrition Study (CHNS) data, this study examines healthcare spending patterns and NCMS benefit distributions in four regions of China and show how the demand-side information can be used to address service-level distortion and improve benefit designs.

3.1.1 China's New Cooperative Medical Scheme (NCMS)

In order to motivate our study's contribution, a clear walk-through of NCMS infrastructure is needed. We illustrate the infrastructure of NCMS in Figure 3.1 and discuss three important components of this scheme.

1) Role of governments

The central government, provincial and county governments in China were involved in the NCMS program and took different responsibilities. At the central level, State Council-level authorities worked with the Ministry of Health to offer annual general guidelines of the NCMS program nationwide (P. H. Brown et al., 2009). There were three basic requirements: 1) the enrollment was voluntary, and all households with valid rural residence status (i.e., *hukou*) were eligible to enroll; 2) the NCMS had to put priority on catastrophic illnesses treatment coverage; 3) The annual risk premium per person should not be less than a minimum amount set by the central government at the beginning of the enrollment year.[15] At the provincial level, health bureaus added more details to the general guidelines, such as administrative arrangements and thresholds of reimbursement rules. The provincial authorities monitored claims and program evaluation data in an NCMS information system and reported summary information to the national NCMS information system. County governments were in charge of the day-to-day NCMS program implementation. Decisions on fund utilization, daily operations, and designs of benefit plans came from a county-level committee, involving officials from the local government and health department. Each county had an NCMS office to implement this program and collect data for the NCMS information system monitored by the province government. The NCMS evaluation criteria varied by provinces, but in general, a county's NCMS program was considered to be successful if its NCMS participation rate (*note*: not utilization rate) exceeded a

[15] The minimum risk premium is constant cross provinces. In some wealthy counties, actual risk premiums were higher than the minimum amount, but most counties set NCMS risk premium as the minimum risk premium required by the central government.

threshold (i.e., 85%), annual NCMS fund was almost break-even[16], and most participants were satisfied[17].

2) Funding sources

The NCMS was a government-led public health program whose funding mainly came from all levels of governments. In 2009, the annual minimum NCMS risk premium was RMB 100 (~ USD 15) per person, which consisted of RMB 20 individual contribution and RMB 80 government subsidies. To encourage NCMS program recruitment efforts, the central government only gave matching subsidy to those counties whose NCMS program achieved a minimum of 80% population coverage (Hou, Van de Poel, Van Doorslaer, Yu, & Meng, 2014; You & Kobayashi, 2009). The percentage of subsidies from the central government were higher for poorer regions (e.g., western and central regions) than the eastern region, which is relatively wealthier (Lei & Lin, 2009). About 1% of the NCMS funding came from other sources, and those sources are not consistent over time and across regions (K. Zhu et al., 2017). For example, the individual contribution of impoverished residents can be covered by a Medical Financial Assistance program (Barber & Yao, 2011), and some state-owned enterprises may also contribute to the funding pool.

3) Reimbursement procedure

To qualify for NCMS reimbursements, participants had to visit those health facilities pre-approved by the county office. There were two standard procedures of reimbursements: participants in some counties had to pay their entire medical bills to the health facility first and

[16] Theoretically, the break-even condition means total fund revenue is equal to the total program cost. In reality, a small fund surplus of 5-10% of the total NCMS fund was maintained. According to "Plan on Recent Priorities in Carrying Out the Reform of the Medical and Health Care System (2009–2011)" of the State Council, a county's NCMS fund surplus should be controlled at less than 15% of the total fund in the same year (Gu, 2011).

[17] Satisfaction rates are determined by random on-site interviews of NCMS patients in local health facilities.

then submit receipts and other necessary documents to NCMS office for reimbursements; in other counties, enrollees only needed to pay the amount of the medical bills after NCMS qualified amount was deducted.

[Figure 3.1 to be here]

This infrastructure gave county governments a significant amount of autonomy towards the NCMS implementation and benefit designs as long as they followed basic guidelines set by the central and provincial governments. Researchers found considerable variations in the NCMS benefit designs such as reimbursement rules and service level coverages by region (Barber & Yao, 2011; P. H. Brown et al., 2009; P. Liu et al., 2018). Despite the wide range of different benefit designs, the NCMS service coverage was generally quite limited. The literature has documented evidence of large deductibles, low ceilings, complicated reimbursement procedures, high coinsurance rates, small budgets available for NCMS (Wagstaff et al., 2009; You & Kobayashi, 2009) and high financial and political pressure on county governments (L. Zhang et al., 2010) as the main reasons.

3.1.2 Adverse selection in the health insurance market

For voluntary health insurance with a flat-rate risk premium similar to NCMS premium structure, the problem of adverse selection has been widely recognized. According to Rothschild and Stiglitz (1978), adverse selection occurs when insurers are uncertain about the health risk levels of their enrollees. The asymmetric information enables high-risk individuals to purchase more generous insurance contracts designed for low-risk individuals and causes non-social optimal increases in the total plan cost. Additionally, healthier people expecting to use fewer services may leave the market and further increase the financial risk borne by insurers.

We suspect there are adverse selections in the NCMS for the following reasons. First, despite the high enrollment rates of the NCMS, the households who filed claims only accounted for 3.3% of the total participants (You & Kobayashi, 2009). This evidence signaled a very low utilization rate of the NCMS. High deductibles prevented enrollees from claiming NCMS benefits if their health spending was not significantly high, which was evident by the observation that reimbursement rates increased as health spending went up. Therefore, wealthier households benefited more from this program than the poor, and those with high medical expenditures (i.e., those likely to be less healthy) were more likely to utilize the program than healthier participants (Lei & Lin, 2009; Wagstaff et al., 2009; Yip & Hsiao, 2009). Second, accumulating empirical evidence has suggested that individuals' health status significantly affected their NCMS enrollment decisions (Wagstaff et al., 2009; Wang, Zhang, Yip, & Hsiao, 2006; W. Xu et al., 2016; You & Kobayashi, 2009). For example, Wagstaff et al. (2009) found that households with a higher portion of members having chronic diseases were more likely to participate in the NCMS. A similar rural medical scheme established by a Harvard research team in rural China also revealed that enrolled individuals had worse health status than non-enrolled individuals (Wang et al., 2006). Moreover, public mistrust of the government-run insurance program was also found to be a participation barrier for those relatively healthy rural residents (Yip & Hsiao, 2009; Zhong, 2011). Furthermore, most farmers did not understand the concept of health insurance and treated NCMS risk premiums as a form of government taxation (Lei & Lin, 2009).

3.1.3 Adverse selection consequence: service-level distortion

Adverse selection can induce inefficient service coverages (Layton, Ellis, McGuire, & Van Kleef, 2017). That is, the insurers distort the service-level plan benefits to discourage utilization from high-cost people and attract low-risk participants. This strategy is also called "indirect selection" (Breyer, Bundorf, & Pauly, 2011). Evidence of service-level distortion has been widely found in Medicare (J. Brown, Duggan, Kuziemko, & Woolston, 2014; Cao & McGuire, 2003; Carey, 2017; Newhouse, Price, Hsu, McWilliams, & McGuire, 2015), marketplaces (Geruso, Layton, & Prinz, 2016; McGuire, Newhouse, Normand, Shi, & Zuvekas, 2014), and employer-based insurance (Eggleston & Bir, 2009).

Frank, Glazer, and McGuire (2000) originated this line of literature that measures service-level distortion. They derived and calculated the optimal shadow prices across different information sets and risk-adjustment methods to quantify under- and over-provision of a health service. They found that a more predictable health service tends to be underprovided (i.e., exhibit relatively high shadow prices) and risk-adjusted premiums can mitigate the distortion. Ellis and McGuire (2007) developed a selection index to measure the incentive to increase shadow prices (i.e., under-provide certain services) based on a sample of Medicare beneficiaries. The selection index brought additional insights by relating the incentive of service-level distortion with service characteristics. It showed that a healthcare service was likely to be under-provided when its spending was highly predictable and correlated with total health spending, and its demand elasticity affected the magnitude of distortion. Among four healthcare plans, Ellis, Jiang, and Kuo (2013) found that the traditional comprehensive plan was the least selective/distorted in service coverages. Similar analyses were conducted in private health insurance exchanges, McGuire et al. (2014) revealed that care for cancer, mental health, and substance abuse was most vulnerable to under-provision.

The extensive decentralization of the NCMS infrastructure puts the incentives of county governments (i.e., the primary implementers) at a crucial position for the program effectiveness. Given the fixed premium per person (i.e., fixed revenue inflow), local governments' costs minimization motivations were salient, especially in those counties with income and health disparities. One of the predominant ways to control costs is through service-level distortion, as found in many other healthcare plans. Those county governments could distort service provisions in many ways such as: making the reimbursement procedure cumbersome, restricting benefits to those without local *Hukou* status, or limiting hospital choices. Acting similarly to private insurers, county governments might directly under-cover the health services that involve a significant amount of medical expenses (e.g., treating chronic diseases) and ignore their residents' disease profiles and healthcare needs (P. H. Brown et al., 2009; Yip & Hsiao, 2009). Alternatively, county governments might provide "too much" coverage for health services demanded by those relatively healthy people. Either of these distorts the efficiency level of the NCMS benefit plans and contradict the primary goal of closing health disparity gaps. Next, we will utilize the service distortion literature methods and CHNS data to quantify the existence and degree of distortion in NCMS services.

3.2 Theoretical Model

This section characterizes how local governments determine the NCMS benefit plans by adopting a modified principal-agent model from Frank et al. (2000). Frank et al. (2000) assumed that the insurers were profit-driven, and there was no information asymmetry regarding insurance benefit expectation between insurers and participants. To better serve the NCMS specific contextual background, we modify the objective function to assume that local governments aim to break-even, instead of profit maximizing, because the central government's evaluation of the local NCMS programs discourages large fund deficits or surplus as stated before. Furthermore, our model reflects that governments are less informed than participating households regarding the households' expected healthcare expenditure (i.e., expected NCMS claims/benefits) to fit the asymmetric information flow existing in the NCMS design. The timeline of the contract is as follows: local governments firstly design the benefit plan of the NCMS based on their predictions of households' insurance benefits and participation decisions. Next, the household decides whether to participate in the NCMS given the observed benefit plan. Finally, enrolled households incur medical expenditures, and the NCMS covers a portion of household healthcare spending.

3.2.1 Agents problem: participation decisions

Before deciding whether to participate in the NCMS, households (i.e., agents) do not know their future health status. Thus, they are uncertain about their expected healthcare spending and how much reimbursement they will receive through NCMS. To simplify the model, we treat a household i as a single agent who would experience only two possible health outcomes in the next year: being unhealthy with a probability of λ_i and being healthy with a probability of 1-λ_i. This probability is household i's private knowledge. Household i's expected benefit of service s is a weighted average of the benefits associated with the two potential health outcomes

$$\widehat{m}_{is} = \lambda_i \overline{m}_{is} + (1 - \lambda_i)\underline{m}_{is},\ \ 0 < \lambda_i < 1 \text{ for } i = 1, 2, \ldots, N \tag{3.1}$$

where $\overline{m}_{is}$ refers to the NCMS benefit of service s received if the household is the unhealthy type in the next year, and $\underline{m}_{is}$ is the benefit associated with a healthy household. Let $\widehat{m}_i = [\widehat{m}_{i1,}\ \widehat{m}_{i2}, \ldots, \widehat{m}_{iS}]$ be a set of expected NCMS benefits over a total of S services ($s = 1, 2, 3, \ldots, S$). The utility of participating in the NCMS for household i is

$$u_i(\widehat{m}_i) = v_i(\widehat{m}_i) +\ \mu_i - c_i, \text{ where } v_i(\widehat{m}_i) = \textstyle\sum_s v_{is}(\widehat{m}_{is}). \tag{3.2}$$

The first component $v_i(\widehat{m}_i)$ is the total valuation of the expected NCMS benefit. If the valuation is assumed to be additive, $v_i(\widehat{m}_i)$ is a sum of household i's valuation of insurance benefit across services. We assume that $v_{is}'(.) > 0$, $v_{is}''(.) < 0$, and the valuation of expected benefit from service s, $v_{is}(.)$, is independent of the benefits from other health services. The second component μ_i indicates household-specific valuation independent of the services from the NCMS,[18] and c_i is household i's cost of enrolling and obtaining insurance benefits. For instance, it includes household's out-of-pocket premium payments and time costs involved with enrolling and reimbursement claim filling. For simplicity, we assume that c_i is not type-dependent: in other words, the costs of enrolling and claims filing do not vary across health status.[19]

Suppose the household does not participate in the NCMS, it stays merely uninsured,[20] so the household's utility will be the expected reservation utility of the form: $\hat{u}_i^0 = \lambda_i \cdot \bar{u}_i^0 + (1-\lambda_i)\underline{u}_i^0$. If $u_i(\widehat{m}_i) > \hat{u}_i^0$, the household will participate in the NCMS; thus, the probability of participation is

$$Prob(u_i(\widehat{m}_i) > \hat{u}_i^0) = Prob\left(\mu_i > \hat{u}_i^0 + c_i - v_i(\widehat{m}_i)\right) \tag{3.3}$$

Assuming μ_i follows a certain cumulative distribution function F_i, we can express the probability as

$$Prob(u_i(\widehat{m}_i) > \hat{u}_i^0) = 1 - \mathrm{F}[\hat{u}_i^0 + c_i - v_i(\widehat{m}_i)] \equiv n_i(\widehat{m}_i, \hat{u}_i^0, c_i) \tag{3.4}$$

---

[18] For example, μ_i captures household i's unobserved trust and perception of government programs. If the household does not trust the county government ($\mu_i \ll 0$), its overall utility of participating can be low even when NCMS' benefit plan is desirable or the cost of enrolling is minimal. Later on, we assume a distribution of μ_i across households because this term is the uncertain component in the utility.

[19] The validity of this assumption may depend on reimbursement procedures. The cost is likely to be type-independent if insured households receive immediate reimbursement directly from health facilities, because the cost of obtaining an insurance benefit is not related to total medical spending (Zhong, 2011). However, if the reimbursement requires the enrollees to submit receipts case-by-case, unhealthy households may encounter a higher cost because they need to make more insurance claims. But assuming type-independence will not lose generality.

[20] In China, most rural residents cannot afford other private health insurance plans (Jin et al., 2016), so the reservation utility is defined as the utility obtained when the household does not purchase any health insurance. Because a healthy household has a lower probability of needing to pay substantial medical expenses, the reservation utility of the healthy type is higher than unhealthy type ($\bar{u}_i^0 < \underline{u}_i^0$).

This equation indicates that a household is more likely to participate in the NCMS if it expects a higher plan benefit ($\frac{\partial n_i}{\partial \widehat{m}_i} > 0$), or a lower reservation utility ($\frac{\partial n_i}{\partial \widehat{u}_i^0} < 0$), or a lower cost of benefit utilization ($\frac{\partial n_i}{\partial c_i} < 0$).

3.2.2 Shadow price

Following Keeler, Carter, and Newhouse (1998), we use shadow prices to measure the generosity of the plan coverage over different services. In the NCMS program context, the shadow price p_s is defined as a government-assessed threshold: to qualify for NCMS specific service reimbursements, a household needs to have a marginal valuation of the NCMS service benefits equal to or exceed p_s:

$$v'_{is}(\widehat{m}_{is}) = \frac{dv_{is}}{d\widehat{m}_{is}} \geq p_s \tag{3.5}$$

When a county government announces NCMS reimbursement rules, it implicitly sets shadow prices for different services from the supply side.

Solving equation (3.5) by setting it to equilibrium status, we can see that the optimal equilibrium level of $\widehat{m}_{is}$ for a given p_s can be derived from solving the inverse demand function $\widehat{m}_{is}^*(p_s) = v_{is}'^{-1}(p_s)$. To demonstrate how shadow prices determine service coverage, we illustrate an example in Figure 3.2 which shows demand curves of two households: a healthy household i and an unhealthy household j. Due to the health status, the unhealthy household j demands more of service s than the healthy household $\boldsymbol{i}$. Given the same shadow price p_s^0, this example shows that the unhealthy household receives more insurance benefits because it has a higher demand for service s than the healthy type. If the plan wants to be less generous regarding service s, one strategy is to increase the shadow price to p_s^1 to make both households receive less benefit. This can be done by decreasing the NCMS reimbursement rate or increasing the deductible of service s. In sum, the higher the p_s, the less coverage of service s provided by the NCMS.

[Figure 3.2 to be here]

3.2.3 Principal problem

1) Objective function and shadow price under perfect information

It is crucial to have a contextually specific objective function selected to capture the essence of insurers' decision making. The previous literature often used two types of objectives: for regulators, studies usually assumed welfare maximization (Einav, Finkelstein, & Cullen, 2010; Layton et al., 2017), while for private insurers, researchers typically assumed profit maximization (Ellis & McGuire, 2007; Frank et al., 2000; McGuire et al., 2014).

Given NCMS program infrastructure, efficient management of insurance funds was an important political interest of local governments. First, local governments had to avoid insurance fund deficits to sustain the program. Counties were required to reserve a 5-10% of fund revenue annually to pay for administration and cover unforeseen costs (L. Zhang et al., 2010). On the other hand, the State Council discouraged huge surpluses of insurance funds in order to strengthen NCMS's mission of financial protection, because excessive fund surplus reduced the total reimbursement and benefit enrollees could claim (Gu, 2011). Based on those two points, it is reasonable to depict that a county government is seeking to minimize the expected surplus and deficits, which has the lower bound of break-even (no surplus or deficit), instead of profit or total welfare maximization.

To fit into this NCMS specific goal, we will frame the principal's (i.e., a county government's) objective as minimizing expected fund surplus and deficit through breakeven seeking. The principal will achieve the goal by setting a vector of shadow prices $\boldsymbol{p} = [p_1, p_2, \dots, p_S]$ across the total of S services. The objective function under perfect information[21] can be defined as[22]

$$\min_{\boldsymbol{p}} Q = \pi(\boldsymbol{p})^2 \tag{3.6}$$
$$= \left\{ \sum_i [\underbrace{n_i(\hat{m}_i(\boldsymbol{p}), \hat{u}_i^0, c_i)}_{Probability\ of\ participation} \times \underbrace{\pi_i(\hat{m}_i(\boldsymbol{p}), r_i)}_{Plan\ net\ profit\ from\ household\ i}] \right\}^2$$
$$= \left\{ \sum_i \left[n_i(\hat{m}_i(\boldsymbol{p}), \hat{u}_i^0, c_i) \times (r_i - \sum_s \hat{m}_{is}(p_s)) \right] \right\}^2$$

where r_i denotes the risk premium collected from household i and is considered the NCMS program earnings. The total insurance benefits received by the household i across all services through NCMS reimbursement, $\sum_s \hat{m}_{is}(p_s)$, are the major part of NCMS program costs. We assume the administration costs are independent of the insurance funds and are constant across counties and states; thus, the plan's expected net profit associated with household i is $r_i - \sum_s \hat{m}_{is}(p_s)$.

After solving the first-order condition $\frac{dQ}{dp_s} = 0$ (see Appendix CA), we obtain the break-even shadow price of service s for household i:

$$p_s^* = \frac{\sum_i n_i \hat{m}_{is}}{\sum_i \mathrm{F}_i' \hat{m}_{is}(r_i - \sum_s \hat{m}_{is})} = \frac{\sum_i n_i \hat{m}_{is}}{\sum_i \mathrm{F}_i' \hat{m}_{is} \pi_i}, \text{ where } \mathrm{F}_i' = \frac{dn_i}{dv_i} = \frac{dF_i}{dv_i} \tag{3.7}$$

This break-even shadow price will determine the specific service-level coverages to achieve the government's break-even objective. From the government's perspective, the numerator

[21] Under the first-best condition (i.e., perfect information flow), the county government and the households will have the same expectation of the households' future insurance benefits.

[22] Instead of taking the absolute values, we minimize the squared sums of excessive surplus to simplify derivations without losing generality.

represents the total expected cost of covering service s. Higher service costs will mean higher shadow prices that result in less coverage for the specific healthcare service. The denominator characterizes the marginal benefit of increasing participation. Since π_i is the expected net profit from enrolling the household i, the local government will cover service s more for those low-risk enrollees, whose associated net profit, π_i , is positive. Theoretically, the denominator is positive because the need for an annual reserve fund requires the overall program to be profitable to sustain this minimal financial needs, and therefore shadow prices should be positive.

Socially optimal condition implies that the household's marginal benefits from additional services provided are the same across all available services (i.e., households' medical needs across all services are equally met) (Frank et al., 2000). Therefore, the break-even shadow prices should be equalized among different services:

$$p_s^* = v_{is}'(\widehat{m}_{is}) = v_{is'}'(\widehat{m}_{is'}) = p_{s'}^* \text{ for } s \neq s' \;\; \forall s = 1, 2, 3, \ldots, S \tag{3.8}$$

If equation (3.7) and (3.8) are both satisfied, the local government not only uses insurance fund efficiently but also maximizes households' valuation of the NCMS benefit plan (i.e., supply and demand sides both reached optimum at the equilibrium shadow prices). On the other hand, service-level distortion occurs when $p_s^* \neq p_{s'}^*$. Specifically, the ratio $\frac{p_s^*}{p_{s'}^*} > 1$ means service s is under-covered than service s', and $\frac{p_s^*}{p_{s'}^*} < 1$ means service s is over-covered than service s'. In either case, the local governments does not achieve the social optimal condition.

Figure 3.3 provides an example of service-level distortion. Suppose the break-even shadow prices are p_s^* and $p_{s'}^*$. In this case, the shadow price of service s is higher than service s'. It implies a household's marginal valuation from additional NCMS benefit of service s is higher than the other service, and service s is under-covered. If the NCMS benefit plan increases coverage on the more desired service and reduces the coverage on the less desired one by reallocating the insurance reimbursement (Δm), the household's valuation could get higher until shadow prices between two services are equalized, as shown in the solid line, which is the social optimal condition. In the first-best case, the local government can achieve this win-win solution via adjusting the risk premium r_i for each participant. In reality, since NCMS programs used flat-rate risk premiums for households within the same county, local governments were unlikely to achieve this condition.

[Figure 3.3 to be here]

2) Information problem: how population characteristics affect the shadow price

In the real world, there is asymmetric information flow between principal and agents embedded in the NCMS implementation: i.e., the government's expected NCMS program reimbursement amounts may not be the same as the households' expected benefits. Under this second-best scenario, the household's expected healthcare expenditure on service s, $\widehat{m}_{is}(p_s) = \lambda_i \overline{m}_{is} + (1 - \lambda_i)\underline{m}_{is}$, is the household's private information, which is unknown to the local government. The principal (i.e., the local government) must estimate/guess several parameters: λ_i, $\overline{m}_{is}$, and $\underline{m}_{is}$ and calculate its own expected program benefit costs.

Based on the feasibility and availability of information to the local government[23], we assume that the county government would predict the probability of a household being unhealthy based on the percentage of unhealthy population in that county, λ: $\lambda_i \approx \lambda$. To estimate the medical expenditure under healthy and unhealthy situations for household i, $\overline{m}_{is}$ and $\underline{m}_{is}$, we assume that the county government will first utilize census data and historical insurance

[23] There are obviously many ways the government can predict NCMS reimbursements. To simplify the theoretical model, we just pick one way to illustrate here.

expenditure to predict a baseline insurance benefit m_{is}^B for household i based on its observed demographic characteristics (e.g., age and gender). Because these demographic characteristics are most accessible and widely used by insurers to predict insurance benefits. Then the county government will adjust m_{is}^B by a service-specific NCMS benefit allocation between healthy and unhealthy groups in a county. As shown in Figure 3.2, unhealthy households usually obtain a higher insurance benefit than healthy households. To approximate the adjustment decision of the local government, we introduce parameter θ_s to capture the discrepancy in NCMS benefit of service s between two health scenarios and assume the government approximates $\overline{m}_{is} \approx \bar{\theta}_s m_{is}^B$ and $\underline{m}_{is} \approx \underline{\theta}_s m_{is}^B$. Based on all those parameters, the local government's prediction of household i's insurance benefit is

$$\widehat{m_{is}(p_s)} \equiv \lambda\overline{\theta}_s m_{is}^B(p_s) + (1-\lambda)\underline{\theta}_s m_{is}^B(p_s) \tag{3.9}$$

$$= \left[\lambda\overline{\theta_s} + (1-\lambda)\underline{\theta_s}\right] m_{is}^B(p_s) = \Theta_s m_{is}^B(p_s),$$

$$\text{where } \Theta_s = \left[\lambda\overline{\theta_s} + (1-\lambda)\underline{\theta_s}\right]$$

The second-best shadow price derived is (see Appendix CB for detailed derivations)

$$p_s^{second} = \frac{\sum_i n_i m_{is}^B}{\sum_i \mathrm{F}_i' m_{is}^B (r_i - \sum_s \Theta_s m_{is}^B)} \tag{3.10}$$

The second-best shadow price formula, equation (3.10), provides additional insights into how population health status and associated health spending affects the distortion of shadow prices. We will illustrate those insights through two implications:

Implication 1: Holding other factors constant, the NCMS plan will cover less of service *s* if a county has a higher portion of unhealthy households as compared to other counties. It can be shown through the partial derivative of shadow price with respect to the unhealthy population proportion: $\frac{\partial p_s^{second}}{\partial \lambda} = \frac{\partial p_s^{second}}{\partial \Theta_s} \times \frac{\partial \Theta_s}{\partial \lambda} > 0$. Due to the concern of running at a deficit, a local government has the incentive to increase the price of (or undercover) health service s if a higher portion of its counties' residents is unhealthy. This implication partially explains why reimbursement rates of the same service vary across counties (P. H. Brown et al., 2009; L. Zhang et al., 2010; Y. Zhang, Chen, Zhang, & Zhang, 2014). For example, in Shaanxi province, Danjiangkou County reimbursed 65% of inpatient expenditure at township hospitals, while the corresponding reimbursement rate in Meixian County was 90% (Y. Zhang et al., 2014).

Implication 2: Given an NCMS benefit plan, we can calculate the second-best shadow prices of different services based on equation (3.10). The difference between $\frac{p_s^{second}}{p_{s'}^{second}}$ and 1 indicates the severity of service-level distortion. Holding other factors the same, the service with the largest benefit difference between the two types of households will experience the most severe upward distortion in the shadow price as compared to other services. Use $\Delta\theta_s = \overline{\theta_s} - \underline{\theta_s}$ to denote the benefit differences between the two types of households, we will have $\frac{\partial p_s^{second}}{\partial \Delta\theta_s} = \frac{\partial p_s^{second}}{\partial \Theta_s} \cdot \frac{\partial \Theta_s}{\partial \Delta\theta_s} > 0$. This condition shows that health service s tends to be under-covered when unhealthy households demand much more insurance benefits than healthy households.

3.3 Data and Variables Description

In section 3.2, we have illustrated theoretically what to expect when asymmetric information is present and how service coverage will be distorted accordingly. This section will turn attention to empirically quantify the NCMS service-level distortion and illustrate how risk premium adjustments can help mitigate the problem.

3.3.1 Data

This study used the China Health and Nutrition Survey (CHNS) data (in 2009 and 2011). The CHNS was a collaborative project between the Carolina Population Center at the University of North Carolina at Chapel Hill and the Chinese National Institute for Nutrition and Health. This survey followed a multistage, random cluster process to draw a sample of Chinese households from nine provinces and three municipalities and collected their nutrition, health behaviors, outcomes, and socioeconomic status. These provinces and municipalities[24] were divided into four regions (X. Xu, Byles, Shi, & Hall, 2015): 1) Northeast China: Heilongjiang and Liaoning; 2) East Coast: Shandong and Jiangsu; 3) Central China: Henan, Hubei, and Hunan and 4) Western China: Chongqing, Guangxi, and Guizhou. Since many emigrants had to obtain NCMS benefits from their county of origin, they may not receive NCMS benefits in the province of residence reported in this survey. As there was no information in CHNS that can identify emigrants and their county of origin, we chose to use region as the level of analysis instead of province to minimize the potentials of counting the medical expenditures claimed wrongly to the province of residence instead of the province of origin. This decision was based on the evidence that the majority of Chinese emigrants were migrating across provinces but within regions (Shen & Liu, 2016).

3.3.2 Key variable construction

The following subsection describes key variables needed to compute regional shadow prices.

[24] We exclude Beijing from Central China and Shanghai from East Coast because these two municipalities' benefit designs and risk premiums are quite high and not comparable with other provinces in the same region. Sensitivity analyses shows that this exclusion does not change the main results of this study.

(1) Lump-sum insurance benefits: Since we are studying NCMS program efficacy, it is important to know the insurance benefits households received from NCMS program only (i.e., excluding coverage from other insurances). However, the CHNS data only collected percentages of health spending covered by all insurance benefits. To get NCMS-only benefits, we excluded non-NCMS participants and those NCMS participants who purchased additional health insurance in the expenditure estimation.

(2) Health status indicators: Identifying healthy and unhealthy households' healthcare spending and estimating the probability of being healthy all need the information for health status classification. The CHNS did not contain direct health status indicators in their 2009 and 2011 waves. We developed two indicators based on available health information. The first indicator was self-reported and based on the following question on the CHNS: "Has a doctor ever told you that you suffer from high blood pressure/diabetes/myocardial infarction/stroke/ cancer/asthma?" Thus, a healthy person was the person answering "no" to all these questions; otherwise, we placed the person in the unhealthy category. We recognized that this measure might only reflect the population's health status conditioning on them seeking medical services. The second indicator utilized physical examination results (objectively assessed by a physician), including blood pressure, height, weight, and health conditions. We defined those unhealthy participants as having high blood pressure, or being obese (BMI>28), or having at least one disability health conditions.[25]

[25] High blood pressure was identified if one's recorded systolic blood pressure was 140 mmHg or more, or diastolic blood pressure was 90 mmHg or more. The Chinese BMI cut-off point for obesity was obtained from (Chen, 2008). Disability status satisfied at least one of these conditions: angular stomatitis, goiter, blind one or both eyes, lose one or both arms, lose one or both legs.

(3) Types of healthcare services: The CHNS did not provide detailed service-level spending or diagnosis codes that could be used to derive micro-level service-specific spending. Consistent with the classification of health services provided by Berndt et al. (2000) and Qian, Pong, Yin, Nagarajan, and Meng (2009), CHNS provided four mutually exclusive aggregate level healthcare services: self-treatment/informal care, outpatient treatments, inpatient treatments, and preventive health services (i.e., health examination, screening tests). Because the NCMS only covered formal health care, we excluded the category of self-treatment/informal care when predicting insurance benefits.

To calculate regional shadow prices, we face two data limitations. The CHNS only collected health spending for the past four weeks while the NCMS program coverage lasted for one year. Furthermore, limited one-month healthcare spending resulted in more zero and missing spending reported. To better identify true zero from missing values, we utilized the answer to the following question: "During the past four weeks, have you been sick or injured?" For those answering "no" to this question, we assumed that they genuinely did not have medical care need and assigned zero to their one-month healthcare spending. Otherwise, we treated the one-month healthcare spending information as missing. To obtain yearly healthcare benefits estimates for the households, we followed the literature (Davis & You, 2011; Horvitz & Thompson, 1952) to predict the other eleven months' healthcare spending on different services for each individual and summed up along with the individual's reported one-month spending to get the total one-year spending. The detailed methods are described in the Empirical Method section.

3.3.3 Summary statistics

We report summary statistics by region and the associated one-way ANOVA group mean tests in Table 3.1. Due to the needs of historical data, we used wave 2009 data to predict the government's expected healthcare spending and the service-level distortion for 2011. Therefore we only present descriptive statistics of the data in the wave of 2011 here.[26] Monthly health expenditures are statistically similar across regions. All regions face similar patterns across service types: inpatient and outpatient services are the two types that cost the highest and have the largest standard deviations. On average, the monthly total health spending ranges from RMB 102 in East Coast to 224 in Northeast.

[Table 3.1 to be here]

For inpatient and outpatient services, there are substantial and statistically significant differences in NCMS service-level coverage percentages across regions, even though the total insurance benefits are similar across regions. NCMS programs in the Central region cover a lot more healthcare spending for preventive health services than other regions. In contrast, the East Coast region's NCMS program coverage percentages for all services are the lowest among all regions, and its average monthly benefit is the lowest as well, with only RMB 21.15 in 2011. Since the East Coast region was considered wealthy, local governments received much lower central contribution than other regions and were mainly responsible for financing NCMS (Barber & Yao, 2011). Lack of central government's matching subsidy may make these local governments more concerned about the financial sustainability of the NCMS and therefore restrict NCMS benefit to a minimal level. The limited service coverage and total benefits also match the smallest percentage of reimbursed farmers in the region: according to You and Kobayashi (2009), only 3% NCMS enrollees in the East Coast region received reimbursement from NCMS, while the percentage of reimbursed enrollees was highest (3.7%) in Central region

[26] Summary statistics of variables in the wave of 2009 are not reported but available upon request.

Across regions, the NCMS participating individuals on average have different age, education level, and annual household income. With an average of 44 years old, Western China respondents are younger than those in other regions and have obtained the lowest level of education. We also find the East Coast has much higher average annual household incomes (i.e., on average RMB 43,551), whereas the Central region has the lowest average household income (i.e., RMB 31,604).

3.4 Empirical Method

To calculate the shadow prices in equation (3.10), we need to know the value of $\mathrm{F}_i' = \frac{dn_i}{dv_i} = \frac{dF_i}{dv_i}$, which is the probability density function of the household-specific valuation μ_i. Following Frank et al. (2000), we assume a uniform distribution of μ_i, so $\mathrm{F}_i' = 1$. Next, we discuss how to calculate five theoretical parameters from the data.

3.4.1 Population health status

The parameter λ is approximated as the portion of unhealthy population in a region. For each region, the benefit adjustment by health status, $\underline{\theta}_s$, is calculated as the ratio of healthy group median insurance benefit of service s over the population median benefit of this service.[27]Likewise, $\overline{\theta}_s$ is the ratio of unhealthy group median insurance benefit over the population median benefit.

[27] Using median is to accommodate the skewed benefit distributions.

3.4.2 Baseline NCMS benefit

A government can predict a household's baseline NCMS benefit in multiple ways. With the feasibility and practicality in mind, we are presenting one way here: i.e., the government will predict it based on a household's observable characteristics through census type of data. To form the prediction, the government needs to have historical data on individuals' healthcare spending covered by the NCMS and their characteristics and aggregate to household level. However, our data contains sizable missing values in the answers to the percent of spending covered by insurance. Among those respondents reporting non-zero healthcare spending, percentages of missing insurance coverage information are 61% for inpatient services, 32% for outpatient services, and 49% for preventive services. To address this issue, we first predict the total healthcare spending amount based on observable characteristics of the individual, then use the province average of insurance coverage percentage to get the individual insurance covered healthcare spending amount for those with the missing percentage information. In each household, we add up individual insurance benefits as the household NCMS benefit. Additionally, we do not know what information the government has available to use for household benefits prediction. Thus, two information sets are used to check how the asymmetry of information influences the degree of distortion: a less-information set and a more-information set. The actual functional forms and variables in each set are discussed in the following section of model specifications.

The next question is what empirical method to use for the prediction model estimation. A variety of empirical models have been used in the literature to predict medical expenditures, such as ordinary least squares (OLS) models (Chang, Lee, & Weiner, 2010; Chang & Weiner, 2010), weighted least squares models (Pope et al., 2004; Zhao et al., 2005), two-part models (Ellis et al., 2013; McGuire et al., 2014; H. Yu, 2017), generalized linear models (GLM) (Ellis & McGuire, 2007), and quantile regressions (Babiarz, Miller, Yi, Zhang, & Rozelle, 2012; Kowalski, 2016). We first predicted individual total health spending using OLS, GLM, two-part models, and quantile regression and selected a two-part model as it produced the lowest Bayesian information criterion.

The preferred two-part model consists of a probit model in part 1 and a conditional exponential model in part 2[28]. Under the more-information set, it is specified as

$$\Pr\left(y_{s,2011} > 0 \,\middle|\, \boldsymbol{DE}, \mathrm{y}_{\mathrm{s},2009},, \boldsymbol{R}, \boldsymbol{M}; \boldsymbol{\alpha_1}\right) = \Phi(\boldsymbol{DE}, \mathrm{y}_{\mathrm{s},2009}, \boldsymbol{R}, \boldsymbol{M}; \boldsymbol{\alpha_1}) \tag{3.11}$$

$$\mathrm{E}\left(y_{s,2011} \,\middle|\, y_{s,2011} > 0, \boldsymbol{DE}, \mathrm{y}_{\mathrm{s},2009}, \boldsymbol{R}, \boldsymbol{M}; \boldsymbol{\alpha_2})\right] = \exp(\boldsymbol{DE}, \mathrm{y}_{\mathrm{s},2009}, \boldsymbol{R}, \boldsymbol{M}; \boldsymbol{\alpha_2})$$

The dependent variable $y_{s,2011}$ is an individual's health spending on service s in the past four week in 2011. $\boldsymbol{DE}$ is a vector of individual-level demographic variables, including age, age squared, a sex indicator, completed years of education, household net income, and interactions between the sex and age. We also include $\mathrm{y}_{\mathrm{s},2009}$: individual i's prior spending on service s. Finally, regional indicators ($\boldsymbol{R}$) and monthly indicators ($\boldsymbol{M}$) are used to control regional and time effects. The error terms are assumed to be correlated within communities. The less-information set is constructed by excluding $\mathrm{y}_{\mathrm{s},2009}$, years of education, household income, and monthly indicator from the full set of explanatory variables. Since CHNS only collected individual health spending for four weeks, we estimate this model using available monthly spending and then predict counterfactual spending at these months not collected by the survey. The individual's annual spending on service s is a sum of reported spending over CHNS survey month and

[28] We conducted a GLM family test as described by Manning and Norton (2013), and the results showed that the variance exceeds mean, suggesting that a log link with a gamma distribution of the errors in part 2 fitted our data the best.

predicted monthly spending in the other eleven months. This model was estimated using STATA 14.0 (StataCorp LP, 2015).

3.4.3 Risk premium

Before implementing the annual NCMS program, the local government determines shadow prices based on expected financing level. We consider three types of risk premiums: 1) minimum risk premium adjusted by past financing performance; 2) risk premium calculated as the average actual NCMS benefit per individual in 2009; 3) risk-adjusted premium based on the Ambulatory Care Group (ACG) algorithm. Because the participation was at the household level, we aggregate individual risk premiums to their household total.

The first type of risk premiums is supply-side driven and represents actual NCMS financial condition. Under NCMS policy design, the flat-rate risk premium of a local NCMS program was driven by the minimum risk premium set by the central government, but the final collected funds were affected by multiple reasons (e.g., local economy, fiscal budget). For example, in 2011, the central government announced the minimum risk premium to be RMB 230 per NCMS enrollee, but the actual risk premiums collected varied across counties. At the beginning of 2011, although local governments were uncertain about the actual NCMS financing level, their past fund collection experience led to an informed expectation. The first type of risk premiums illustrates this scenario in Table 3.2. We obtained the actual average risk premium collected by each province in 2010 from the Census and got the ratio of the actual amount over the minimum risk premium required by the central government (RMB 150). As the minimum requirement increased to RMB 230 in 2011, we calculate the ex-ante risk premium in each province by multiplying 230 with the corresponding ratio assuming the same percentage of deviation from the minimum requirement over time. As shown in Table 3.2, the expected risk premiums are very close to the actual risk premiums collected in 2011.

[Table 3.2 to be here]

The second type of risk premiums utilizes demand-side information, which is the households' NCMS benefit. If break-even is the objective of the principal, the risk premium should be close to average NCMS benefit claimed by NCMS participants. However, only the insurance claims in previous years are available to local governments at the time of designing benefit packages for 2011. As there was no CHNS data in 2010, we calculate regional average NCMS benefit in 2009 as the second type of risk premium. This risk premium incorporates the cost of actual utilization of the NCMS program.

Based on NCMS benefit information, the third type of risk premiums is risk-adjusted following the ACG algorithm (Weiner et al., 1996). The ACG risk-adjustment approach classified *International Classification of Diseases, 9th Revision, Clinical Modification* (ICD-9-CM) diagnosis codes into distinct Ambulatory Diagnosis Groups (ADG) and assigned individuals to one or more ADGs based on their previous diagnoses. The original study regressed individual annual medical expenditure on four dummy variables: male, years over 65, ever disabled and Medicaid eligibility, and indicators of 13 ADG groups. The ACG-adjusted risk premium was the annual total Medicare expenditure predicted using individual characteristics and diagnoses documented in the last year.

This study could not apply the ACG algorithm directly because there was very limited diagnosis information in the CHNS, and available diagnosis information was not coded into the ICD-9-CM system. However, following the insight of the ACG method, we estimate a risk-adjustment model and use the predicted 2011 NCMS benefits as the third type of risk premium.

The first step is to classify CHNS disease history information into four disease categories, as shown in Table 3.3.

[Table 3.3 to be here]

Then, we specify a risk-adjustment model as follows:

$$m_i = \sum_s m_{is} = \gamma_0 + \gamma_1 d_{>65} + \gamma_2 male + \gamma_3 d_{>65} * male + \gamma_4 disable + \gamma_5 d_{>65} * disable + \boldsymbol{\gamma_6 DC} + \boldsymbol{\gamma_7 M} + \boldsymbol{\gamma_8 R} + \nu_i \quad (3.12)$$

Individuals' total NCMS benefits are regressed on indicators of years over 65 ($d_{>65}$), being a male ($male$), disability status ($disable$), and their interaction terms. The risk-adjustment model also controls for indicators of disease categories ($\boldsymbol{DC}$), months ($\boldsymbol{M}$), and regions ($\boldsymbol{R}$). Since respondents only report NCMS benefit for the past four weeks, we predict a counterfactual m_i at these months not covered by the survey period. For the month when m_i is reported, we keep the original values. An individual's ACG risk premium (i.e., annual NCMS benefit) is a sum of monthly predicted m_i.

3.5 Results and Discussion

This section starts with an overview of population health status and healthcare service characteristics, which affect governments' selective incentives as suggested in the theoretical model. Next, we assess NCMS benefit distortions between services and across regions and discuss how risk premium adjustments can provide insights into program improvement.

3.5.1 Population characteristics

The upper panel of Table 3.4 reports population health status and benefit distribution. For each region, we calculate the percentage of unhealthy respondents as an approximation of λ.

Results based on self-reported health status are pretty consistent with examination-based results. About 17% of respondents in the East Coast region report suffering from chronic diseases, and they tend to be sickest and have high percentages of high blood pressure, obesity, or disability. The Western region again has the lowest rate of the unhealthy population according to self-reported diseases (10%) and physical examination results (26%), perhaps because this sub-group are younger on average.

[Table 3.4 to be here]

Next, we calculate the median insurance benefits of the healthy and unhealthy groups classified by self-reported indicators[29] and construct $\underline{\theta_s}$ and $\overline{\theta_s}$. The parameter $\Delta\theta_s = \overline{\theta_s} - \underline{\theta_s}$ is calculated under two scenarios: 1) when governments can predict 2011 NCMS benefits with the more-information set and 2) when they can predict with less information. Under the less information set, we do not find a substantial group difference in insurance benefits for preventive health services ($\Delta\theta \approx 0$). The government anticipates that healthy people demand slightly more benefit on preventive health services than unhealthy people. On the other hand, inpatient and outpatient benefits are more attractive to unhealthy households than the other group. Specifically, the discrepancy of benefit ($\Delta\theta$) for inpatient treatments is approximately twice of that for outpatient treatments. For outpatient expenditure, more-informed governments expect a larger $\Delta\theta$ between two groups compared with the less-informed scenario.

[29] We also calculated $\underline{\theta_s}$ and $\overline{\theta_s}$ based on exam-based indicators, and the differences in median benefits between two groups are minimal ($\Delta\theta \approx 0$) , so we use the self-reported health status to calculate selection indices.

3.5.2. Health service characteristics

Besides population health status, selective incentive also depends on predictability and predictiveness of a service (Ellis & McGuire, 2007; McGuire et al., 2014). The coefficient of variation of predicted benefits on service s, $cov(\widehat{m}_{is})$, is the measurement of predictability: the more predictable the spending on a health service, the larger the $cov(\widehat{m}_{is})$. The predictiveness measurement, $\rho_{\mathrm{m_{is}},\pi_i}$, is the correlation coefficient between predicted service benefits and plan profit.

Table 3.5 reports $cov(\widehat{m}_{is})$ and $\rho_{\mathrm{m_{is}},\pi_i}$ under two information assumptions. The left panel shows the less-information scenario. All services are not quite predictable ($cov(\widehat{m}_{is}) \approx 1$). With additional information (i.e., past spending, education, household income, and month of utilization), spending on preventive services is most predictable in the right panel. Large negative correlations ($\rho_{\mathrm{m_{is}},\pi_i} \leq -0.7$) are observed for inpatient services because NCMS emphasized reimbursement for inpatient care. Outpatient serves are moderately predictable and negatively correlated with plan profits. Next, we show results of shadow prices and explain how population health status and service characteristics play a role in service-level distortions.

[Table 3.5 to be here]

3.5.3. Benefit distortions between services

The second-best shadow prices are calculated according to equation (3.10). Under current NCMS financial-level (i.e., the first type of risk premiums), calculated shadow prices are negative in Northeast, Central, and Western regions due to predicted negative plan profits. This finding is consistent with studies by L. Zhang et al. (2010) and P. H. Brown et al. (2009), which suggested that several counties' NCMS programs faced risks of budget deficits. Although our result may underestimate plan profit since local governments may finance the NCMS program through state-owned companies or use the reserve funds saved in previous years, the financial sustainability of NCMS can be a substantial pressure for poorer counties (W. Xu et al., 2016; L. Zhang et al., 2010).

By following Frank et al. (2000), the first type of risk premiums is adjusted upward until all shadow prices in a region become positive, and the minimum adjustments are reported in Table CC-1. There is no upward adjustment in the East Coast region because their NCMS benefit plans have expected fund surplus. To reveal service-level distortions, we normalize the shadow prices of inpatient services to one in each region; thus, all other shadow prices are relative to this category. Inpatient service is chosen as the baseline because all NCMS programs cover this type of services, while outpatient and preventive services are not covered in some places. Note that we cannot compare relative shadow prices between regions since the baselines are not the same for each region.

Figure 3.4 plots relative shadow prices by region (see Table CC-1 for standard errors). If shadow prices are socially desirable (i.e., marginal benefits are the same across all services), all relative shadow prices should be close to 1 within a region. Therefore, we draw a vertical line at $p_s = 1$ and use an asterisk to indicate those values significantly different from one at 5%. The upper panel of Figure 3.4 illustrates relative shadow prices when local governments are less informed. With the first type of risk-premiums, the pattern of service-level distortions is pretty consistent in four regions. The shadow prices of outpatient services are distorted upwards, meaning the NCMS undercovers outpatient services the most, while preventive services are generally over-covered than other services. One possible reason is that some NCMS programs offered a general physical examination for free to those enrollees who did not claim any NCMS benefit in that year (Lei & Lin, 2009). Under the less information set, preventive services are more appealing to healthy households than unhealthy households ($\Delta\theta < 0$ in Table 3.4). According to adverse selection theory, insurers might provide "too much" coverages of preventive services to attract good risk. The distortions of service-level coverages are reduced as the asymmetry of information between local governments and household become smaller. For instance, over-provision of preventive service benefits does not occur if local governments have more information to predict households' insurance benefit, as indicated by $cov(\widehat{m}_{is})$ in Table 3.5. In the more-information case, shadow prices of outpatient services are not significantly different from 1 in three regions. One explanation is that, although the NCMS emphasizes on inpatient care for catastrophic illness, informed local governments would realize residents' demand for outpatient coverages. As a result, their NCMS benefit plans provide more outpatient benefits rather than over-supply general physical examinations.

[Figure 3.4 to be here]

How do shadow prices change if the government can adjust risk premium based on prior program costs? To illustrate this scenario, we employ the second type of risk premiums (i.e., the sample average insurance benefit in 2009) and report the relative shadow prices in the block of "premium 2". Although most shadow prices in the less-information scenario are close to 1, the Western region still has lower shadow price of preventive services than other services, meaning that their NCMS programs could do better to meet local residents' need if the plan reduces benefit on preventive health services and uses the saved funds to increase the reimbursements on the inpatient or outpatient treatments. By considering individual characteristics and past disease history, ACG risk adjustment ("premium 3") outperforms the first type but is not better than the second cost-adjusted premium in terms of reducing the service-level distortion. As shown in the bottom of Figure 3.4, all shadow prices are close to one, but outpatient care is slightly under-covered no matter which risk adjustment is used.

3.5.4. Service-level coverages between regions

The relative shadow prices in Figure 3.4 are not comparable between regions because they are relative to the shadow prices of inpatient services varying by region. Another notable investigation would be to examine the equity of service coverage between regions. In the following analyses, we normalize Central China's shadow prices to be one and calculate relative shadow prices of each type of services.

Figure 3.5 to 3.7 plot each service type’s relative shadow prices in a comparison between regions. We firstly look at the less information scenario under the supply-side driven financing level (i.e., premium 1). The Northeast region has much higher shadow prices than other regions, but these values are not significantly different from one due to large standard errors (see Table CC-2 in Appendix). In Figure 3.5, we find the East Coast has a significantly higher shadow price of preventive services than the reference region, indicating East Coast’s counties generally under-cover preventive services compared with counties in Central China. The underlying reason may be the difference in population health status. East Coast region has the highest portion of unhealthy population. As a result, for counties in the East Coast, the incentive to attract healthy households by over-covering preventive services may not be as strong as other regions. This observation also supports our theoretical implication that NCMS plan tends to be less generous in a sicker county.

[Figure 3.5 to be here]

Significant shadow prices suggest that the local governments in Northeast undercover inpatient services the most (Figure 3.6), while the Western region tends to undercover outpatient services more than other regions (Figure 3.7). According to a report of NCMS benefit designs by Barber and Yao (2011), many counties separated NCMS funds into a social pooling account and households medical savings accounts (MSAs). While counties in other regions tended to reimburse outpatient services through the social pooling account, the most common reimbursement model in the West region only reimbursed outpatient services from households MSAs. The budget of MSA was very limited and not sufficient for formal medical services (Lei & Lin, 2009). One possible reason for under-covering outpatient treatments in the Western region is that these county governments are poorer (Ping, 2003). Our empirical analyses reveal the highest risk of fund deficits in this region. To conserve funds, county governments in Western China have to restrict NCMS coverage more than counties in other regions due to a small funding pool.

[Figure 3.6 to be here]

Compared with the first type of risk premiums, two risk-adjustment systems seem to enlarge the regional gaps in service-level coverages. If governments set the risk premium as the average NCMS benefit in 2009, Western and East Coast will continue to be less generous than other regions. ACG risk adjustment is helpful to reduce the between-region inequity of insurance benefits, but cannot address the lack of NCMS coverages in the Western region under less information scenario.

[Figure 3.7 to be here]

Risk adjustment appears to be more effective in reducing service-level distortions within a region, but not in addressing the inequality of service coverages between regions. This finding reveals the trade-off between efficiency and equity of service-level coverage. If the goal is to allocate the NCMS benefits more efficiently between services, risk premiums should be adjusted to reduce local governments' incentive to engage in "indirect selection." Adjusting by average NCMS benefit is slightly better than ACG method because the former is easier to implement and has similar results as the latter. However, this strategy may increase inequity in health coverage between regions. If the social priority is to equalize insurance benefits between regions, a minimum lump-sum adjustment on the actual NCMS risk premium works the best. As a final point, the ACG risk adjustment is a compromise between these two goals. It reduces distortions between services within a region and does not induce severe inequality between regions as the second type. One practical drawback is that the ACG algorithm charges more from the sickest population, which may violate the NCMS's objective of increasing social welfare. Additionally, applying ACG risk premium may increase the cost of implementing the NCMS and discourage participation.

3.6. Conclusion

This study investigated service-level coverages of the NCMS program and made three contributions to the literature. First, we developed a theoretical framework to show how population health status and health service characteristics affected local governments' selective incentives. Second, using the 2009 and 2011 CHNS data, we revealed service-level distortions in different regions of China and showed how governments behaved under different information assumptions. Third, different risk-adjustment systems were explored to provide additional details on how to improve the efficiency and equity of NCMS benefit designs.

Results suggested that sample average insurance benefit from the NCMS program exceeded the actual risk premium collected by the program, which revealed potential fund deficits in three regions. This finding highlighted the challenge of financial sustainability faced by the NCMS program, which may explain its discontinuity. Although the NCMS program was merged with URRMI in 2018, we suspect similar problems persist in the new health program if governments' cost-control incentives are not appropriately addressed. The empirical results from CHNS data supported our theoretical implications: less-informed NCMS plan tends to over-cover the services preferred by healthy people, such as preventive services. Although the primary objective of the NCMS was to reduce medical impoverishment, our results suggested that those local governments facing risk of fund deficits had a strong incentive to under-cover outpatient services. This finding may explain why the NCMS failed to provide financial protection to the rural population in the short run, even with high participation rates.

In addition to the supply-side risk premium mimicking the actual NCMS financing level, we evaluated two alternative risk-adjustment methods. One risk premium was flat-rate but reflected previous total NCMS cost, and the other was the ACG risk premium adjusted by individual characteristics and disease history. Both methods attenuated service-level distortions within a region but had the potential to enlarge inequity of service coverages between regions. As China is in the process of unifying rural and urban insurance system, policymakers may need to consider the balance of efficiency and equity in benefit designs to meet residents' healthcare demand. Also, the integration of NCMS and URRMI should be accommodated with a suitable financing level. Otherwise, service-level distortions continue to exist, and public health insurance may benefit the urban wealthy more by sacrificing the welfare of rural poor.

This study has limitations that suggest directions for future research. First, health spending data were reported for only the past four weeks; thus, we may have introduced measurement errors when predicting individuals' monthly health spending and summing them up to annual spending. Further studies using the annual data could produce better estimates. Second, the types of health services reported in the CHNS survey were very rough. Detailed classification of health services is required to provide more insights into the NCMS benefits designs. Researchers may consider using claims data from the NCMS information system if they can access this information. Third, this study treated a household as a single agent. This assumption may be relaxed in future studies to consider potential a conflict of interest between family members. Another concern is that health services may be correlated. For instance, consuming preventive services now may reduce future inpatient treatment costs. However, the correlations between health services are not obvious in a cross-sectional analysis like our case, and it may be considered in dynamic models. Finally, due to the small sample size, we could not run separate empirical models by region; thus, we added regional dummies to capture the heterogeneity in health spending between regions. This model specification may underestimate regional variations. If the sample size is sufficiently large, separated regressions in each region are more appropriate. Nevertheless, the general conclusion of this study still stands: in those areas with income and health disparities, government-led insurance programs have incentives to undercover the health services that residents who are sicker demand most, and appropriate risk adjustments are required to reduce the distortions.

Appendix CA. Solve the first-best optimal shadow price

Solving for the optimal shadow prices in

$$\min_{\boldsymbol{p}} \pi(\boldsymbol{p})^2 = \left\{\sum_i \left[n_i(\widehat{m}_i(\boldsymbol{p}), \hat{u}_i^0, c_i) \times (r_i - \sum_s \widehat{m}_{is}(p_s))\right]\right\}^2$$

We can take the first-order condition

$$\frac{d\pi^2}{dp_s} = 2\pi \sum_i [\frac{dn_i}{dp_s} \times \pi_i - n_i \widehat{m}'_{is}] = 0,$$

$$where\ \pi_i = r_i - \textstyle\sum_s \widehat{m}_{is},\ \widehat{m}'_{is} = \frac{d\widehat{m}_{is}}{dp_s}$$

This equation can be solved if either of the following two conditions is satisfied:

1) $\pi = \sum_i [n_i \times \pi_i] = 0$

This condition can be ruled out because although a local government aims to break-even in NCMS funds, the probability-weighted total program surplus/deficit is unlikely to be zero with the uncertainty of n_i in reality.

2) $\sum_i [\frac{dn_i}{dp_s} \times \pi_i - n_i \widehat{m}'_{is}] = 0$

According to Equation (3.4) and (3.5),

$$\frac{dn_i}{dp_s} = \frac{dn_i}{dv_i} \times \frac{dv_i}{d\widehat{m}_{is}} \times \frac{d\widehat{m}_{is}}{dp_s} = \mathrm{F}'_i p_s \widehat{m}'_{is}$$

Let the elasticity of service s be $e_s = \frac{p_s \widehat{m}'_{is}}{\widehat{m}_{is}}$, and therefore $\widehat{m}'_{is} = e_s \widehat{m}_{is}/p_s$. Then $\frac{dn_i}{dp_s} = \mathrm{F}'_i e_s \widehat{m}_{is}$

Plug it into the above condition, we get

$$\sum_i [\mathrm{F}'_i e_s \widehat{m}_{is} \pi_i - \frac{n_i e_s \widehat{m}_{is}}{p_s}] = 0$$

Assuming p_s and e_s are the same for all households, we multiply both sides by $\frac{p_s}{e_s}$, the first-order condition becomes $p_s \sum_i \mathrm{F}'_i \widehat{m}_{is} \pi_i - \sum_i n_i \widehat{m}_{is} = 0$

So the first-best optimal shadow price is

$$p_s^* = \frac{\sum_i n_i \widehat{m}_{is}}{\sum_i \mathrm{F}_i' \widehat{m}_{is}(r_i - \sum_s \widehat{m}_{is})}$$

Appendix CB. Calculate second-best shadow prices with government uncertainty

$$\min_{\boldsymbol{p}} \pi(\boldsymbol{p})^2 = \left\{\sum_i \left[n_i\left(\widehat{m_i(\boldsymbol{p})}, \hat{u}_i^0, c_i\right) \times (r_i - \sum_s \widehat{m_{is}(p_s)})\right]\right\}^2$$

$$= \left\{\sum_i \left[n_i(\Theta_1 m_{i1}^B(p_1), \Theta_2 m_{i2}^B(p_2), \dots, \Theta_S m_{iS}^B(p_S), \hat{u}_i^0, c_i)\right.\right.$$

$$\left.\left.\times \left(r_i - \sum_s \Theta_S m_{is}^B(p_s)\right)\right]\right\}^2$$

Again, take the first-order condition.

$$\frac{d\pi^2}{dp_s} = 2\pi \sum_i [\frac{dn_i}{dp_s} \times \ (r_i - \sum_s \Theta_s m_{is}^B) - \Theta_s n_i m_{is}^B{}'] = 0, \;\; m_{is}^B{}' = \frac{dm_{is}^B}{dp_s}$$

This condition can be satisfied if

1) $\pi(\boldsymbol{p}) = \sum_i [n_i \times \pi_i] = 0$. This condition can be ruled out for the same reason discussed in Appendix A

2) $\sum_i [\frac{dn_i}{dp_s} \times (r_i - \sum_s \Theta_s m_{is}^B) - \Theta_s n_i m_{is}^B{}'] = 0$

According to Equations (3.4) and (3.9), we have $n_i\left[\widehat{m_i(\boldsymbol{p})}, \hat{u}_i^0, c_i\right] = 1 - \mathrm{F}[\hat{u}_i^0 + c_i - \sum_s v_{is}(\Theta_S m_{is}^B)]$ and assume $v_{is}(.)$ is independent of other services $s' \neq s$.

$$\frac{dn_i}{dp_s} = \frac{dn_i}{dv_{is}} \times \frac{dv_{is}}{d\Theta_s m_{is}^B} \times \frac{d\Theta_s m_{is}^B}{dp_s} = \mathrm{F}_i' p_s \Theta_s m_{is}^B{}'$$

Let the baseline elasticity of service s be $e^s = \frac{p_s m_{is}^{B\prime}}{m_{is}^B}$, and therefore $m_{is}^B{}' = e_s m_{is}^B / p_S$. Then $\frac{dn_i}{dp_s} = \mathrm{F}_i' \Theta_s e_s m_{is}^B$. Plug it into the first-order condition

$$\frac{d\pi}{dp_s} = \sum_i [\mathrm{F}_i' \Theta_s e_s m_{is}^B \times (r_i - \sum_s \Theta_s m_{is}^B) - \frac{\Theta_s n_i e_s m_{is}^B}{p_s}] = 0$$

Given p_s, Θ_S and e_s are the same for all households. We multiply both sides by $\frac{p_s}{e_s \Theta_S}$, and the first-order condition becomes

$$p_s \sum_i [\mathrm{F}_i' m_{is}^B \times (r_i - \sum_s \Theta_s m_{is}^B)] - \sum_i n_i m_{is}^B \;= 0$$

So, the optimal shadow price in the second-best case is $p_s^{second} = \frac{\sum_i n_i m_{is}^B}{\sum_i \mathrm{F}_i' m_{is}^B \cdot (r_i - \sum_s \Theta_s m_{is}^B)}$

Appendix CC. Relative shadow prices results

Table CC-1. Service-level relative shadow prices under two information sets and risk-adjustment systems by region[a,b]

Type of services	Mini. Adj (RMB)	Less-information set			More-information set		
		Preventive services	Inpatient services	Outpatient services	Preventive services	Inpatient services	Outpatient services
Part 1: Expected NCMS risk premium in 2011 with adjustments[c]							
Northeast	40	0.47	1.00	3.46	1.30	1.00	1.82
		(9.63)	-	(1.74)	(0.44)	-	(1.45)
Central	150	0.95	1.00	1.26	1.25	1.00	1.87
		(0.03)	-	(0.01)	(0.11)	-	(3.59)
Western	290	0.48	1.00	1.17	0.98	1.00	1.04
		(0.04)	-	(0.05)	(0.12)	-	(0.10)
East	0	0.92	1.00	1.44	1.23	1.00	1.59
Coast		(0.03)	-	(0.03)	(0.10)	-	(0.15)
Part 2: Risk premium adjusted by the regional mean benefit in 2009							
Northeast		0.93	1.00	1.07	1.02	1.00	1.10
		(0.01)	-	(<0.01)	(0.05)	-	(0.03)
Central		0.95	1.00	1.06	1.02	1.00	1.10
		(0.01)	-	(<0.01)	(0.29)	-	(0.02)
Western		0.71	1.00	1.04	0.97	1.00	1.04
		(0.02)	-	(0.01)	(0.26)	-	(0.05)
East		0.94	1.00	1.09	1.05	1.00	1.15
Coast		(0.02)	-	(<0.01)	(0.05)	-	(0.05)
Part 3: Risk premium adjusted by disease groups and disability status							
Northeast		1.10	1.00	1.16	0.99	1.00	1.11
		(0.01)	-	(0.01)	(0.81)	-	(0.04)
Central		1.11	1.00	1.16	1.18	1.00	1.15
		(0.01)	-	(<0.01)	(4.73)	-	(0.05)
Western		1.01	1.00	1.16	1.01	1.00	1.01
		(0.02)	-	(0.01)	(0.12)	-	(0.06)
East		1.14	1.00	1.22	1.02	1.00	1.15
Coast		(0.02)	-	(0.05)	(0.05)	-	(0.06)

Note: [a]All shadow prices are relative to the category of inpatient services; thus, the shadow prices for the category of inpatient services are normalized to 1.00 in all cases.
[b] Standard errors are in parentheses. We draw 99.9% CHNS respondents randomly in each region and predict $\widehat{m}_{is}$ using equation (3.11) and province averages of insurance coverage percentage. Then, relative shadow prices are calculated in each bootstrap sample. After repeating this process by 1000 times, the standard error is calculated as the standard deviations of bootstrap relative shadow prices.

[c]2011 NCMS risk premium was adjusted upward with a minimum amount to obtain positive shadow prices, and the minimum amounts are reported.

Table CC-2. Between-region relative shadow prices under two information sets and risk-adjustment systems by service[a,b]

Type of services	Less-information set: Central	Northeast	Western	East Coast	More-information set: Central	Northeast	Western	East Coast
	Expected NCMS risk premium in 2011 with adjustments[c]							
Preventive	1.00	4.83	1.63	2.23	1.00	0.85	0.24	0.66
	-	(46.63)	(0.15)	(0.14)	-	(13.11)	(3.89)	(11.04)
Inpatient	1.00	4.56	3.23	2.30	1.00	1.16	0.51	1.00
costs	-	(0.89)	(0.34)	(0.12)	-	(5.26)	(2.14)	(4.47)
Outpatient	1.00	12.71	3.01	2.62	1.00	0.97	0.19	0.63
costs	-	(7.90)	(0.38)	(0.16)	-	(11.82)	(2.61)	(9.30)
	Risk premium adjusted by the regional mean benefit in 2009							
Preventive	1.00	2.38	7.93	4.40	1.00	2.15	4.59	4.00
	-	(0.06)	(0.47)	(0.16)	-	(1.18)	(3.87)	(2.20)
Inpatient	1.00	2.43	10.66	4.45	1.00	2.17	5.05	3.98
costs	-	(0.05)	(0.58)	(0.13)	-	(0.84)	(1.58)	(1.09)
Outpatient	1.00	2.47	10.54	4.61	1.00	2.18	4.79	4.16
costs	-	(0.05)	(0.60)	(0.15)	-	(0.93)	(2.03)	(1.55)
	Risk premium adjusted by disease groups and disability status							
Preventive	1.00	1.32	3.30	3.06	1.00	1.15	2.39	2.68
	-	(0.17)	(0.59)	(1.24)	-	(1.49)	(3.16)	(1.69)
Inpatient	1.00	1.33	3.63	2.96	1.00	1.25	2.45	2.78
costs	-	(0.17)	(0.59)	(1.01)	-	(0.32)	(3.40)	(0.96)
Outpatient	1.00	1.33	3.63	3.16	1.00	1.20	2.09	2.79
costs	-	(0.20)	(0.51)	(1.92)	-	(0.48)	(4.56)	(1.46)

Note: [a]All shadow prices are relative to those prices in Central China; thus, the shadow prices in Central China are normalized to 1.00 in all cases.

[b] Standard errors are in parentheses. We draw 99.9% CHNS respondents randomly in each region and predict $\widehat{m}_{is}$ using equation (3.11) and province averages of insurance coverage percentage. Then, relative shadow prices are calculated in each bootstrap sample. After repeating this process by 1000 times, the standard error is calculated as the standard deviations of bootstrap relative shadow prices.

[c] 2011 NCMS risk premium was adjusted upward with a minimum amount to obtain positive shadow prices.

References

Babiarz, K. S., Miller, G., Yi, H., Zhang, L., & Rozelle, S. (2012). China's new cooperative medical scheme improved finances of township health centers but not the number of patients served. *Health Affairs, 31*(5), 1065-1074.

Barber, S. L., & Yao, L. (2011). Development and status of health insurance systems in China. *The International journal of health planning and management, 26*(4), 339-356.

Berndt, E. R., Cutler, D. M., Frank, R. G., Griliches, Z., Newhouse, J. P., & Triplett, J. E. (2000). Medical care prices and output *Handbook of health economics* (Vol. 1, pp. 119-180): Elsevier.

Breyer, F., Bundorf, M. K., & Pauly, M. V. (2011). Health care spending risk, health insurance, and payment to health plans *Handbook of health economics* (Vol. 2, pp. 691-762): Elsevier.

Brown, J., Duggan, M., Kuziemko, I., & Woolston, W. (2014). How does risk selection respond to risk adjustment? New evidence from the Medicare Advantage Program. *American Economic Review, 104*(10), 3335-3364.

Brown, P. H., De Brauw, A., & Du, Y. (2009). Understanding variation in the design of China's new co-operative medical system. *The China Quarterly, 198*, 304-329.

Cao, Z., & McGuire, T. G. (2003). Service-level selection by HMOs in Medicare. *Journal of health economics, 22*(6), 915-931.

Carey, C. (2017). Technological change and risk adjustment: benefit design incentives in medicare Part D. *American Economic Journal: Economic Policy, 9*(1), 38-73.

Carrin, G., Ron, A., Hui, Y., Hong, W., Tuohong, Z., Licheng, Z., . . . Qicheng, J. (1999). The reform of the rural cooperative medical system in the People's Republic of China: interim experience in 14 pilot counties. *Social Science & Medicine, 48*(7), 961-972.

Chang, H.-Y., Lee, W.-C., & Weiner, J. P. (2010). Comparison of alternative risk adjustment measures for predictive modeling: high risk patient case finding using Taiwan's National Health Insurance claims. *BMC health services research, 10*(1), 343.

Chang, H.-Y., & Weiner, J. P. (2010). An in-depth assessment of a diagnosis-based risk adjustment model based on national health insurance claims: the application of the Johns Hopkins Adjusted Clinical Group case-mix system in Taiwan. *BMC medicine, 8*(1), 7.

CHEI. (2009). *China National Health Accounts Report 2008*. Retrieved from Beijing:

Chen, C. M. (2008). Overview of obesity in Mainland China. *Obesity reviews, 9*, 14-21.

CHSI. (2006). *China's New Rural Cooperative Medical Scheme: Studies of Development and Effectiveness. 2005 Research Reports on Pilots of New Rural Cooperative Medical Scheme (in Chinese). Center for Health Statistics and Information, Ministry of Health, Beijing*.

Davis, G. C., & You, W. J. F. P. (2011). Not enough money or not enough time to satisfy the Thrifty Food Plan? A cost difference approach for estimating a money–time threshold. *36*(2), 101-107.

Eggleston, K., & Bir, A. (2009). Measuring selection incentives in managed care: evidence from the Massachusetts state employee insurance program. *Journal of Risk and Insurance, 76*(1), 159-175.

Einav, L., Finkelstein, A., & Cullen, M. R. (2010). Estimating welfare in insurance markets using variation in prices. *The quarterly journal of economics, 125*(3), 877-921.

Ellis, R. P., Jiang, S., & Kuo, T.-C. (2013). Does service-level spending show evidence of selection across health plan types? *Applied Economics, 45*(13), 1701-1712.
Ellis, R. P., & McGuire, T. G. (2007). Predictability and predictiveness in health care spending. *Journal of health economics, 26*(1), 25-48.
Frank, R. G., Glazer, J., & McGuire, T. G. (2000). Measuring adverse selection in managed health care. *Journal of health economics, 19*(6), 829-854.
Geruso, M., Layton, T. J., & Prinz, D. (2016). *Screening in contract design: evidence from the Aca health insurance exchanges*. Retrieved from
Gu, X. (2011). 6. Official Start of Medical and Health Care System Reform in China *The China Society Yearbook, Volume 5* (pp. 105-120): BRILL.
Horvitz, D. G., & Thompson, D. J. J. J. o. t. A. s. A. (1952). A generalization of sampling without replacement from a finite universe. *47*(260), 663-685.
Hou, Z., Van de Poel, E., Van Doorslaer, E., Yu, B., & Meng, Q. (2014). Effects of NCMS on access to care and financial protection in China. *Health economics, 23*(8), 917-934.
Jin, Y., Hou, Z., & Zhang, D. J. P. O. (2016). Determinants of health insurance coverage among people aged 45 and over in China: who buys public, private and multiple insurance. *11*(8), e0161774.
Keeler, E. B., Carter, G., & Newhouse, J. P. (1998). A model of the impact of reimbursement schemes on health plan choice. *Journal of health economics, 17*(3), 297-320.
Kowalski, A. (2016). Censored quantile instrumental variable estimates of the price elasticity of expenditure on medical care. *Journal of Business & Economic Statistics, 34*(1), 107-117.
Layton, T. J., Ellis, R. P., McGuire, T. G., & Van Kleef, R. (2017). Measuring efficiency of health plan payment systems in managed competition health insurance markets. *Journal of health economics, 56*, 237-255.
Lei, X., & Lin, W. (2009). The new cooperative medical scheme in rural China: Does more coverage mean more service and better health? *Health economics, 18*(S2), S25-S46.
Liu, P., Guo, W., Liu, H., Hua, W., & Xiong, L. (2018). The integration of urban and rural medical insurance to reduce the rural medical burden in China: a case study of a county in Baoji City. *BMC health services research, 18*(1), 796.
Liu, Y. (2004). Development of the rural health insurance system in China. *Health policy and planning, 19*(3), 159-165.
Manning, W., & Norton, E. (2013). *Modeling Health Care Costs and Counts.* Paper presented at the iHEA World Congress in Sydney, Australia.
McGuire, T. G., Newhouse, J. P., Normand, S.-L., Shi, J., & Zuvekas, S. (2014). Assessing incentives for service-level selection in private health insurance exchanges. *Journal of health economics, 35*, 47-63.
Meng, Q., & Xu, K. (2014). Progress and challenges of the rural cooperative medical scheme in China. *Bulletin of the World Health Organization, 92*(6), 447-451.
Newhouse, J. P., Price, M., Hsu, J., McWilliams, J. M., & McGuire, T. G. (2015). How much favorable selection is left in Medicare Advantage? *American journal of health economics, 1*(1), 1-26.
NHFPC. (2015). *National statistical report on health and family planning development*. Retrieved from
NHFPC. (2016). *The National Health and Family Planning Commission of the People's Republic of China, Explanation of the guideline on the integration of basic medical*

insurance for urban and rural residents (in Chinese). Retrieved from http://www.nhc.gov.cn/tigs/s7847/201601/b5acdfc387a94ac3a287d475cde4fb22.shtml.

Ping, X. (2003). The Selection of Financing System of Medical and Health Service in the Rural Areas-a View Based on the Expending Behavior in Medical and Health Work for China's Peasants [J]. *Management World, 11*, 006.

Pope, G. C., Kautter, J., Ellis, R. P., Ash, A. S., Ayanian, J. Z., Iezzoni, L. I., . . . Robst, J. (2004). Risk adjustment of Medicare capitation payments using the CMS-HCC model. *Health care financing review, 25*(4), 119.

Qian, D., Pong, R. W., Yin, A., Nagarajan, K., & Meng, Q. (2009). Determinants of health care demand in poor, rural China: the case of Gansu Province. *Health policy and planning, 24*(5), 324-334.

Rothschild, M., & Stiglitz, J. (1978). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information *Uncertainty in economics* (pp. 257-280): Elsevier.

Shen, J., & Liu, Y. (2016). Skilled and less-skilled interregional migration in China: A comparative analysis of spatial patterns and the decision to migrate in 2000–2005. *Habitat International, 57*, 1-10.

Sidel, V. W. (1993). New lessons from China: equity and economics in rural health care. *Am J Public Health, 83*(12), 1665-1666.

StataCorp LP. (2015). Stata data analysis and statistical Software 14.0.

Su, D., Chen, Y., Gao, H., Li, H., Chang, J., Jiang, D., . . . Chen, Z. (2019). Effect of integrated urban and rural residents medical insurance on the utilisation of medical services by residents in China: a propensity score matching with difference-in-differences regression approach. *BMJ open, 9*(2), bmjopen-2018-026408.

Su, M., Si, Y., & Zhou, Z. (2017). The effects of China's three basic health insurance schemes on the equity of health-related quality of life: a cross-sectional survey using coarsened exact matching. *The Lancet, 390*, S73.

Wagstaff, A., Lindelow, M., Jun, G., Ling, X., & Juncheng, Q. (2009). Extending health insurance to the rural population: an impact evaluation of China's new cooperative medical scheme. *Journal of health economics, 28*(1), 1-19.

Wang, H., Zhang, L., Yip, W., & Hsiao, W. (2006). Adverse selection in a voluntary Rural Mutual Health Care health insurance scheme in China. *Social Science & Medicine, 63*(5), 1236-1245.

Weiner, J. P., Dobson, A., Maxwell, S. L., Coleman, K., Starfield, B. H., & Anderson, G. F. (1996). Risk-adjusted Medicare capitation rates using ambulatory and inpatient diagnoses. *Health care financing review, 17*(3), 77.

Wong, V. C., & Chiu, S. W. J. I. J. o. P. S. M. (1997). Health-care reforms in the People's Republic of China: strategies and social implications. *10*(1/2), 76-92.

Xu, W., Cai, G. J., Li, G. N., Cao, J. J., Shi, Q. H., & Bai, J. (2016). Age or health status: which influences medical insurance enrollment greater? *Journal of global health, 6*(2).

Xu, X., Byles, J. E., Shi, Z., & Hall, J. J. (2015). Evaluation of older Chinese people's macronutrient intake status: results from the China Health and Nutrition Survey. *British Journal of Nutrition, 113*(1), 159-171.

Yip, W., & Hsiao, W. C. (2009). Non-evidence-based policy: how effective is China's new cooperative medical scheme in reducing medical impoverishment? *Social Science & Medicine, 68*(2), 201-209.

You, X., & Kobayashi, Y. (2009). The new cooperative medical scheme in China. *Health policy, 91*(1), 1-9.

Yu, B., Meng, Q., Collins, C., Tolhurst, R., Tang, S., Yan, F., . . . Liu, X. (2010). How does the New Cooperative Medical Scheme influence health service utilization? A study in two provinces in rural China. *BMC health services research, 10*(1), 116.

Yu, H. (2017). China's medical savings accounts: an analysis of the price elasticity of demand for health care. *The European Journal of Health Economics, 18*(6), 773-785.

Zhang, L., Cheng, X., Liu, X., Zhu, K., Tang, S., Bogg, L., . . . Tolhurst, R. (2010). Balancing the funds in the New Cooperative Medical Scheme in rural China: determinants and influencing factors in two provinces. *The International journal of health planning and management, 25*(2), 96-118.

Zhang, Y., Chen, Y., Zhang, X., & Zhang, L. (2014). Current level and determinants of inappropriate admissions to township hospitals under the new rural cooperative medical system in China: a cross-sectional study. *BMC health services research, 14*(1), 649.

Zhao, Y., Ash, A. S., Ellis, R. P., Ayanian, J. Z., Pope, G. C., Bowen, B., & Weyuker, L. (2005). Predicting pharmacy costs and other medical costs using diagnoses and drug claims. *Medical care*, 34-43.

Zhong, H. (2011). Effect of patient reimbursement method on health-care utilization: evidence from China. *Health economics, 20*(11), 1312-1329.

Zhu, K., Zhang, L., Yuan, S., Zhang, X., & Zhang, Z. (2017). Health financing and integration of urban and rural residents' basic medical insurance systems in China. *International journal for equity in health, 16*(1), 194.

Zhu, N., Ling, Z., Shen, J., Lane, J., & Hu, S. J. B. o. t. W. H. O. (1989). Factors associated with the decline of the Cooperative Medical System and barefoot doctors in rural China. *67*(4), 431.

Figures and Tables

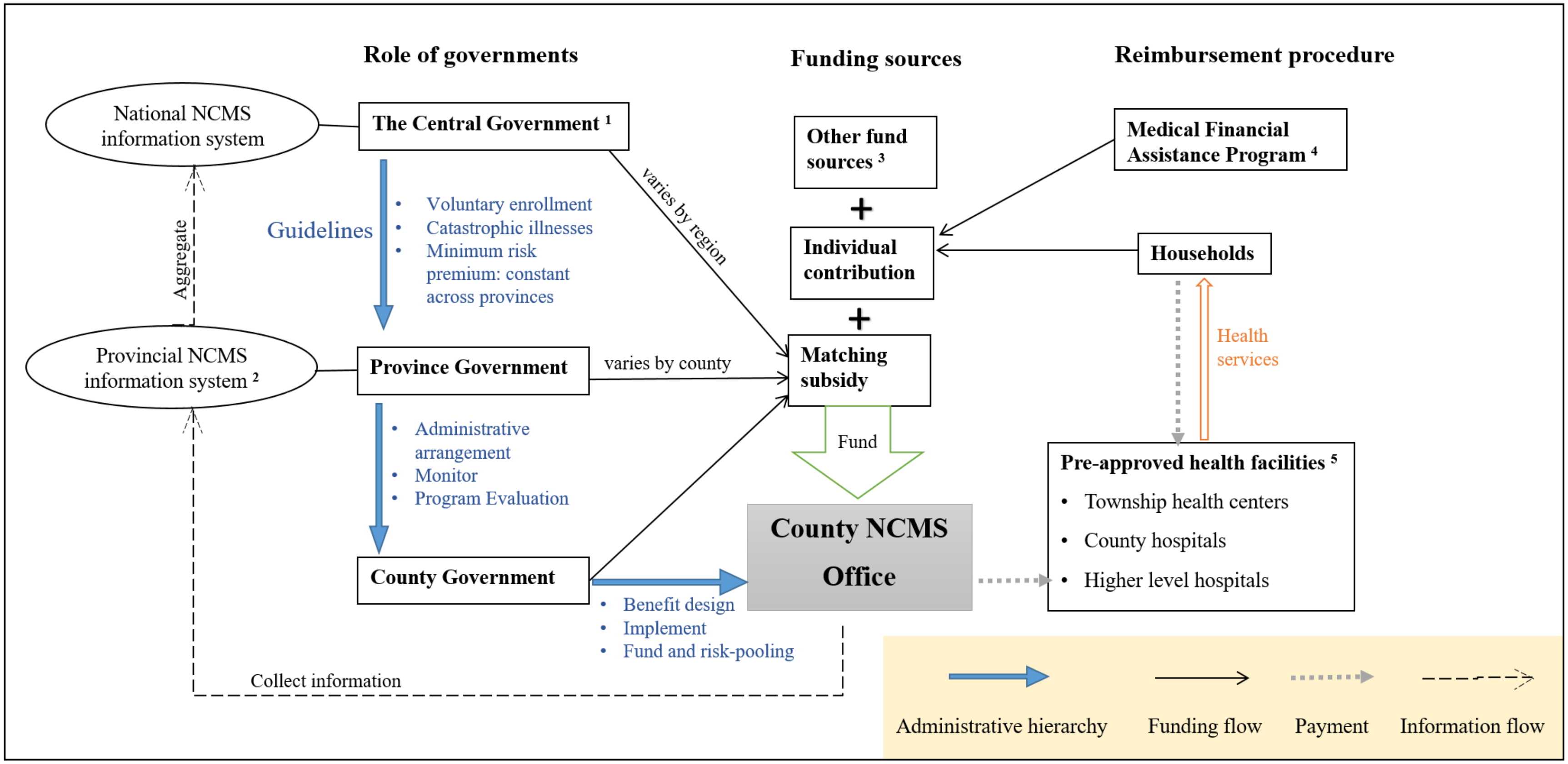

Figure 3.1 Illustration of China's NCMS infrastructure.

Note: [1] State Council-level authority offered leadership and worked with the Ministry of Health to develop general guidelines.
[2] The NCMS information system collected each county's socioeconomic data, participants' health spending and insurance benefits, NCMS fund management information, and evaluation of program performance.
[3] Local state-owned enterprises and social organizations were encouraged to contribute to the NCMS funds, but their portion was tiny and often ignored in previous studies.
[4] Medical Financial Assistance program covered individual contributions for the impoverished and vulnerable households, including those with *Wu Bao*, *Te Kun* or *Di Bao* status.

[5] Reimbursement rates varied by types of health facilities and by total medical expenditures

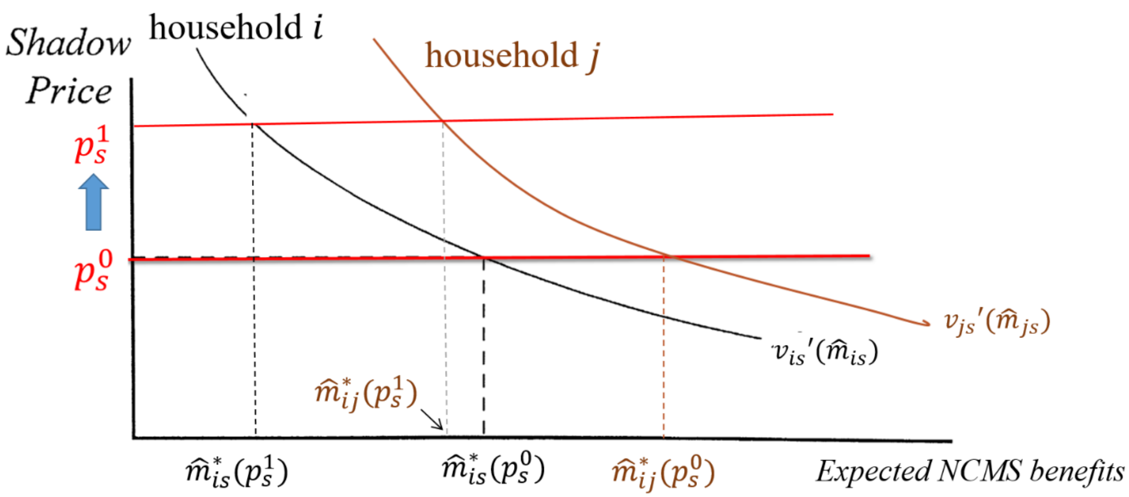

Figure 3.2 Illustration of shadow price and service coverage

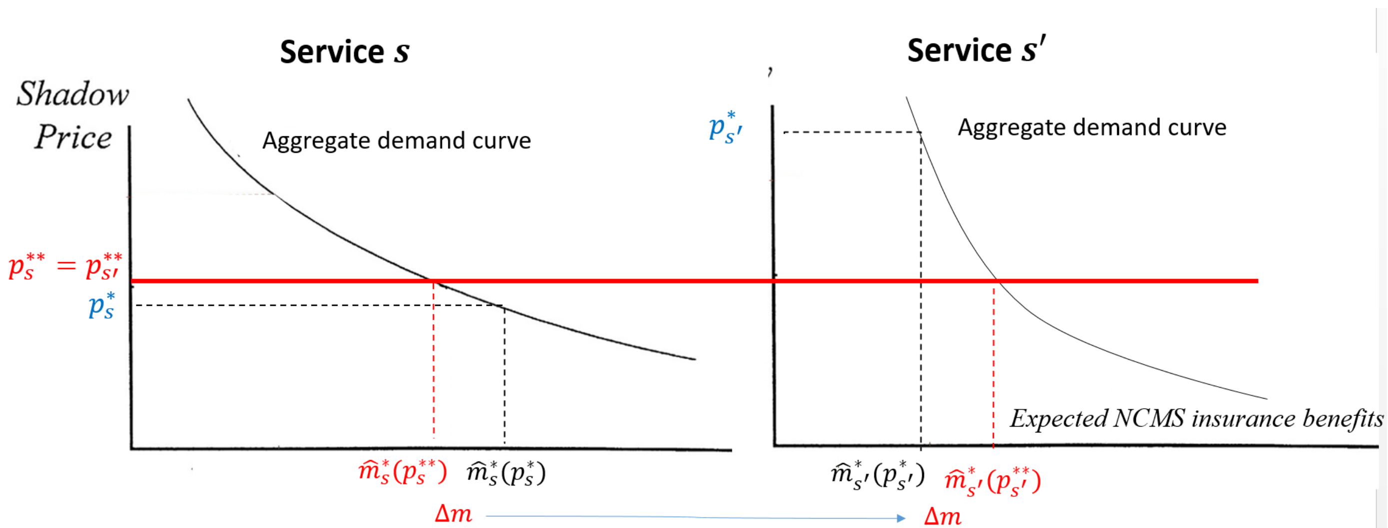

Figure 3.3 Illustration of service-level distortion and social optimal condition

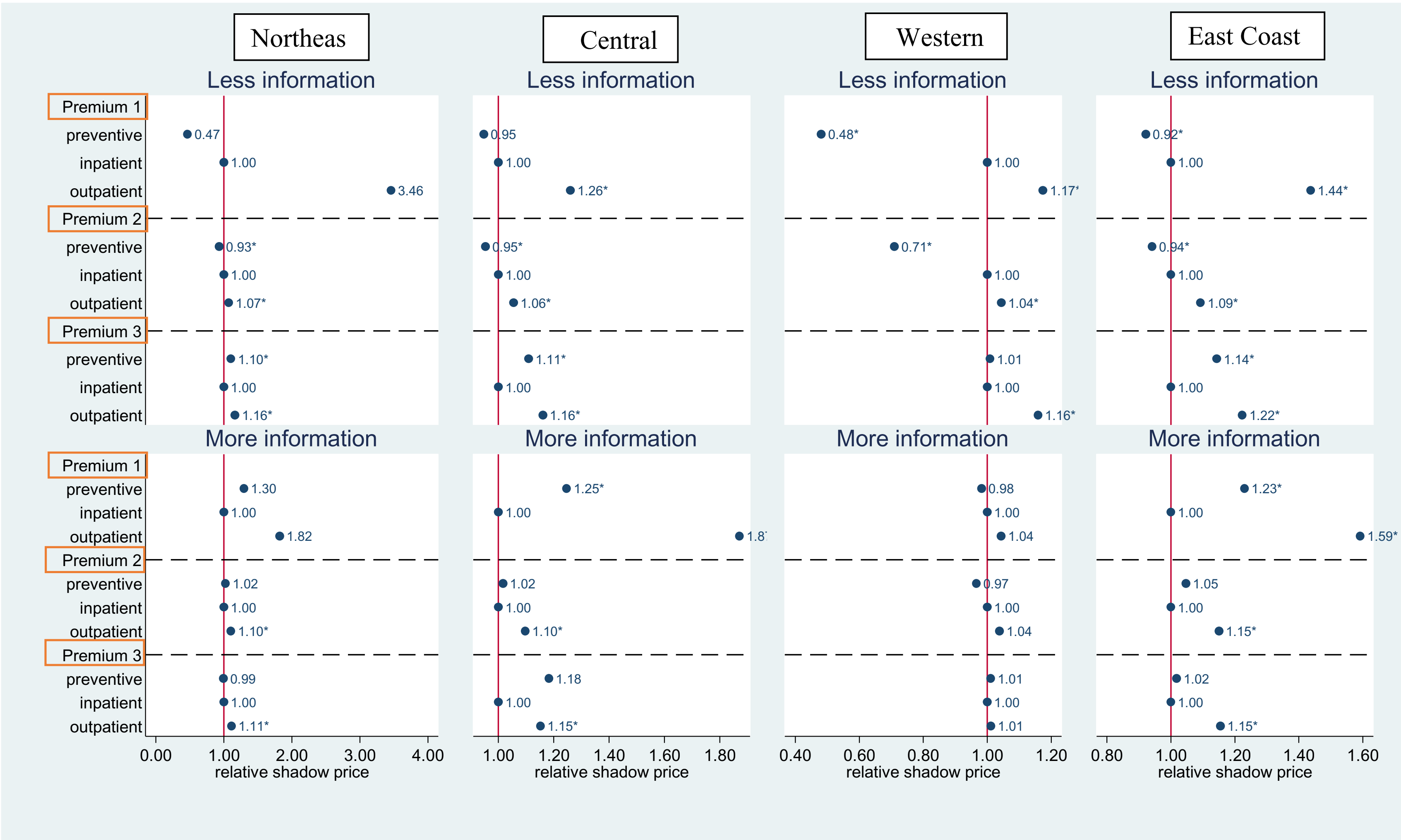

Figure 3.4 Service-level relative shadow prices under two information sets and risk-adjustment systems by region

Note: the shadow prices for the category of inpatient services are normalized to 1, and other relative shadow prices with an asterisk are significantly different from 1.

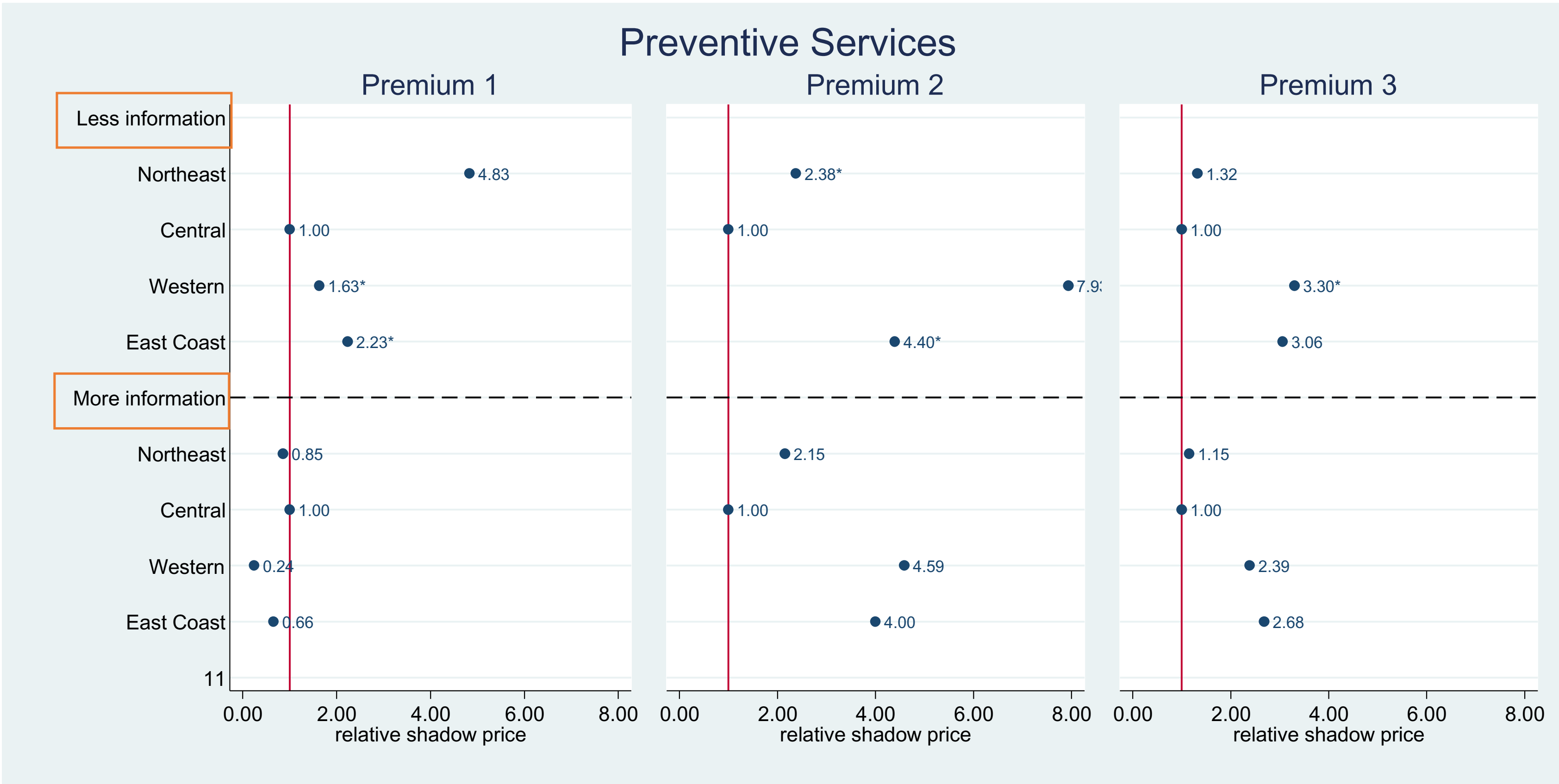

Figure 3.5 Relative shadow prices of preventive services under two information sets and risk-adjustment systems

Note: All shadow prices in Central China are normalized to 1, and other relative shadow prices with an asterisk are significantly different from 1.

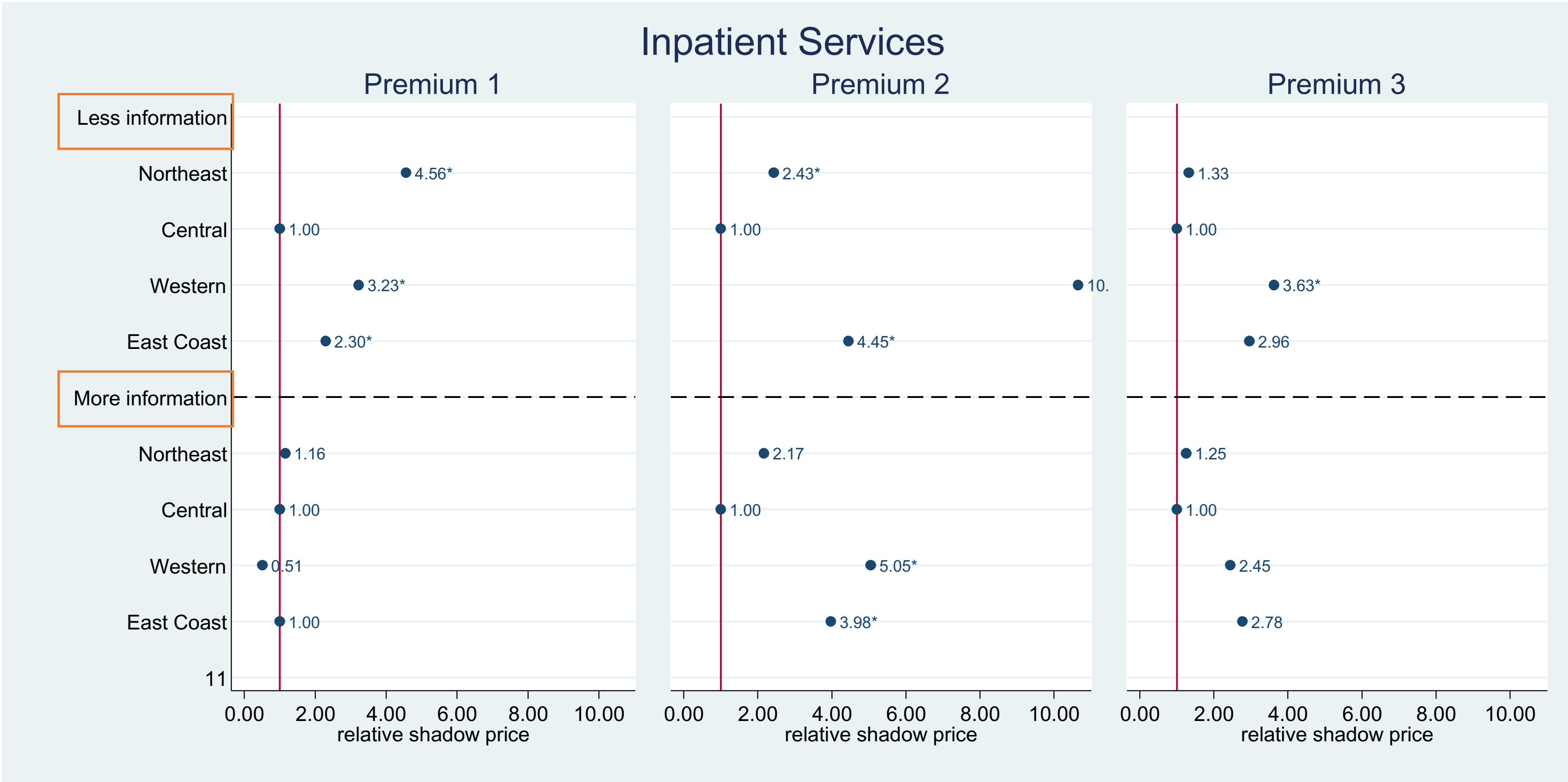

Figure 3.6 Relative shadow prices of inpatient services under two information sets and risk-adjustment systems

Note: All shadow prices in Central China are normalized to 1, and other relative shadow prices with an asterisk are significantly different from 1.

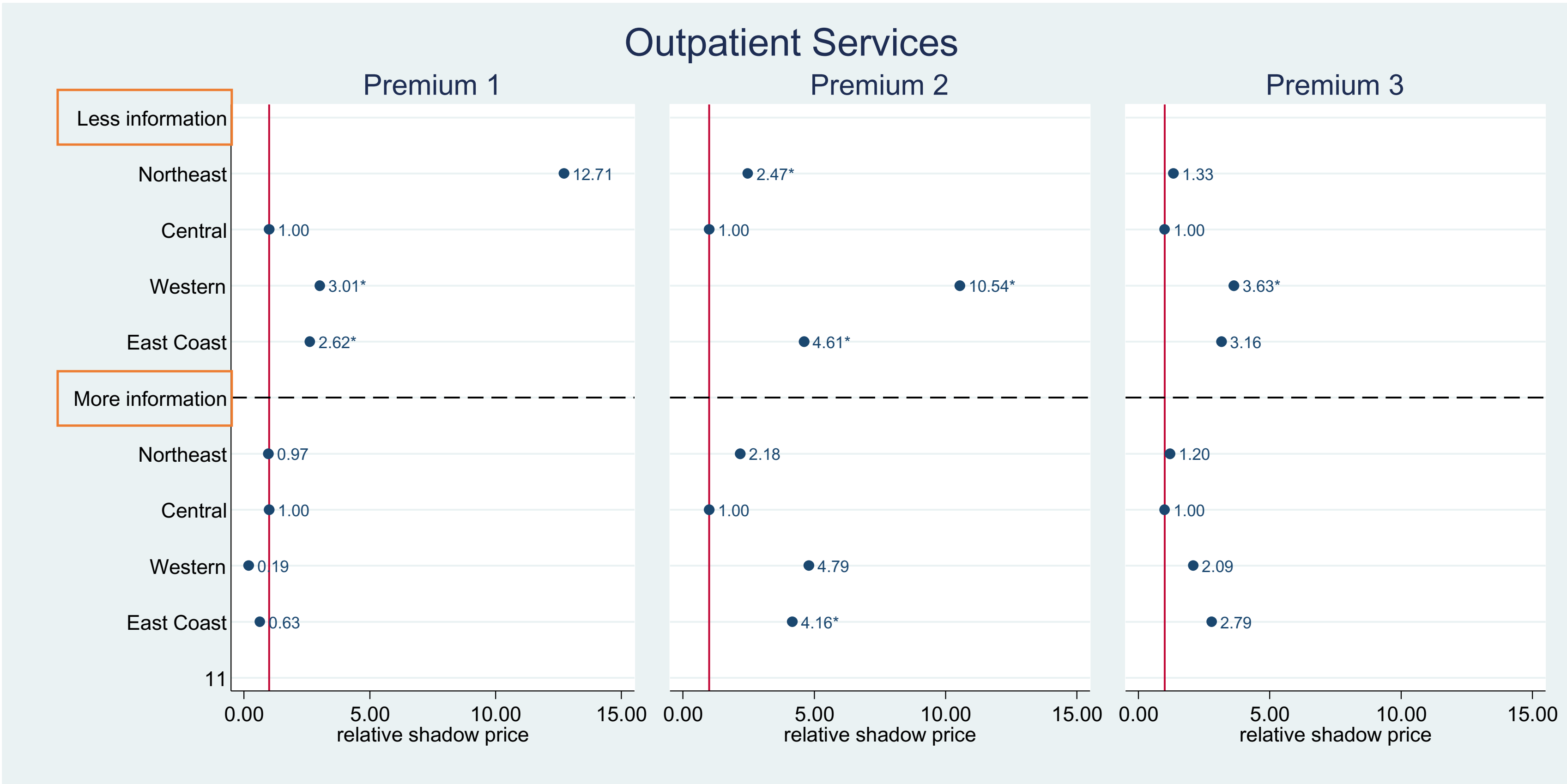

Figure 3.7 Relative shadow prices of outpatient services under two information sets and risk-adjustment systems

Note: All shadow prices in Central China are normalized to 1, and other relative shadow prices with an asterisk are significantly different from 1.

Table 3.1 Summary statistics for selected key variables (2011)

Region		Northeast	Central	Western	East Coast	p-value
Healthcare Spending (RMB) in the past four weeks						
Inpatient services	Mean	132.02	104.23	77.33	47.87	0.623
	Std. Dev.	1932.93	1898.44	1405.77	679.45	
Outpatient services	Mean	77.28	69.88	46.69	47.80	0.733
	Std. Dev.	1139.29	1353.27	391.09	413.46	
Preventive services	Mean	0.96	1.78	1.73	0.18	0.229
	Std. Dev.	18.49	28.62	23.28	5.73	
Total health spending	Mean	223.94	182.77	142.58	102.44	0.420
	Std. Dev.	2240.34	2327.83	1463.42	790.69	
% covered by NCMS						
Inpatient services	Mean	52.86	55.36	37.15	33.36	0.036
	Std. Dev.	26.80	29.69	28.30	25.78	
Outpatient services	Mean	13.00	13.25	14.01	4.65	0.032
	Std. Dev.	26.07	25.77	26.67	14.06	
Preventive services	Mean	7.69	15.43	7.27	2.58	0.092
	Std. Dev.	27.74	34.44	22.31	14.37	
Benefit of NCMS (RMB) in the past four weeks						
Total benefit	Mean	78.48	70.03	38.81	21.15	0.282
	Std. Dev.	897.30	1182.16	666.03	271.53	
Demographic characteristics						
Age	Mean	46.39	46.20	43.75	49.42	<0.001
	Std. Dev.	15.94	19.17	20.81	17.81	
Gender	Mean	0.52	0.55	0.52	0.56	0.135
	Std. Dev.	0.50	0.50	0.50	0.50	
Years of education	Mean	6.79	6.31	5.76	5.96	<0.001
	Std. Dev.	3.16	3.95	3.85	4.12	
Total net household income	Mean	38,371	31,604	32,142	43,551	<0.001
	Std. Dev.	38,337	44,781	31,871	42,299	
Number of Observations		1119	2186	2563	990	

Note: Beijing and Shanghai are excluded in our analysis because their NCMS benefit designs and risk premiums are not comparable with other provinces in the same regions.

Table 3.2 Predicted and actual risk premiums in 2011 by province.

Province： region	2010	% of r_{min}	2011 predicted	2011 actual collected
Liaoning: Northeast China	158.4	105.60%	242.9	234.9
Heilongjiang: Northeast China	151.2	100.80%	231.8	230.6
Jiangsu: East Coast	192.0	128.00%	294.4	273.0
Shandong: East Coast	135.2	90.13%	207.3	256.2
Henan: Central China	150.6	100.40%	230.9	231.4
Hubei: Central China	150.3	100.20%	230.5	235.2
Hunan: Central China	141.2	94.13%	216.5	231.1
Guangxi: Western China	150.4	100.27%	230.6	230.6
Guizhou: Western China	146.4	97.60%	224.5	225.4
Chongqing: Western China	141.5	94.33%	217.0	232.0
Minimum risk premium (r_{min})	150		230	

Source: provinces' actual average risk premiums were obtained from the tables of Conditions of New Cooperative Medical System by Region in China Statistical Yearbook, 2010-2012 (http://www.stats.gov.cn/english/Statisticaldata/AnnualData/)

Table 3.3 Classification of CHNS disease history information

Disease Categories	**Description**	**Examples of CHNS recorded diseases**
DC1	Time-limited, major	myocardial infarction, stroke
DC2	Chronic medical	high blood pressure, diabetes
DC3	Malignancy	all types of cancers
DC4	Others	asthma, bone fracture

Table 3.4 Benefit distribution between healthy and unhealthy households

Type of services	Less-information set: Northeast	Central	Western	East Coast	More-information set: Northeast	Central	Western	East Coast
Percent of the unhealthy population (λ)								
Self-reported	0.17	0.13	0.10	0.18				
Exam-based	0.37	0.36	0.26	0.40				
Discrepancy in insurance benefit ($\Delta\theta_s = \overline{\theta_s} - \underline{\theta_s}$)								
Preventive	-0.04	-0.03	-0.04	-0.03	0.19	0.19	0.19	0.19
Services	(0.07)	(0.05)	(0.07)	(0.05)	(0.09)	(0.10)	(0.09)	(0.09)
Inpatient	0.32	0.33	0.32	0.33	0.28	0.28	0.28	0.29
costs	(0.05)	(0.05)	(0.05)	(0.05)	(0.03)	(0.03)	(0.03)	(0.03)
Outpatient	0.15	0.14	0.13	0.15	0.27	0.24	0.28	0.25
costs	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)

*Note: S*tandard errors are in parentheses. We draw 99.9% CHNS respondents randomly in each region and calculate $\Delta\theta_s$ in each bootstrap sample. After repeating this process by 1000 times, the standard error is calculated as the standard deviations of bootstrap estimates of $\Delta\theta_s$.

Table 3.5 Predictability and predictiveness of three services.

Type of services	Less-information set: Northeast	Central	Western	East Coast	More-information set: Northeast	Central	Western	East Coast
Preventive services								
$cov(\widehat{m}_{is})$	1.09	1.09	1.09	1.10	12.73	12.76	12.90	11.46
	(0.03)	(0.03)	(0.04)	(0.04)	(3.18)	(2.51)	(2.73)	(2.85)
ρ_{m_{is},π_i}	-0.31	-0.33	-0.35	-0.23	-0.13	-0.11	-0.38	-0.08
	(0.03)	(0.02)	(0.03)	(0.02)	(0.10)	(0.09)	(0.10)	(0.10)
Inpatient services								
$cov(\widehat{m}_{is})$	0.60	0.60	0.59	0.60	5.98	5.11	3.62	8.00
	(0.01)	(0.01)	(0.01)	(0.01)	(3.90)	(3.76)	(4.35)	(3.71)
ρ_{m_{is},π_i}	-0.65	-0.61	-0.89	-0.48	-0.70	-0.71	-0.82	-0.78
	(0.02)	(0.01)	(<0.01)	(0.01)	(0.08)	(0.07)	(0.11)	(0.08)
Outpatient services								
$cov(\widehat{m}_{is})$	0.94	0.93	0.92	0.93	7.14	5.11	8.32	5.94
	(0.00)	(0.00)	(0.00)	(0.00)	(0.46)	(0.29)	(0.19)	(0.41)
ρ_{m_{is},π_i}	-0.71	-0.72	-0.67	-0.58	-0.67	-0.68	-0.38	-0.59
	(<0.01)	(0.01)	(0.01)	(0.01)	(0.10)	(0.09)	(0.06)	(0.08)

*Note: S*tandard errors are in parentheses. We draw 99.9% CHNS respondents randomly in each region and predict $\widehat{m}_{is}$ using equation (3.11) and province averages of insurance coverage percentage in each bootstrap sample. After repeating this process by 1000 times, the standard error is calculated as the standard deviations of bootstrap $cov(\widehat{m}_{is})$ and ρ_{m_{is},π_i}