THE UNIVERSITY OF CHICAGO

THREE ESSAYS IN ENERGY AND ENVIRONMENT AND DEVELOPMENT ECONOMICS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE IRVING B. HARRIS
GRADUATE SCHOOL OF PUBLIC POLICY STUDIES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

BY
YUVRAJ PATHAK

CHICAGO, ILLINOIS
JUNE 2020

I dedicate this dissertation to the people who matter the most:

To my parents, Maya Pathak and Ved Prakash Pathak, who supported my dreams and gave me everything they could.

To Moire Corcoran: thank you for your love and patience.

TABLE OF CONTENTS

LIST OF FIGURES . . . vi

LIST OF TABLES . . . vii

ACKNOWLEDGMENTS . . . ix

ABSTRACT . . . x

1 WOMEN'S POLITICAL RESERVATION, EARLY CHILDHOOD DEVELOPMENT, AND LEARNING IN INDIA . . . 1
- 1.1 Introduction . . . 1
- 1.2 Data, Identification, and Empirical Strategy . . . 6
  - 1.2.1 The Reservation Policy and the Cohort of Interest . . . 6
  - 1.2.2 Random Rotation of Reservation and Balance . . . 10
  - 1.2.3 Empirical Specification . . . 16
- 1.3 Long-Term Impacts of Early Exposure to Reservation . . . 18
- 1.4 Impacts on Other Cohorts . . . 24
- 1.5 Conclusion . . . 27
- References . . . 29

2 HOUSEHOLD EMPLOYMENT AND CHILD DEVELOPMENT . . . 35
- 2.1 Introduction . . . 35
- 2.2 Background on NREGA . . . 38
  - 2.2.1 Administrative Structure in Rural India . . . 38
  - 2.2.2 National Rural Employment Guarantee Act (NREGA) . . . 38
- 2.3 Data . . . 39
  - 2.3.1 Pre-program Characteristics of Early and Late Districts . . . 41
- 2.4 Econometric Framework . . . 44
  - 2.4.1 Cross section OLS . . . 44
  - 2.4.2 Difference in Differences (DID) . . . 45
  - 2.4.3 Child fixed effects . . . 49
- 2.5 Results . . . 50
  - 2.5.1 Cross Section Results: Early vs Late Districts . . . 50
  - 2.5.2 DID Results . . . 52
  - 2.5.3 Child fixed effects . . . 54
  - 2.5.4 Importance of Early Childhood . . . 56
- 2.6 Discussion . . . 57
  - 2.6.1 Food Intake . . . 59
  - 2.6.2 Time Allocation . . . 60
- 2.7 Conclusion . . . 62
- References . . . 63

3 SAVING FOR TOMORROW: COAL SUPPLY DISTORTIONS, STOCKPILING, AND POWER OUTAGES IN INDIA . . . 65
3.1 Introduction . . . 65
3.2 Background on Coal and Power sectors in India . . . 71
3.2.1 Brief History of Coal Allocation in India . . . 71
3.2.2 *Coal-gate*: Judicial decision led to future uncertainty . . . 74
3.3 Theoretical Framework . . . 79
3.4 Data . . . 85
3.5 Empirical Strategy . . . 89
3.6 Results . . . 93
3.6.1 Short term results . . . 93
3.6.2 Longer term results . . . 105
3.7 Welfare Analysis . . . 114
3.7.1 Change in Welfare: Lower Bound . . . 116
3.7.2 Change in Welfare: Upper Bound . . . 117
3.8 Robustness Checks . . . 119
3.9 Conclusion . . . 120
3.10 APPENDIX A . . . 122
3.11 APPENDIX B . . . 125
3.12 APPENDIX C . . . 136
3.13 APPENDIX D . . . 137
References . . . 144

LIST OF FIGURES

1.1 Timeline of election and survey rounds and children's age 8
1.2 Nutritional status and test scores in 2009 for children born in GPs reserved in 1995 versus 2006. 19

2.1 Early vs Late Districts: Average GDP over time prior to NREGA roll out 46
2.2 Early vs Late Districts: Average GDP per capita over time prior to NREGA roll out 47
2.3 Early vs Late Districts: Average number of Primary Health Care centers over time prior to NREGA roll out 48
2.4 Early vs Late Districts: Average Female Literacy Rate over time prior to NREGA roll out 48
2.5 Early vs Late Districts: Average Infant Mortality Rate (IMR) over time prior to NREGA roll out 49
2.6 Effect of NREGA on Height for Age by age in year 2006 56
2.7 Effect of NREGA on Height for Age by age category in year 2006 57

3.1 Total power supply by generators and demand by utilities 72
3.2 An illustration of how *coalgate* led to policy uncertainty 76
3.3 Map of all coal fired power plants in India 88
3.4 Average daily outages due to coal shortage by treatment and control 92
3.5 Monthly treatment effect on capacity under outage 96
3.6 Monthly treatment effect on outages reported due to fuel shortage 98
3.7 Time trend of ATT estimate of uncertainty on total coal received 100
3.8 Evolution of treatment effect on end of month coal stock 104
3.9 Time trend of ATT estimate of effect on outages 109
3.10 Welfare Analysis: An illustrative example 115
3.11 Increased generation cost due to uncertain fuel supply 117
D.1 Loss in producer surplus due to uncertain fuel supply - September 2014 137
D.2 Loss in producer surplus due to uncertain fuel supply - October 2014 138
D.3 Loss in producer surplus due to uncertain fuel supply - November 2014 139
D.4 Loss in producer surplus due to uncertain fuel supply - December 2014 140
D.5 Loss in producer surplus due to uncertain fuel supply - January 2015 141
D.6 Loss in producer surplus due to uncertain fuel supply - February 2015 142
D.7 Loss in producer surplus due to uncertain fuel supply - March 2015 143

LIST OF TABLES

1.1 Selection into sample: characteristics of villages, households, and children with and without election data . . . 11
1.2 Descriptive Statistics in 2002 for the GPs reserved in 1995 and 2006 . . . 13
1.3 Descriptive Statistics in 2002 for the GPs reserved in 2001 and 2006 . . . 15
1.4 Differential long-term effect of 1995 reservation on test scores and nutritional status in 2009: children in utero during 1995–2001 reservation (12–17 months in fall 2002) . . . 20
1.5 Medium and Short-term effects of 1995 reservation on nutritional and health status in 2007 and 2002: children in utero during 1995–2001 reservation (12–17 months in fall 2002) . . . 21
1.6 Differences between 1995 and 2006 reservation groups on schooling outcomes in 2009: children in utero during 1995–2001 reservation (12–17 months in fall 2002) 22
1.7 Correlation of tests scores with educational variables . . . 23
1.8 Differential long-term effect of 1995 reservation on test scores and nutritional status in 2009: children born after 1995–2001 reservation (6–11 months in fall 2002) . . . 24
1.9 Medium-term effects of 1995 reservation on nutritional and health status in 2007 and 2002: children born after 1995–2001 reservation (6–11 months in fall 2002) . 25
1.10 Differential long-term effect of 2001 reservation on test scores and nutritional status in 2009 . . . 26

2.1 Pre-program Characteristics: Child characteristics in 2002 . . . 41
2.2 Pre-program Characteristics: Vaccination and Prenatal care variables in 2002 . . 42
2.3 Pre-program Characteristics: Mother and Household Characteristics in 2002 . . 43
2.4 Cross-section: Height for Age Z-Score . . . 51
2.5 DID: Height for Age Z-Score . . . 53
2.6 Child fixed effects: Height for Age Z-Score . . . 55
2.7 Effect of NREGA on HFA by Age Category . . . 58
2.8 Mechanism: Food Intake . . . 60
2.9 Mechanism: Time Allocation . . . 61

3.1 Baseline means comparison of plants facing no uncertainty and plants facing uncertainty due to SC decision . . . 91
3.2 Short term effect of uncertainty on daily generation capacity under outage . . . 94
3.3 Short term effect of uncertainty on capacity under outage due to reported fuel shortage . . . 97
3.4 Short term effect of uncertainty on monthly receipts of coal . . . 99
3.5 Short term effect of uncertainty on monthly consumption of coal . . . 102
3.6 Difference in difference estimates of the effect of uncertainty on monthly precautionary coal savings . . . 103
3.7 Short term effect of uncertainty on electricity generation . . . 106
3.8 Longer term effect of mining contract cancellation on daily generation capacity under outage . . . 108

3.9 Longer term effect of mining contract cancellation on monthly receipts of coal . 110
3.10 Longer term effect of mining contract cancellation on monthly consumption of coal 111
3.11 Longer term effect of mining contract cancellation on monthly coal saving 112
3.12 Longer term effect of mining contract cancellation on electricity generation . . . 113
A.1 Short term effect of uncertainty on daily generation capacity under outage (FULL SAMPLE) 122
A.2 Short term effect of uncertainty on capacity under outage due to reported fuel shortage (FULL SAMPLE) 123
A.3 Longer term effect of mining contract cancellation on daily generation capacity under outage (FULL SAMPLE) 124
B.1 Short term effect of uncertainty on daily generation capacity under outage (BOOTSTRAP SE) 125
B.2 Short term effect of uncertainty on capacity under outage due to reported fuel shortage (BOOTSTRAP SE) 126
B.3 Short term effect of uncertainty on monthly receipts of coal (BOOTSTRAP SE) 127
B.4 Short term effect of uncertainty on monthly consumption of coal (BOOTSTRAP SE) 128
B.5 Difference in difference estimates of the effect of uncertainty on monthly precautionary coal savings (BOOTSTRAP SE) 129
B.6 Short term effect of uncertainty on electricity generation (BOOTSTRAP SE) . . 130
B.7 Longer term effect of mining contract cancellation on daily generation capacity under outage (BOOTSTRAP SE) 131
B.8 Longer term effect of mining contract cancellation on monthly receipts of coal (BOOTSTRAP SE) 132
B.9 Longer term effect of mining contract cancellation on monthly consumption of coal (BOOTSTRAP SE) 133
B.10 Longer term effect of mining contract cancellation on monthly coal saving (BOOTSTRAP SE) 134
B.11 Longer term effect of mining contract cancellation on electricity generation (BOOTSTRAP SE) 135
C.1 Effect of uncertain fuel supply on coal plants vs gas plants 136

ACKNOWLEDGMENTS

I would like to thank my dissertation committee: Ryan Kellogg, Koichiro Ito, and Steve Cicala for their incredible support throughout my dissertation research work. I want to give special thanks to two anonymous mentors and friends without whose support I would not have been able to write this dissertation. I would like to thank my friends and colleagues at University of Chicago for their honest feedback and comments. I received generous support from the Becker Friedman Institute and Tata Center for Development for my research work. This dissertation also received financial support from Fuel Freedom Foundation. All errors are mine alone.

ABSTRACT

This dissertation critically examines the impact of public policies in the fields of development economics and energy and environment economics. The first two chapters study the impact of two different nation-wide policies on child development in India. The third chapter demonstrates how an inefficient fuel allocation policy leads to power outages in India.

Chapter 1 (co-authored with Karen Macours, Paris School of Economics) studies the impact of women's political empowerment in the form of reserved seats in local government on long-term child outcomes. We find that having a women legislator in early years of a child's life improves the health and nutritional environment in-utero and during the first few years of life, and this leads to long-term impacts on children's learning.

In chapter 2, I study the effect of household employment shocks on the human capital of young children. Specifically, I exploit the phased roll-out of the National Rural Employment Guarantee Act of India to investigate the effect of caregiver employment on child height for age z-score. Using child level fixed effects, I show the importance of investment of caregivers' time on child development − exposure to positive employment shocks because of NREGA causes a 0.25 standard deviation decline in height for age z-score over a period of 12 years. The adverse effect on children less than five years old at the time of exposure to NREGA is twice that of older children. Time allocation of children shows that when caregivers become employed, children share the load of household duties which could be driving the long-term adverse health outcome.

Chapter 3 demonstrates how regulatory uncertainty can cause large welfare losses by distorting firms' incentives and giving rise to inefficient production. Specifically, I analyze the production decisions of power plants in India that rely on state-regulated supply of coal for fuel. A court-ordered future reallocation of mining contracts in 2014 led to an unexpected increase in uncertainty about future coal supply for a subset of plants while leaving other plants with long-term supply contracts unaffected. I use this quasi-experimental variation in a difference-in-difference framework and a unique dataset linking coal mines and power plants

to estimate the effect of future regulatory uncertainty on power production. I show that affected power plants under-report their generation capacity available for power generation and begin stockpiling fuel for future periods. The behavior of these power plants is driven by precautionary saving motive, and I provide empirical evidence that power plants began stockpiling coal by reducing consumption, even as the supply of coal remains unchanged. In the short run, this precautionary saving driven stockpiling led to a 7% reduction in electricity generation. The negative impact on power production is persistent and the effects last for over 3 years in the long run. Using data on plants' marginal cost of production, I compute the short-run welfare cost of this regulatory uncertainty to be between 0.3 and 1.5 billion dollars.

CHAPTER 1

WOMEN'S POLITICAL RESERVATION, EARLY CHILDHOOD DEVELOPMENT, AND LEARNING IN INDIA

1.1 Introduction

Early childhood experiences are often thought to be crucial determinants of children's cognitive and non-cognitive development (Heckman 2008). Many children in the developing world are exposed to a multitude of risk factors, hampering their cognitive development from very early on (Grantham-McGregor et al. 2007). This is thought to lead to subsequent lack of human capital accumulation, resulting in poverty and continued inequality. Increasingly, policies therefore focus on improving health, nutrition, and educational outcomes during early childhood. And as there typically is a strong relationship between mother's characteristics and children's cognitive and educational outcomes, policies addressing constraints on the mothers of these children are often believed to be particularly promising (World Bank 2011). This is even more so if women have stronger preferences for human capital investments (as hypothesized, for instance, by Thomas [1990, 1993]), and policies that increase women's decision making power could potentially have large implications for the human capital of the next generation. Yet very little is known about the longer-term impact of such policies on children's human capital outcomes.

This article contributes by analyzing the impact of women's representation in elected bodies on children's learning and nutritional status more than 10 years later and tests whether political reservation for women resulted in improvements in children's longer-term learning and nutritional outcomes. Understanding the impact of female reservation policy on children's human capital accumulation is of high interest given existing evidence on children's delays in physical growth, cognition, and learning outcomes in India. Stunting rates, for instance, have remained extremely high in India despite economic growth (Deaton and Dreze 2009; Jayachandran and Pande 2015), and lack of learning in school has been highlighted as a particular challenge (Pritchett and Beatty 2012; Muralidharan 2013).

The 73rd Amendment of the Constitution of India of 1992 prescribes that one-third of the seats in all local councils, as well as one-third of the leadership positions, must be reserved for women[1]. These local councils can affect children's outcomes through their responsibilities for public good provision in the Gram Panchayat (GP, the lowest administrative level). They also decide the allocation of private benefits, such as housing or toilet construction, public employment, or food subsidies for the poor. Exploiting the random allocation of the seats reserved for women in local elections, several studies have shown that reservations had important impacts on policy choices in the short term. Chattopadhayay and Duflo (2004) show that female reservation increased investment in infrastructure related to the expressed development priorities of women in West Bengal and Rajasthan. Duflo and Topalova (2004) further show for a sample covering 24 states that GPs reserved for women invest more in public goods and that the measured quality of these goods is at least as high as in non-reserved GPs. This is particularly the case for drinking water infrastructure. Bhalotra and Clots-Figueras (2011) find that female leaders are more likely to invest in rural health infrastructure than men. And Beaman et al. (2009) show that reservation for women not only has led to a decrease in bias among voters against women candidates but also has resulted in a subsequent increase in the percentage of female local leaders. More broadly, political reservation for the Scheduled Caste, Scheduled Tribe, and women has been linked

1. Seats in councils and leadership positions were also reserved for the two disadvantaged minorities in India, Scheduled Castes and Scheduled Tribes, in the form of mandated representation proportional to each minority's population share in each district.

to more policy influence for these groups (Pande 2003). Pande and Ford (2011) further show the effectiveness of quotas in changing attitudes toward women leaders and giving them more policy influence. And exposure to female leaders can raises parents' aspirations for their girls (Beaman et al. 2012) and increase women's participation in political processes (Beaman et al. 2010). Reservation can also otherwise affect women's outcomes, with evidence by Jha, Nag, and Nagarajan (2011), for instance, documenting how improvements in water availability lead to higher participation of women in the labor market and more productive work. Ghani, Kerr, and O'Connell (2013) find a positive impact on female entrepreneurship.

Focusing more specifically on child-related indicators, the available evidence suggests short-term positive impacts, but also substantial heterogeneity of such impacts. Clots-Figueras (2012) shows that female leaders have a positive effect on primary education attainment of children in urban India, though not in rural areas. Beaman et al. (2012) suggest that education investments might be particularly important for girls, as having a GP reserved for a female leader reduces the gender gap in school attendance. And Kumar and Prakash (2012) similarly find positive impacts of women leaders on child mortality in Bihar, but the impacts are found only for the wealthier quintiles of the population. Less is known about impacts on children's health and nutrition status during early childhood.

More generally, not all the available evidence points to positive impacts. Ban and Rao (2008) show that women's reservation in four southern states, including Andhra Pradesh, does not lead to more investments congruent with women's preferences. Bardhan, Mookherjee, and Para Torrado (2010) further show that female reservation can lead to worse targeting of private benefits within villages and, as such, potentially limit the impact of the public good investments. Yet Besley, Pande, and Rao (2005) find no such impact on targeting. Sathe et al. (2013) find better service delivery in villages with female leaders who have been in power for 3.5 years, but worse in those where female leaders only recently came to power. And relatedly, Afridi, Iversen, and Sharan (2013) find more leakage in National Rural Employment Guarantee Act programs with female leadership but improvement in governance as female leaders gain more experience. Deininger et al. (2015) also found that quality of service provision might be worse in the short run but better in the long run. Finally, Iyer et al. (2012) show how female reservation increased reporting of crime against women, which they interpret as a positive development.

Taken together the literature hence suggests a relatively wide variety of mechanisms through which the reservation policy could affect children's human capital. If female leaders indeed favor policies beneficial to human capital, they could increase investments in education and health care infrastructure or vaccination programs. They may also affect children's food intake, for instance, by increasing the number of households benefiting from the public food distribution system. Increased bargaining power resulting from higher earning potential can enhance some of these effects. But giving women political power does not necessarily lead to better human capital investments, as it is not automatically true that women would favor such investments (Duflo 2012). Indeed, both Chattopadhayay and Duflo (2004) and Ban and Rao (2008) show that female and male preferences can be quite heterogeneous between GPs, so that one would not necessarily expect reservation to lead to similar investments in all GPs. And if reservation has positive impacts on women's labor market outcomes, this may come at the cost of time spent for child care. That said, even if women focus more directly on their own needs, such as their access to food and preventive health care, this could result in indirect benefits for children, in particular through better protection during the critical in utero period. Any longer-term impacts may well depend then on the ages at which children were exposed to potential benefits from having a female elected leader.

This article builds on the existing literature and contributes by showing long-term impacts of reservation of seats for women in the GP elections on children's learning and nutritional outcomes in rural Andhra Pradesh. We do not aim to attribute these results to specific policy choices by female leaders, but instead analyze potential critical periods during early childhood in which such policies can be particularly effective. Specifically, we analyze children's test scores at age 8 comparing children exposed to reservation in utero and very early in life with children exposed after age 5. Using the three waves of the Young Lives panel data set that follows children from 2002 to 2009, we show that women's political reservation improved nutritional and health outcomes very early in the children's life, and we measure the longer-term impacts on children's test scores capturing both math and language achievement and on their nutritional status. We are able to attribute causality to the reservation policy in the Panchayat as the seats to be reserved were randomly allocated through a rotating mechanism. Specifically, we focus on a cohort of children (12–17 months in 2002) who were exposed to reservation in utero if their GP was randomly selected to receive reservation between 1995 and 2001, while they benefited from reservation only after the age of 5 if their GP was randomly selected for reservation between 2006 and 2011. We then compare the outcomes for this cohort when they are 8 years old between GPs reserved in 1995 versus 2006. This allows drawing lessons regarding the relative impact of being exposed to reservation in utero and very early in life.

The results show that children in GPs that were randomly reserved for women prior to the child's birth have significantly better outcomes on test scores in 2009 when compared to GPs that were reserved after the child turned 5. As such, this article also contributes to the literature on the long-term impacts of early-life interventions (Barnett 1995; Garces, Thomas, and Currie 2002; Walker et al. 2005; Maluccio et al. 2009; Barham 2012) and the critical periods for such interventions (Barker 1990; Adair 1999; Doyle et al. 2009; Almond and Currie 2011; Barham, Macours, and Maluccio 2013).

The article is structured as follows: Section II discusses the data sources and provides details about the reservation policy, the random rotation of reserved seats, and the balance for the cohort of interest. It further discusses identification and the empirical specification. Section III then shows the long-term impact of having a woman legislator on child learning and nutritional outcomes, comparing exposure in utero and very early in life with exposure after age 5. Section IV provides further evidence for other cohorts, and Section V presents conclusions.

1.2 Data, Identification, and Empirical Strategy

1.2.1 The Reservation Policy and the Cohort of Interest

We use the Young Lives Survey (YLS) data set for Andhra Pradesh, a three wave panel data set (2002, 2007, and 2009) focused on childhood poverty covering 20 mandals in six rural districts[2]. The YLS survey instrument consists of a child questionnaire, a household questionnaire, and a community questionnaire. YLS follows 2011 children who were between 6 and 17 months of age in 2002[3]. For the initial sampling, each mandal was divided into four geographical areas, and one village was randomly selected in each area. YLS then randomly sampled 100 households with a 1-year-old child in each mandal. Our overall target sample in particular consists of 1,547 children, as we exclude 464 children from urban areas, where the GP reservation is irrelevant (the local election system being different for urban areas).

The Panchayat is a system of village-level (Gram Panchayat), mandal-level (Panchayat Samiti), and district-level (Zilla Parishad) councils, with membership determined through local elections. Their main responsibility is the administration of local public goods. Each mandal consist of various GPs, and each GP encompasses between 1,000 and 10,000 individuals in a group of villages (between one and 15). In 1992, the 73rd Amendment of the Constitution of India gave new powers to the Panchayats and provided that one-third of the seats in all Panchayat councils, as well as one-third of the leadership positions, must be reserved for women. Seats and leadership positions were also reserved for the two disadvantaged minorities in India, Scheduled Castes and Scheduled Tribes, in the form of mandated

2. The YLS sampled six rural districts, two districts in each of three agroclimatic regions, and selecting one poor and one non-poor district in each region: Srikakulam and West Godavari in the coastal region, Anantapur and Cuddapah in Rayalaseema, and Mahboobnager and Karimnagar in Telangana. The mandal or block is the administrative unit just below the district. After ranking all the mandals sites in these districts on the basis of a selected set of indicators of economic and human development and infrastructure, 20 mandals were sampled from these districts.

3. In parallel, YLS also sampled a second group of 1,000 children who were, on average, 8 years old in 2002. We do not use this second sample given the focus on early childhood. More generally, the YLS initiative collected data on 12,000 children in four countries over 15 years. It is led by a team in the DFID at the University of Oxford in association with research and policy partners.

representation proportional to each minority's population share in each district[4].

In Andhra Pradesh, for all practical purposes, this act came into effect with the Andhra Pradesh Panchayat Raj Act of 1994, which mandates the elections for the seat of the sarpanch, or the head of the GP. For each subsequent election, one-third of the total seats were reserved for women, and this was to be done by rotation[5]. Only women candidates could contest the election for these seats, thereby ensuring a woman sarpanch in the reserved GPs, without any exception. The first Panchayat or GP elections after the Andhra Pradesh Panchayati Raj Act were held in March 1995. The next elections took place in July–August of 2001 and 2006. Given the rotation rule, a GP reserved in 1995 could not be reserved in 2001 or 2006, and by 2006 every GP had been reserved once. There are hence three treatment episodes relevant for this article: 1995– 2001, 2001–6, and 2006–11. The State Election Commission typically announced the dates for the elections 1–2 months in advance, and the term of the newly elected sarpanch started immediately after each election.

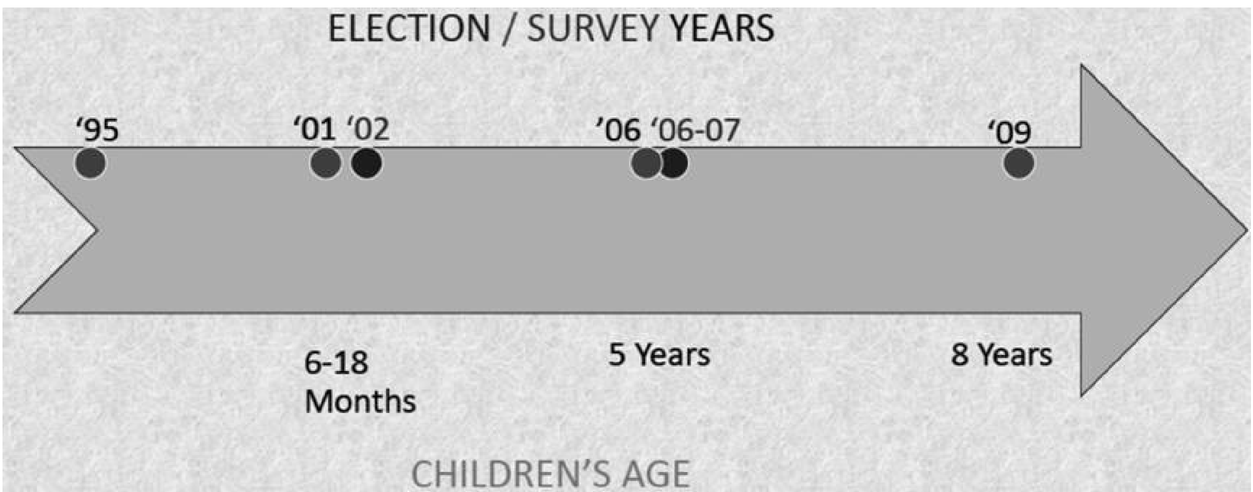
The first round of the YLS was conducted from September to December 2002, approximately 1 year after the sarpanch elected in 2001 came into power and 7 years after the first election with reservation. See figure 1.1 for a timeline of elections and data collection. At that point the children in the sample were 6–17 months old. The second round of data was collected from January to July 2007, that is, 6–12 months after the 2006 election. The third round of YLS was mostly conducted in the second half of 2009, that is, more than 3 years after the 2006 election. The YLS hence does not allow characterizing the situation prior

4. In this article, we use the term "reservation" specifically for the reservation of the sarpanch position for women. We do not consider the reservation for minorities for lack of exogenous variation in this group.

5. The Andhra Pradesh Panchayat Raj Act of 1994 mandates that the office of sarpanch of the GPs in the state be reserved in the following manner: (a) Of the total number of GPs in the state, the state election commissioner shall reserve a number of GPs for Scheduled Castes (SC) and Scheduled Tribes (ST), such that the proportion of reserved GPs to total GPs reflects the ratio of the SC/ST population to the total population. (b) This reservation of seats for SC/ST has to be done by rotation, which implies that a GP reserved for SC/ST in the first election year is not to be reserved for SC/ST in the next election year. An exception to this case, though, is possible in theory subject to the constraint that the total number of seats reserved must reflect the proportion of population of SC/ST in the state. (c) A minimum of one-third of the GPs reserved for SC/ST must be reserved for women. This is also to be done by rotation. (d) A minimum of one-third of the GPs, including the ones reserved for SC/ST, are to be reserved for women. This is also to be done by rotation (see http://www.apsec.gov.in/).

to the start of the reservation policy. Given the randomized rotation of the reservation, we do not need such baseline information for identification. The lack of a baseline does imply that only the characteristics unlikely to have been affected by the reservation can be used as control variables. The major advantage of the YLS data is that they allow us to track the medium to longer-term impacts of the reservation policy at the child level and notably to investigate to what extent impacts on health and nutritional gains, very early in life, are reflected in longer-term impacts on learning and nutritional status.

Figure 1.1: Timeline of election and survey rounds and children's age



For identification purposes the key feature of the reservation policy is that the reserved seats were randomly allocated. It is this random allocation that most other papers on women's reservation in India—as reviewed above—have also relied on for identifying the causal impact of the reservation. Ban and Rao (2008) describe the process of assignment of women reservation to GPs in some states of South India, including Andhra Pradesh, as follows: GPs to be reserved for SC/ST in the first step and GPs to be reserved for women in the next step are drawn by arranging the GPs in descending order of the ratio of the population of the respective category for which it is to be reserved. Then the first, fourth, seventh, and so on numbered GPs on the list were assigned reservation for the 1995 election. The second, fifth, eighth, and so on GPs are assigned reservation for the 2001 election. And the third, sixth, ninth, and so on GPs are assigned reservation in the 2006 election. Note that this rule implies that GPs reserved in 1995, on average, should have slightly higher female population ratios but that conditional on these slight differences, assignment is random. This may nevertheless be a potential concern for identification if higher female population ratios are also related through other mechanisms to children's longer-term outcomes. We return to this point below.

Given that the YLS data sampled children aged 6–17 months in the fall of 2002, reservation had been in place approximately 6 years prior to their birth in the GPs reserved from 1995 to 2001. While they had limited direct exposure to reservation after birth, they may have continued to benefit from policies implemented during reservations. In addition, for the 12–17-month-olds, reservation was still in place when they were in utero. Children who were between 6 and 11 months old in the fall of 2002, on the other hand, had limited in utero exposure. In contrast, in the GPs reserved from 2001 to 2006, reservation starts almost at the same time or just before the children in the sample are born, and reservation was in place until they were about 5 years old. As it is likely that the results of reservation on policy implementation were not immediate, they may not have benefited much in utero or very early in life. Finally, in the GPs reserved from 2006 to 2011, reservation took place when children were older than 5, that is, at the moment they were entering primary school age.

Hence there are potentially important differences between cohorts as well and GPs in terms of being affected by reservation. In particular, we focus on children aged 12–17 months in the first round of the YLS survey: they were exposed to reservation in utero if their GP was randomly selected to receive reservation in the 1995 election, while they benefited from reservation after the age of 5 only if their GP was randomly selected for the 2006 reservation. By comparing the outcomes of this cohort between 1995 and 2006 reservation GPs in 2009, we can hence draw lessons regarding the relative impact of being exposed to reservation in utero and very early in life. By the time the third round of the survey is conducted in 2009, the children in the sample are about 8 years old. This allows evaluating the longer-term impacts of the reservation policy and, in particular, analyzing whether the differential timing of exposure leads to longer-term differences in learning and nutritional outcomes.

1.2.2 Random Rotation of Reservation and Balance

We first use data from the 1991 census to verify that villages in GPs reserved in 1995, 2001, and 2006 were indeed similar prior to reservation. To do so, we merge the 1991 census data with the GP election reservation data for Andhra Pradesh for the 660 villages in mandals included in the rural YLS sample. The election data are freely available and come from the Andhra Pradesh State Election Commission. We were able to obtain election data for the years 2001 and 2006. As expected, there are no GPs that were reserved for a female sarpanch in both years. Indeed, since the election reservation is done on a rotation basis and each election had female reservation for one-third of the GPs, these data also implicitly contain the election reservation data for the 1995 election (the GPs that did not receive reservation in 2001 and 2006 must have received reservation in 1995). Because of differences in village names between the election and the census data, the merge between the two data sources is imperfect, but we manage to merge 289 out of 660. For those that can be merged, we find very few differences between villages reserved in 1995 and 2006 and notably no clear differences in health, water, or education infrastructure[6]. As expected, the female population ratio in the GPs reserved in 1995 is slightly larger than in the GPs reserved in 2006. The difference is small (0.501 compared to 0.493) though significant at the 1% level. Overall, these results confirm previous findings in the literature regarding the random rotation of female reservation.

The YLS cannot be merged with the census data for confidentiality reasons[7]. We hence

6. We compared the 1995 reservation villages with the 2006 reservation villages for 84 characteristics covering demographics (including female literacy and labor force participation), education, health, water infrastructure, power, and other services. Three variables are significant at the 5% level and one at 10%. Similar results are obtained for the 2001–6 comparison or when using villages of all mandals in Andhra Pradesh. Detailed results are available from the authors.

7. The election reservation data were merged with the YLS data set by the YLS team, and the names of

Table 1.1: Selection into sample: characteristics of villages, households, and children with and without election data

Variable	YLS GP without Election Data	YLS GP with Election Data	Difference	p-value
		Viilage-Level Data		
Observations	23	59		
Total land area in the village	1,875	1,708	166.6	.71
Total arable land in the village	1,147	989.2	157.8	.50
Total irrigated land in the village	255.7	349.1	-93.5	.48
Total population	1,677	1,809	-132.2	.68
Distance to the district/regional capital (in km)	75.77	69.89	5.88	.52
Hindi one of the most widely spoken languages	.435	.576	-.141	.25
Marathi one of the most widely spoken languages	.304	.254	.050	.65
BC the largest ethnic group	.565	.586	-.021	.87
SC and/or ST the largest ethnic group	.261	.310	-.050	.66
Christians one of the major religious groups	.652	.322	.330***	.01
Muslims one of the major religious groups	.391	.508	-.117	.35
Handicrafts one of the main economic activities	.217	.220	-.003	.98
Construction one of the main economic activities	.348	.339	.009	.94
Trade one of the main economic activities	.609	.492	.117	.35
Ratio of adult females/adult males	.499	.498	.001	.90
Average number of adult women in the household (>15 years)	1.773	1.791	-.018	.82
		Household and Child-Level Data		
Observations	346	1,201		
Male child	.552	.529	.023	.35
Child's age (in months)	11.99	11.82	.168	.54
Only child born in 2002	.367	.371	-.004	.88
Age of biological mother	22.97	23.68	-.713*	.06
Mother's ethnicity (dummy for BC)	.526	.457	.069	.36
Mother's ethnicity (dummy for SC)	.237	.193	.044	.34
Mother's ethnicity (dummy for ST)	.127	.192	-.064	.29
Household head completed primary	.315	.326	-.011	.77
Highest grade mother completed in school	2.443	2.181	.262	.54
Mother had no schooling	.656	.716	-.060	.29

Note: P-values are clustered at the GP level. Results are similar when p-values are clustered at the mandal level. Urban locations are excluded. * $p < .1$. *** $p < .01$.

cannot repeat this exercise for the specific villages in the YLS sample. But given the random selection of villages into the YLS sample, there is no apriori reason to believe these results would be different. That said, as was the case for the census data, not all the YLS villages could be merged with the election reservation data (election data were merged for 59 out of 82 villages), and this imperfect match could create imbalances. To verify this, we use both the YLS village and household surveys of 2002 and focus on characteristics unlikely to have changed because of reservation policies[8]. Table 1.1 first analyzes whether villages that could not be merged to the election data differ from those for which the reservation status is known. Overall, there are few significant differences, with only one of the 16 village-level variables as well as one out of 10 from the household survey being significantly different. Overall, these results suggest no large selection bias into the sample.

Table 1.2 then shows the same characteristics by the reservation status of the GP to which the village belongs[9]. It shows in particular that the villages that were reserved in 1995 and 2006 are very similar in observed characteristics, with the exception of the average number of female adults per household in the village[10]. Also at the household and child levels, variables unlikely to be affected by the reservation policies such as mother's age and ethnicity, child's gender, and birth order seem to be well balanced, suggesting there was no selective attrition between the 1995 and 2006 reservation villages[11]. Broadly similar results

---

the GPs were removed to guarantee confidentiality and assure anonymity.

8. We exclude variables with very little variation in the data. For instance, we do not report whether agriculture is reported as one of the main income sources because this is the case in all villages.

9. Because of the YLS sampling design, every village in the sample corresponds to a different GP. We hence use village and GP interchangeably when discussing the results.

10. In contrast to the census data, the sample shows balance on the estimated female population ratio. The reason could just be the small number of villages or that the female population ratio in the YLS is a noisy estimate for the true female population ratio, as it is based on only the households in the sample.

11. Given the wide scope of potential impacts of reservation documented in the literature, the set of variables unlikely to have been affected is limited to demographic and village population characteristics. Note that even if reservation may well affect female education, we include parental education in these balance checks both because of its importance for children's human capital and because average ages of mothers and the extremely low level of education (with 70% of mothers having no education at all) suggest that their education levels observed in 2001 likely reflect 1995 education levels.

Table 1.2: Descriptive Statistics in 2002 for the GPs reserved in 1995 and 2006

Variable	GP with 1995 Reservation	GP with 2006 Reservation	Difference	p-value
		Viilage-Level Data		
Observations	20	27		
Total land area in the village	1,724	1,606	118.3	.83
Total arable land in the village	1,000	940.5	59.76	.83
Total irrigated land in the village	266.6	418.8	-152.3	.43
Total population	1,719	1,979	-260.3	.54
Distance to the district /regional capital (in km)	67.88	66.93	.956	.94
Hindi one of the most widely spoken languages	.450	.630	-.180	.23
Marathi one of the most widely spoken languages	.200	.333	-.133	.32
BC the largest ethnic group	.632	.556	.076	.62
SC and/or ST the largest ethnic group	.316	.333	-.018	.90
Christians one of the major religious groups	.400	.185	.215	.11
Muslims one of the major religious groups	.400	.556	-.156	.30
Handicrafts one of the main economic activities	.150	.222	-.072	.54
Construction one of the main economic activities	.300	.370	-.070	.62
Trade one of the main economic activities	.500	.556	-.056	.71
Ratio of adult females/adult males	.495	.492	.003	.74
Average number of adult women in the household (>15 years)	1.658	1.828	-.170**	.04
p-value joint significance test				.49
		Household and Child-Level Data (12–17 Months Old)		
Observations	225	301		
Male child	.498	.561	-.064	.12
Child's age (in months)	14.44	14.52	-.083	.56
Only child born in 2002	.391	.349	.042	.30
Age of biological mother	23.78	23.75	.032	.93
Mother's ethnicity (dummy for BC)	.431	.462	-.031	.75
Mother's ethnicity (dummy for SC)	.209	.176	.033	.54
Mother's ethnicity (dummy for ST)	.227	.183	.044	.65
Household head completed primary	.458	.350	.108	.12
Highest grade mother completed in school	2.959	2.332	.627	.28
Mother had no schooling	.613	.695	-.082	.22
p-value joint significance test				.32

Note: P-values are clustered at the GP level. Urban locations are excluded. ** $p < .05$.

are obtained when focusing on children 6–11 months old, though for this age group there is an imbalance on children's age. We will therefore control for monthly age fixed effects in all estimations. Note also that while infant mortality is not negligible for the population of interest, there are no significant differences between the 1995 and 2006 reservation villages in the likelihood that the mother has lost a previous child[12]. Hence mortality selection does not seem to be a major concern. Finally, attrition rates within the YLS matched sample of the 1995 and 2006 reservation villages are very low, with only 4% of children originally sampled in 2002 missing test scores or height-for-age z-scores in 2009, and this attrition is uncorrelated to reservation year (p-value .43).

However, more imbalances are observed when comparing the 2001 and 2006 reservation villages, with notably relatively large differences in the level of education of the mothers (see table 1.3). The differences are negative and so are unlikely to be a result of reservation policies. Given the importance of parental education for children's human capital (Behrman and Rosenzweig 2005; Carneiro, Meghir, and Parey 2012), this imbalance is a potential concern. It probably results from the relatively small sample of villages with 2001 reservation. In the sample of 59, 20 were reserved in 1995, only 12 in 2001, and 27 in 2006. The low number in 2001 suggests that there is a proportional large share of the 2001 reservation villages among the 23 that could not be matched (as in expectation one-third of all villages should be reserved in each of the years). Given these results, the comparison of the 2001 reservation villages with the 2006 villages should be interpreted with more caution. The main focus of the article will hence be on the comparison between the 1995 and 2006 reservation villages[13].

Finally, we note that while the implementation of many policies affecting child outcomes

12. On average, 12% of mothers have lost a child, but the difference between the two groups is 20.005 with standard error 0.047.

13. Of the total of 1,547 children 6–17 months old in the data set, 1,201 are in mandals with reservation data. Among them, 384 are in GPs reserved in 1995, 240 in GPs reserved in 2001, and 577 in GPs reserved in 2006. For children 12–17 months old (the main group of interest), there are 225 are in GPs reserved in 1995, 159 in GPs reserved in 2001, and 301 in GPs reserved in 2006.

Table 1.3: Descriptive Statistics in 2002 for the GPs reserved in 2001 and 2006

Variable	GP with 2001 Reservation	GP with 2006 Reservation	Difference	p-value
		Viilage-Level Data		
Observations	12	27		
Total land area in the village	1,913	1,606	307.0	.57
Total arable land in the village	1,087	940.5	146.7	.62
Total irrigated land in the village	307.0	418.8	-111.8	.64
Total population	1,511	1,979	-468.2	.36
Distance to the district/regional capital (in km)	79.42	66.93	12.49	.38
Hindu one of the most widely spoken languages	.667	.630	.0370	.83
Marathi one of the most widely spoken languages	.167	.333	-.167	.30
BC the largest ethnic group	.583	.556	.0278	.88
SC and/or ST the largest ethnic group	.250	.333	-.0833	.61
Christians one of the major religious groups	.500	.185	.315**	.05
Muslims one of the major religious groups	.583	.556	.0278	.88
Handicrafts one of the main economic activities	.333	.222	.111	.48
Construction one of the main economic activities	.333	.370	-.0370	.83
Trade one of the main economic activities	.333	.556	-.222	.21
Ratio of adult females/adult males	.514	.492	.0220*	.08
Average number of adult women in the household (>15 years)	1.874	1.834	.0395	.72
		Household and Child-Level Data (12–17 Months Old)		
Observations	136	301		
Male child	.485	.561	-.076*	.09
Child's age (in months)	14.74	14.52	.224	.10
Only child born in 2002	.404	.349	.056	.26
Age of biological mother	24.51	23.75	.758	.19
Mother's ethnicity (dummy for BC)	.529	.462	.068	.60
Mother's ethnicity (dummy for SC)	.191	.176	.015	.85
Mother's ethnicity (dummy for ST)	.191	.183	.008	.94
Household head completed primary	.213	.350	-.137	.14
Highest grade mother completed in school	.963	2.332	-1.370**	.02
Mother had no schooling	.866	.695	.171**	.01

Note: P-values are clustered at the GP level. Urban locations are excluded. ** $p < .05$.

happens at the GP level, budgets for such policies are determined at the mandal level, and separate reservation rules were used for the mandal authorities. We will therefore include mandal fixed effects in all estimations. GPs with 1995 and 2006 reservation status cover 16 mandals, and on average, there are 4.3 GPs per mandal. The number of GPs reserved per mandal, on average, was 1.5 in 1995.

1.2.3 Empirical Specification

We estimate simple intent-to-treat regressions, comparing outcomes of children born in GPs with reserved seats for women between 1995 and 2001 with children born in GPs with seats reserved from 2006 to 2011. Specifically, we estimate

$$Y_{igm} = \beta_0 + \beta_1 T95_{gm} + X_{igm}\gamma + \Sigma\delta\eta_m + \varepsilon_{igm} \tag{1.1}$$

where Y_{igm} is the outcome of interest for child i in GP g and mandal m. Given the randomized rotation of the reservation, the estimates of β_1 measure the differential impacts of having been exposed to the 1995 reservation as compared to the 2006 reservation. As such they capture the impact of exposure in utero and very early in the child's life, as compared to after the age of 5. We will refer to this as "early" versus "late" exposure to reservation. The vector of child-level controls X includes child's gender and age (using monthly age dummies)[14]. Given the lack of baseline, we add only controls that likely reflect baseline information: mother's age, ethnicity (dummy variables for ST, SC, and other backward caste [BC]), and education (highest grade level achievement and a dummy for no education). As table 1.2 showed imbalance in the average number of adult females per household, we

14. There are a few children in the data set younger than 6 months or older than 17 months, despite the fact that sampling was meant to include only children from 6 to 17 months old. As we use age-month fixed effects, we round the age of 5-month-old children to 6 months and those of more than 17 months down to 17 months. Results are robust to exclusion of these observations.

also control for this variable[15]. Mandal fixed effects are included to control for all the unobservables that are fixed at the mandal level, η_m.

Our main interest is in the estimation of this regression for children aged 12–17 months in 2002, as this is the group of children born before the end of the reservation if they were born in a GP reserved in 1995. We also separately estimate the same specification for children aged 6–11 months in 2002 to look at differential impacts for this younger cohort, for whom exposure due to the timing of the elections was more limited. By using separate estimations for the two cohorts, we allow the mandal fixed effects and other controls to separately control for any factors that may be specific for these age groups[16].

The outcomes of primary interest are the longer-term impacts (in 2009) of reservation on children's test scores. The 8-year-old children were given the Early Grade Reading Assessment test, which measures the basic skills required for literacy acquisition in early years of schooling. It includes recognizing letters of the alphabet, reading simple words, understanding sentences and paragraphs, and listening with comprehension[17]. They were also given a math test with basic computing exercises appropriate for 8-year-old children, as well as the Peabody Picture Vocabulary Test, a popular test for assessing receptive vocabulary, often used as a proxy for cognition. The z-scores of these tests were calculated by subtracting the mean and dividing by the standard deviation. We also calculate the average of these tests to account for multiple inferences. The second main outcome we focus on is the children's nutritional status, in particular, their height-for-age z-scores and stunting levels, which can be tracked through the three rounds of the survey.

15. All results are robust to controlling in addition for the estimated female population ratio. Results are also robust to adding controls for the share of SC or ST.

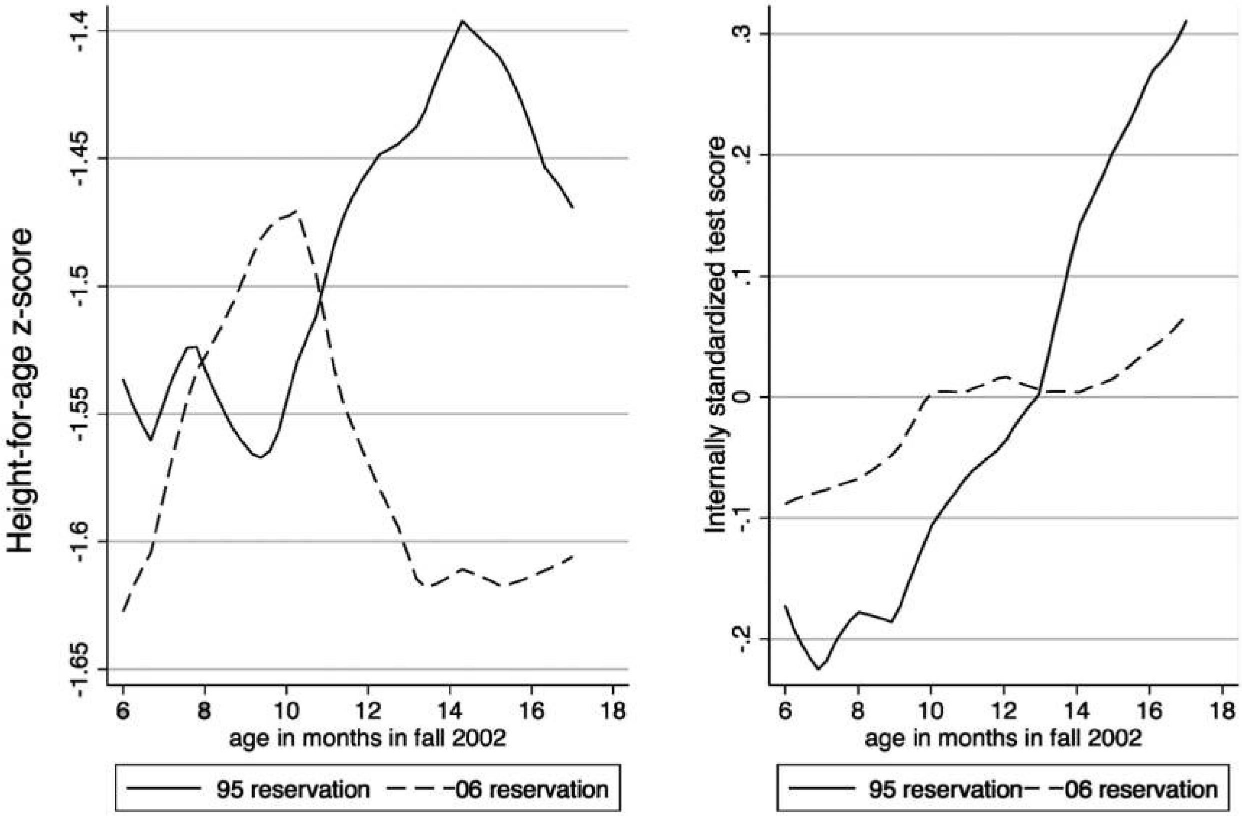
16. Evidence presented by Brainerd and Menon (2015) indicates, for instance, that religion can be differentialy related to nutrition for these two age groups.

17. See the US Agency for International Development's "Early Grade Reading" (https://www.eddataglobal.org/reading/index.cfm).

1.3 Long-Term Impacts of Early Exposure to Reservation

Prior to estimating the full model with all the controls and fixed effects, figure 1.2 first provides the raw evidence for the main outcomes of interest. The graph shows the height-for-age z-scores and the average test scores for all children in 2009. The figure clearly shows that the oldest children born in early reservation GPs are doing better than children of the same age born in late reservation GPs. This holds for both height-for-age z-scores and test scores. The difference is positive for children aged 12–13 months or older, and learning differences are the largest for the oldest children. This is consistent with the oldest children benefiting both from in utero exposure and from better early life health care. For the youngest children, the differences for height-forage are less clear, while the test score results suggest they might be doing worse in early reservation GPs.

Figure 1.2: Nutritional status and test scores in 2009 for children born in GPs reserved in 1995 versus 2006.



Notes: Nutritional status and test scores in 2009 for children born in GPs reserved in 1995 versus 2006. Results of kernel-weighted local polynomial regression height-for-age z-scores and average test scores on age in months in 2002, separately for the 1995 and 2006 reservation groups. Average test scores are calculated as the mean of five internally standardized test scores measuring word recognition, reading fluency, listening comprehension, mathematics skills, and receptive vocabulary. Color version available as an online enhancement.

Table 1.4 then shows the main estimates for children 12–17 months old in 2002, comparing outcomes of children born in GPs reserved from 1995 to 2001 with children from GPs reserved from 2006 to 2011. Hence these estimates compare the main cohort of interest as they were exposed in utero and during very early childhood in the early reservation GPs but not in the late reservation GPs. The results show that 8-year-old children born in 2001 in GPs that have been randomly exposed to reservation in the 1995–2001 period have significantly better test outcomes than 8 year olds who had reached the age of 5 before their GPs were exposed to reservation. The differences are positive for all tests, and on average, children in early reservation GPs have test scores that are 0.19 standard deviations (SD) higher than those of the same cohort of children in late reservation GPs. In addition, children in early reservation GPs have height-for-age z-scores that are 0.26 SD larger, a sizable impact given the low average height-for-age z-scores in the population of -1.60. The probability of stunting is also lower, though not significantly so[18]. Overall, these differences are substantial, in particular given that the 2006 reservation may still have affected the children in late reservation GPs.

Table 1.4: Differential long-term effect of 1995 reservation on test scores and nutritional status in 2009: children in utero during 1995–2001 reservation (12–17 months in fall 2002)

	Word Recognition	Reading Fluency	Listening Comprehension	Math	Receptive Vocabulary	Average z-score Tests	Height for-age Score	Stunting
Reservation in 1995–2001 period	.191***	.250**	.248*	.026	.222*	.188**	.258**	-.064
	(.065)	(.10)	(.12)	(.095)	(.11)	(.085)	(.12)	(.046)
Observations	499	500	501	499	497	502	501	501
R^2	.23	.22	.22	.28	.21	.29	.10	.05
Mean in 2006 reservation GPs	-.016	.016	-.003	.151	.033	.035	-1.602	.339

Note: Standard errors (in parentheses) are clustered at the GP level. All estimations include only children in GPs who were reserved either in 1995 or in 2006. Controls include child's gender and monthly age dummies, dummy indicating whether child is first-born, mother's age, mother's ethnicity (dummies for SC, ST, and BC), mother's highest educational grade completed and dummy for no schooling, average adult females in household in GP, and mandal fixed effects. * $p < .1$. ** $p < .05$. *** $p < .01$.

Table 1.5 shows that these outcomes reflect nutritional and health gains made earlier in the same children's lives[19]. Indeed, in 2007 (when the children were 5–6 years old) and in 2002 (when they were 12–17 months old) they already had higher height-for-age z-scores and a significantly lower probability of stunting. The impact on stunting is large, indicating a reduction of 12 percentage points (of a mean of about 40) for children born in early

18. There is no significant heterogeneity in results for learning or nutrition status by gender, caste, or mother's education (results are available from the authors).

19. Table 1.5 specifically focuses on indicators of health status that are directly observable (anthropometric measurement) or verifiable (vaccinations were verified on a vaccination card, where available).

reservation GPs. The same children early on also showed a significantly higher probability of having received the major vaccinations by 2002 and hence were better protected against childhood diseases from very early on. Note also that the GPs with 2006 reservation had not yet had any reservation by the time of the 2002 survey. This allows us to interpret results for 2002 as the absolute short-term impact of being born in a GP exposed to reservation for the last 6 years.

Table 1.5: Medium and Short-term effects of 1995 reservation on nutritional and health status in 2007 and 2002: children in utero during 1995–2001 reservation (12–17 months in fall 2002)

	Health Status in 2007		Health Status in 2002					
	Height for-age z-score	Stunting	Height-for-age z-score	Stunting	BCG (Tuberculosis)	Measles	Polio	No. of Vaccinations (out of 3)
Reservation in 1995–2001 period	.176*	-.127**	.216	-.118**	.034	.086**	.063**	.182**
	(.10)	(.060)	(.16)	(.049)	(.030)	(.041)	(.024)	(.080)
Observations	500	500	500	500	520	520	520	520
R^2	.11	.07	.18	.15	.08	.14	.12	.12
Mean in 2006 reservation GPs	-1.763	.423	-1.562	.400	.913	.794	.957	2.670

Note: Standard errors (in parentheses) are clustered at the GP level. All estimations include only children in GPs who were reserved either in 1995 or in 2006. Controls include child's gender and monthly age dummies, dummy indicating whether child is first-born, mother's age, mother's ethnicity (dummies for SC, ST, and BC), mother's highest educational grade completed and dummy for no schooling, average adult females in household in GP, and mandal fixed effects. The last column sums up vaccinations of BCG, measles, and polio. There are 20 observations with missing values for anthropometric outcomes because of missing measurements and trimming of values below -5 SD and above +5 SD. * $p < .1$. ** $p < .05$.

Overall, the results in tables 1.4 and 1.5 hence suggest that reservation in 1995–2001 led to an early advantage in health and nutrition status of children born toward the end of this period and that this advantage was largely maintained from 2002 to 2007 and 2009. When these children reach the age of 8 (in 2009), they also show substantially better test scores, consistent with better health and nutrition in early childhood leading to cognitive gains. These gains are still observed in 2009, despite the fact that in the early reservation GPs (with reservation in 1995–2001), reservation had not been in place already for 8 years, while reservation in fact was in place at that point in the late reservation GPs (with reservation in 2006–11).

Table 1.6: Differences between 1995 and 2006 reservation groups on schooling outcomes in 2009: children in utero during 1995–2001 reservation (12–17 months in fall 2002)

	Do You Receive a Midday Meal at School?	Has [Name] Ever Dropped Out of School?	Grade Child was in 2009	Child's Subjective Status: Where on the ladder Do you Feel You Personally Stand at the Moment?
Reservation in 1995–2001 period	-.126***	.0475**	-.177	.661**
	(.045)	(.019)	(.11)	(.26)
Observations	500	490	474	500
R^2	.36	.11	.21	.23
Mean in 2006 reservation GPs	.628	.007	2.887	4.574

Note: Standard errors (in parentheses) are clustered at the GP level. All estimations include only children in GPs that were reserved either in 1995 or in 2006. Controls include child's gender and monthly age dummies, dummy indicating whether child is first-born, mother's age, mother's ethnicity (dummies for SC, ST, and BC), mother's highest educational grade completed and dummy for no schooling, average adult females in the household in the GP, and mandal fixed effects. ** $p < .05$. *** $p < .01$.

To further understand these findings, table 1.6 shows a number of schooling outcomes. Previous literature on female reservation has indicated that female leaders often invest in schooling or school meals, and GPs with reservation in place in 2006 might therefore, apriori, be expected to have a schooling advantage. Table 1.6 indeed provides some indications of such advantages, as the children in the early GPs are significantly less likely to have school meals and are significantly more likely to have dropped out of school. They also appear to be in a lower grade (though not significantly so). This helps us to further interpret the test score results in table 1.4. Children exposed to early reservation have higher test scores despite apparent disadvantages in a number of schooling indicators in 2009[20]. Children's

20. Note, however, that reservation policies in early reservation GPs may well have increased educational quality during the years of reservation and beyond in ways that are not captured in the data. Children

own perceptions of their performance are also significantly higher for the early than for the late reservation group, consistent with their higher test scores. Overall, the results suggest that reservation policies may have the highest potential for human capital accumulation when affecting the life circumstances of children very early in their young lives.

Table 1.7: Correlation of tests scores with educational variables

	Average z-score Tests	
	Coefficient	Standard Error
Mother's years of schooling	.074***	.004
Child's grade in 2009	.355***	.008
Has [name] ever dropped out of school?	-.641***	.008

Note: Each line shows the coefficient of a separate bivariate regression. Standard errors are clustered at the GP level. *** $p < .01$.

To put the magnitude of the test score results in context, table 1.7 shows the coefficients of a set of bivariate regressions of test scores on schooling indicators. The effect on test scores of 0.19 SD amounts to about half a grade of schooling of the child and to about 2.5 years of schooling of the mother. Given that this is a differential result, this suggests a sizable long-term impact of being exposed to reservation in early childhood. We can also compare the magnitude with other evidence on early childhood exposure to better nutrition or health environments. The differential impact of 0.19 SD is slightly higher than the 0.15 SD differential impact found in the study by Barham et al. (2013), who compare early versus late exposure to a 3-year conditional cash transfer program in Nicaragua after 10 years. On the other hand, it is smaller than typical impacts of intensive interventions combining nutrition, health care, and stimulation in other developing countries that can show short-term impacts of more than 1 SD (Engle et al. 2007). Evidence on impacts of such interventions when children are in primary school is scarcer. A potentially interesting comparison is, however,

born in GPs with earlier reservations could then have benefited for a longer period of such higher quality. When they entered the schooling system, the quality improvements may already have materialized, while for children born in 2006 reservation GPs, the quality improvements would not have been there when they entered.

the well-known Jamaica early childhood home visiting program (Grantham-McGregor et al. 1997). This intensive intervention combining nutrition and stimulation in early childhood showed very small gains at age 7, even if large impacts were found both immediately after the 2-year intervention (when children were about 4 years old) and again when children were adults (0.6 SD; see Walker et al. 2011). If this means that cognitive gains are hard to observe for this age group, it could imply that our estimates provide a lower bound for the longer-term impacts of female reservation.

1.4 Impacts on Other Cohorts

Table 1.8: Differential long-term effect of 1995 reservation on test scores and nutritional status in 2009: children born after 1995–2001 reservation (6–11 months in fall 2002)

	Word Recognition	Reading Fluency	Listening Comprehension	Math	Receptive Vocabulary	Average z-score Tests	Height for-age Score	Stunting
Reservation in 1995–2001 period	.006	-.010	-.033	.093	-.104	-.008	.131	-.051
	(.077)	(.073)	(.087)	(.079)	(.10)	(.061)	(.10)	(.046)
Observations	413	414	418	418	419	420	421	421
R^2	.28	.24	.26	.41	.31	.38	.23	.17
Mean in 2006 reservation GPs	-.090	-.071	-.046	-.083	.041	-.057	-1.615	.357

Note: Standard errors (in parentheses) are clustered at the GP level. All estimations include only children in GPs that were reserved either in 1995 or in 2006. Controls include child's gender and monthly age dummies, dummy indicating whether child is first-born, mother's age, mother's ethnicity (dummies for SC, ST, and BC), mother's highest educational grade completed and dummy for no schooling, average adult females in the household in the GP, and mandal fixed effects.

To further understand the relationship between exposure during early childhood and long-term learning and nutritional gains, table 1.8 shows results for the cohort of children that was born right after the 1995–2001 reservation, that is, the children aged 6–11 months in the fall of 2002. For this cohort, and in contrast to those children exposed in utero, we find no long-term differential effects on either average test scores or height-for-age z-scores. While the results of the two cohorts are not statistically different from each other, the contrast nevertheless is striking. Table 1.9 further shows that the short-term results on nutrition and health indicators are also not comparable with the older cohort. There are, in particular, no significant impacts on vaccination or stunting levels. This suggests that the policies causing short-term impacts of reservation on children's health in utero and during very early childhood did not persist after reservation was no longer in place.

Table 1.9: Medium-term effects of 1995 reservation on nutritional and health status in 2007 and 2002: children born after 1995–2001 reservation (6–11 months in fall 2002)

	Health Status in 2007		Health Status in 2002					
	Height for-age z-score	Stunting	Height-for-age z-score	Stunting	BCG (Tuberculosis)	Measles	Polio	No. of Vaccinations (out of 3)
Reservation in 1995–2001 period	.157*	-.034	.098	-.017	-.025	.026	-.017	-.016
	(.086)	(.041)	(.15)	(.050)	(.026)	(.053)	(.018)	(.075)
Observations	424	424	427	427	434	434	434	434
R^2	.19	.13	.25	.16	.12	.23	.14	.18
Mean in 2006 reservation GPs	-1.792	.424	2.967	.261	.949	.598	.953	2.500

Note: Standard errors (in parentheses) are clustered at the GP level. All estimations include only children in GPs that were reserved either in 1995 or in 2006. Controls include child's gender and monthly age dummies, dummy indicating whether child is first-born, mother's age, mother's ethnicity (dummies for SC, ST, and BC), mother's highest educational grade completed and dummy for no schooling, average adult females in household in GP, and mandal fixed effects. The last column sums up vaccinations of BCG, measles, and polio. Observations with missing values for anthropometric outcomes are due to missing measurements and trimming of values below -5 SD and above +5 SD. * $p < .1$.

The results in table 1.10 are further in line with these findings. Table 1.10 shows the results for the comparison of the children born in GPs reserved from 2001 to 2006 versus GPs reserved from 2006 to 2011. The results for the 12–17-month olds show that children born before the 2001 reservation started in their GP do not show any significant differences compared to children in the 2006 GPs. Indeed point estimates are negative, though not significant. On the other hand, the results for 6–11-month-olds suggest that children born after the start of the reservation in the 2001 GPs, and potentially having benefited from reservation in their GPs the next 5 years, have significantly higher test scores and nutritional status compared to those in the GPs with reservations in 2006. As indicated earlier, these comparisons should be interpreted with caution, given the evidence of imbalance presented in table 1.3 and the overall low number of GPs reserved in 2001. While all estimates control for the variables that were found not to be balanced, it is possible that other unobservable characteristics are not balanced. But even if the evidence is less definite, the pattern in data further supports the main findings of this article, which indicate that having been exposed to reservation very early in life can lead to long-term learning and nutritional gains.

Table 1.10: Differential long-term effect of 2001 reservation on test scores and nutritional status in 2009

	Word Recognition	Reading Fluency	Listening Comprehension	Math	Receptive Vocabulary	Average z-score Tests	Height for-age Score	Stunting
				12-17 Months in Fall 2002				
Reservation in 2001-06 period	-.072	-.219*	-.254**	-.184	.060	-.135	-.061	-.001
	(.092)	(.11)	(.10)	(.12)	(.12)	(.094)	(.11)	(.057)
Observations	417	417	420	417	415	420	420	420
R^2	.22	.22	.36	.39	.23	.35	.13	.06
Mean in 2006 reservation GPs	-.016	.016	2.003	.151	.033	.035	-1.602	.339
				6-11 Months in Fall 2002				
Reservation in 2001-06 period	.235*	.207**	-.003	.145*	.186**	.167**	.238*	-.117*
	(.13)	(.096)	(.078)	(.078)	(.073)	(.065)	(.12)	(.060)
Observations	364	364	368	364	367	369	368	368
R^2	.17	.35	.38	.46	.36	.46	.26	.17
Mean in 2006 reservation GPs	-.090	-.071	-.046	-.083	.041	-.057	-1.615	.357

Note: Standard errors (in parentheses) are clustered at the GP level. All estimations include only children in GPs that were reserved either in 2001 or in 2006. Controls include child's gender and monthly age dummies, dummy indicating whether child is first-born, mother's age, mother's ethnicity (dummies for SC, ST, and BC), mother's highest educational grade completed and dummy for no schooling, average adult females in household in GP, and mandal fixed effects. * $p < .1$. ** $p < .05$.

1.5 Conclusion

This article shows long-term impacts of reservation of local political seats for women on learning and nutritional outcomes in rural Andhra Pradesh. It relies on three rounds of the Young Lives panel data set to analyze the impact of exposure to political reservation during different stages of young children's life. We are able to attribute causality to the reservation policy in the Panchayat as the leadership positions to be reserved were randomly allocated through a rotating mechanism. Using this exogenous variation, we show that reservation policies can be particularly effective in increasing learning outcomes when they occur very early during a child's life. The results further show that political reservation improved the health and nutritional environment in utero and during the first years of life, which can help explain the long-term impacts on children's learning.

The data used in this article do not allow analyzing in more detail which specific policy decisions resulting from the female reservation are driving the long-term impacts. The available evidence in the literature suggests, moreover, that policies favored by female leaders differ across GPs on the basis of heterogeneity of female preferences and needs, so that different policies might have been affected in different places[21]. The results in this article reveal that, on average, these policy choices led to better health status of children in 2002, which continued in 2007, and translated in higher learning by 2009. While the article does not investigate the causal pathway, we note that these results are at least consistent with female leaders prioritizing key constraints for children's human capital, constraints that may well differ from GP to GP. The results in this article further suggest that any potential advantages of being exposed to reservation-related policies affecting schooling during primary school years did not offset the early nutrition and health gains[22].

21. As the first round of YLS data was collected only in 2002, it is not well suited to analyze specific policy choices during the early reservation period from 1995 to 2001. Analysis of a wide range of possible intermediate outcomes for the later reservation periods does not provide support for one particular mechanism. Similarly, data on village infrastructure changes in the different periods do not reveal any robust changes, consistent with reservation leading to different investments in different GPs.

22. An alternative interpretation of the findings could be the potential decrease in effectiveness of reser-

These findings go beyond the existing evidence on reservation policies by shedding light on their potential longer-term impacts. They suggest in particular that empowering female leaders might be especially important for improving the health and nutrition of very young children and that this in turn can lead to substantial longer-term impacts. The evidence also suggests that gains from interventions early in a child's life can be observed many years later, indicating that evaluations of policies targeting early childhood may benefit from a sufficiently long horizon. From a broader policy perspective, the results provide support for interventions that improve children's conditions very early in their lives and suggest that when targeting improvements in learning outcomes, it can be worth considering interventions targeting early childhood in addition to policies directly targeting primary school children.

---

vation policies to affect policy choices. When the reservation first came into place in 1995, it might have had a shock effect, and the impact of reservation might have diminished in subsequent cycles in 2001 and 2006. The reason could be that other political actors might have been able to respond to the initial shock of reservation and possibly made it ineffective in affecting policy choices by the time the third group of GPs benefited from reservation.

REFERENCES

[1] Adair, Linda S. 1999. “Filipino Children Exhibit Catch-Up Growth from Age 2 to 12 Years.” Journal of Nutrition 129, no. 6:1140–48.

[2] Afridi, Farzana, Vegard Iversen, and M. R. Sharan. 2013. “Women Political Leaders, Corruption and Learning: Evidence from a Large Public Program in India.” Discus- sion Paper no. 13-02, Indian Statistical Institute, Kolkata.

[3] Almond, Douglas, and Janet Currie. 2011. “Human Capital Development before Age Five.” In Handbook of Labor Economics, vol. 4b, ed. Orley Ashenfelter and David Card, 1315–1486. Amsterdam: Elsevier.

[4] Ban, Radu, and Vijayendra Rao. 2008. “Tokenism or Agency? The Impact of Women's Reservations on Village Democracies in South India.” Economic Development and Cultural Change 56:501–30.

[5] Bardhan, Pranab, Dilip Mookherjee, and Monica Parra Torrado. 2010. “Impact of Political Reservations in West Bengal Local Governments on Anti-Poverty Target- ing.” Journal of Globalization and Development 1, no. 1:art. 5.

[6] Barham, Tania. 2012. “The Effects of Child Health Interventions on Later Life Cognitive Functioning: Evidence from Matlab in Bangladesh.” American Economic Journal: Applied Economics 4, no. 1:245–73.

[7] Barham, Tania, Karen Macours, and John Maluccio. 2013. “Boys' Cognitive Skill Formation and Physical Growth: Long-Term Experimental Evidence on Critical Ages for Early Childhood Interventions.” American Economic Review Papers and Proceedings 103, no. 3:467–71.

[8] Barker, David J. P. 1990. “The Fetal and Infant Origins of Adult Disease.” British Medical Journal 301, no. 6761:1111.

[9] Barnett, Steven. 1995. "Long Term Effects of Early Childhood Programs on Cogni- tive and School Outcomes." Future of Children 5:25–50.

[10] Beaman, Lori, Raghabendra Chattopadhyay, Esther Duflo, Rohini Pande, and Petia Topalova. 2009. "Powerful Women: Does Exposure Reduce Bias?" Quarterly Jour- nal of Economics 124, no. 4:1497–1540.

[11] Beaman, Lori, Esther Duflo, Rohini Pande, and Petia Topalova. 2010. "Political Reservation and Substantive Representation: Evidence from Indian Village Councils." In India Policy Forum, 2010, ed. Suman Bery, Barry Bosworth, and Arvind Panagariya. Washington, DC: Brookings Institution Press and National Council of Applied Economic Research.

[12] ———. 2012. "Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India." Science 335, no. 6068:582–86.

[13] Behrman, Jere R., and Mark R. Rosenzweig. 2005. "Does Increasing Women's Schooling Raise the Schooling of the Next Generation? Reply." American Economic Review 95, no. 5:1745–51.

[14] Besley, Timothy, Rohini Pande, and Vijayendra Rao. 2005. "Participatory Democracy in Action: Survey Evidence from South India." Journal of the European Economic Association 3, nos. 2–3:648–57.

[15] Bhalotra, Sonia, and Irma Clots-Figueras. 2011. "Health and the Political Agency of Women." American Economic Journal: Economic Policy 6, no. 2:164–207.

[16] Brainerd, Elizabeth, and Nidhiya Menon. 2015. "Religion and Health in Early Childhood: Evidence from South Asia." Population and Development Review 41, no. 3: 439–63.

[17] Carneiro, Pedro, Costas Meghir, and Matthias Parey. 2012. “Maternal Education, Home Environments and the Development of Children and Adolescents.” Journal of the European Economic Association 11, no. S1:123–60.

[18] Chattopadhyay, Raghabendra, and Esther Duflo. 2004. “Women as Policy Makers: Evidence from a Randomized Policy Experiment in India.” Econometrica 72, no. 5: 1409–43.

[19] Clots-Figueras, Irma. 2012. “Are Female Leaders Good for Education? Evidence from India.” American Economic Journal: Applied Economics 4, no. 1:212–44.

[20] Deaton, Angus, and Jean Dreze. 2009. “Food and Nutrition in India: Facts and Interpretation.” Economic and Political Weekly 44, no. 7:42–65.

[21] Deininger, Klaus, Songqing Jin, Hari Nagarajan, and Xia Fang. 2015. “ Does Female Reservation Affect Long-Term Political Outcomes? Evidence from Rural India.” Journal of Development Studies 51, no. 1:32–49.

[22] Doyle, Orla, Colm P. Harmon, James J. Heckman, and Richard E. Tremblay. 2009. “Investing in Early Human Development: Timing and Economic Efficiency.” Eco- nomics and Human Biology 7, no. 1:1–6.

[23] Duflo, Esther. 2012. “Women Empowerment and Economic Development.” Journal of Economic Literature 50, no. 4:1051–79.

[24] Duflo, Esther, and Petia Topalova. 2004. “Unappreciated Service: Performance, Perceptions, and Women Leaders in India.” Unpublished manuscript, Massachusetts Institute of Technology.

[25] Engle, P. L., M. M. Black, J. R. Behrman, M. Cabral de Mello, P. J. Gertler, L. Kapiriri,

[26] R. Martorell, M. Eming Young, and the International Child Development Steering Group. 2007. “Strategies to Avoid the Loss of Developmental Potential in More than 200 Million Children in the Developing World.” Lancet 369, no. 9557:229–42.

[27] Garces, Eliana, Duncan Thomas, and Janet Currie. 2002. “Longer-Term Effects of Head Start.” American Economic Review 92, no. 4:999–1012.

[28] Ghani, Ejaz, William Kerr, and Steven O’Connell. 2013. “Political Reservations and Women’s Entrepreneurship in India.” Policy Research Working Paper no. 6307, World Bank, Washington, DC.

[29] Grantham-McGregor, S., Y. Cheung, S. Cueto, P. Glewwe, L. Richter, B. Strupp, and the International Child Development Steering Group. 2007. “Developmental Poten- tial in the First 5 Years for Children in Developing Countries.” Lancet 369:60–70. Grantham-McGregor, S., S. Walker, S. Chang, and C. Powell. 1997. Effects of Early Childhood Supplementation With and Without Stimulation on Later Development in Stunted Jamaican Children.” American Journal of Clinical Nutrition 66, no. 2:247–53.

[30] Heckman, James J. 2008. “Schools, Skills and Synapses.” Economic Inquiry 46, no. 3: 289–324.

[31] Iyer, Lakshmi, Anandi Mani, Prachi Mishra, and Petia Topalova. 2012. “The Power of Political Voice: Women’s Political Representation and Crime in India.” American Economic Journal: Applied Economics 4, no. 4:165–93.

[32] Jayachandran, Seema, and Rohini Pande. 2015. “Why Are Indian Children So Short? The Role of Birth Order and Son Preference.” Working Paper no. 21036, National Bureau of Economic Research, Cambridge, MA.

[33] Jha, Raghbendra, Sharmistha Nag, and Hari K. Nagarajan. 2011. “Political Reservations, Access to Water and Welfare Outcomes: Evidence from Indian Villages.” ASARC Working Paper no. 2011/15, Australia South Asia Research Centre, Aus- tralian National University.

[34] Kumar, Santosh, and Nishit Prakash. 2012. "Political Decentralization, Women's Reservation and Child Health Outcomes: A Case Study of Rural Bihar." Working Pa- per no. 2012-19, Department of Economics, University of Connecticut.

[35] Maluccio, John A., John Hoddinott, Jere R. Behrman, Reynaldo Martorell, Agnes R. Quisumbing, and Aryeh D. Stein. 2009. "The Impact of Improving Nutrition dur- ing Early Childhood on Education among Guatemalan Adults." Economic Journal 119, no. 537:734–63.

[36] Muralidharan, Karthik. 2013. "Priorities for Primary Education Policy in India's 12th Five Year Plan." In India Policy Forum, 2012–13, ed. Shah Shakhar and Arvind Panagariya, 1–46. New Delhi: Sage.

[37] Pande, Rohini. 2003. "Can Mandated Political Representation Increase Policy Influence for Disadvantaged Minorities? Theory and Evidence from India." American Economic Review 93, no. 4:1132–51.

[38] Pande, Rohini, and Deanna Ford. 2011. "Gender Quotas and Female Leadership: A Review." Background paper for World Development Report on Gender. https://openknowledge.worldbank.org/bitstream/handle/10986/9120/WDR2012-0008.pdf.

[39] Pritchett, Lant, and Amanda Beatty. 2012. "The Negative Consequences of Over- Ambitious Curriculum in Developing Countries." HKS Faculty Research Working Paper no. RWP12-035, Kennedy School of Government, Harvard University.

[40] Sathe, Dhanmanjiri, Stephan Klasen, Jan Priebe, and Mithila Biniwale. 2013. "Does Having a Female Sarpanch Promote Service Delivery for Women and Democratic Participation of Women? Evidence from Maharashtra, India." Discussion Paper no. 138, University of Goettingen.

[41] Thomas, Duncan. 1990. "Intra-household Resource Allocation: An Inferential Approach." Journal of Human Resources 25, no. 4:635–64.

[42] ———. 1993. “The Distribution of Income and Expenditure within the House- hold.” Annales d’Economie et de Statistique 29:109–35.

[43] Walker, Susan P., Susan M. Chang, Christine A. Powell, and Sally M. Grantham-McGregor. 2005. “Effects of Early Childhood Psychosocial Stimulation and Nutri- tional Supplementation on Cognition and Education in Growth-Stunted Jamaican Children: Prospective Cohort Study.” Lancet 366:1804–7.

[44] Walker, Susan P., Susan M. Chang, Marco Vera-Hernandez, and Sally Grantham- Mc-Gregor. 2011. “Early Childhood Stimulation Benefits Adult Competence and Reduces Violent Behavior.” Pediatrics 127:849–57.

[45] World Bank. 2011. World Development Report 2012: Gender Equality and Development. Washington, DC: World Bank.

CHAPTER 2

HOUSEHOLD EMPLOYMENT AND CHILD DEVELOPMENT

2.1 Introduction

In this paper, I study the effects of household employment shocks on children's human capital development. The setting for the paper is one of the world's largest public employment guarantee programs, National Rural Employment Guarantee Act (NREGA), in India. India passed the National Rural Employment Guarantee Act (NREGA) in 2005 to provide guaranteed employment to rural households. This study attempts to study the effect of NREGA on child human capital measured by health outcomes. I seek to shed light on the following questions: How important is caregiver's time investment in a child's development? Additionally, is this time investment most critical during the early childhood period, defined in the literature as less than five years old?

Presence of primary caregivers has been shown to be beneficial for the long-term development of children. Cunha et al. (2006) show that time, especially time spent with mother, is an important input for child development and that this time investment is most crucial in early childhood. In this paper, I study the effect of a decrease in time investment in children caused by caregivers' employment through the NREGA program. Specifically, I look at the effect of this employment shock on child long-term health using child height for age z-score. An unanticipated change in caregivers' employment involves two changes pertinent to investment in human capital: 1) an increase in household income because of the NREGA job wages, and 2) substitution of time away from investment in the children and in to the labor market. Because of these two simultaneous changes, the direction of the net effect of caregivers' employment is unclear.

The setting for this study is the rural areas of the state of Andhra Pradesh (AP) in India. Despite impressive economic growth, India has one of the highest rates of stunting (Pande and Jayachandran (2015)), with rural areas experiencing markedly higher rates of stunting compared to urban areas. Deaton and Dreze (2009) show that the relatively higher rate of stunting in rural areas has persisted, despite an increase in calorie intake in these rural areas over time.

There are many reasons why increased rates of stunting is a cause for concern, but one in particular is the correlation between adult height and human capital stock . Using nationally representative data, Deaton (2008) shows that adult height is a strong predictor of future labor market returns. Hence it is crucial to understand the causes of stunting. This paper aims to contribute to this by documenting the persistence of negative effects in height for age caused by a large scale welfare policy.

In addition, this paper aims to contribute to the literature on the importance of time investments in children, specifically in early childhood before the age of five. Early childhood interventions have been shown to be crucial for long-term human capital development. Interventions in the United States, such as the Perry Preschool Project and the Abecedarian Project, show that early childhood investments are critical for the development of human capital (Heckman (2008)). However, there is limited evidence of the impact of interventions in early childhood in developing countries. In case of India, Macours and Pathak (2017) show that exposure to female leaders in early childhood can lead to lasting effects on cognitive and health outcomes. This paper aims to add to this literature by studying the effects of increased employment on child health outcomes.

In order to estimate the causal impact of employment shocks on child human capital formation, I exploit the staggered roll-out of NREGA over time. The program was implemented across India covering all the districts in three phases over time; the first phase in 2006, second phase in 2007 and the third and final phase in 2008. I am not the first to exploit this variation in program roll-out timing; other recent literature studying the effects of NREGA, such as Imbert and Papp (2015), Shah and Steinberg (wp, 2016), Deininger et al (2016), also relied on the staggered roll-out of the program. I use the same variation in policy roll-out over time, and potentially provide a tighter econometric framework by employing an individual fixed effects framework to estimate causal effects of NREGA on health outcomes.

The data for this study comes from the Young Life Survey (YLS) which is a longitudinal household survey taking place in 2002, 2007, 2009, and 2014 in the state of AP. Households are selected to participate in the survey if they have a child that is aged between 6 and 18 months at the time of the first survey, in 2002. The survey collects data on child characteristics including anthropometric measures , health indicators and cognitive tests. It also consists of detailed data on mother's characteristics such as education and health.[1]

There are two crucial aspects of the data which make it ideal for this research study: 1) YLS sample districts received NREGA at different times which allows for variation in policy over time, and 2) The sample children are between 6 to 18 months old in 2002, or 4.5 to 5.5 years at the time of roll out of the first phase of the program. This allows for studying the differential effect of parental employment shocks by child age.

The main finding of the paper is that NREGA led to a long-term decline in child human capital - the program led to a 0.25 standard deviation decline in height for age z-score over a period of eight years from 2006 to 2014. This effect is larger for younger children ; the decline being a 0.4 standard deviation as compared to a 0.2 standard deviation for older children.

Lastly , this paper also contributes to the growing literature of effects of NREGA on child development. Islam and Sivasankaran (2014) study the effect of NREGA on child labor in India. Shah and Steinberg (2015) find that NREGA leads to higher test scores for older children aged 13 years or more. Interestingly, they find no such effect for younger children. Maity (2015) shows that NREGA led to changes in time allocation of children . Afridi et al (2016) show that the jobs welfare program led to lowering of school drop-out rate.

To the best of my knowledge, this is the first paper to study the effect of households' employment during the early childhood period. It is also the first to study the long-term

1. YLS also collected employment data for NREGA for each household member in the sample. Due to data quality issues, the current version does not take advantage of this information. The next version of the paper will use the cleaned-up employment data for a richer analysis.

effects of NREGA on human capital.

The rest of this chapter is organized as follows. Section 2 presents institutional details on NREGA. Section 3 describes the data used in the paper. Section 4 describes the econometric framework. Section 5 describes the results, section 6 has a discussion of the findings of the paper, and section 7 concludes.

2.2 Background on NREGA

2.2.1 Administrative Structure in Rural India

The administrative structure in India is multi-tiered. The lowest level of administrative structure is the village. Above villages, lies the Gram Panchayat (GP) which encompasses between 1,000 and 10,000 individuals in a group of villages. [2] Most of the federal and state government schemes are implemented at the GP level. The main responsibility of the GP is to administer the local public goods. The GPs fall under the administrative block called a mandal, which in turn are a part of a district .

2.2.2 National Rural Employment Guarantee Act (NREGA)

NREGA of India was enacted in 2005 with the objective of providing rural households with a safety net by providing up to 100 days of guaranteed employment. In the case employment is not provided, the households are entitled to compensation in the form of unemployment allowance. The scheme was intended to provide subsistence to poor households in rural India in the off-agriculture season or during lean times.

NREGA employment is assigned at the household level. To apply for employment under NREGA, the household must register with the local government at the GP level for a job card. A single job card is assigned to the household and the job card contains the names and photographs of all household members.

2. A GP can contain between 1 to 15 villages.

The kind of jobs provided under NREGA involve unskilled manual labor. Some examples of this are digging up temporary water reserves and painting street signs. One salient feature of the program is that it requires one-third of beneficiaries to be women.

The wage rate for NREGA program employment is the state minimum wage. This implies that those workers employed by NREGA were previously unemployed[3]. Hence, I argue that if the households receive employment under NREGA, then this comes directly at the cost of time spent at home rather than substitution from other forms of employment. If these households have a young child, then the employment of caregivers would come at the cost of time away from taking care of the young child.

The program was rolled out in three phases over time. The districts were ranked in ascending order of '*backwardness*' index, and the first 200 districts began the guaranteed jobs program in February 2006. The second phase began in April 2007 and included an additional 130 districts, and third phase, of 285 districts began in April 2008.

2.3 Data

I use the Young Lives Survey (YLS) data set for Andhra Pradesh, a 4-wave panel data set (2002, 2007, 2009 and 2014) focused on childhood poverty covering 20 mandals in 6 rural districts. The YLS survey instrument consists of a child questionnaire, a household questionnaire, and a community questionnaire. YLS follows n=2011 children that were between 6 and 18 months of age in 2002. For the initial sampling, each mandal was divided in 4 geographical areas, and one village was randomly selected in each area. YLS then randomly sampled 100 households with a one-year old child in each mandal. [4]

3. This is a weak assumption, and can be tested using the NREGA program data. I plan to do this in the next version of the paper.

4. YLS sampled 6 rural districts, two districts in each of three agro-climatic regions and selecting one poor and one non-poor district in each region: Srikakulam and West Godavari in the Coastal region; Anantapur and Cuddapah in Rayalaseema; and Mahboobnagar and Karimnagar in Telangana. The mandal or block is the administrative unit just below the district. After ranking all the mandals sites in these districts on the basis of a selected set of indicators of economic, human development and infrastructure, 20 mandals were sampled from these districts. In parallel, YLS also sampled a second group of 1000 children that were

The rural sample consists of 1506 children. After dropping observation where the age of the child was less than 6 months or more than 18 months in 2002, the sample size reduces to 1491. In the first round of the survey, data on height is missing for 17 children, which makes the final sample size for this study equal to 1474.

The first round of the YLS was conducted from September 2002 to December 2002, when the children were aged 6 to 18 months. The second round of data was collected from January 2007 to July 2007, approximately a year after first phase of NREGA came into effect. The third and fourth rounds of the YLS were conducted in 2009 and 2014 respectively. All the districts in the country, and hence in the sample, had received the rural jobs guarantee program before the third round of survey.

The timing of the YLS round 1 (2002) allows me to look at the pre-program characteristics of the sample across different phases of NREGA, as this was prior to the roll out of the first phase of the program. In round 2 (2006) data, the phase 1 districts (early districts from here on) have already received the program, but the phase 2 and phase 3 districts (late districts from here on) are not yet covered. All districts in the sample are covered under NREGA by the time round 3 (2009) and round 4 (2014) data was collected.

The YLS data is ideal for the purposes of studying early childhood development and the long run effects of interventions in early childhood. One major advantage of the long panel is that it allows me to track the medium to longer-term impact of employment shocks during the early childhood. NREGA impacted children early in life and I look at its effect on later health outcomes to study the persistence of the effects of the program.

The main outcome of interest is the child's health status as measured by internally standardized height for age z-score. This health measure is available for all four rounds, and hence allows me to study the change in outcome of interest before and after the program.

on average 8 years old in 2002. I do not use this second sample given the focus on early childhood. More generally, the YLS initiative collected data on 12,000 children in 4 countries over 15 years. It is led by a team in the Department of International Development at the University of Oxford in association with research and policy partners.

Table 2.1: Pre-program Characteristics: Child characteristics in 2002

	(1) Early Districts (Phase 1)	(2) Late Districts (Phase 2 & 3)	(3) Δ(Early-Late)	(4) p-value	(5) N
Child Characteristics					
Gender (Female 0, Male 1)	0.536	0.521	0.015	0.494	1491
Age in months	11.864	11.881	-0.017	0.938	1491
Dummy for being the only child	0.504	0.571	-0.067	0.245	1491
Starting month for consuming solid foods	6.985	7.202	-0.217	0.688	1195
Weight in Kg.	7.770	7.757	0.013	0.895	1475
Height in inches	71.702	70.751	0.950	0.131	1474
Z-score of weight for age	-1.658	-1.664	0.007	0.960	1475
Z-score of height for age	-1.302	-1.676	0.374	0.132	1474

*** $p<0.01$, ** $p<0.05$, * $p<0.1$ Standard error clustered at the mandal level.

Z-score is calculated by subtracting the sample mean and dividing by the sample standard deviation for each round of the survey .

2.3.1 Pre-program Characteristics of Early and Late Districts

Table 2.2 shows that there is no statistical difference between the early and late districts in the level of prenatal care investment. Prenatal care is a critical input in the health of young children, so any difference between the early and late districts would indicate that one group had a potential head start. No statistically significant differences in prenatal care is suggestive of similar child care investment prior to the program roll-out between the early and late districts. In addition, the children in both the early and late districts are equally

Table 2.2: Pre-program Characteristics: Vaccination and Prenatal care variables in 2002

	(1) Early Districts (Phase 1)	(2) Late Districts (Phase 2 & 3)	(3) Δ(Early-Late)	(4) p-value	(5) N
Dummy for whether received the following vaccination, Yes -1 No-0					
BCG vaccine	0.901	0.946	-0.044	0.067	1490
Measles vaccine	0.717	0.718	-0.001	0.984	1491
Polio vaccine	0.945	0.968	-0.023	0.227	1490
Prenatal care dummy variables for child's mother, Yes-1 No-0					
Any antenatal visits during pregnancy?	0.854	0.941	-0.086	0.158	1470
>2 tetanus injections during antenatal visits?	0.971	0.976	-0.005	0.743	1296
Given iron folic tablets during antenatal visits?	0.951	0.921	0.030	0.428	1295
Took iron folic tablets for at least 3 months?	0.941	0.900	0.041	0.401	1218

*** p<0.01, ** p<0.05, * p<0.1 Standard error clustered at the mandal level. BCG vaccine is taken for prevention of tuberculosis.

likely to receive the measles and polio vaccines. There is a small difference between the two group in the fraction of children receiving BCG (Bacillus Calmette-Guérin) vaccination meant for prevention of tuberculosis; early districts children were 4% more likely to have been vaccinated. In my later analyses, I control for the prenatal and vaccination variables in Table 2.2, in addition to the child characteristics in Table 2.1.

Finally, Table 2.3 shows a comparison of mother and household characteristics for early and late districts. Mother's characteristics are one of the key determinants for child human capital. The average years of formal education are extremely low for both the early and late districts; average mother in early districts had 1.8 years of formal education while the

Table 2.3: Pre-program Characteristics: Mother and Household Characteristics in 2002

	(1) Early Districts (Phase 1)	(2) Late Districts (Phase 2 & 3)	(3) Δ(Early-Late)	(4) p-value	(5) N
Mother's characteristics					
Age in years	23.630	23.357	0.273	0.454	1481
Years of formal education	1.786	2.738	-0.951	0.011	1480
Household characteristics					
Housing quality index	0.410	0.430	-0.020	0.740	1491
Consumer durable index	0.131	0.141	-0.010	0.612	1491
Wealth index	0.341	0.304	0.038	0.177	1491
Does dwelling have electricity?	0.833	0.640	0.193	0.011	1491
Household size	5.816	5.175	0.641	0.058	1491
Ethnicity, Yes -1 No -0					
Scheduled caste	0.257	0.115	0.142	0.009	1491
Scheduled Tribe	0.106	0.334	-0.228	0.189	1491
Backward Caste	0.478	0.461	0.017	0.912	1491

*** $p<0.01$, ** $p<0.05$, * $p<0.1$ Standard error clustered at the mandal level.

late districts fare slightly better at 2.7 years of formal education. Though this difference is statistically significant, it is not necessarily a meaningful difference for child's human capital investment, given the extremely low levels of education. Nonetheless, I control for mother's education in the all specifications, in addition to the child and household controls mentioned above.

I start with comparing the early and late districts prior to the roll-out of NREGA in 2006. As mentioned above, survey wave 1 (2002) allows for the comparison between the early and the late districts prior to the roll out of NREGA. Table 2.1 showed the comparison between the mean of the child characteristics in the year 2002. The children are on average 12 months old in 2002. Children in the early and late districts look similar in terms of gender composition, age, birth order, and other characteristics, with all the differences in mean values being statistically indistinguishable from zero. The early districts seem to have an advantage in height for age in 2002, but the difference is not statistically significant.

2.4 Econometric Framework

I use height for age (hfa) as proxy for human capital. The internalized z-score for hfa is then the main outcome of interest to measure the impact of employment shocks on children's human capital.

To identify the effect of NREGA on the outcomes of interest I employ three different empirical strategies. I describe the three methods, the required assumptions and their advantages and disadvantages below:

2.4.1 Cross section OLS

In the first specification, I employ a straightforward cross-sectional comparison between the early and late districts. The linear relationship estimated is:

$$Y_{id} = \delta NREGA_d + \kappa_{id} + \gamma_{md} + \varepsilon_{id} \tag{2.1}$$

where Y_{id} is the outcome of interest for child i in district d. Variable $NREGA_d$ is the treatment variable dummy which equals 1 for NREGA phase 1 districts and 0 for NREGA phase 2 and 3 districts. Coefficient of interest is δ which measures the effect of exposure to NREGA on the outcome of interest. This is estimated separately for rounds 2, 3, and 4 of the survey. κ_{id} represents pre-program child characteristics such as gender, age, prenatal care and vaccination, mother's characteristics such as age, ethnicity and education and household characteristics such as household size, wealth index and weather-related crop failure shock indicator. γ_{md} is the mandal fixed effects to account for unobservable characteristics at the mandal level. Replacing mandal fixed effects with district level fixed effects does not change the results. The standard errors are clustered at the mandal level as it is the highest rural administrative unit at which NREGA was implemented. Clustering the standard errors at district level yields similar results.

The *treatment* variable in this specification, $NREGA_d$ can be interpreted as an interaction of actual NREGA roll-out and a child age dummy variable that takes the value 1 if the children receive NREGA at age less than 5 years. The late districts are considered comparison group even though these districts end up getting the program later. The appeal of such a cross sectional comparison is that it allows to study the short, medium and long term effect of employment using the waves of survey data in years 2007, 2009 and 2014. However, the estimates from the cross-sectional relationship are potentially biased as Covariance($NREGA_d$, ε_{id}) is potentially not equal to 0. The estimates could be driven by the unobservable differences between the districts. In other words, since the early and late districts are plausibly not as if randomly assigned, the estimates could be driven by the unobservable differences between the early and late districts other than the assignment.

2.4.2 Difference in Differences (DID)

To avoid making the strong assumption that NREGA assignment is exogenous, I estimate the effect of NREGA off within-district variation by controlling for district fixed effects along with year fixed effects. The linear relationship estimated now becomes:

$$Y_{idt} = \delta NREGA_{dt} + \phi_{idt} + \kappa_{id} + \gamma_d + \lambda_t + \varepsilon_{idt} \tag{2.2}$$

where Y_{idt} is the outcome of interest for child i in district d and year t. Variable $NREGA_{dt}$ is the treatment indicator variable which equals 1 if a district has NREGA in year t. Therefore, the treatment indicator is equal to 0 for all districts in wave 1 (year

2002), equal to 1 for early districts and 0 for late districts in wave 2 (2007), and equal to 1 for all districts in wave 3 (2009) and wave 4 (2014). ϕ_{idt} consists of time variant controls such as child and mother ages and weather shocks. κ_{id} includes the same pre-program child, mother, and household characteristics as equation 2.1. γ_d and λ_t represent district and year fixed effects respectively.

The DID identifying assumption is that, prior to the NREGA phase 1 roll-out, the early and late districts have similar trends in the outcome variable over time. Due to lack of available data on height for age at the district level, I am unable to directly test this assumption.

However, using the census data, I can test for pre-trends in determinants of height for age: economic factors: district GDP and per capita GDP; health care infrastructure: number of primary health care centers (PHC); health outcomes infant mortality rate (IMR); and female education: female literacy rate. Figures 2.1 to 2.5 show trends in each of these factors prior to NREGA roll-out.

Figure 2.1: Early vs Late Districts: Average GDP over time prior to NREGA roll out

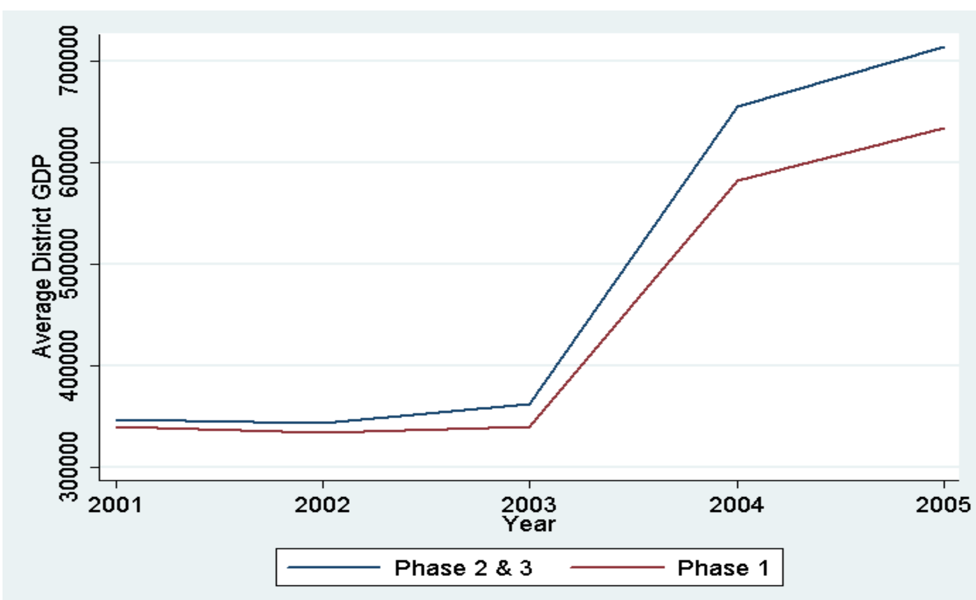
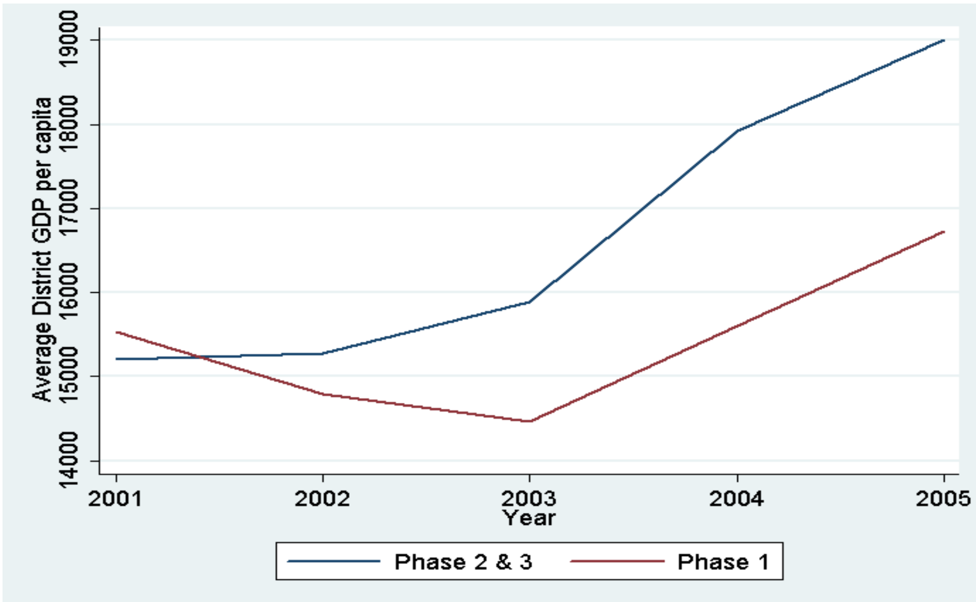


Figure 2.2: Early vs Late Districts: Average GDP per capita over time prior to NREGA roll out



Economic status of a district is plausibly correlated with investment in health which can drive the health outcomes for children. I check for trends in early and late districts' economic status measured by their GDP and GDP per capita. Figures 2.1 and 2.2 show that early and late districts have parallel trends in these wealth indicators, district GDP and per capita GDP respectively, prior to the roll our NREGA phase 1. Education level of the caregiver has been shown to be strong predictor of height for age and I show that there does not seem to be any differential time trend in literacy level between the early and late districts. Figure 2.4 shows that the pre-trends in female literacy rates are similar over time. The time trend for total literacy rate (not shown) is identical.

Figure 2.3: Early vs Late Districts: Average number of Primary Health Care centers over time prior to NREGA roll out

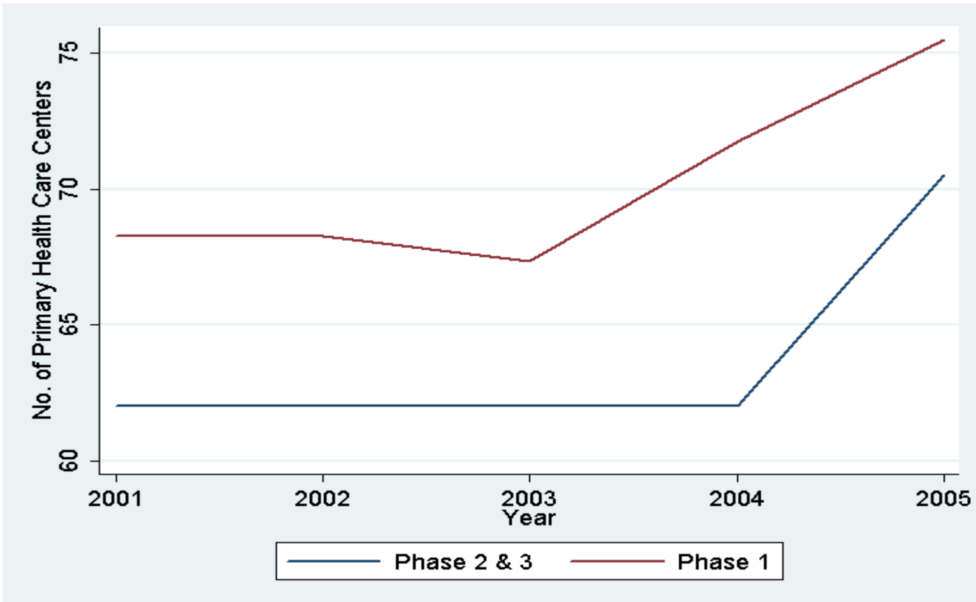
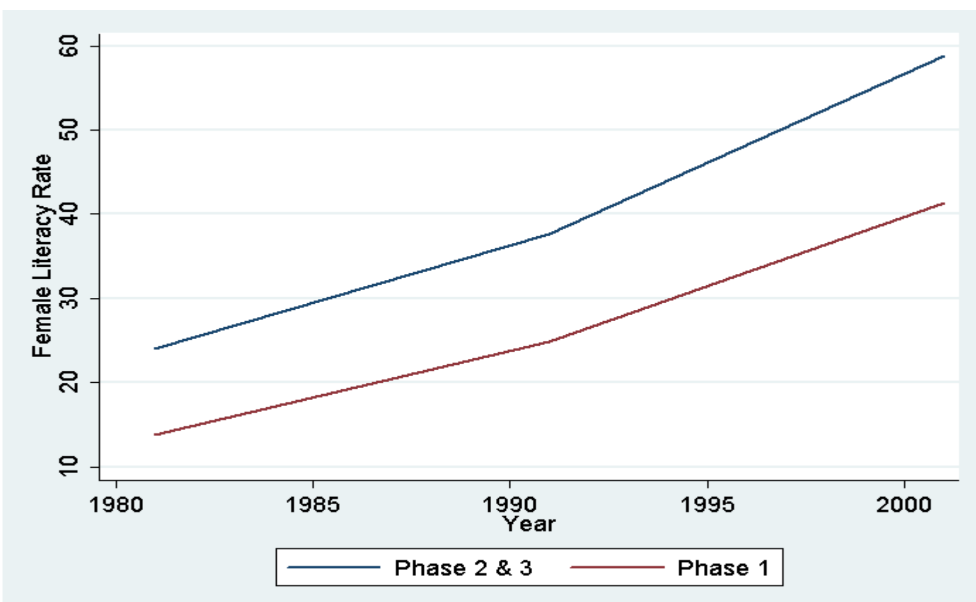
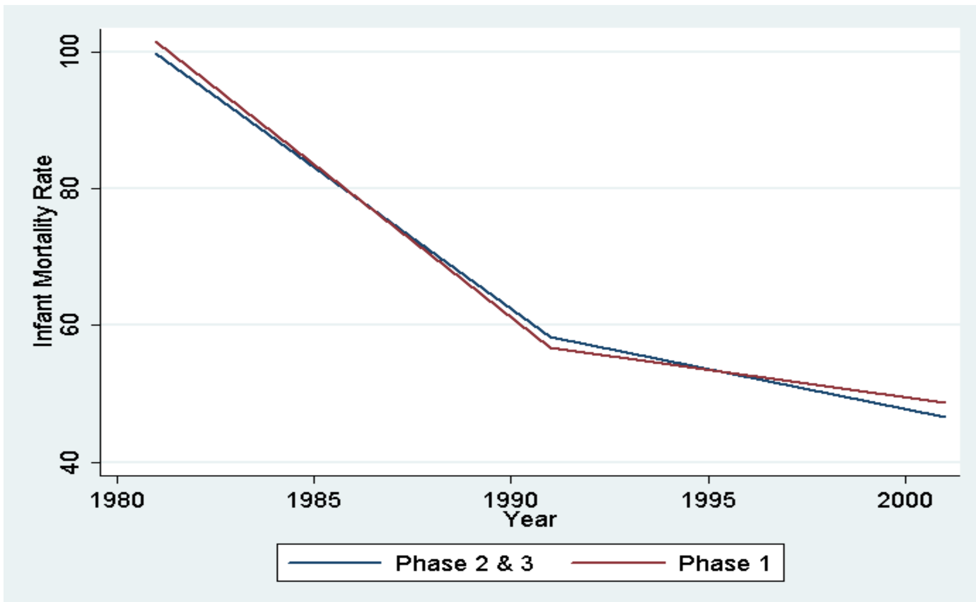


Figure 2.4: Early vs Late Districts: Average Female Literacy Rate over time prior to NREGA roll out



Finally, and perhaps most crucially, figure 2.5 shows that the Infant Mortality Rate (IMR) for under 1 year of age, is the same in level and trend over time for early and late districts. This suggests that there is no selection due to death in the children across the two groups. Comparison of health infrastructure for early vs late districts is shown in figure 2.3 and shows the same trend as previous figures.

Figure 2.5: Early vs Late Districts: Average Infant Mortality Rate (IMR) over time prior to NREGA roll out



To the extent that trends in height for age are driven by income, health inputs, mother's education and selection by survival, it is reasonable to assume that the early and late districts have parallel trends in height for age prior to the introduction of NREGA phase 1.

2.4.3 Child fixed effects

The DID specification does not account for unobserved heterogeneity at the individual level. This unobserved heterogeneity could arise from unobserved differences among households such as differences in altruism, or utility derived from investment in children, or other behavioral factors that affect the human capital investment decision.

I account for these potential differences by using a child fixed effects estimation. This is equivalent to using household fixed effects as the sample has one child per household. This fixed effects approach controls for all time invariant child characteristics that could potentially affect height for age. This allows me to identify the effect of employment shocks on the outcome of interest using within child variation.

I estimate the following linear relationship:

$$Y_{idt} = \delta NREGA_{dt} + \phi_{idt} + \kappa_{id} + \gamma_d + \lambda_t + \alpha_i + \varepsilon_{idt} \tag{2.3}$$

This is same as equation 2.2, with the addition of child fixed effects, α_i.

All the child and household level time invariant characteristics such as child's gender, prenatal care indicators, immunization indicators, and ethnicity get absorbed in the child fixed effects.

If the district indicator for child location does not change over the duration of the panel, then the district indicator would get subsumed by the fixed effects. However , not employing district fixed effects would implicitly assume that there is no migration and could potentially confound the estimates. I estimate the child fixed effects both with and without the district fixed effects. The results remain largely unchanged , which is reassuring. This is partially driven by the low rates of migration between the districts.

2.5 Results

2.5.1 Cross Section Results: Early vs Late Districts

Table 2.4 shows estimation results from equation 2.1 above. The main independent variable is the early vs late district binary variable, $NREGA_d$, that takes the value 1 for the early phase and 0 for the late phases of the program roll-out. The OLS estimates of differential effect of NREGA in equation 2.1 will be biased if the roll-out of NREGA is correlated with unobservable characteristics at the district level. The implicit assumption here is that the NREGA roll out was done as if randomly. This is a strong assumption and the results should hence be interpreted with caution. With this caveat in mind, I begin analyzing the effect of NREGA over time with the cross section estimates in Table 2.4.

Table 2.4: Cross-section: Height for Age Z-Score

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable: Height for Age z-score							
Short Term (Year 2007)							
NREGA	-0.181**	-0.333***	-0.0835***	-0.363***	-0.569***	-0.592***	-0.605***
	(0.0852)	(0)	(0.0189)	(0.0255)	(0.0693)	(0.0692)	(0.0719)
p value	0.0368	0	3.01e-05	0	0	0	0
Observations	1,421	1,421	1,421	1,420	1,410	1,410	1,410
R-squared	0.006	0.113	0.115	0.120	0.136	0.147	0.147
Medium Term (Year 2009)							
NREGA	-0.139	-0.567***	-0.539***	-0.553***	-0.523***	-0.529***	-0.531***
	(0.0885)	(0)	(0.0172)	(0.0198)	(0.0630)	(0.0638)	(0.0636)
p value	0.121	0	0	0	0	0	0
Observations	1,462	1,462	1,462	1,461	1,451	1,451	1,451
R-squared	0.003	0.102	0.106	0.109	0.126	0.131	0.132
Long Term (Year 2014)							
NREGA	-0.100	-0.943***	-0.933***	-0.811***	-1.168***	-1.172***	-1.174***
	(0.0747)	(0)	(0.0109)	(0.0170)	(0.0724)	(0.0724)	(0.0711)
p value	0.184	0	0	0	0	0	0
Observations	1,452	1,452	1,452	1,450	1,440	1,440	1,440
R-squared	0.002	0.095	0.096	0.099	0.113	0.131	0.134
Child Controls			X	X	X	X	X
Immunization				X	X	X	X
Mother Controls					X	X	X
HH Controls						X	X
Shocks							X
District FE		X	X	X	X	X	X

Notes: *** p<0.01, ** p<0.05, * p<0.1 Standard error clustered at the mandal level are in the parentheses. Results remain unchanged when standard errors are clustered at the district level

Panel A shows the short-term effect of NREGA in the early districts as compared to the late districts on average height for age z-scores of children as measured in year 2007. Panel B and Panel C represent the medium and long-term effect of NREGA as measured in the years 2009 and 2014 respectively.

These results show a persistent effect of NREGA exposure on early districts as compared to the late districts. Panel A shows that NREGA roll out led to a 0.5 standard deviation decrease in height for age in children in the early districts as compared to the late districts. This negative effect is similar in magnitude in the medium-term results in panel B, and the adverse effect of NREGA exacerbates to -1.1 standard deviation decrease in height for age z-scores in the long run (Panel C).

This result is striking given that the late districts also receive the NREGA after a gap of few years. In other words, the late group also receives the treatment with the only difference that the children in late districts are older when their households receive jobs under NREGA in comparison to the early districts. Hence, one interpretation of these results is that the effect of caregivers' employment is most crucial for children in the earlier years, which are crucial for development.

Given the strong assumption involved, however, these results should only be treated as directional. I come back to the early childhood mechanism later in the paper.

2.5.2 DID Results

To avoid making the strong assumptions required in the cross-section analysis, I exploit the panel dimension of the data and estimate the DID framework described above in equation 2.2. The main identifying assumption here is that in the absence of NREGA, the early and the late districts would have similar trends over time in height for age of the sample children, commonly referred to as the parallel trend assumption in the literature.

The children in the sample are, on average, 1 year old in 2002, which is the only pre-program round of data available. Hence, I cannot directly test the parallel trends assumption

Table 2.5: DID: Height for Age Z-Score

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable: Height for Age z-score							
			Panel A: FULL SAMPLE				
NREGA	-0.164**	-0.229***	-0.236***	-0.235***	-0.232***	-0.233***	-0.249***
	(0.0692)	(0.0711)	(0.0715)	(0.0715)	(0.0718)	(0.0721)	(0.0766)
p value	0.0198	0.00183	0.00142	0.00149	0.00177	0.00172	0.00167
Observations	5,824	5,824	5,824	5,818	5,769	5,769	5,358
R-squared	0.004	0.086	0.090	0.094	0.105	0.112	0.117
			Panel B: GIRLS				
NREGA	-0.249***	-0.220**	-0.224**	-0.225**	-0.227**	-0.228**	-0.266**
	(0.0766)	(0.0979)	(0.0981)	(0.0978)	(0.0989)	(0.0987)	(0.102)
p value	0.00167	0.0277	0.0247	0.0240	0.0244	0.0237	0.0111
Observations	2,733	2,733	2,733	2,696	2,696	2,491	2,491
R-squared	0.117	0.118	0.119	0.122	0.136	0.142	0.149
			Panel C: BOYS				
NREGA	-0.266**	-0.241***	-0.250***	-0.248***	-0.241***	-0.242***	-0.242***
	(0.102)	(0.0874)	(0.0873)	(0.0874)	(0.0880)	(0.0888)	(0.0906)
p value	0.0111	0.00713	0.00527	0.00573	0.00760	0.00786	0.00906
Observations	3,091	3,091	3,085	3,073	3,073	2,867	2,867
R-squared	0.149	0.128	0.134	0.137	0.145	0.154	0.156
Child Controls			X	X	X	X	X
Immunization				X	X	X	X
Mother Controls					X	X	X
HH Controls						X	X
Shocks							X
District FE		X	X	X	X	X	X
Year FE		X	X	X	X	X	X

Notes: *** p<0.01, ** p<0.05, * p<0.1 Standard error clustered at the mandal level are in the parentheses. Results remain unchanged when standard errors are clustered at the district level for height for age. However, as discussed in the section describing empirical strategy, the parallel trends in GDP, GDP per capita, health infrastructure, Infant Mortality Rate, and education levels are assuring and demonstrate that assumption of parallel trends in outcome variable is reasonable.

Table 2.5 shows the DID estimates from equation 2.2. The DID estimates show that over a period of 12 years, from 2002 to 2014, the NREGA employment shock to the household led to a 0.25 standard deviation decline in the height for age z-score. The specification in Table 2.5 gradually controls for various child, mother, and household level characteristics. Column 7, in addition to these controls, also accounts for weather induced crop failure shocks and is the preferred specification. The effect is slightly worse for girls, -0.27 standard deviation, than boys who experience a decrease of 0.24 over the same period.

2.5.3 Child fixed effects

The DID framework identifies the effect of NREGA off the within district variation. However, potential differences between children from different families could still be biasing the results. These differences could arise from unobserved differences in family structure or family behavior that could affect the investment in human capital. To address this issue, equation 2.3 controls for child fixed effects in addition to the controls of equation 2.2. Hence the effect of NREGA is now identified off within child variation.

The estimates from fixed effects estimation are in Table 2.6. The preferred specification in column 7 shows that the overall effect of NREGA is a 0.27 decrease in height for age z-score. The size of this effect is almost equal to the DID estimates, which is reassuring. Gender heterogeneity results are also close to the DID results in magnitude as well.

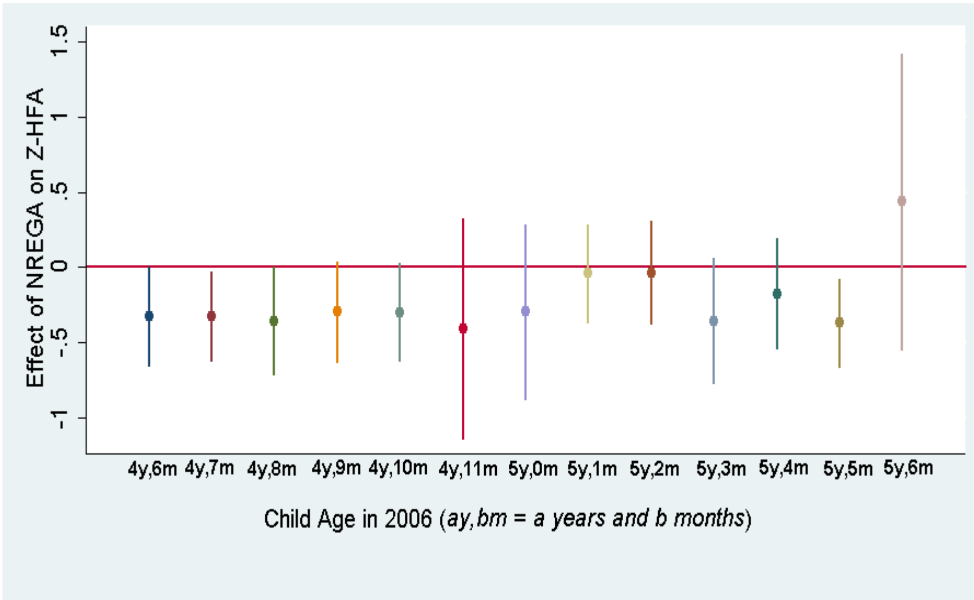
Table 2.6: CHILD FIXED EFFECTS: HEIGHT FOR AGE Z-SCORE

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable: Height for Age z-score							
			Panel A: FULL SAMPLE				
NREGA	-0.180**	-0.227***	-0.228***	-0.226***	-0.227***	-0.227***	-0.269***
	(0.0791)	(0.0724)	(0.0725)	(0.0726)	(0.0728)	(0.0728)	(0.0758)
p value	0.0254	0.00241	0.00230	0.00253	0.00255	0.00255	0.000652
Observations	5,764	5,764	5,764	5,758	5,709	5,709	5,304
R-squared	0.010	0.024	0.024	0.024	0.025	0.025	0.028
			Panel B: GIRLS				
NREGA	-0.269***	-0.198**	-0.198**	-0.198**	-0.203**	-0.203**	-0.248**
	(0.0758)	(0.0958)	(0.0966)	(0.0966)	(0.0978)	(0.0978)	(0.104)
p value	0.000652	0.0419	0.0441	0.0441	0.0410	0.0410	0.0193
Observations	2,705	2,705	2,705	2,668	2,668	2,466	2,466
R-squared	0.028	0.035	0.035	0.035	0.038	0.038	0.043
			Panel C: BOYS				
NREGA	-0.248**	-0.252***	-0.254***	-0.251***	-0.246***	-0.246***	-0.286***
	(0.104)	(0.0894)	(0.0892)	(0.0896)	(0.0902)	(0.0902)	(0.0926)
	0.0193	0.00607	0.00560	0.00636	0.00768	0.00768	0.00270
Observations	3,059	3,059	3,053	3,041	3,041	2,838	2,838
R-squared	0.043	0.022	0.022	0.022	0.022	0.022	0.024
Child Age			X	X	X	X	X
Immunization				X	X	X	X
Mother Education					X	X	X
Mother Age					X	X	X
HH Controls						X	X
Shocks							X
District FE		X	X	X	X	X	X
Year FE		X	X	X	X	X	X
Child FE	X	X	X	X	X	X	X

Notes: *** p<0.01, ** p<0.05, * p<0.1 Standard error clustered at the mandal level are in the parentheses. Results remain unchanged when standard errors are clustered at the district level.

2.5.4 Importance of Early Childhood

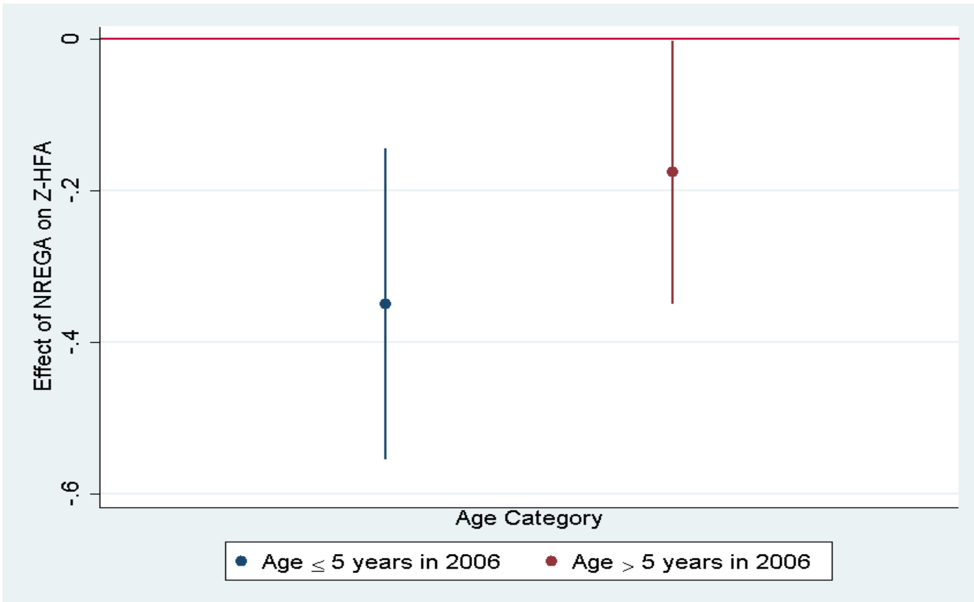
Figure 2.6: Effect of NREGA on Height for Age by age in year 2006



Early childhood investment has been shown to be crucial for human capital development. In the setting of this paper, the children in the sample were 4.5 to 5.5 years old when NREGA was first rolled out in 2006. Getting an NREGA job and the resulting time shock could assume more importance for the younger children. Given that this a crucial age for child development, it is important to see if the employment shocks had a differential impact by age. The persistence of cross section results suggested that the effect of employment shocks could be most severe for younger children. I further explore this using my preferred child fixed effects specification from column 7, table 2.6.

Figure 2.6 shows that the effect is homogeneous and consistently negative for children less than or equal to 5 years of age in 2006. The effect seems to taper off for children that were older than five at the time of NREGA roll-out and is even positive for some age groups.

Figure 2.7: Effect of NREGA on Height for Age by age category in year 2006



To bring out a clean analysis of heterogeneous effect of NREGA by age, I divide my sample into two age group categories (less than or equal to five years), and others (greater than five years), I estimate equation 2.3 separately for these two categories. Figure 2.7 confirms what is suggestive in figure 2.6: NREGA impact was worse for children of the age group less than 5 years. Table 2.7 shows that the effect is twice as large (-0.4) as compared to the children of age greater than 5 years (-0.2).

This is evidence of the importance of time investments in early childhood for human capital development. This striking result shows that the effect of shocks to the time investment at the age of less than five are persistent over a long period of time. This finding is one of the main contributions of this paper as it contributes to our understanding of the factors important during early childhood.

2.6 Discussion

The treatment implies a negative shock in the time investment in the child's human capital along with an increase in the household income. As argued above, it is not obvious which

Table 2.7: Effect of NREGA on HFA by Age Category

Dependent Variable: Height for Age z-score		
	Age $\leq$ 5 years in 2006	Age $>$ 5 years in 2006
NREGA	-0.349***	-0.175**
	(0.103)	(0.0871)
p value	0.00104	0.0472
Observations	2,922	2,411
R-squared	0.078	0.014

Notes: *** p<0.01, ** p<0.05, * p<0.1 Standard error clustered at the mandal level are in the parentheses. Results remain unchanged when standard errors are clustered at the district level. All regressions control for child's age, mother's age and education, household wealth, ethnicity, electricity access, district FE, year FE, and child FE direction the net effect will be. The results of this paper showing a long term negative impact on the height for age z-score, though surprising, is not counter-intuitive. There are two reasons for that: first, NREGA program provided employment at minimum wage, hence the effect on household income was potentially small. Second, the age heterogeneity result shows that the importance of caregivers' time investment is much higher for younger children. In poor households, the caregiver time investment may be even more important than in other settings as caregivers' time investment is arguably one of the only core investments.

To summarize, the mechanisms behind the result is clear: the increase in household income is unable to negate the effect of decreases in time investment by the caregiver on children's long term human capital.

But, how does the absence of caregiver manifest in health outcomes? It is important to understand what the absence of time investment by the caregiver implies. I explore this by looking at two potential factors that are plausibly correlated with caregivers' presence at home.

2.6.1 Food Intake

As hypothesized earlier, caregivers' employment could affect the child's health in many ways. If the child is dependent on the caregiver for food and nutrition, then one potential mechanism could be that young children in early districts received fewer meals and hence potentially less nutrition as compared to the late districts. If this deprivation occurs at earlier ages, this could lead to a serious negative impact on the development of the child.

In table 2.8, I investigate if NREGA led to children getting fewer meals. The YLS asks whether a child has consumed a meal in the last 24 hours at different times of the day: before breakfast, breakfast, between breakfast and lunch, lunch, between lunch and dinner, dinner, after dinner.

I use these to construct two versions of total meals consumed during the day. Version 1 simply adds up the binary variables for before breakfast, breakfast, between breakfast and

Table 2.8: MECHANISM: FOOD INTAKE

VARIABLES	(1) Number of Meals during the day (ver 1)	(2) Number of Meals during the day (ver 2)
NREGA	0.606	0.786
	(0.409)	(0.503)
p value	0.142	0.122
Observations	5,387	5,387
R-squared	0.083	0.082

Notes: *** p<0.01, ** p<0.05, * p<0.1 Standard error clustered at the mandal level are in the parentheses. Results remain unchanged when standard errors are clustered at the district level. All regressions control for child's age, mother's age and education, household wealth, ethnicity, electricity access, district FE, year FE, and child FE. Column 2 includes all meals counted in column 1 and meals between midday and evening. Result for total number of meals during a 24 hour cycle (not shown) are similar.

lunch, and lunch. Version 2 adds between lunch and dinner meal consumed indicator to version 1.

The results for version 1 and 2 of the total meals consumed during day are in column 1 and 2 respectively of table 2.8. The results suggest there is no statistically significant difference between the number of meals consumed during the day caused by NREGA. There is no difference between the total number of meals in a 24-hour day (not shown). However, this result is not very insightful in understanding the nutrition mechanism as number of meals is a poor proxy for nutritional intake. Unfortunately, due to data limitations, I am unable to directly test for nutritional intake differences.

2.6.2 Time Allocation

Another consequence of the employment of the caregiver could be that, due to substitution of time towards employment, the child now is now forced to allocate time towards household activities. This could include helping with household chores or taking care of younger siblings or older grandparents. If the children in the household are required to substitute away from activities that are conducive to their development, then this could lead to long term

Table 2.9: Mechanism: Time Allocation

	(1)	(2)	(3)	(4)
	Hours spent doing the following activities in the last 24 hours:			
	Sleeping	Household Chores	Unpaid Work	Paid Work
NREGA	-0.164*	0.137***	-0.0556	-0.0216
	(0.0982)	(0.0486)	(0.0427)	(0.0229)
p value	0.0984	0.00595	0.197	0.350
Observations	3,757	3,756	3,756	3,757
R-squared	0.273	0.293	0.022	0.009

Notes: *** p<0.01, ** p<0.05, * p<0.1 Standard error clustered at the mandal level are in the parentheses. Results remain unchanged when standard errors are clustered at the district level. All regressions control for child's age, mother's age and education, household wealth, ethnicity, electricity access, district FE, year FE, and child FE. Household chores includes time spent taking care of family members.

disadvantages in health outcomes. If the caregivers' employment places the burden of taking acre of household members or doing household chores instead of investing in their own human capital then this might be detrimental for their health.

I use the questions on time allocation of children to investigate this hypothesis. Rounds 2, 3, and 4 of the YLS ask about children's time allocation in terms of number of hours spent sleeping, taking care of other household members, doing household chores, work outside home for pay and otherwise, and playing.

Table 2.9 shows that the caregivers' employment altered the time allocation of the sample children. On average, NREGA led to children spending an additional hour per week[5] on household activities such as household chores or caring for other household members. This is in line with the findings of Maity (2015).

One implication of this result is that these children are substituting away from activities that augment their health such as sleeping. Column 1 suggests that NREGA was associated with children spending less time sleeping on average.

5. 0.14 hours per day

Together, these results suggest that when the household members get employed through NREGA jobs, children allocate some of their time in helping with household chores. This substitution away from potential human capital investment activities could further help us understand the long run negative effect on health outcome.

2.7 Conclusion

Using the Young Life Survey, I show that NREGA led to a long-term decline in the human capital of children. The cross-section results suggest that this decline is persistent over time.

Using differences in difference and child level fixed effects estimation, I estimate the effect of employment through NREGA off within child variation. Using census data, I show that the time trends in wealth, health infrastructure, child mortality, and mother's education level across the early and late districts are parallel. The parallel trends in all these factors that are deterministic of human capital (as shown in literature), allow me to construct differences in difference estimate of the effect of NREGA.

To further account for heterogeneity at the child level, I use a child fixed effects framework to estimate the effect off within child variation over a period of 12 years. This allows me to account for time-invariant unobserved heterogeneity at the child level. The estimates from the two specifications are almost the same, which is reassuring for the assumption in the DID design.

The age heterogeneity results are crucial. I show that the effect of caregivers' employment is twice as large for children of age less than 5 years at the time of the program rollout as compared to children of age more than 5 years. This is a crucial result and underscores the importance of age less than 5 years for investment in child human capital. The persistence of the negative effect of the employment shocks shows that the early childhood period is crucial from the human capital investment point.

REFERENCES

[1] Afridi, Farzana, Mukhopadhyay, Abhiroop and Sahoo, Soham. 2016. "Female labor force participation and child education in India: evidence from the National Rural Employment Guarantee Scheme." IZA Journal of Labor & Development 5, no. 1 (2016): 7.

[2] Cunha, Flavio, J. Heckman, James, Lochner, Lance and Dimitriy V. Masterov. 2006. "Interpreting the evidence on life cycle skill formation." Handbook of the Economics of Education 1 (2006): 697-812.

[3] Cunha, Flavio, and J. Heckman, James. 2008. "Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation." Journal of human resources 43, no. 4 (2008): 738-782.

[4] Deaton, Angus, and Drèze, Jean. 2009. "Food and nutrition in India: facts and interpretations." Economic and political weekly (2009): 42-65.

[5] Deaton, Angus. 2008. "Height, health, and inequality: the distribution of adult heights in India." The American economic review 98, no. 2 (2008): 468.

[6] Deininger, Klaus, Nagarajan, Hari K. and K. Singh, Sudhir. 2016. "Short-term effects of India's employment guarantee program on labor markets and agricultural productivity." (2016).

[7] Imbert, Clement, and Papp, John. 2015. "Labor market effects of social programs: Evidence from india's employment guarantee." American Economic Journal: Applied Economics 7, no. 2 (2015): 233-263.

[8] Islam, Mahnaz, and Sivasankaran, Anitha. 2014. "How does Child Labor respond to changes in Adult Work Opportunities? Evidence from NREGA." In International Conference on MGNREGA Impact, Indira Gandhi Institute of Development Research, Mumbai, October. 2014.

[9] Jayachandran, Seema, and Pande, Rohini. 2015. "Why are Indian children so short?. No. w21036. National Bureau of Economic Research, 2015.

[10] Maity, Bipasha. 2015. "Rural Employment Guarantee, Women's Participation and Children's Welfare: Evidence from India." (2015).

[11] Pathak, Yuvraj and Macours, Karen. 2017. "Women's Political Reservation, Early Childhood Development, and Learning in India," Economic Development and Cultural Change, 2017.

[12] Shah, Manisha, and Millett Steinberg, Bryce. 2015. "Workfare and Human Capital Investment: Evidence from India. No. w21543. National Bureau of Economic Research, 2015.

CHAPTER 3

SAVING FOR TOMORROW: COAL SUPPLY DISTORTIONS, STOCKPILING, AND POWER OUTAGES IN INDIA

3.1 Introduction

Access to electricity is strongly correlated with economic development across countries (Source: World Bank (2015)). In India, economic development relies upon cheap power generated from vast reserves of inexpensive coal expected to last more than 150 years. Despite a resurgence of interest in renewable sources of energy, coal-based power plants still supply over 80% of the total power generated in India[1]. If energy is the engine of growth, then the importance of coal for the Indian economy cannot be overstated. It is likely the reason that coal and power sectors are tightly regulated by the Government of India (GOI). Both sectors suffer from extremely limited access to markets and from persistent unmet demand. In such heavily regulated sectors with incomplete markets, uncertainty about future supply of factors of production can have a pronounced effect on forward-looking firms' production decisions and output.

In this paper, I describe how uncertainty surrounding the Indian government's announced changes to its coal allocation policy led to an increase in power plants' uncertainty about *future* coal supply. Using a difference-in-difference framework, I establish that increased regulatory uncertainty in the coal sector caused inefficient production by power plants and led to an increase in power outages. In demonstrating how distortions in coal supply can affect outcomes in the power sector, I highlight the role that supply side inefficiencies play in causing power outages in India. This paper contributes to the growing academic interest in studying the prevalence of power outages in India that has primarily focused on understanding the consequences of power outages (Alcott, Collard-Wexler and O'Connell (2016)) and demand side factors that cause outages (Mahadevan (2019)).

1. Source: Ministry of Power, India

Despite the impressive expansion of access to the power grid, shortages in power supply continue to plague electricity provision in India[2] with power outages widespread across the country. Both demand as well as supply side factors are to blame. On the demand side, the per-unit electricity prices paid by end consumers are regulated to be lower than the price at which power distribution companies or utilities purchase electricity from the power plants[3]. Since the utilities are engaged in a loss-making operation, they simply resort to power cuts. The amount of load shedding is planned in advance, and the utilities submit their expected requirement of power, net of planned power cuts, to the generators[4].

On the supply side, the generation capacity actually exceeds the amount needed to generate enough power to meet that the assessed demand of power distribution companies [5]. However, the generators suffer from frequent outages that result in failure to supply enough power to meet the energy requirement of the utilities, which is already net of planned power cuts [6]. This oddity of excess capacity and shortage of power is a salient feature of the power sector in India and implies inefficiencies on the supply side that might be affecting power generation and contributing to power outages.

I emphasize one such inefficiency in the form of distortions in coal supply due to regulatory uncertainty that affects plants' incentives to generate power. In 2014, the Supreme Court (SC) of India declared 214 coal block mining contracts illegal due to alleged corruption and ordered that the mining leases be voided at a future date[7]. Specifically, the court ordered the GOI to formulate rules for reallocation of the voided coal mining leases by March 2015.

2. Annual Reports, Ministry of Power, India

3. I use power plants and generators interchangeably throughout this text

4. I am careful to refer to demand of power by utilities as requirement. The reason is to distinguish the utilities' requirement of power, which already factors in the planned load shedding, with actual demand for electricity in India which is an abstract concept and numerically unknown. This is in line with the nomenclature used in the official Indian data and policy documents.

5. Source: Central Electricity Authority, India

6. Source: Load Generation Balance Reports, Central Electricity Authority, 2013-2017

7. The full Supreme Court of India Judgement on *coal-gate* can be accessed at `www.prsindia.org/sites/default/files/bill_files/SC_judgement_aug.pdf`

This order led to the beginning of a period of ambiguity about future coal supply that would last until, in theory, a new policy to allocate the cancelled mines came into force in March 2015. This landmark judicial decision led to unprecedented increase in uncertainty for the power plants subject to the ruling—that is, those firms with mining leases that were declared illegal. The rest of the plants had secure long-term fuel supply agreements from the state and were immune to the effects of the SC judgement. I exploit this quasi-experimental variation in future uncertainty about fuel supply faced by plants, and I use a difference-in-difference empirical strategy to demonstrate how the incentives and constraints that power plants face cause inefficient generation and lead to large economic losses. The empirical results of my paper are derived using a novel dataset that I assembled linking the universe of coal mines to power plants in India and allows for a rich analysis.

The Supreme Court's decision to order a reallocation of mining contracts came in response to the irregularities in the initial awarding of mines, dubbed in popular media as the *coal-gate* scandal. This judicial intervention in policy making remains the largest of its kind in India's history and was widely unexpected. I show that the affected plants responded to that intervention by declaring generation capacity under outage as a means to engage in precautionary saving of the generation fuel (coal) even before the cancellation of mines became reality. The incentive for plants to engage in precautionary saving comes from fixed charge component of the power purchase contracts (PPA) called capacity payments, which is contingent upon plants having at least 85% of their capacity available for generation on average. In order to be ready to generate power, plants need to have a minimum amount of coal. So, when faced with possibility of supply shocks in the future, they try to smooth their consumption of coal by saving for the uncertain future.

To analyze the effects of the mining contracts reallocation that induced future uncertainty, I use a straightforward two-period model to theorize how uncertainty about future fuel supply might affect power plants' production behavior. I show that if plants believe that future shortages are likely, they engage in precautionary saving and curtail production.

Second, using a difference-in-difference framework and rich panel data on plant input and output, I derive empirical estimates of the impact of regulatory uncertainty on power plants' production behavior using a number of different measures of output. The empirical strategy relies on quasi-experimental variation in uncertainty faced by plants as a result of the SC decision and compares the plants affected by the SC decision to the plants that had secure long-term fuel supply from state. The outcome of interest is the output at the plant level, which I measure both indirectly with data on daily capacity under outage (MW), monthly coal delivery, and consumption and stockpiling (Metric Tonnes per month) and directly with data on monthly electricity generation (GWh per month).

The empirical results show that regulatory uncertainty has a significant negative impact on firms' output and the direction of the empirical estimates is consistent with the model's predictions. Future uncertainty about coal supply causes plants both to declare their generation capacity to be under outage due to coal shortage and to stockpile coal even as their fuel deliveries see no change. This partial shutdown, induced by precautionary saving motive, resulted in a statistically significant and economically large loss of 7% of monthly output as measured by monthly electricity generation. The long-term results show that these effects persist over a long period of time (more than 3 years).

The shutting down of power plants not only causes reduction in aggregate output, but also increases the total cost of generation, because, on account of their proximity to the coal mines, the plants affected by the mining cancellation had lower generation cost relative to unaffected firms. To estimate the increased cost of generation, I simulate least-cost dispatch supply curves and show that the change in plants' production decisions led to an increase of approximately 0.3 billion dollars[8] in the total cost of generation. I estimate the value of lost generation using the results of Alcott, Collard-Wexler and O'Connell (2016) showing the impact of outages on manufacturing plants revenue. The estimated net value of lost generation load is approximately equal to 1.5 billion dollars. These two estimates form the

8. 290 Million dollars

lower and upper bound of the economic cost of regulatory uncertainty. The actual welfare loss is a convex combination of the two estimated bounds.

This paper, revealing the role of uncertainty in production decisions of power plants in India contributes to the literature on energy and development, precautionary savings, and regulatory frictions. To my knowledge, this is the first study to show how supply side constraints affect the production decisions of power plants in India. In doing so, this paper adds to the growing interest in understanding the electricity sector in India (Chan, Cropper and Malik (2014), Alcott, Collard-Wexler and O'Connell (2016)). This paper builds upon the relatively thin literature showing micro evidence-based effects of precautionary saving (Fuchs-Schündeln and Schündeln (2005), Fuchs-Schündeln (2008), Giavazzi and McMahon (2012)), and sheds light on the importance of precautionary saving in incomplete markets where uncertainty can play a big role in firms' production decisions.

In demonstrating the role of precautionary saving due to uncertain fuel supplies on plants' production behavior, this paper distinguishes itself from ongoing research on supply side of India's electricity sector which looks at market power (Nick Ryan (2018) *working paper*) and cost of non-market based dispatch (Burlig, Jha and Preonas (2019) *working paper*). Both of these studies focus on the Indian electricity exchange, IEX, which accounts for 5% of the total power transacted in India.

Recent literature on empirically measuring the effect of policy uncertainty on economic output by Baker et al (2016) has developed a unique index to measure uncertainty (Economic Policy Uncertainty Index) and provides robust estimates of its effect on macro output. I complement this literature by providing micro founded causal estimates of the impact of uncertainty on the output of a firm operating in incomplete markets. More recently, Bloom and co-authors (Bloom et al, working paper, 2019) show that Brexit induced uncertainty caused up to 5% reduction in firms' productivity in the UK. This paper complements prior studies by examining the role of policy uncertainty in incomplete markets specific to highly regulated coal and power sectors in India.

The focus of this paper is to study the behavior of power plants faced with regulation-induced future uncertainty about fuel sourcing. In doing so, it contributes to the understanding of how procurement of fuel by power plants is affected by regulatory frictions (Cicala (2015)). My paper also speaks to the broader energy economics literature with a twin focus on natural resource allocation and the electricity sector. In showing the uncertainty due to allocation woes in the coal sector in India, it emphasizes the challenges of natural resource allocation as a public good for both private and public consumption.

Lastly, this work contributes to our understanding of regulatory obstacles faced by firms in developing countries. Uncertainty arising from incomplete markets or regulatory frictions are not unique to power plants in India and affect a wide spectrum of firms in developing countries. This paper shows how such constraints can affect firms' decision making in non-market set-ups common in developing countries.

The end of this introduction is followed by section 2, in which I give a detailed description of the background of coal sector in India and the origin of uncertainty for some of the coal-fired power plants. Section 3 provides a theory model to anchor the empirical results. I report the data sources used in this paper and the information contained in those datasets in section 4. In section 5, I describe the estimation strategy and discuss the assumptions needed for identification. In section 6, I examine the short term and longer term results of my analysis. I estimate the cost of regulatory uncertainty in section 7. Finally, I discuss robustness checks in section 8, and section 9 concludes.

3.2 Background on Coal and Power sectors in India

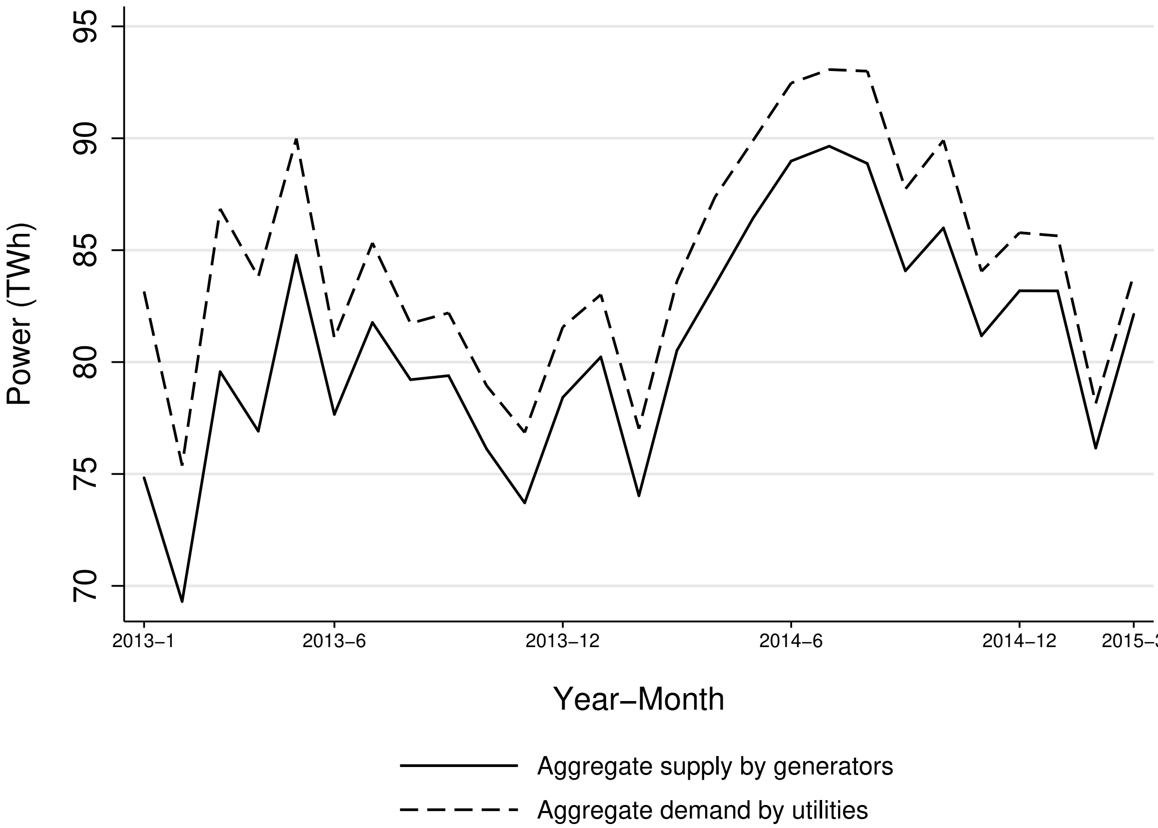
3.2.1 Brief History of Coal Allocation in India

Power generation in India is dominated by coal power plants[9], and a vast majority of the coal burnt by these plants comes from mines within India. In the year 2014, almost 95% of the coal consumed by the power plants came from Indian coal mines with the rest being sourced from imports[10]. However, thermal power plants in India frequently declare their generation capacity to be under outage causing the total supply of electricity to consistently fall short of the power requirement of the utilities. Figure 3.1 shows the aggregate supply of power by the generators and the requirement or assessed demand of power by the utilities. The trend shows that generators are unable to meet even the already rationed amount of power required by the utilities which weakly increases the blackouts at the consumer level.

9. Coal fired plants account for $\approx 80\%$ of the total electricity generated in India. Source: Ministry of Power, 2017

10. In terms of volume of coal consumed, in the year 2017, the total coal consumed in India was 800 million Metric Tonnes (MT) of which the power sector's share was close to 500 million MT (Provisional Coal Statistics 2017, Ministry of Coal).
Indian coal is 3-4 times cheaper than the alternative imported coal from Indonesia, South Africa or Australia

Figure 3.1: Total power supply by generators and demand by utilities



Notes: This figure shows the total supply and demand of power in TWh as reported in monthly power supply position reports by the Central Electricity Authority of India. The trend shows that generators are unable to meet the power requirements of the utilities.

The supply of coal used by these power plants has historically been controlled by the state through a public-sector utility company, Coal India Limited (CIL). The coal mines are a property of the state and, initially, private sector participation was not permitted. As India's economy grew, the demand for coal increased, and the inefficiency of allocating coal through a state-owned non-profit maximizing monopoly started showing up in the form of unmet demand for coal.

To ease the demand pressure on CIL, in the year 1993, India's Ministry of Coal starting allocating individual coal mine blocks to industries with exclusive extraction rights[11]. These

11. A natural solution to this problem would have been to open the coal sector to private sector and allow commercial mining. Market forces would then dictate that the firm with highest willingness to pay for coal

exclusive mining contracts were awarded for use of coal for a specific power plant or in a few cases a specific set of power plants, and any diversion of coal was prohibited. These mines are referred to as captive use mines or, simply, captive mines. The contract also specified the amount of coal a firm could extract based on their projected need of the existing plant. Unconsumed extracted coal was supposed to go back to the state at pre-notified fixed rate.

The awarding of captive mines for exclusive use of coal created two broad categories of power plants – 1) Plants that sourced coal from the state through long term fuel supply agreement (FSA), and 2) Plants that at least partially relied on their captive mine contract for fuel supply. The cancellation of captive mines thus only affected the latter category of plants.

The process of awarding captive mine contracts was slightly different for private firms and state-owned firms. Both state-owned and privately-owned utilities had to apply for a contract to the central government and a screening committee comprised of GOI cabinet ministers and other policy makers made the decision whether to award the contract or not. However, in some cases state owned utilities could also directly get a coal block allocated without going through the screening committee[12]. Neither the public sector firms nor the private firms owned the mines but only had the rights to extract coal from it.

The central issue with the allocation of captive mines in this manner was the absence of a defined rule to determine who gets the mining leases and who doesn't. This lack of transparency in awarding the mining leases made the allocation process seemingly arbitrary and impartial. And, this potentially corrupt allocation process of captive mining contracts is what formed the basis of *coal-gate* scam. In the year 2012, the Comptroller and Auditor General (CAG), the chief audit authority in India, published a report questioning the procedure followed in allocating coal blocks for captive use since the year 1993. This report formed the

would receive fuel for their operations. However, there was limited political support in the 90's for opening up coal sector to the market.

12. It is unclear how some state utilities got mining leases without going through the screening committee. I suspect that there were non-market forces at play but do not have any evidence in support

basis of a public interest litigation against the Government of India in the Supreme Court which led to the eventual de-allocation of these captive mines.

3.2.2 Coal-gate: Judicial decision led to future uncertainty

On August 25th 2014, the Supreme Court of India declared all captive coal block allocations made since 1993, total of 214 coal blocks, illegal - both to private and public entities. However, the mines that were producing coal were allowed to stay functional until March 2015, making March 31st 2015 the effective mining contract cancellation date for plants sourcing coal from the producing captive mines. As a part of the same judgment, the Government of India was asked to formulate a new policy to reallocate the cancelled mining leases by March 2015[13].

The SC decision to intervene in the allocation of state-owned natural resources was a widely unanticipated development, and it remains to date one of the biggest judicial intervention, in India.

For the power plants affected by the SC judgement, the immediate impact was unprecedented uncertainty about future supply of coal. These plants that were going to lose their captive mines, now depended on the outcome of reallocation of cancelled coal blocks after March 2015 for fuel supply. The resulting uncertainty about future fuel supply was starker for plants that were already receiving coal from their producing captive mines. At the time of the August 2014 SC judgment, 42 out of the 214 coal blocks were under production and 26 coal blocks were ready for extraction to begin. The remaining canceled coal blocks had not started production yet.

The power plants linked to the non-producing captive mines had been given temporary fuel supply contracts (called linkages) by the state as a stop gap arrangement until they can start extraction from their captive mine. Importantly, these 'temporary' supply arrangements came with the possibility of an extension beyond their end date. The August 2014

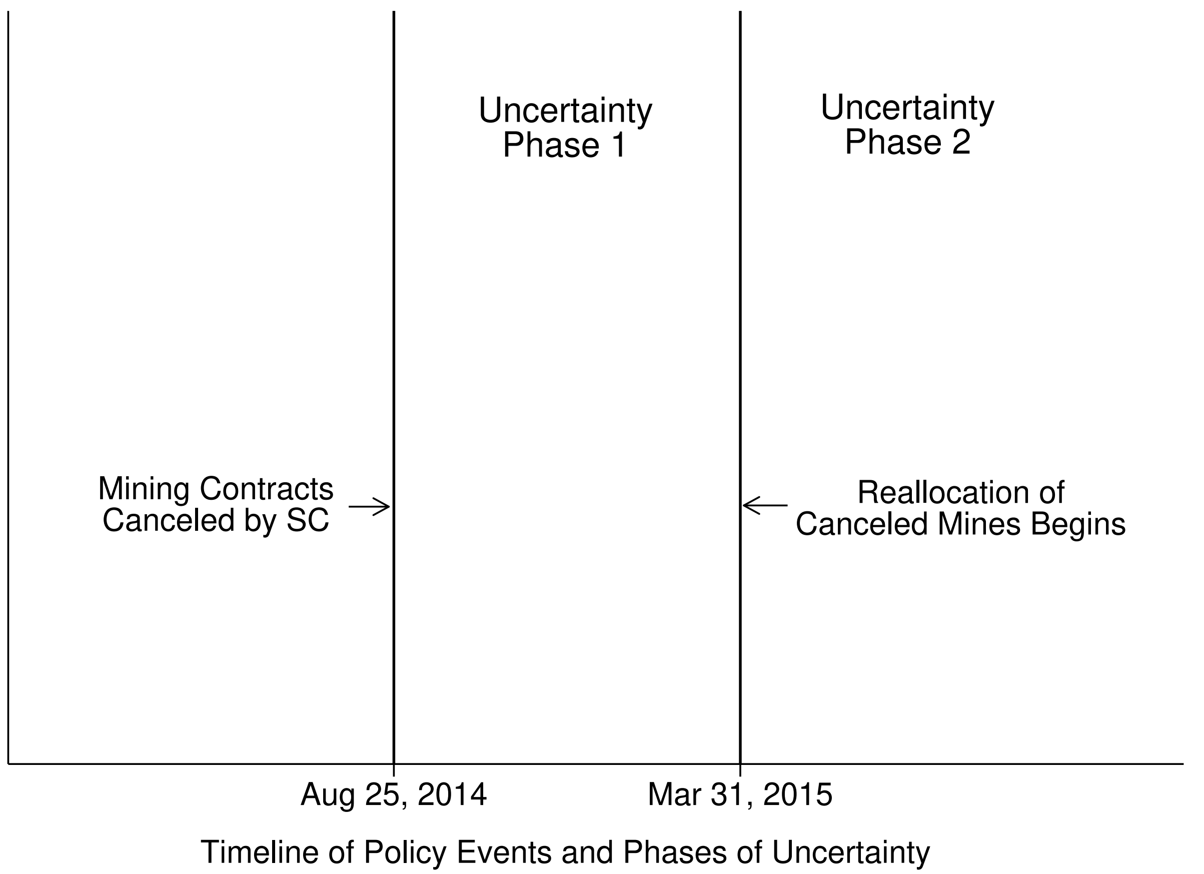
13. Coal Mines (Special Provisions) Act, 2015

judgment canceled both the non-producing captive mine allocations as well as the temporary linkages that were made against the non-producing captive mine allocation. However, the temporary linkages were restored shortly afterwards subduing the uncertainty about future coal supply for the corresponding plants.

The process to reallocate the cancelled coal block started in March 2015 and involved a need-based allotment to some public generation companies and an auction mechanism for private and public generation companies. However, the reallocation of these coal blocks did not yield the desired results due to regulatory interference and the design of the new mining contracts. As a result, more than three years after the SC decision, by the end of 2017, less than half of the cancelled mining contracts had been reallocated. For the plants that lost their mining contract in 2015 and had not received an alternate supply of coal, this implied a continuation of uncertainty about future fuel deliveries.

In the rest of the paper, I refer to the time period between August 2014 and March 2015 as the first phase of uncertainty. This phase one of uncertainty should have ended in March 2015 when the new policy to reallocate the cancelled coal blocks came into force. But the unsuccessful reallocation of cancelled blocks after March 2015 meant that uncertainty about future supply continued in the coal sector. I refer to the period after March 31 2015 as the second phase of uncertainty. The timeline of the evolution of this uncertainty is presented in figure 3.2 followed by the relevant details for each phase.

Figure 3.2: An illustration of how *coalgate* led to policy uncertainty



Notes: This figure illustrates the timeline of the cancellation of mining contracts by the Supreme Court (SC) of India and subsequent uncertainty about future coal supply. The first vertical line represents the day SC declared the mining contracts illegal. Second vertical line represents the end of first phase of uncertainty when the reallocation of cancelled mining blocks began. The time period between the two vertical lines represents the first phase of uncertainty. Second phase of uncertainty starts after the second vertical line.

Uncertainty Phase 1 & High vs. Low Uncertainty facing Plants

The first phase of uncertainty refers to the time period time between the SC decision to declare mining contracts illegal and the effective cancellation date, March 31st 2015. The rationale behind the SC verdict to allow the plants to have access to their respective captive mines until March 2015 was to give the plants some 'breathing room' and minimize disruption. The 'breathing room' however created a unique period of future uncertainty where power plants still had access to their existing supply of coal but knew that supply of coal was uncertain after March 31 2015.

For the plants linked to a non-producing coal mines, the SC judgment had two important consequences. First, they lost their right to extract coal in the future from their captive mine. Second, the temporary fuel linkage granted to them against the captive mine allocation was also annulled. The temporary fuel linkages were however quickly restored by Ministry of Coal and stayed in effect for the entire post-cancellation period covered in this study[14].

Since the plants linked to non-producing mines had a renewal clause in their temporary fuel supply arrangements, these plants were, to a large extent, immune to the consequences of whether or not they receive a new captive mining lease after March 2015. On this basis, I argue that these plants faced lower future uncertainty as compared to the plants linked to producing mines. In the rest of the paper, I refer to these plants as low uncertainty plants.

By contrast, the plants linked to producing mines had no such backstop for future fuel supply. These plants faced an uncertain future as securing future supplied at least partially relied upon the outcome of reallocation of cancelled mines after March 2015. I argue that these plants faced higher uncertainty and refer to them as high uncertainty plants in the rest of the paper.

Uncertainty Phase 2

After the cancellation of mining contracts, the GOI started allocation of coal blocks afresh in a new two-pronged policy starting march 2015. The private sector generation companies were invited to submit bids for coal mine leases to be awarded through reverse bid auctions[15]. In these auctions, the generators were asked to submit their extraction costs as bids and the

14. At first, the temporary linkages were restored until March 2015. Source: File No. 23014/3/2014-CPD, Ministry of Coal, India.
Then, they were again renewed till May 2015. Source: File No. 23011/106/2014-CPD, Ministry of Coal, India.
They were renewed a third time in June 2015 until March 2016. Source: File No. 23011/19/2015-CPD, Ministry of Coal, India.
Finally, these temporary fuel supplies were renewed again in 2016 and stayed in effect till at least the end of 2017, spanning the entire time period of this paper.

15. Reverse bid auctions were conducted to award mining leases to power generation companies. First price auctions were conducted for steel and cement manufacturers.

utility with lowest extraction cost was to win the lease to extract coal. This was done to keep the electricity prices in check, as fuel cost is the largest component of the variable cost of power generation[16]. A separate non-market mechanism was followed to apportion mining leases to the public sector generation companies. The state-owned generation companies were awarded captive mines on the basis of their unmet demand for coal and distance from the said mine[17].

The reverse bid auctions offer a first glimpse into the extent of uncertainty faced by the power plants. In all the auctions conducted for awarding captive mines to the power sector, the winning bids quoted *zero* extraction costs and, in addition, bid to pay a royalty to the state. In other words, the winning power plant owners bid *negative* extraction costs in order to win captive mining leases. The rationale behind this, from the power plants' perspective, was that the total bid amounts would be a complete pass-through as fixed charge in their respective power purchase contracts (PPAs) and hence came at no cost to them. This would turn out to not be the case.

In a strange decision, the GOI decided to cap the pass through of the auction bids after the auctions had concluded. This decision, which was later found to have been against the spirit of the law by the highest courts, essentially altered the conditions under which the auction was conducted and created a disincentive for the utilities to extract coal from their newly won captive mines[18]. Consequently, the demand for captive mines hit bottom and by the end of year 2017, only 89 of the 214 cancelled blocks had been reallocated, out of which supply of coal had only started from 28 mines. The total coal production form these

16. Electricity prices in India are rate regulated and determined at the state level by the respective state regulators. A lower extraction cost would imply a lower fuel cost which in turn would reduce the cost of generation and the final price paid the end consumers.

17. The explicit criteria mentioned in the official policy documents states two factors - requirement of coal and distance of the end use plant(s) from coal mine. The implicit objectives were to maximize generation of electricity and minimize cost of generation.

18. At least 3 generation companies subsequently went to the courts and won against the GOI. SC allowed the utilities to annul the mining lease won in the auctions and receive their money back from the GOI.

reallocated mines stood at half of the pre-cancellation level in 2014[19]. This shows that the reallocation of cancelled coal blocks was largely unsuccessful and, to a large extent, did not abate the uncertainty regarding coal allocation and supply.

It wasn't until the SHAKTI[20] policy announced in 2017 to auction longer term fuel supply agreements that the uncertainty surrounding coal supply was finally put to rest. The first auctions to award long term supply contracts were conducted in September of year 2017 and the fuel supply contracts were signed in December 2017. The time period of this paper includes data until December 2017 and hence covers the entire evolution of policy uncertainty regarding coal supply to the power plants in India.

3.3 Theoretical Framework

In this section, I provide a conceptual framework to describe how future uncertainty affects firms' production behavior. The aim of the model is simply to provide theoretical foundation for the empirical results to follow.

How might uncertainty about future supply of coal affect power plants operating in incomplete market in India? The answer to this question lies in the way power supply contracts called Power Purchase Agreements (PPAs) are structured. As per the terms in the PPA, the generators' payoff consists of a fixed component (capacity payments) and variable component (fixed charge). The capacity payment is conditional upon declared availability to generate and the plants need to have at least 85 % of their contracted generating capacity to be available in order to receive this fixed component of their payment. This capacity payment is not contingent upon the plants actually generating power. When faced with an uncertain coal supply in the future, the pursuit of capacity payment would then create an incentive for the plants to engage in precautionary saving of coal to ensure enough fuel for future periods when fuel deliveries are going to be unreliable. This way the plants can avoid

19. Source: Provisional Coal Statistics 2015, Ministry of Coal, India

20. Scheme to Harness and Allocate Koyla(*hindi* for coal) Transparently in India

having to declare outages due to fuel shortage and ensure at least 85 % of their capacity is available for generation.

I present a simple two period model to show how uncertainty about future supply of coal impacted the plants affected by the mine cancellations. The model is akin to a "divide the pie" problem where the size of pie is not known with certainty. For profit maximizing plants, the dominant strategy involves a production schedule that allows the plants to receive capacity payments in every period. The model assumes that the representative plant follows such a production schedule to receive capacity payments in every period and that it is this incentive that drives the fuel consumption smoothing.

Power plants are assumed to be risk-neutral, forward looking firms that maximize the net present value of profits. In a cost-plus regulated setting such as electricity sector in India, the unit *price* of output is determined by the regulator and is fixed. The per unit cost of fuel is also similarly pre-determined and fixed. In such a scenario, the profit maximization exercise is equivalent to output maximization.

A representative power plant generates power for two periods. At the beginning of period 1, power plant finds out that its exclusive mining contract has been cancelled. This leads the plant to believe that the fuel supply in period 2 might vary from the fuel supply in period 1.

Let the per period production function Y of a representative power plant affected by future uncertainty of supply of coal be:

$$Y_t = g(K_t, C_t, L_t)$$

Y_t represents the output of production in period t using K_t amount of capital, consuming C_t amount of fuel and utilizing L_t amount of labor. t denotes the period t and takes the values 1 and 2. I assume that the marginal product of consumption of fuel, C, is greater than zero, that is more coal produces more power.

The consumption of fuel C_t is determined by the inflow of fuel in period t and the amount of fuel left over from the previous period. Let the amount of fuel received by the representative plant in period t be denoted by F_t.

The cancellation of mines at a future date implies that an affected power plant faces uncertainty about the amount of fuel that it will receive in period 2, while the fuel deliveries for period 1 remain unaffected.

Let the uncertainty faced in period 2 be characterized by the possible state of period 2, $s \in S$, which occurs with probability π_s. State $s \in S$ of period 2 implies a supply shock of δ_s such that $\delta_s = N(0, \sigma^2)$. Thus, ex-ante, plants believe that they might experience a positive or a negative supply shock.

The model progresses as follows: At the beginning of period 1, plant receives F_1 amount of fuel and learns that the fuel supply for period 2 is uncertain due to the cancellation of the captive mines. The fuel supply for the two periods of the plants affected by uncertainty can be written as:

$$
\begin{aligned}
F_1 &= F_1 \quad \text{, and} \\
F_2 &= F_1 + \delta_s \quad \textit{for a given state } s{\in}S \textit{ in period 2}
\end{aligned}
$$

Then, after receiving F_1 and learning that F_2 is subject to uncertainty, the plant divides up the fuel *pie* sized $F_1 + E_1[F_2]$ into consumption bundles for the two time periods, C_1 and C_2. The consumption bundles at the beginning of period 1 can be written in terms of the fuel supply as follows:

$$
\begin{aligned}
C_1 &\leq F_1 \quad \text{, and} \\
C_2 &= F_2 + F_1 - C_1
\end{aligned}
$$

Before going to the plant's objective function, let us pause and look at how the capacity payment incentive affects fuel consumption. Let the amount of coal required to sustain 85% of generating capacity of this plant be Ω. Then, at the beginning of period 1, plant wants to allocate consumption bundles, C_1 and C_2, such that $C_1 \geq \Omega$ and $C_2 \geq \Omega$. Writing C_2 in terms of C_1 and combining the two constraints in a single expression shows how the capacity payment incentive induced fuel stock requirements affect consumption:

$$\Omega \leq C_1 \leq F_2 + F_1 - \Omega$$

The above expression shows that as the available capacity requirement increases, driving up the value of Ω, the per period consumption of coal will become more constrained. More importantly, the basic intuition behind precautionary saving is clear from the expression above: If plants expect fuel shortages in the future, then as realization of F_2 decreases, the right hand side constraint tightens up leading to decrease in C_1. In this framework, the change in expected future fuel supply is being potentially driven by changes in both first and second moment of future fuel supply and the model is unable to differentiate between the two. With this caveat, I show below how uncertain future supply can lead to precautionary saving and affect output.

In this paper's setting, the output (electricity) and input (coal) prices are pre-determined by the regulator and fixed. Thus, the power plant's profit maximization objective is equivalent to maximizing quantity of output. The cost per unit of fuel, labor and capital are held fixed and as such do not change the nature of the maximization problem. To focus on the intuition of the results, I do not include the costs in the optimization exercise. Including the cost functions do not alter the results in any way. The expected utility of the plant can be written as:

$$
\begin{aligned}
EV &= Y_1 \ + \ \ E_1[Y_2] \\
&= g(K_1, C_1, L_1) \ + \ \ E_1[g(K_2, C_2, L_2)] \\
&= g(C_1) \ + \ \ E_1[g(C_2)] \ \ \textit{holding K and L fixed} \\
&= g(C_1) \ + \ \ \sum_{s \in S} [\pi_s \ g(2F_1 - C_1 \ + \ \delta_s)]
\end{aligned}
$$

Subject to these two constraints, plant will maximize the net present value of its expected profit inflow. I assume away the per period capacity payment as it is fixed and take the discount rate to be one. The maximization problem takes the form of the following Lagrangian:

$$
\begin{aligned}
&\max_{C_1, C_2} [\quad g(C_1) + E_1[g(C_2] + \lambda_1(C_1 - K) + \lambda_2(C_2 - K) \quad] \\
\text{or, } &\max_{C_1} [\quad g(C_1) \ + \sum_{s \in S} [\pi_s g(2F_1 - C_1 + \delta_s)] + \lambda_1(C_1 - K) + \lambda_2(F_1 + F_2 - C_1 - K) \quad]
\end{aligned}
$$

Differentiating the above expression with respect to C_1;

$$
\begin{aligned}
0 \quad &= \quad g'(C_1) - \sum_{s \in S} [\quad \pi_s g'(2F_1 + \delta_s - C_1)] + \lambda_1 - \lambda_2 \\
\text{or, } \ g'(C_1) \quad &= \quad \sum_{s \in S} [\quad \pi_s g'(2F_1 + \delta_s - C_1)] + \lambda_2 - \lambda_1
\end{aligned}
$$

Here, λ_t represents the shadow value of consuming marginal unit of coal above the threshold K in period t. For fixed values of electricity and coal, λ_1 should be equal to λ_2 as the value of producing an extra unit of power is the same in both periods.

Thus, we can rewrite the above first order condition as:

$$
\begin{aligned}
g'(C_1) &= \sum_{s\in S}[\ \pi_s\ g'(2F_1\ -\ C_1\ +\ \delta_s)\ ] + \lambda_2 - \lambda_1 \\
&= \sum_{s\in S}[\ \pi_s\ g'(2F_1\ -\ C_1\ +\ \delta_s)\ ] \\
&= g'(\sum_{s\in S}[\ \pi_s\ (2F_1\ -\ C_1\ +\ \delta_s)\ ])
\end{aligned}
$$

The above expression implies:

$$
\begin{aligned}
C_1 &= \sum_{s\in S}[\ \pi_s\ (2F_1\ -\ C_1\ +\ \delta_s)\ ] \\
\text{or, } C_1 &= F_1\ +\ \frac{\sum_{s\in S}(\pi_s\delta_s)}{2} \\
&\leq F_1\ \textit{if}\ \sum_{s\in S}\pi_s\delta_s \leq 0
\end{aligned}
$$

The last expression shows how future uncertainty could lead firms to stockpile coal. If a plant affected by the cancellation of mines believes ex-ante that the future fuel deliveries will be negatively affected, then it will start conserving coal in the present even *before* the supply gets affected. The high uncertainty plants faced a greater likelihood of fuel shortage in the future as compared to the low uncertainty plants that had renewable temporary fuel supply.

This leads to two important predictions for the empirical results:

1. Uncertainty in future supply of coal will lead to precautionary savings in the current period if plants believe uncertainty to devolve into a 'bad state'.

2. Plants facing a more uncertain fuel supply will reduce their consumption of coal more than the plants facing relatively lower uncertainty. Hence, more uncertain the faced,

greater the loss in output.

In the next sections I discuss the datasets used and the empirical strategy employed to provide causal evidence in support of both these predictions.

3.4 Data

This paper uses a number of publicly available datasets and one confidential dataset rich in information about coal mines and power plants in India. Combining these different datasets, I form a monthly panel linking fuel from coal mines to electricity produced by power plants. Ultimately, the dataset contains information on coal-fired plants' fuel source, fuel delivery, consumption and stock, and measures of production outcome such as declared outages and power generation.

From a data perspective, the innovation in this paper is that, using a confidential dataset linking coal mines to power plants, I am able to map all thermal power plants to their respective source(s) of coal. Being able to link the power plants to their corresponding captive mine allows me to identify the exact power plants that were affected by the SC decision to cancel the mining contracts.

Details on the source of the datasets, information contained in them and variables of interest follows:

Coal Mines

The first source of data are the annual reports published by the Ministry of Coal, India. The annual reports contain detailed information on the production status of every coal block in the country. Two pieces of information important for this study come from these reports – 1) the reports for years 2014 and 2015 identify the coal blocks that were canceled by the Supreme Court, and 2) it provides information on which coal blocks were producing coal at the time of the contract cancellation. The reports also list the utilities or the firms that owned the mining rights to these mines, however, it does not identify the exact beneficiary power plant (called the End-Use Plant)[21].

Linking Captive Coal Mines to Power Plants

I use confidential data obtained from Central Electricity Authority (CEA) and Ministry of Coal, India that identifies the power plants linked to every captive mine allocated to the power sector. Using this, I can identify the plants that faced uncertainty over future supply of coal as a direct consequence of losing access to their captive mines. The rest of the power plants sourced coal through long term fuel supply contracts and hence were not affected by the uncertainty[22].

Combining this dataset with the information on producing captive mines from the Ministry of Coal annual reports, I can categorize plants as facing high uncertainty or low uncertainty.

Thermal Power Plants' Outcomes and Characteristics

For information on power plants, I rely on reports published by the Central Electricity Authority (CEA). CEA collects and disseminates rich data on power plants and electricity sector in India through various publications and reports.

The daily outage reports provide information on the capacity declared to be under outage by all thermal, hydro and nuclear power plants. Data on outage declared by coal-fired plants comes from these reports. The outage reports, by its very nature, only contain information on plants that report outage on a given day. Therefore, for the plants missing from the report on any given day, I impute zero capacity under outage.

21. Captive mine allocations were made against a specific plant or set of plants. The mine lessee was contractually obligated to only use the extracted coal for the End-Use plant. This applied to all firms, generation companies, steel plants, cement manufacturers etc.

22. A handful of coastal power plants are set up based on use of imported coal which is lower in ash content than Indian coal. These plants were also unaffected by the SC decision to cancel mining contracts.

Data on monthly receipts and consumption of coal comes from the monthly coal reports published by CEA. For every month, these reports provide data on total coal received and consumed at the plant level in thousand Metric Tonnes (MT). The generation reports published by CEA give information on power produced at all power plants at a monthly frequency. From these reports, I use the data for coal fired plants.

Other variables of interest are plant characteristics such as plant age and name plate capacity. To compute plants' age, I use the commissioning date provided in their annual list of stations. A plant's age is simply calculated as the difference between given year and the year of their commissioning. Information on plants' nameplate generating capacity comes from the generation reports.

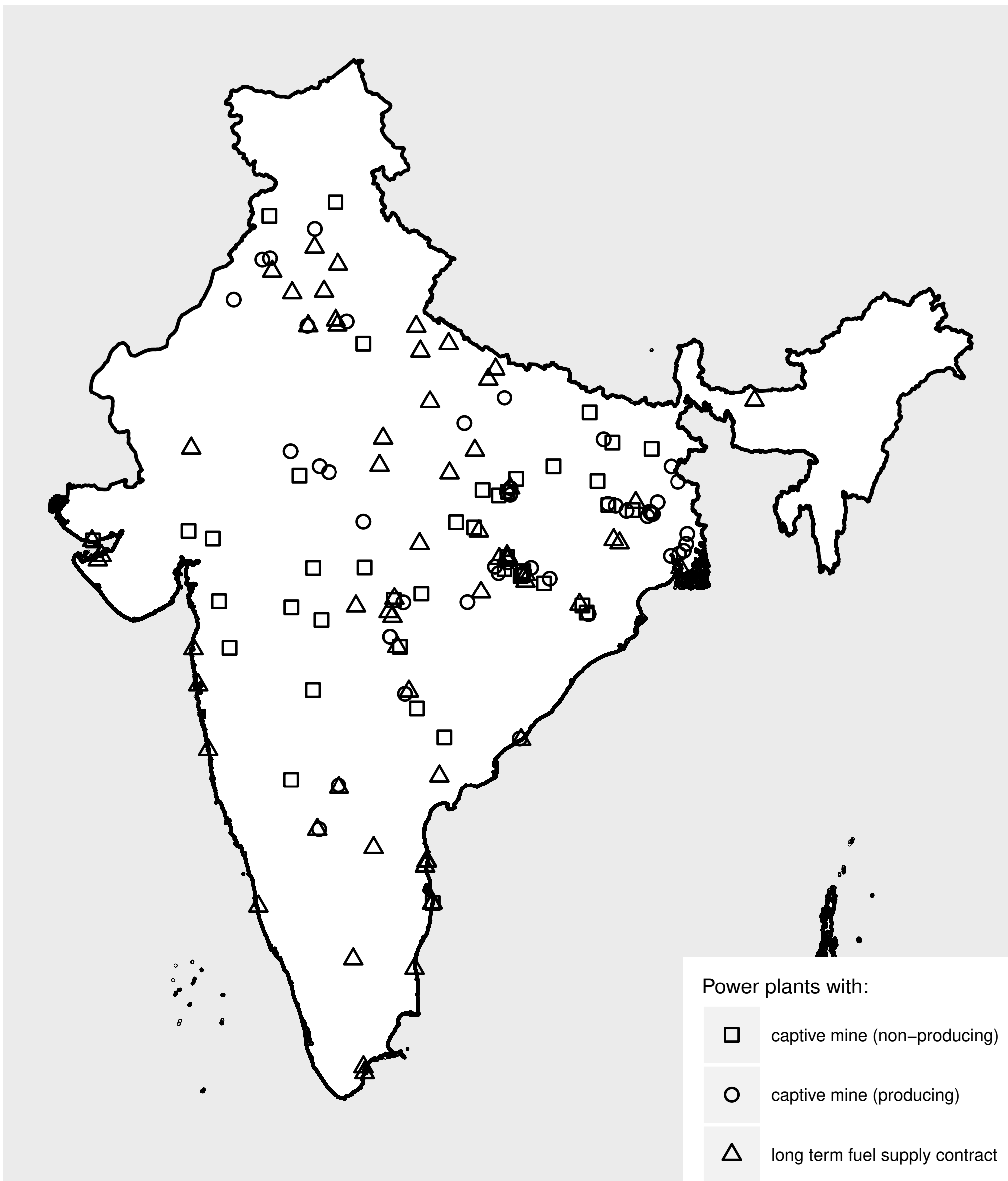
Monthly and Daily Panels

Using these sources of data, I construct two panels for coal fired power plants – a daily panel for capacity reported to be under outage and a monthly panel for coal receipts, consumption and generation of power. Both panels start from first month in 2013 and end in last month of 2017, covering a period of 60 months. In July 2014, prior to the SC decision affecting the power plants, there were 140 coal fired plants that were supplying electricity to the grid. The monthly coal statements however only contain data on coal receipts and consumption for 124 of these plants. I restrict my main sample to these 124 plants for which I have complete information on all outcomes of interest. My main sample then consists of a slightly unbalanced panel of 124 power plants for the time period 2013 to 2017.

In addition to these, I also collect data on distance between coal mines and power plants, coal grade and coal field specific fuel prices, and heat rates for all power plants based on multiple sources. Using this information, I am able to compute marginal cost of generation for all the plants in the sample.

Figure 3.3 shows the locations of the entire sample of coal fired plants on a map of India. The plants are categorized by how they were affected by the future uncertainty resulting from cancellation of mining contracts – plants unaffected by the cancellation of mines, plants that faced high uncertainty as they were linked to producing mines, and plants that faced low uncertainty. The spatial distribution of the plant shows that both the "treated" and "control" samples draw plants from across the country and are spatially widely distributed.

Figure 3.3: Map of all coal fired power plants in India



Notes: This figure shows the spatial distribution of all coal fired power plants on a map of India. The plants are categorized based on where they received coal from: non-producing captive mine (temporary fuel supply form state), producing captive mine, and long term fuel supply form the state. The map shows that power plants are widely dispersed throughout India.

3.5 Empirical Strategy

Difference-in-Difference

I employ difference-in-difference estimation method to estimate the impact of the policy shock induced future uncertainty on power plants' behavior, The SC judgment inducing policy uncertainty was an unanticipated shock to the power plants but it might still be correlated with some unobservables at the plant level. I avoid assuming that the treatment and control plants are balanced on observables as well as unobservables, and instead rely on difference in differences estimation framework with plant fixed effects. By controlling for time invariant plant characteristics, the effect of policy uncertainty on plants' outages is estimated off within plant variation in the outcome of interest.

The linear relationship for estimating effect of policy uncertainty on outages then is:

$$Y_{jt} = \beta_1 1.(cancellation_j) \times 1.(post_t) + X_{jt} + \phi_j + \kappa_t + \varepsilon_{jt} \quad (3.1)$$

where Y_{jt} is the outcome of interest for a plant j. $1.(cancellation_j)$ is an indicator variable that takes the value 1 for plants that had their captive mines declared illegal and faced uncertainty over future supply of coal as a result, and 0 for plants that were unaffected by the SC decision as they did not have any captive coal mines. $1.(post_t)$ equals 1 for the time period after the cancellation of contracts and 0 before. And, β_1 is the coefficient on the interaction of these two indicator functions and represent the estimate of average treatment effect on treated (ATT). X_{jt} represents a plant's time varying characteristics namely age and generation capacity. ϕ_j, j dummies; one for each plant, represent plant fixed effects and absorb plant time invariant characteristics such as location. κ_t represent sample month-year fixed effects.

To explore heterogeneity by level of uncertainty, I estimate a version of the main specification where 1.($cancellation_j$) indicator is interacted with a binary variable *producing mine* which takes the value 1 / 0 for plants that were linked to producing / non-producing captive mines and faced high / low uncertainty respectively after the loss of their mining leases.

I argue that the cancellation of mining contracts induced uncertainty about future supply of coal. Based on this, I test the hypothesis that supply constraints in the form of policy shock induced future uncertainty decrease power plants' current output. I test this using direct measure of output, power generated, and indirect measures of output such as generation capacity under outage and fuel consumption. Given the objective of this paper on explaining why outages exist, the primary outcome of interest is the daily reported capacity under outage at the plant level.

The estimation strategy assumes that the plants that had captive coal mines ("treated") and plants that did not ("control") are valid comparison groups. Since both groups of power plants consist of thermal power plants use coal to produce a homogeneous product, electricity, and operate under the same jurisdiction, this is a reasonable assumption.

To understand how the two groups of plants compare, table 3.1 shows the comparison between "treated" and "control" plants across all the relevant variables for the month of July, 2014. Even though the empirical strategy does not require equality of means at baseline, the similarity between the two groups of plants across all characteristics except age shows that the groups of plants were similar at the baseline. The one variable where difference is significant is the age of the plants. This is by design as most of the captive mine allocation were given out to newer plants that were commissioned in the 1990s. After controlling for plants fixed effects, the mean age difference is no longer significant. All the regression estimates will control for plant age among other characteristics.

The second identifying assumption is specific to the difference-in-difference empirical strategy, and assumes that in the absence of the contract cancellation, the "treated" and "control" plants would have had similar trends in the outcomes of interest over time. While this assumption is fundamentally untestable, I can show that prior to the SC decision, the

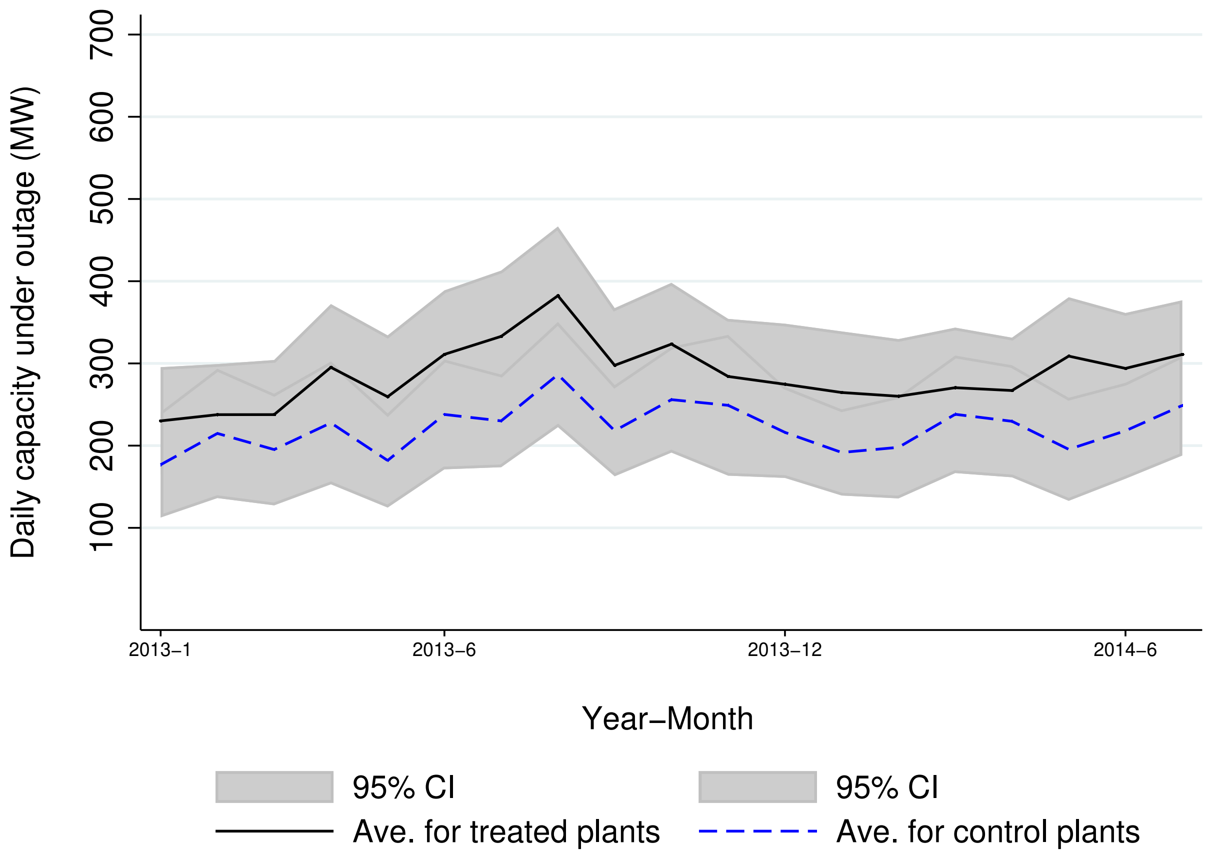
Table 3.1: Baseline means comparison of plants facing no uncertainty and plants facing uncertainty due to SC decision

Variable	(1) Captive mine *treated (T)*	(2) Long term contract *control (C)*	(3) Difference *(C) - (T)*	(4) Difference *(with Plant FE)*
Total Monthly Coal Receipt (in MT)	329.694 (328.095)	379.362 (325.483)	49.669 (59.688)	80.000 (0.000)
Coal Receipt from State (in MT)	264.373 (257.512)	329.000 (310.986)	64.627 (55.307)	103.000 (0.000)
Monthly Coal Consumption (in MT)	331.565 (313.725)	403.017 (333.826)	71.453 (59.233)	-51.000 (0.000)
Monthly Generation Target (in MWh)	478.419 (498.622)	521.276 (458.622)	42.857 (87.388)	-135.000 (87.388)
Monthly Electricity Generation (in MWh)	506.923 (503.615)	548.451 (484.935)	41.527 (90.251)	-65.200 (90.251)
Annual Electricity Generation till date (in MWh)	419.551 (482.250)	464.602 (478.069)	45.051 (87.700)	-105.510 (0.000)
Plant Capacity (in MW)	1,042.540 (899.633)	1,179.405 (842.154)	136.865 (159.003)	-0.000 (0.000)
Plant Age (in years)	16.903 (15.953)	22.948 (14.496)	6.045** (2.780)	-6.000 (0.000)
Observations	62	58	120	120

Notes: This table presents the results for difference in means for various plant level characteristics in the month prior to the start of first phase of uncertainty. Column 1 and 2 present the mean of "control" and "treated" plants respectively. Column 3 report the difference in mean between the two groups. Column 4 reports the mean after controlling for plant fixed effects. Robust standard errors clustered at the plant level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

two groups of thermal plants had parallel trends. Figure 3.4 shows the pre trends for the average daily capacity under outage for plants facing uncertainty and plants unaffected by the cancellation of mining contracts. A visual examination of the trends for the two groups of plants before the day of SC judgement shows that the parallel trend assumption is justified here.

Figure 3.4: Average daily outages due to coal shortage by treatment and control



Notes: This figure shows the pre-trends for the raw means of daily capacity under outage (MW / day) with 95 % confidence interval for the "treated" and "control".group of power plants before the mine cancellations. The time trends of capacity under outage for the two group of plants are parallel before the SC judgement.

A third assumption required is the stable unit treatment value assumption (SUTVA) which requires that the uncertainty faced by the plants that lost their exclusive mining contracts should not have affected the behavior of plants that were unaffected by the SC decision. If the plants facing uncertainty as a result of the policy shock changed their production behavior, then this might have affected the control plants' production behavior - perhaps at the behest of a state regulator. However, this is of limited concern as the extent to which the control plants can ramp up generation is constrained by their total capacity and availability of fuel via their fuel supply contracts. Still, this is an important disclaimer for the results that follow.

3.6 Results

The results section is organized in two subsections – short term results and longer term results. Short term results refer to the uncertainty phase one period of seven months, which started on August 25 2014 when SC announced that the captive mining contracts will be effectively cancelled on March 31 2015.

Longer term results cover the time period starting at the same time as the short term and lasting till end of 2017 when the market based allocation of long term fuel supply contract starts. All the regressions results are estimated using the empirical specification discussed in previous section and control for plant fixed effects, plant's age and total generation capacity and month, year and month by year fixed effects.

3.6.1 Short term results

Plants respond to future uncertainty by declaring partial outages

I start by estimating the impact of supply side frictions in the form of uncertainty about future fuel deliveries on plant level outages. The SC decision to take away plants' exclusive mines made the future supply of coal uncertain and the plants responded by declaring partial outages. This decision to reduce production of power was made to smooth consumption of fuel in the face of future uncertainty.

I report estimated effect on plant level outages for the 124 plants for which coal receipt and consumption data is available. The results for the universe of plants that started operations before the policy shock are similar and documented in appendix A.

Table 3.2 reports the ATT estimates of the effect of policy uncertainty on generation capacity outage. Columns 1 compares the plants affected by the SC decision to all the other plants and the estimates show that the resulting uncertainty about fuel supply caused a daily increase of 46 MW of capacity under outage. The results are statistically significant and lie within the 95% confidence interval.

Table 3.2: SHORT TERM EFFECT OF UNCERTAINTY ON DAILY GENERATION CAPACITY UNDER OUTAGE

	Capacity under Outage (MW/day)	
	(1)	(2)
cancellation	45.84**	
	(23.15)	
cancellation × *(producing mine)*		68.95**
		(29.17)
cancellation × *(non-producing mine)*		14.38
		(25.96)
Observations	96,396	96,396
R-squared	0.043	0.045
Number of plants	124	124
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X
N (producing mine)		35
N (non-producing mine)		26

Notes: This table presents the results for short term effect of uncertainty on daily outages at the power plant level. Column 1 compares all treated plant units to all control plant units. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Robust standard errors clustered at the plant level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

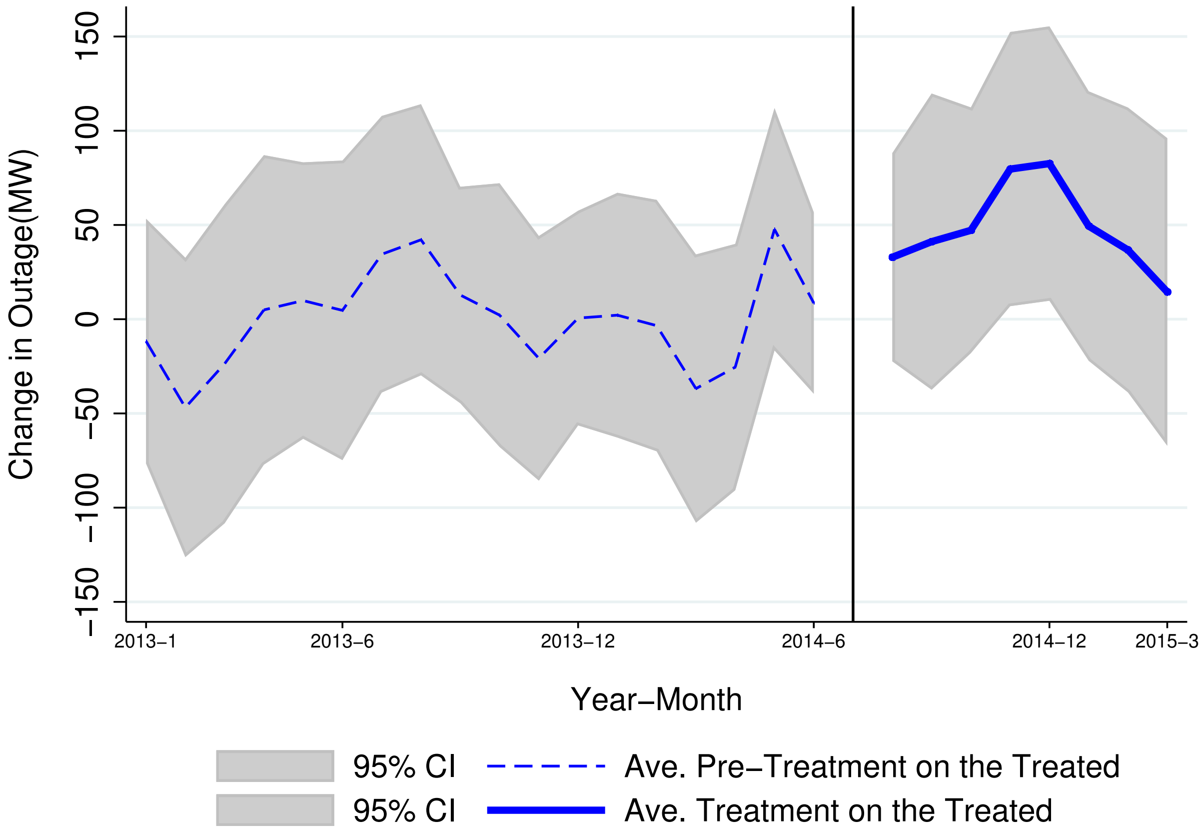
To explore heterogeneity in treatment effect, columns 2 interacts the *cancellation* indicator variable with a binary for inferred level of uncertainty faced by the plants, *producing mine*, which takes the value 1(0) for plants that had producing (non-producing) mining contract cancelled. Among the plants affected by SC decision, the plants that were receiving coal from the producing mines at the time of cancellation faced higher uncertainty over future coal supply as compared to the plants that were linked to non-producing mines.

The estimates for the high uncertainty facing plants are larger than the overall results as well as the plants that faced less uncertain future fuel supply. Column 2 estimates report that on average, plant linked to producing mines reported 69 MW of daily capacity under outage in anticipation of future fuel shortages. Based on the theoretical framework section, this suggests that the plants linked to producing mines believed future fuel shortages to be a much more likely outcome than the plants linked to non-producing mines. This is not surprising given that the non-producing mine linked plants had the temporary fuel supply agreements which acted as a backstop. Overall, these results show that the estimated result for the full sample in column 1 is largely due to the plants that were more vulnerable to future fuel shortages.

Further, we can see how the treatment effect on capacity under outage evolves over time. Figure 3.5 plots the monthly estimates of ATT in an event study style graph. The y axis plots the coefficient corresponding to the interaction term of *cancellation* and a monthly binary which takes the value 1 for a given month and 0 otherwise[23]. And the x axis represents a month-year. *cancellation* interacted with the indicator for the month prior to the SC decision, July 2014, is the omitted variable and all the estimates are relative to this omitted variable. The monthly evolution of the treatment effect shows that the uncertainty due to mine cancellations causes an increase in capacity under outage and the partial shutdown lasts for the entire first phase of the uncertainty.

23. Using daily or weekly binaries produces a similar trend but a noisy picture.

Figure 3.5: Monthly treatment effect on capacity under outage



Notes: This figure is an event study plot of the evolution of the ATT estimate of the effect of mining contract cancellation on generation capacity under outages due to future fuel uncertainty. The estimated coefficient is plotted along with 95 % confidence interval constructed using standard errors clustered at the plant level. Plants face uncertain future due to future mine cancellation and respond by declaring partial outages to conserve fuel.

The outage dataset also contains information on the reason behind a plant declaring outages. The reasons provided for outages are self-reported by plants and it is up to the plants to report the cause of outage in their words. To further uncover the mechanism behind the estimated effects on outages, I utilize the valuable information contained in the reported reasons for the capacity outages. I construct a new dependent variable for capacity under outages to only include the magnitude of capacity reported to be under outage due to "fuel shortage" or "coal shortage" in the data. Table 3.3 reports the regression estimates for the effect of policy shock induced uncertainty with the dependent variable being the capacity reported to be under outage due to fuel shortage. Column 1 reports that in comparison to "control" group of plants, the "treated" plants declared 35 MW/day of capacity to be under outage reported due to fuel shortage. The coefficient is statistically significant and, importantly, the magnitude of the coefficient is comparable to the estimate for capacity under outage due to all reasons in table 3.2.

Table 3.3: Short term effect of uncertainty on capacity under outage due to reported fuel shortage

	Capacity under Outage due to reported coal shortage (MW/day)	
	(1)	(2)
cancellation	35.19* (18.13)	
cancellation × *(producing mine)*		46.92** (23.68)
cancellation × *(non-producing mine)*		19.24 (21.28)
Observations	96,396	96,396
R-squared	0.028	0.029
Number of plants	124	124
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X
N (producing mine)		35
N (non-producing mine)		26

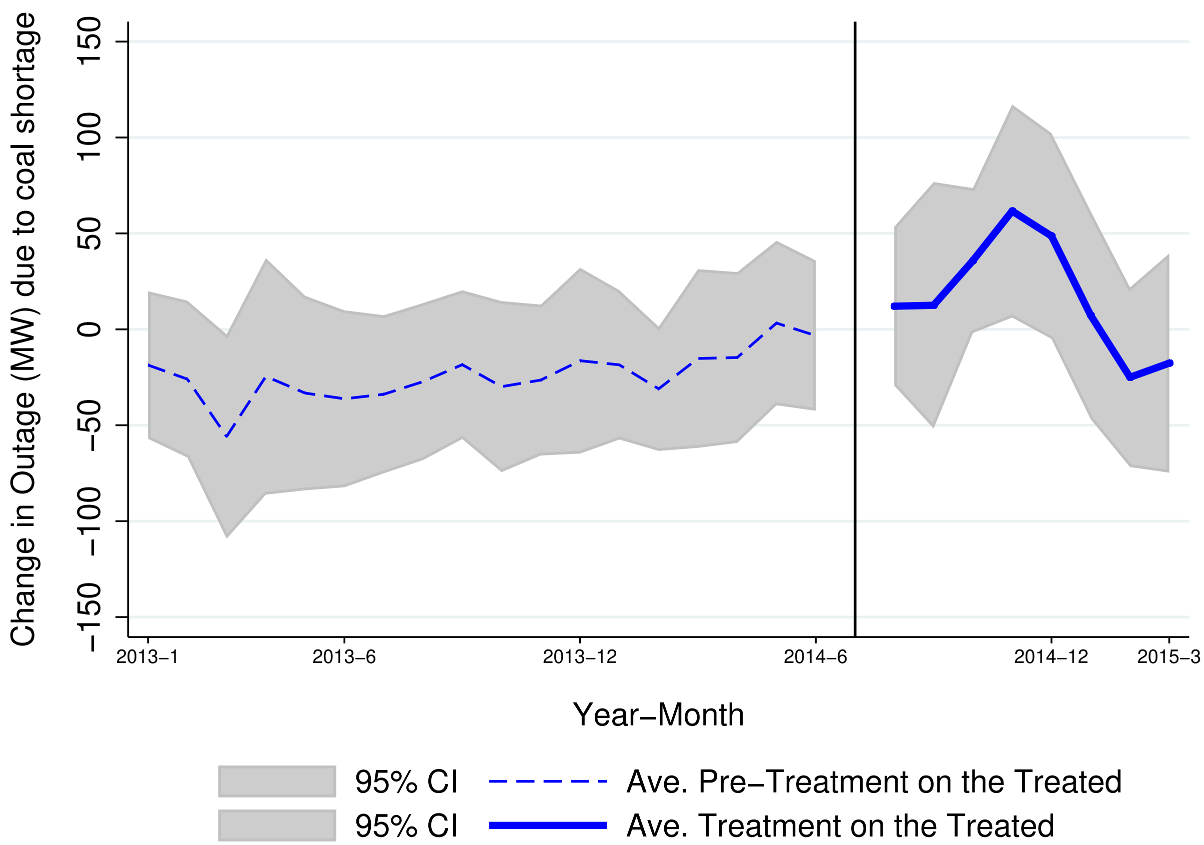
Notes: This table presents the results for short term effect of uncertainty on daily outages reported due to coal shortage at the power plant level. Column 1 compares all treated plant units to all control plant units. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Robust standard errors clustered at the plant level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Figure 3.6 shows analogous version of figure 3.5, but this time only for capacity reported to be under outage due to fuel shortage. The monthly evolution of the ATT estimate follows a similar trend to the total capacity under outage.

At first, this result is surprising given the fact that the cancellation of mines had not come into effect yet for producing mines and hence there was no change in existing coal supply

(I formally test for this below). Interestingly, it points towards the underlying mechanism behind the result on outages that the plants believed the future uncertainty about coal supply to devolve into a negative shock to the coal supply even though their fuel supply had not yet been affected.

Figure 3.6: Monthly treatment effect on outages reported due to fuel shortage



Notes: This figure is an event study plot of the evolution of the ATT estimate of the effect of mining contract cancellation on capacity reported to the regulator to be under outage due to fuel shortage. The estimated coefficient is plotted along with 95 % confidence interval constructed using standard errors clustered at the plant level. Facing uncertainty fuel supply, plants resort to partial shutdown and report lack of fuel even as fuel supplies remain unchanged.

Mine cancellation did not affect fuel deliveries

For the plants receiving coal from their exclusive producing mines, the cancellation of mining contracts did not come into effect until the end of the first phase of uncertainty. And, the low uncertainty plants, linked to non-producing captive mines, had their temporary fuel supply restored after the initial cancellation. This suggests that during the first phase of uncertainty, the SC judgement induced uncertainty did not have any effect on the actual supply of coal to the power plants.

Using the same main estimation framework, I can formally test for the effect of uncertainty on coal receipts, if any. Table 3.4 reports the estimated coefficient of interest for all plants, and high and low uncertainty plants. All three estimated coefficients are statistically indistinguishable from zero. These null results rule out fuel shortages as a potential driver of the effect of SC judgement on determining plants' production behavior in the short run. This lends further support to the underlying mechanism driving plants' behavior being perceived future coal shortages.

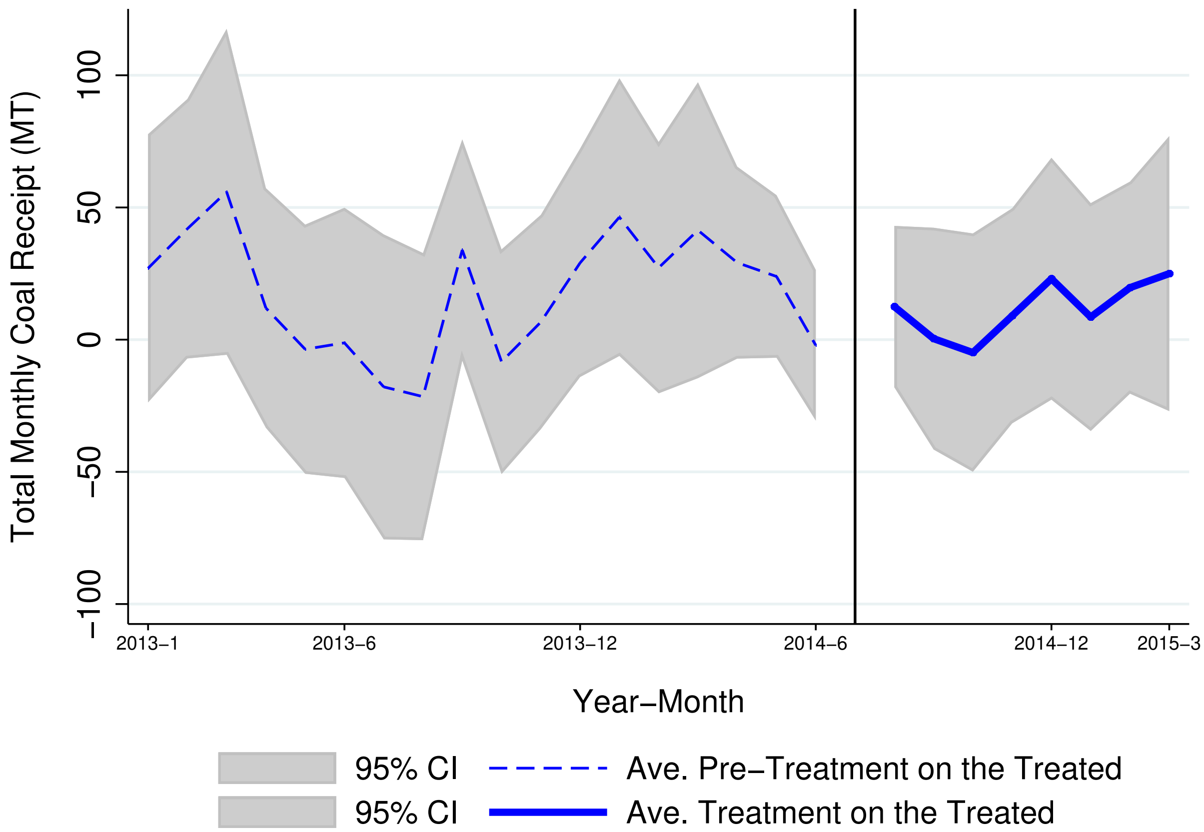
Table 3.4: Short term effect of uncertainty on monthly receipts of coal

	Coal Supply ('000 Metric Tonnes/month)	
	(1)	(2)
cancellation	-4.982	
	(10.03)	
cancellation × *(producing mine)*		-7.423
		(12.18)
cancellation × *(non-producing mine)*		-1.605
		(12.87)
Observations	3,163	3,163
R-squared	0.192	0.193
Number of plants	124	124
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X
N (producing mine)		35
N (non-producing mine)		26

Notes: This table presents the results for short term effect of uncertainty on monthly receipt of coal at the power plant unit level. Column 1 compares all treated plant units to all control plant units. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Robust standard errors clustered at the plant level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Further, figure 3.7 shows the short term trend of the effect of uncertainty on monthly average amount of coal received. The y axis represents the monthly ATT estimate with 95% confidence interval. The x axis represents year-month. The estimated coefficient remains close to zero for the entire duration of the first phase of uncertainty confirming that the first phase of uncertainty saw no differential change in fuel supply as a result of the contract cancellation.

Figure 3.7: TIME TREND OF ATT ESTIMATE OF UNCERTAINTY ON TOTAL COAL RECEIVED



Notes: This figure is an event study plot of the evolution of the ATT estimate of the effect of mining contract cancellation on fuel deliveries. The estimated coefficient is plotted along with 95 % confidence interval constructed using standard errors clustered at the plant level. The fuel supply remained unaffected during first phase of uncertainty.

Uncertainty about future fuel supply led to precautionary saving

The theoretical framework predicts that when faced with uncertainty about future supply of coal, plants will respond by reducing their consumption and engage in precautionary saving of coal. This is indeed the motive behind the plants declaring partial outages of their generation capacity. I can test for this hypothesis this using the monthly coal consumption data at the plant level.

Table 3.5 reports the estimated effect of uncertainty on monthly coal consumed at the plant level. Column 1 estimates show that, on average, there was a reduction in consumption of coal totaling 22,310 Metric Tonnes (MT)[24] per month. The monthly coal receipt data also contain plants' stated requirement for coal in a given month. Using this plants' stated requirement of coal information, a simple back of the envelope calculation suggests that the 22,310 MT of coal would have been sufficient to feed 60 MW of generation capacity for a month. Though admittedly this is a crude bench-marking of coal requirement for power plants, and one that engineers would probably not approve of, I find comfort in the fact that it is consistent with the estimated effect on outages. The significant negative impact on consumption of coal shows that when faced with uncertainty, plants smooth their consumption of coal to reduce the variability in consumption in the future.

Column 2 shows the estimates of the effect of uncertainty for plants facing high uncertainty and low uncertainty. For plants linked to producing mines, the decrease in consumption is 31,150 MT and within the 95% confidence interval constructed using standard errors clustered at the plant level. For plants facing low uncertainty, the coefficient is negative but the null hypothesis that the estimate is equal to zero cannot be rejected. Given the earlier result on increase in power plants' declared outages, this significant decrease in consumption of coal is unsurprising and lends support to the earlier result showing the effect of uncertain fuel supply on plants' production decisions.

The effect of mine cancellations on reduction in coal consumption by power plants is

24. 1 Metric Tonne = 1000 Kilogram

Table 3.5: SHORT TERM EFFECT OF UNCERTAINTY ON MONTHLY CONSUMPTION OF COAL

	Coal Consumption ('000 Metric Tonnes/month)	
	(1)	(2)
cancellation	-22.31**	
	(10.93)	
cancellation × *(producing mine)*		-31.15**
		(12.65)
cancellation × *(non-producing mine)*		-10.08
		(11.97)
Observations	3,163	3,163
R-squared	0.202	0.204
Number of plants	124	124
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X
N (producing mine)		35
N (non-producing mine)		26

Notes: This table presents the results for short term effect of uncertainty on monthly consumption of coal at the power plant unit level. Column 1 compares all treated plants to all control plants. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Robust standard errors clustered at the plant level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

indicative of precautionary saving in light of future uncertainty. Using fuel delivery and consumption information, monthly coal saving can be calculated as the difference between the two. Table 3.6 reports the regression results for effect of uncertainty on monthly coal savings which can be inferred as the level of precautionary savings. Column 1 shows that on average plants added 17,330 MT of coal to their stockpile every month as a result of mine cancellations. The coefficient is algebraically a linear combination of the effect on receipts and consumption and is statistically significant. Column 2 shows that these savings are mainly due to plants that had their producing mines declared illegal. On the back of the estimated effects on receipts and consumption, the results on savings follow mechanically and are expected.

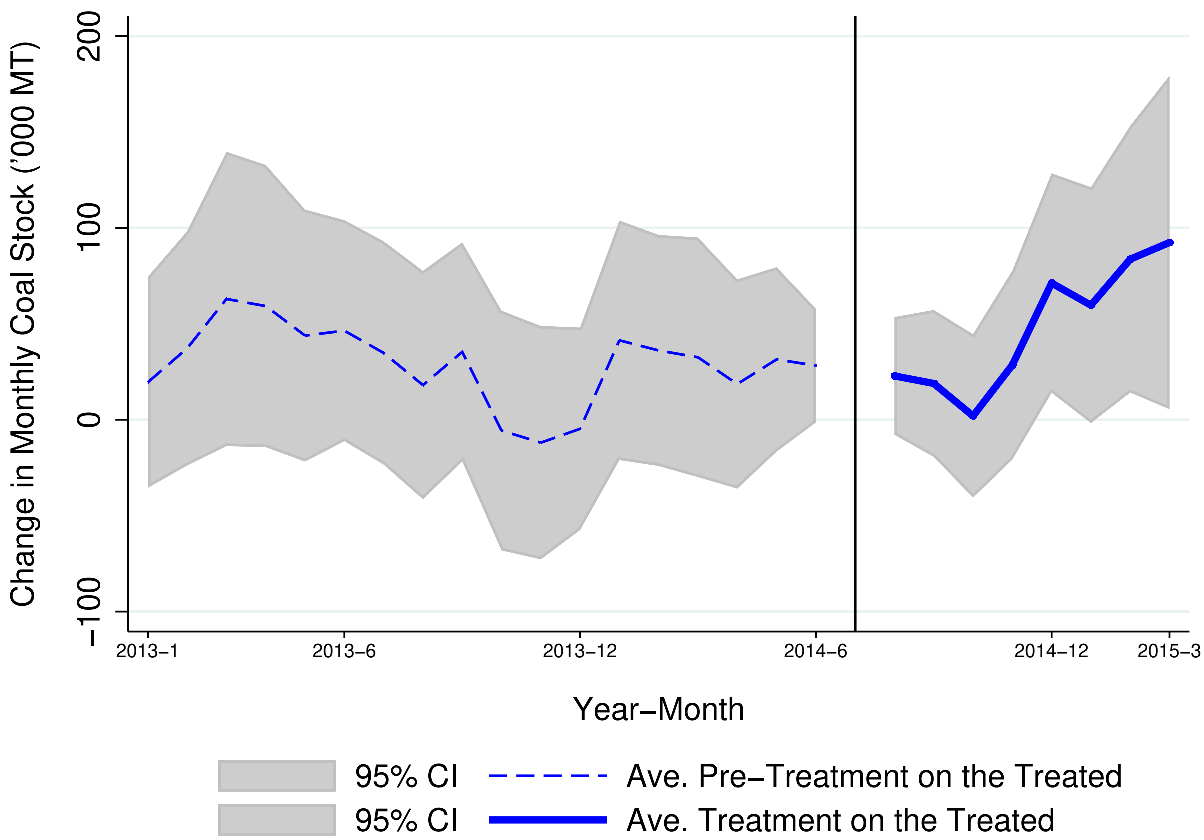
Table 3.6: DIFFERENCE IN DIFFERENCE ESTIMATES OF THE EFFECT OF UNCERTAINTY ON MONTHLY PRECAUTIONARY COAL SAVINGS

	Coal Savings ('000 Metric Tonnes/month)	
	(1)	(2)
cancellation	17.33**	
	(6.789)	
cancellation × *(producing mine)*		23.73***
		(8.184)
cancellation × *(non-producing mine)*		8.477
		(7.757)
Observations	3,163	3,163
R-squared	0.202	0.204
Number of plants	124	124
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X
N (producing mine)		35
N (non-producing mine)		26

Notes: This table presents the results for short term effect of uncertainty on monthly precautionary coal saving at the power plant level. Column 1 compares all treated plants to all control plants. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Robust standard errors clustered at the plant level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The monthly addition to plants' fuel stockpile in the form of precautionary saving also gets reflected in the closing monthly stock of coal. Figure 3.8 plots the monthly ATT estimate for the effect on monthly closing coal stock. The estimated coefficient has a distinct upward trend and shows the stockpiling response of the plants to mining contract cancellation. In the face of potential future fuel shortages, plants start stockpiling fuel for generation by reducing their capacity being utilized for generating power. This shows how future uncertainty can affect firms' output even as the current state remains unchanged.

Figure 3.8: Evolution of treatment effect on end of month coal stock



Notes: This figure is an event study plot of the evolution of the ATT estimate of the effect of mining contract cancellation on closing monthly stock at the plant level. The estimated coefficient is plotted along with 95 % confidence interval constructed using standard errors clustered at the plant level. Power plants curtailed generation capacity utilization to engage in precautionary saving of coal in face of future uncertain supply. As a result, the stockpile of coal increased during the first phase of uncertainty even as the plants declared outage due to coal shortage.

Uncertain future supply led to decrease in current output

Lastly, I look at the effect on the direct measure of plants' output, power generation. Table 3.7 reports the regression estimates for the effect of contract cancellation on monthly generation of power at the plant level. In line with earlier results, the coefficient of interest corresponds to the interaction term of *cancellation* binary and binary for *post* which takes value 0 before the mining contract cancellation and 1 after. Column 1 shows that during the first phase of uncertainty, uncertainty led to an average monthly loss of 36 GWh of electricity. As a sanity check I benchmark these to the result on capacity under outage. The overall generation loss estimates in column 1 correspond to the outages estimate in table 3.2 column 1 at a 70% capacity utilization (or plant load factor). The average capacity utilization in the sample for "treated" plants in the pre-treatment period is 67%. Thus, all the estimated effects on the different measures of output are consistent with each other.

Column 2 shows the results for plants facing high uncertainty and low uncertainty. The generation loss for plants that were linked to producing mines at the time of cancelation was greater in magnitude and statistically significant at 5%. Estimates for plants facing low uncertainty are negative in magnitude but not statistically different from zero. Again, these results are similar to the earlier results on outages and coal consumption.

3.6.2 Longer term results

The longer term results estimate the effect of uncertainty using the entirety of the post cancelation time period, i.e. first phase of uncertainty and the second phase of uncertainty. The first phase of uncertainty ended with a reallocation of some of the canceled mining contracts. At the end of year 2017, less than half of the cancelled coal blocks had been reallocated, and even from these reallocated mines, the total coal produced stood at around half of the pre-cancellation level in the year 2014. Thus, the uncertainty resulting from cancelled mining leases continued after the end of the first phase though arguably at a lower level. The lack of productivity from the reallocated mines indicates that after the end

Table 3.7: Short term effect of uncertainty on electricity generation

	Power Generation (GWh/month)	
	(1)	(2)
cancellation	-36.13**	
	(15.14)	
cancellation × *(producing mine)*		-51.47***
		(19.17)
cancellation × *(non-producing mine)*		-14.89
		(15.44)
Observations	3,163	3,163
R-squared	0.295	0.297
Number of plants	124	124
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X
N (producing mine)		35
N (non-producing mine)		26

Notes: This table presents the results for short term effect of uncertainty on monthly amount of power generated at the plant level. Column 1 compares all treated plants to all control plants. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Robust standard errors clustered at the plant level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

of first phase of uncertainty, the uncertainty stemming from the mines that got reallocated devolved into a ‘bad state’ with potential negative consequences for the corresponding supply of coal. Since the rest of the coal mines remained unallocated, the uncertainty stemming from cancellation of these mines remained. Thus, in the longer term, the “treated” plants continued to face uncertainty about future supply of coal along with potential coal supply constraints. Accordingly, I cautiously interpret the longer term results as the effect of two supply side constraints – policy uncertainty and actual fuel constraints - on the affected plants.

The estimating equations remains the same as the one for short term results. All regressions continue to control for the same time variant and invariant characteristics at the plant level.

Outages

The estimates longer term effect of supply side frictions on capacity under outages are in table 3.8. The full sample result in column 1 shows that the plants affected by the cancellation of mining contracts continue to have a significantly higher capacity under outage as compared to the plants that were not affected by the uncertainty. The coefficient is slightly larger than the corresponding short term estimate and lies within the 95% confidence interval. Similar to the short term results, plants that faced higher uncertainty shutdown more capacity in response.

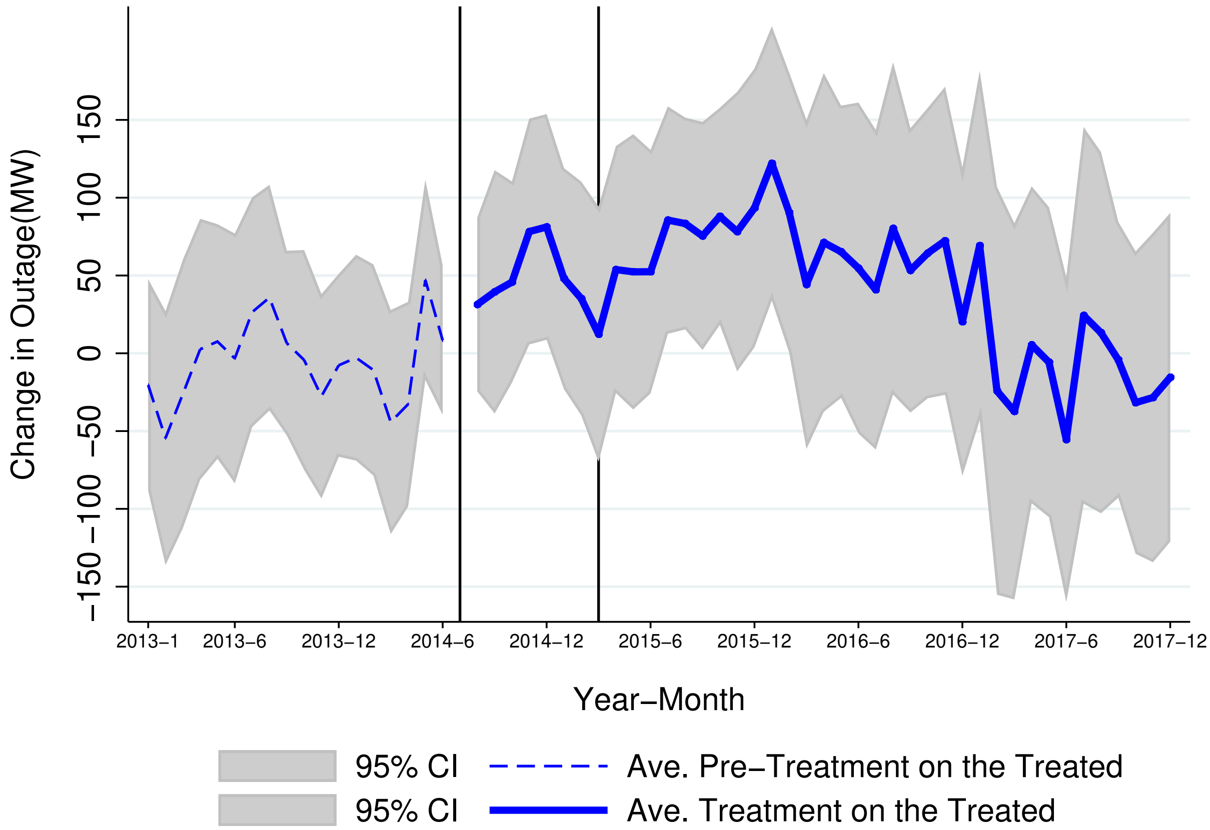
The time trend of estimated treatment effect is shown in figure 3.9 where all the coefficients are estimated relative to the omitted interaction term of the *uncertainty* binary and the day just before the contract cancellation decision. The trend shows that outages persist rise as a result of the supply side frictions and only start to subside towards the end of year 2017 as the long term fuel allocation policy (SHAKTI) is set into motion.

Table 3.8: LONGER TERM EFFECT OF MINING CONTRACT CANCELLATION ON DAILY GENERATION CAPACITY UNDER OUTAGE

	Capacity under Outage (MW/day)	
	(1)	(2)
cancellation	55.34**	
	(25.99)	
cancellation × *(producing mine)*		75.50**
		(35.36)
cancellation × *(non-producing mine)*		27.76
		(27.64)
Observations	211,674	211,674
R-squared	0.064	0.065
Number of plants	124	124
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X
N (producing mine)		35
N (non-producing mine)		32

Notes: This table presents the results for longer term effect of mine cancellation on daily outages at the power plant unit level. Column 1 compares all treated plant units to all control plant units. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Robust standard errors clustered at the plant level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Figure 3.9: Time trend of ATT estimate of effect on outages



Notes: This figure is an event study plot of the longer term evolution of the ATT estimate of the effect of mining contract cancellation on capacity under outage. The estimated coefficient is plotted along with 95 % confidence interval constructed using standard errors clustered at the plant level. The effect of mining contract cancellation on outages at the plant level persists for over three years as some of the uncertainty about future coal supply turned into actual coal shortages in the second phase of uncertainty.

Coal Supply, Consumption and Precautionary Savings

At the start the of the second phase of uncertainty, the mining cancellations go into effect and this combined with the subsequent inefficient allocation of the cancelled blocks leads to a negative impact on coal supply. Table 3.9 shows the estimated effect of mine cancellation on longer term coal supply. In column 1, the coefficient on the uncertainty dummy is now negative and significant and 5%. The coefficients for high and low uncertainty plants are also negative and significant. These results points towards the failure of the reallocation of canceled captive mines after the first phase of uncertainty and suggests at least partial devolution of uncertainty into a "bad state" where fuel supply was negatively affected in the longer run.

Table 3.9: LONGER TERM EFFECT OF MINING CONTRACT CANCELLATION ON MONTHLY RECEIPTS OF COAL

	Coal Supply ('000 Metric Tonnes/month)	
	(1)	(2)
cancellation	-29.73**	
	(13.36)	
cancellation × (producing mine)		-31.93*
		(16.71)
cancellation × (non-producing mine)		-26.71*
		(15.07)
Observations	7,341	7,341
R-squared	0.253	0.253
Number of plants	134	134
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X
N (producing mine)		35
N (non-producing mine)		32

Notes: This table presents the results for longer term effect of mine cancellation on monthly receipt of coal at the power plant unit level. Column 1 compares all treated plant units to all control plant units. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Robust standard errors clustered at the plant level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The longer term estimates of the impact on coal consumption are reported in table 3.10. For the full sample, the coefficient of interest increases in absolute size as compared to the short run and the estimate is precise at 99% confidence interval. The estimated effects for both high and low uncertainty plants increase in absolute size in comparison to short term results and are statistically significant at 5% suggesting an overall increase in supply side constraints.

Results for savings follow a similar pattern as the short term results and are reported in table 3.11. Plants affected by the SC decision to cancel captive mine allocation continue to

Table 3.10: Longer term effect of mining contract cancellation on monthly consumption of coal

	Coal Consumption ('000 Metric Tonnes/month)	
	(1)	(2)
cancellation	-37.85***	
	(13.72)	
cancellation $\times$ *(producing mine)*		-42.36**
		(16.76)
cancellation $\times$ *(non-producing mine)*		-31.67**
		(15.61)
Observations	7,340	7,340
R-squared	0.272	0.272
Number of plants	134	134
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X
N (producing mine)		35
N (non-producing mine)		32

Notes: This table presents the results for longer term effect of mine cancellation on monthly consumption of coal at the power plant unit level. Column 1 compares all treated plants to all control plants. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Robust standard errors clustered at the plant level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

save more coal relative to plants with secure long term fuel supply contracts. The bulk of the savings come from plants that were linked to a producing captive mine at the time of SC judgement on *coal-gate*.

Table 3.11: Longer term effect of mining contract cancellation on monthly coal saving

	Coal Saving ('000 Metric Tonnes/month)	
	(1)	(2)
cancellation	8.026**	
	(3.749)	
cancellation × *(producing mine)*		10.26**
		(4.064)
cancellation × *(non-producing mine)*		4.954
		(3.899)
Observations	7,340	7,340
R-squared	0.076	0.077
Number of plants	134	134
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X
N (producing mine)		35
N (non-producing mine)		32

Notes: This table presents the results for longer term effect of mine cancellation on monthly coal saving at the power plant unit level. Column 1 compares all treated plants to all control plants. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Robust standard errors clustered at the plant level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Power generation

Finally, the estimated longer term effect on power generation lost is reported in table 3.12 and shows the persistence of the estimated effect. The treatment on treated effect for the entire sample in column 1 is larger in absolute magnitude and within the 95% confidence interval. Similar to the previous results, the majority of the generation loss comes from the plants more vulnerable to future coal shortages.

Table 3.12: LONGER TERM EFFECT OF MINING CONTRACT CANCELLATION ON ELECTRICITY GENERATION

	Power Generation (GWh/month)	
	(1)	(2)
cancellation	-47.53**	
	(20.92)	
cancellation × *(producing mine)*		-65.41**
		(27.45)
cancellation × *(non-producing mine)*		-22.97
		(20.71)
Observations	7,341	7,341
R-squared	0.318	0.320
Number of plants	134	134
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X
N (producing mine)		35
N (non-producing mine)		32

Notes: This table presents the results for longer term effect of mine cancellation on monthly amount of electricity generated at the power plant level. Column 1 compares all treated plants to all control plants. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Robust standard errors clustered at the plant level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

3.7 Welfare Analysis

The aim of this section is to estimate the lower and upper bounds of the cost of regulatory uncertainty faced by power plants. As such, this section focuses on measuring the dollar value of welfare loss in short term during uncertainty phase 1. The estimation relies on the assumption that the social planner's objective is to maximize generation output and minimize total cost of generation.

For a power plant, inter-temporal substitution of output is constrained by its fixed generation capacity. The calculations in this section rely on this fact and assume that outages at the plant level are not substitut-able over time. In other words, I assume that an increase in outages today does not necessarily result in decrease in outages tomorrow and represents an absolute increase in total outages. Given this assumption, the welfare estimates should be treated with caution and interpreted as bounds on the *ex-ante* welfare loss due to regulatory uncertainty.

To calculate the change in welfare, I first draw the least-cost supply curve for the actual monthly generation of the sample plants as reported in the data. The least cost supply curve represents the merit order dispatch of generation output and is produced by supplying power from the plants in increasing order of their marginal cost.

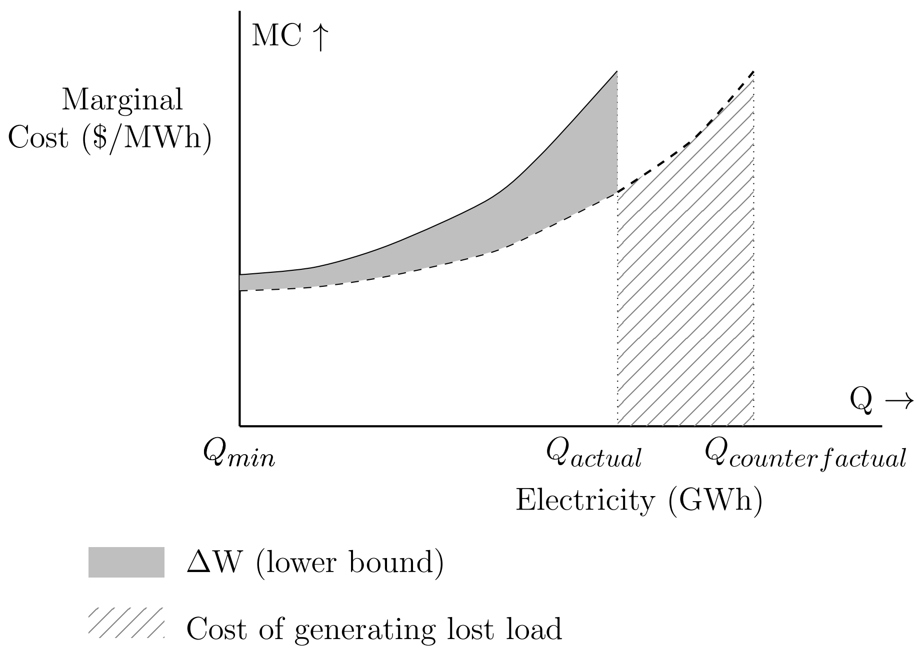
Then, I simulate a counterfactual supply curve which represents the merit order dispatch of counterfactual generation schedule had the "treated" plants not declared outage due to coal shortage when they still had coal. To generate this counterfactual supply curve, I add back to the "treated" plants, generation output from the capacity that was reported to be under outage due to coal shortage[25].

Figure 3.11 illustrates an example of this welfare analysis. Plant's marginal cost to produce electricity is on the y axis and the x axis represents the corresponding quantity of output supplied. The *solid* line represents the supply curve for actual amount of generation

25. I assume capacity utilization rate of 85%

in the data, Q_{actual}. The *dashed* line represents the counterfactual supply curve if the "treated" plants had not declared outage due to coal shortage when they still had coal. The total counterfactual generation is denoted by $Q_{counterfactual}$. The *dashed* line lies below the *solid* line because the "treated" plants had lower generation cost on account of lower fuel transportation as they were in close proximity to the captive coal mines.

Figure 3.10: Welfare Analysis: An illustrative example



Notes: This figure provides an example of the welfare analysis. The solid line represents the total cost minimizing supply curve for actual generation. Least cost supply curve for predicted generation is represented by dashed line. The relevant areas for welfare analysis are shaded and denoted in the legend.

I start by computing the lower bound estimate of cost of regulatory uncertainty by assuming zero generation loss due to regulatory uncertainty. That is, I assume that in the counterfactual scenario, the total generation is the same as the actual generation, and the *dashed* line stops at Q_{actual}. The lower bound welfare loss estimate is thus represented by the gray shaded area between the *solid* and *dashed* supply curves and is equal to the increase in the total cost of generation due to increase in outages reported by "treated" plants. I describe this in detail below:

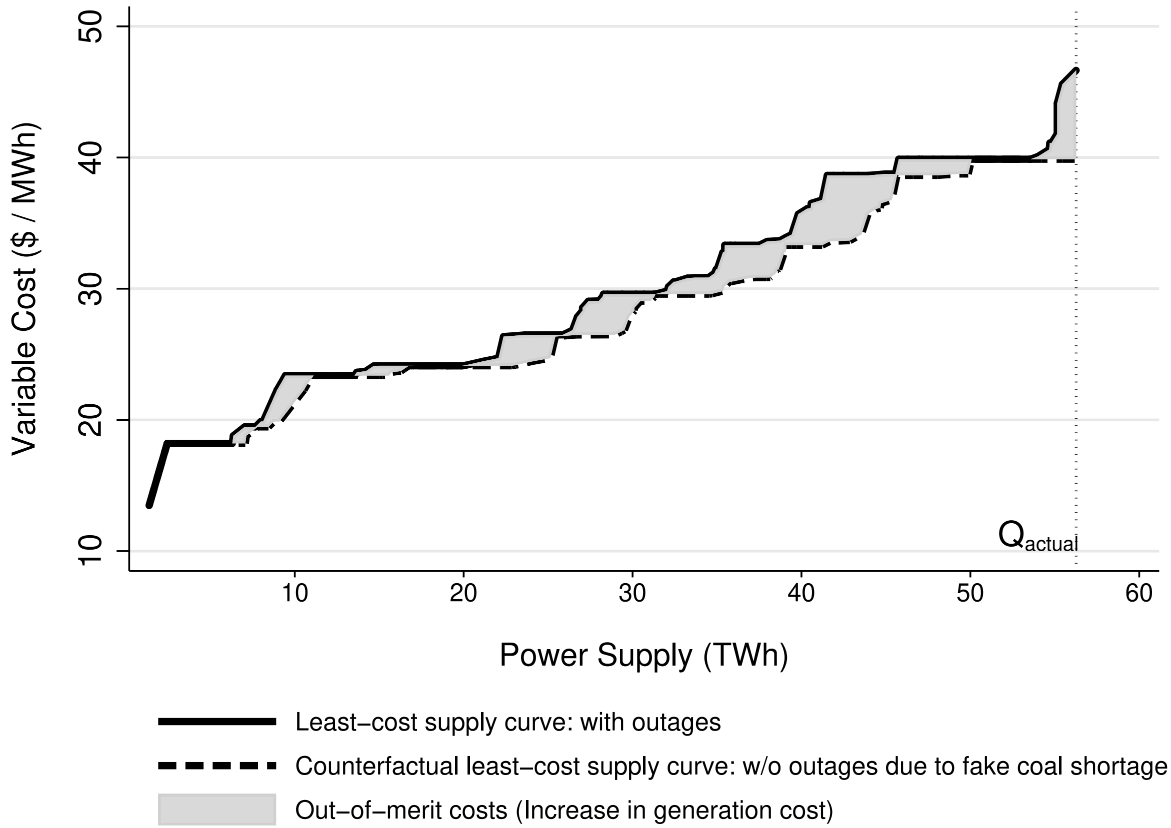
3.7.1 Change in Welfare: Lower Bound

As described above, the lower bound of the welfare loss is estimated to be equal to the increase in total generation cost is due to out-of-merit production.

To estimate the out-of-merit costs, I generate least-cost dispatch supply curves as described above for all 7 months of the first phase of uncertainty from September 2014 to March 2015.

Figure 3.12 shows the corresponding supply curves for the first month of uncertainty after the SC judgement, September 2014. The area between the two supply curves represents the increased generation cost to produce the same amount of electricity and gives the lower bound estimate of the welfare loss due to out-of-merit generation as a result of uncertain fuel supply.

Figure 3.11: Increased generation cost due to uncertain fuel supply



Notes: The figure shows the actual and simulated least-cost dispatch supply curves for the month of September 2014. The area in gray represents the out-of-merit generation costs under least-cost dispatch and totals 72.56 Million US Dollars.

I draw similar supply curve comparisons for all 7 months of the first phase of uncertainty from September 2014 to March 2015. These figures can be found in appendix D. The lower bound of the estimated aggregate cost of regulatory uncertainty equals the sum of the out-of-merit costs across the 7 months of uncertainty phase 1 and equals 291 million dollars

3.7.2 Change in Welfare: Upper Bound

To estimate the upper bound of change in welfare, I assume that the regulatory uncertainty resulted in generation loss equal to $Q_{counterfactual}$–Q_{actual}. Assigning a value v to a unit of generation and denoting marginal cost by MC, the change in welfare can be written as[26]:

$$
\begin{aligned}
\Delta\text{W } \textit{(upper bound)} &= [v \times Q_{c'factual} - \int_{Q_{min}}^{Q_{c'factual}} \text{MC } dq] - [v \times Q_{actual} - \int_{Q_{min}}^{Q_{actual}} \text{MC } dq] \\
&= [v \times (Q_{c'factual} - Q_{actual}) - \int_{Q_{actual}}^{Q_{c'factual}} \text{MC } dq] + \int_{Q_{min}}^{Q_{actual}} \text{MC } dq \\
&= \qquad \textit{net} \text{ value of lost load} \qquad + \ \Delta\text{W } \textit{(lower bound)}
\end{aligned}
$$

The expression above shows that the upper bound of the estimated loss in welfare is equal to the sum of value of lost generation output net of cost and the estimated lower bound.

For estimate of v, I rely on the results of Alcott, Collard-Wexler and O'Connell (2016) that looks at the effect of outages onthe formal sector manufacturing plants in India. In their paper, they estimate that a percentage point increase in gap between supply of power by generators and requirement by utilities leads to a "1.091 percent decrease in revenues" of manufacturing plants in India.

The average revenue in of a manufacturing plant in their sample is 3 Million dollars. Using that and the fact that the average percentage gap between supply and demand was 3.33 % during the first phase of uncertainty, I compute the value of lost generation to be 1.92 Billion dollars.[27]

26. Since the supply curve is discrete, the actual computations involves summations instead of integrals

27. From table 7, page 611, of Allcott et al (2016):

$$
\begin{aligned}
\text{Annual revenue loss caused by} & \\
\text{1 \% increase in outages } &= \$3,000,000 \times 0.0109 \times (613930/19) \\
&= \$987,877,994.20
\end{aligned}
$$

$$
\begin{aligned}
\text{Revenue loss for this paper} &= \$987,877,994.20 \times (1/12 \times 7) \times 3.33 \\
&= \$1,918,235,392.71
\end{aligned}
$$

Using data on marginal cost and estimated generation, I calculate the total cost of generating this lost load to be equal to 718 Million dollars[28] for the 7 months of first phase of uncertainty. This translates into a *net* vaue of lost load equal to 1.2 Billion dollars. Hence the upper bound of the estimated aggregate cost of regulatory uncertainty is approximately equal to 1.5 Billion dollars.

The aggregate cost of regulator uncertainty in the context of this paper lies between (0.3, 1.5) billion dollars.

3.8 Robustness Checks

I conduct three robustness checks to further cement the results. All robustness results are in the appendices attached at the end of this article.

First, appendix A contains the estimates for the effect of mining contract annulment on outages for the entire sample of 147 plants for which data on capacity under outage is available. The size and statistical significance of the estimates is similar to the main results.

Second, the panel data sample used to estimate short term results consists of 124 clusters. To the extent the modest number of clusters affect the estimated standard errors, I compute standard errors clustered at the plant level using bootstrap method. The standard errors do not change much and the significance of the estimated coefficients remains unchanged for all estimations. The results of this exercise can be found in appendix B.

Lastly, the cancellation of coal mines could have potentially affected the coal fired plants that were unaffected by the mine cancellation if the regulator directed these plants to change their production behavior. I find no evidence of this in the results section but as an additional check I compare all coal fired plants to all gas and lignite fired plants. In appendix C, I present the estimates for impact on outages and generation treating all coal fired plants as "treated" units and natural gas and lignite (*brown coal* based thermal power plants as "control" plants.

---

28. \$718,020,309.97

The results for the effect of mine cancellation on capacity under outage and power generated are slightly larger but the economic and statistical significance remains unchanged.

3.9 Conclusion

The main insight from this paper is that regulatory levers on the supply side play a major role in the persistence of power outages in India. The results of this paper point towards various inefficiencies in the coal allocation policy and how they affect the provision of electricity in India. Policy discussions about addressing power outages in India often revolve around addressing inefficiencies at the generation and distribution of electricity. This paper shows that there is a more fundamental issue of allocation of fuel to power plants that needs to be addressed as part of the solution to power outages and provide continuous access to electricity.

And this is important for environment policy as well as it shows that it is critical to evaluate the non-market incentives and constraints that are common in India and other developing countries, in order to make meaningful progress on combating the urgent problem of environmental degradation due to fossil fuels.

The persistence of the effects of mine cancellations on firms' output shows that the effects can last longer than anticipated in tight regulatory environments. The Supreme Court of India perhaps had the right intention when they canceled the mining contracts. And the decision to give the government of India a seven-month window to formulate a new policy can be, ex-ante, argued to be pragmatic. However, the results of this paper show in the absence of an immediate fallback option, scrapping the status quo leads to economic loss.

Finally, the large effects on output due to regulatory uncertainty show that waiting is costly. The cost of uncertainty monotonically increases with the length of time and this underlines the need to avoid delays in policy making. This has lesson for future natural resource allocation issues – achieving first best is desirable but there is cost attached to waiting to arrive at first best.

Description of Appendices

Appendix A contains the results for the effect of mine cancellation on daily capacity under outage for the entire sample of power plants that started operations before 2015.

Appendix B replicates all the results of the paper with bootstrapped standard errors.

Appendix C tabulates the ATT estimates of the effect on outages and generation by comparing all coal fired plants ("*treated*") with gas fired plants ("*control*").

Appendix D shows the monthly out of merit costs from August 2014 to march 2015 for the first phase of regulatory uncertainty.

3.10 APPENDIX A

Table A.1: Short term effect of uncertainty on daily generation capacity under outage (FULL SAMPLE)

	Capacity under Outage (MW/day)	
	(1)	(2)
cancellation	43.53**	
	(20.69)	
cancellation × *(producing mine)*		64.24**
		(27.91)
cancellation × *(non-producing mine)*		20.17
		(22.67)
Observations	110,884	110,884
R-squared	0.042	0.043
Number of plants	147	147
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X
N (producing mine)		35
N (non-producing mine)		34

Notes: This table presents the results for short term effect of uncertainty on daily outages at the power plant level for the full sample for which information on capacity under outage is available. Column 1 compares all treated plant units to all control plant units. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Robust standard errors clustered at the plant level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.2: Short term effect of uncertainty on capacity under outage due to reported fuel shortage (FULL SAMPLE)

	Capacity under Outage due to reported coal shortage (MW/day)	
	(1)	(2)
cancellation	33.25**	
	(16.11)	
cancellation × *(producing mine)*		47.62**
		(22.80)
cancellation × *(non-producing mine)*		17.03
		(17.68)
Observations	110,884	110,884
R-squared	0.025	0.026
Number of station_id	147	147
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X
N (high uncertainty)		35
N (low uncertainty)		34

Notes: This table presents the results for short term effect of uncertainty on daily outages reported due to coal shortage at the power plant level for the full sample for which information on capacity under outage is available. Column 1 compares all treated plant units to all control plant units. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Robust standard errors clustered at the plant level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.3: LONGER TERM EFFECT OF MINING CONTRACT CANCELLATION ON DAILY GENERATION CAPACITY UNDER OUTAGE (FULL SAMPLE)

	Capacity under Outage (MW/day)	
	(1)	(2)
cancellation	48.29**	
	(23.93)	
cancellation × (producing mine)		74.07**
		(34.75)
cancellation × (non-producing mine)		18.68
		(24.07)
Observations	247,484	247,484
R-squared	0.064	0.065
Number of station_id	147	147
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X
N (high uncertainty)		35
N (low uncertainty)		34

Notes: This table presents the results for longer term effect of mine cancellation on daily outages at the power plant unit level for the full sample for which information on capacity under outage is available. Column 1 compares all treated plant units to all control plant units. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Robust standard errors clustered at the plant level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

3.11 APPENDIX B

Table B.1: Short term effect of uncertainty on daily generation capacity under outage (BOOTSTRAP SE)

	Capacity under Outage (MW/day)	
	(1)	(2)
cancellation	45.84**	
	(23.32)	
cancellation × *(producing mine)*		68.95**
		(31.12)
cancellation × *(non-producing mine)*		14.38
		(24.31)
Observations	96,396	96,396
R-squared	0.043	0.045
Number of plants	124	124
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X
N (producing mine)		35
N (non-producing mine)		26

Notes: This table presents the results for short term effect of uncertainty on daily outages at the power plant level. Column 1 compares all treated plant units to all control plant units. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Bootstrapped standard errors clustered at the plant level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B.2: Short term effect of uncertainty on capacity under outage due to reported fuel shortage (BOOTSTRAP SE)

	Capacity under Outage due to reported coal shortage (MW/day)	
	(1)	(2)
cancellation	35.19*	
	(20.94)	
cancellation × *(producing mine)*		46.92
		(29.60)
cancellation × *(non-producing mine)*		19.24
		(22.12)
Observations	96,396	96,396
R-squared	0.028	0.029
Number of plants	124	124
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X
N (producing mine) 32		35
N (non-producing mine) 40		26

Notes: This table presents the results for short term effect of uncertainty on daily outages reported due to coal shortage at the power plant level. Column 1 compares all treated plant units to all control plant units. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Bootstrapped standard errors clustered at the plant level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B.3: Short term effect of uncertainty on monthly receipts of coal (BOOTSTRAP SE)

	Coal Supply ('000 Metric Tonnes/month)	
	(1)	(2)
cancellation	-4.982	
	(9.602)	
cancellation × *(producing mine)*		-7.423
		(13.09)
cancellation × *(non-producing mine)*		-1.605
		(11.40)
Observations	3,163	3,163
R-squared	0.192	0.193
Number of plants	124	124
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X
N (producing mine)		35
N (non-producing mine)		26

Notes: This table presents the results for short term effect of uncertainty on monthly receipt of coal at the power plant unit level. Column 1 compares all treated plant units to all control plant units. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Bootstrapped standard errors clustered at the plant level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B.4: Short term effect of uncertainty on monthly consumption of coal (BOOTSTRAP SE)

	Coal Consumption ('000 Metric Tonnes/month)	
	(1)	(2)
cancellation	-22.31*	
	(11.42)	
cancellation × *(producing mine)*		-31.15**
		(13.50)
cancellation × *(non-producing mine)*		-10.08
		(11.26)
Observations	3,163	3,163
R-squared	0.202	0.204
Number of plants	124	124
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sampe Month-Year FE	X	X
N (producing mine)		35
N (non-producing mine)		26

Notes: This table presents the results for short term effect of uncertainty on monthly consumption of coal at the power plant unit level. Column 1 compares all treated plants to all control plants. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Bootstrapped standard errors clustered at the plant level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B.5: Difference in difference estimates of the effect of uncertainty on monthly precautionary coal savings (BOOTSTRAP SE)

	Coal Savings ('000 Metric Tonnes/month)	
	(1)	(2)
cancellation	17.33** (8.720)	
cancellation × *(producing mine)*		23.73*** (10.25)
cancellation × *(non-producing mine)*		8.477 (8.796)
Observations	3,163	3,163
R-squared	0.202	0.204
Number of plants	124	124
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X
N (producing mine)		35
N (non-producing mine)		26

Notes: This table presents the results for short term effect of uncertainty on monthly precautionary coal saving at the power plant level. Column 1 compares all treated plants to all control plants. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Bootstrapped standard errors clustered at the plant level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B.6: SHORT TERM EFFECT OF UNCERTAINTY ON ELECTRICITY GENERATION (BOOTSTRAP SE)

	Power Generation (GWh/month)	
	(1)	(2)
cancellation	-36.13**	
	(16.87)	
cancellation × *(producing mine)*		-51.47**
		(20.33)
cancellation × *(non-producing mine)*		-14.89
		(16.81)
Observations	3,163	3,163
R-squared	0.295	0.297
Number of plants	124	124
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X
N (producing mine)		35
N (non-producing mine)		26

Notes: This table presents the results for short term effect of uncertainty on monthly amount of power generated at the plant level. Column 1 compares all treated plants to all control plants. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Bootstrapped standard errors clustered at the plant level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B.7: LONGER TERM EFFECT OF MINING CONTRACT CANCELLATION ON DAILY GENERATION CAPACITY UNDER OUTAGE (BOOTSTRAP SE)

	Capacity under Outage (MW/day)	
	(1)	(2)
cancellation	55.34**	
	(23.56)	
cancellation × *(producing mine)*		75.50**
		(34.07)
cancellation × *(non-producing mine)*		27.76
		(26.46)
Observations	211,674	211,674
R-squared	0.064	0.065
Number of plants	124	124
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X
N (producing mine)		35
N (non-producing mine)		32

Notes: This table presents the results for longer term effect of mine cancellation on daily outages at the power plant unit level. Column 1 compares all treated plant units to all control plant units. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Bootstrapped standard errors clustered at the plant level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B.8: Longer term effect of mining contract cancellation on monthly receipts of coal (BOOTSTRAP SE)

	Coal Supply ('000 Metric Tonnes/month)	
	(1)	(2)
cancellation	-29.73**	
	(12.45)	
cancellation × *(producing mine)*		-31.93*
		(16.42)
cancellation × *(non-producing mine)*		-26.71*
		(14.63)
Observations	7,341	7,341
R-squared	0.253	0.253
Number of plants	134	134
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X
N (producing mine)		35
N (non-producing mine)		32

Notes: This table presents the results for longer term effect of mine cancellation on monthly receipt of coal at the power plant unit level. Column 1 compares all treated plant units to all control plant units. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Bootstrapped standard errors clustered at the plant level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B.9: Longer term effect of mining contract cancellation on monthly consumption of coal (BOOTSTRAP SE)

	Coal Consumption ('000 Metric Tonnes/month)	
	(1)	(2)
cancellation	-37.85***	
	(12.80)	
cancellation × *(producing mine)*		-42.36***
		(16.10)
cancellation × *(non-producing mine)*		-31.67**
		(15.47)
Observations	7,340	7,340
R-squared	0.272	0.272
Number of plants	134	134
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X
N (producing mine)		35
N (non-producing mine)		32

Notes: This table presents the results for longer term effect of mine cancellation on monthly consumption of coal at the power plant unit level. Column 1 compares all treated plants to all control plants. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Bootstrapped standard errors clustered at the plant level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B.10: Longer term effect of mining contract cancellation on monthly coal saving (BOOTSTRAP SE)

	Coal Saving ('000 Metric Tonnes/month)	
	(1)	(2)
cancellation	8.026**	
	(3.981)	
cancellation × *(producing mine)*		10.26**
		(4.528)
cancellation × *(non-producing mine)*		4.954
		(3.737)
Observations	7,340	7,340
R-squared	0.076	0.077
Number of plants	134	134
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X
N (producing mine)		35
N (non-producing mine)		32

Notes: This table presents the results for longer term effect of mine cancellation on monthly coal saving at the power plant unit level. Column 1 compares all treated plants to all control plants. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Bootstrapped standard errors clustered at the plant level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B.11: Longer term effect of mining contract cancellation on electricity generation (BOOTSTRAP SE)

	Power Generation (GWh/month)	
	(1)	(2)
cancellation	-47.53***	
	(17.81)	
cancellation × *(producing mine)*		-65.41***
		(24.77)
cancellation × *(non-producing mine)*		-22.97
		(18.33)
Observations	7,341	7,341
R-squared	0.318	0.320
Number of plants	134	134
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X
N (producing mine)		35
N (non-producing mine)		32

Notes:This table presents the results for longer term effect of mine cancellation on monthly amount of electricity generated at the power plant level. Column 1 compares all treated plants to all control plants. In column 2, the treatment is interacted with a dummy for high uncertainty; high certainty dummy takes the value 1 for plants linked to producing mines and 0 for plants that had temporary linkages in lieu of their non-producing mines. All regressions control for plant level time-invariant characteristics. Bootstrapped standard errors clustered at the plant level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

3.12 APPENDIX C

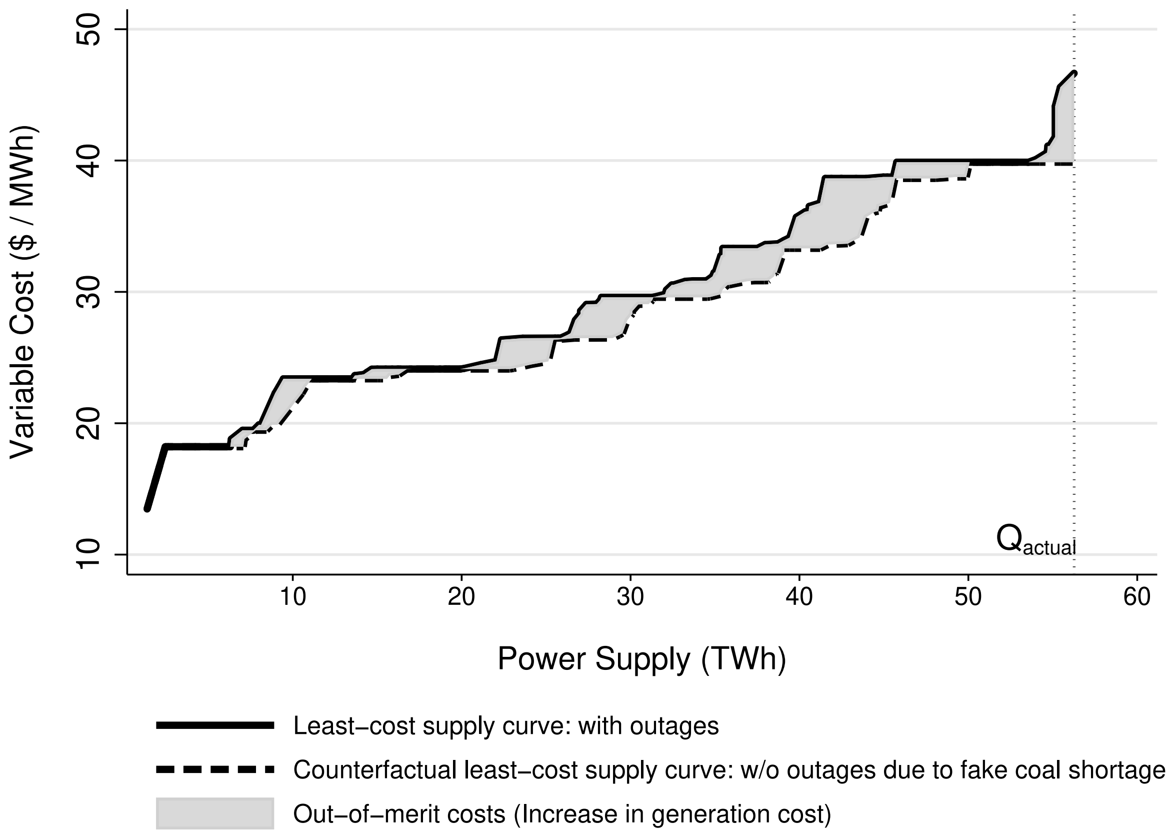
Table C.1: Effect of uncertain fuel supply on coal plants vs gas plants

	Daily outages (MW / day)	Monthly generation (GWh / month)
	(1)	(2)
cancellation	33.83**	19.22**
	(14.63)	(9.117)
Observations	140,352	4,592
R-squared	0.016	0.269
Number of plants	177	176
Plant Age	X	X
Plant Capacity	X	X
Plant FE	X	X
Sample Month-Year FE	X	X

Notes: This table presents the results for short term effect of uncertainty on daily outages and generation at the power plant level for the coal fired plants vs gas fired plants. Column 1 contains the results for effect on daily outages and column 2 provides estimates for monthly generation. All regressions control for plant level time-invariant characteristics. Robust standard errors clustered at the plant level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

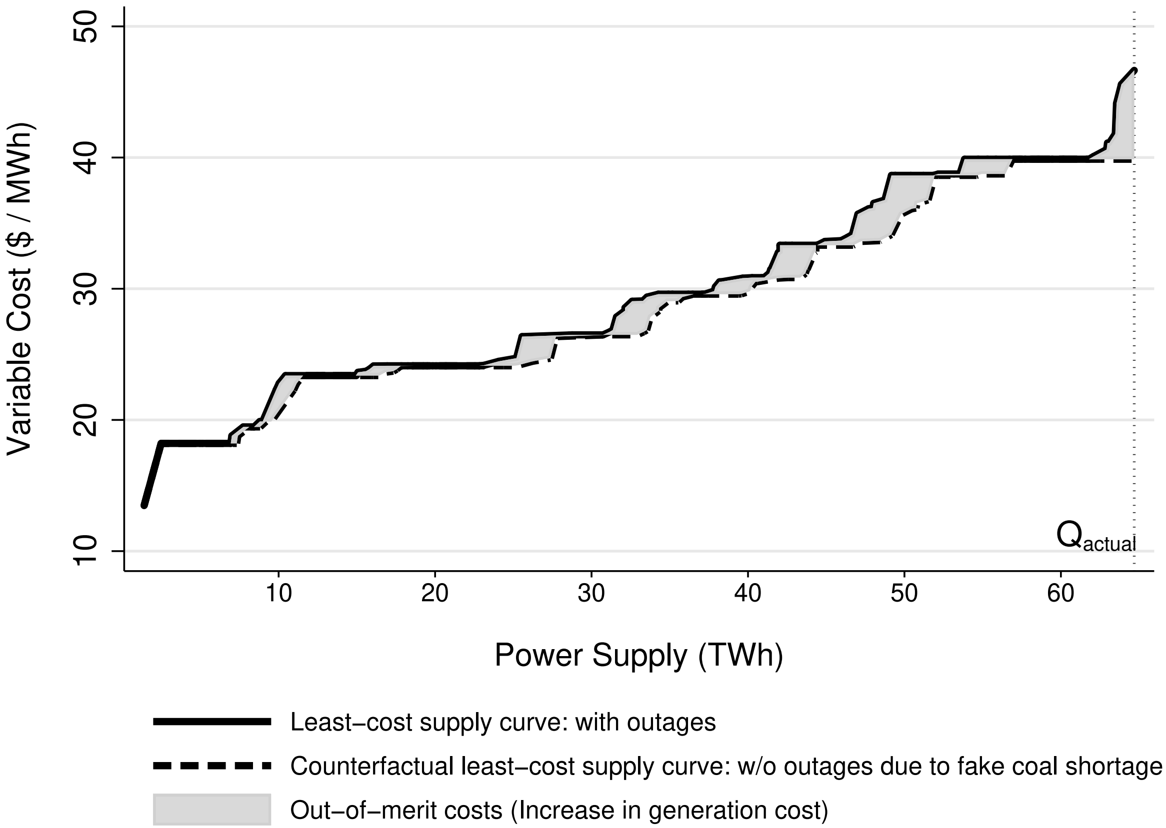
3.13 APPENDIX D

Figure D.1: Loss in producer surplus due to uncertain fuel supply - September 2014



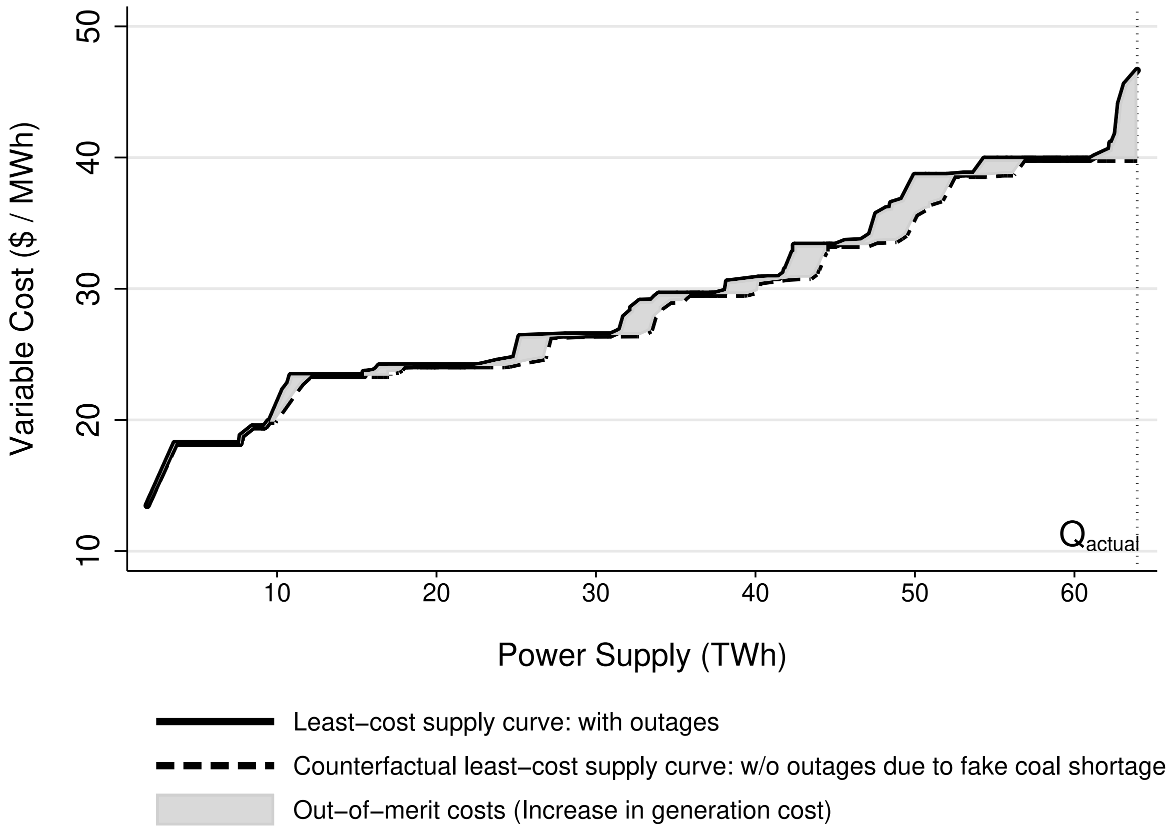
Notes: The figure shows the actual and simulated least-cost dispatch supply curves for the month of September 2014. The area in gray represents the out-of-merit generation costs under least-cost dispatch and totals 72.56 Million US Dollars.

Figure D.2: Loss in producer surplus due to uncertain fuel supply - October 2014



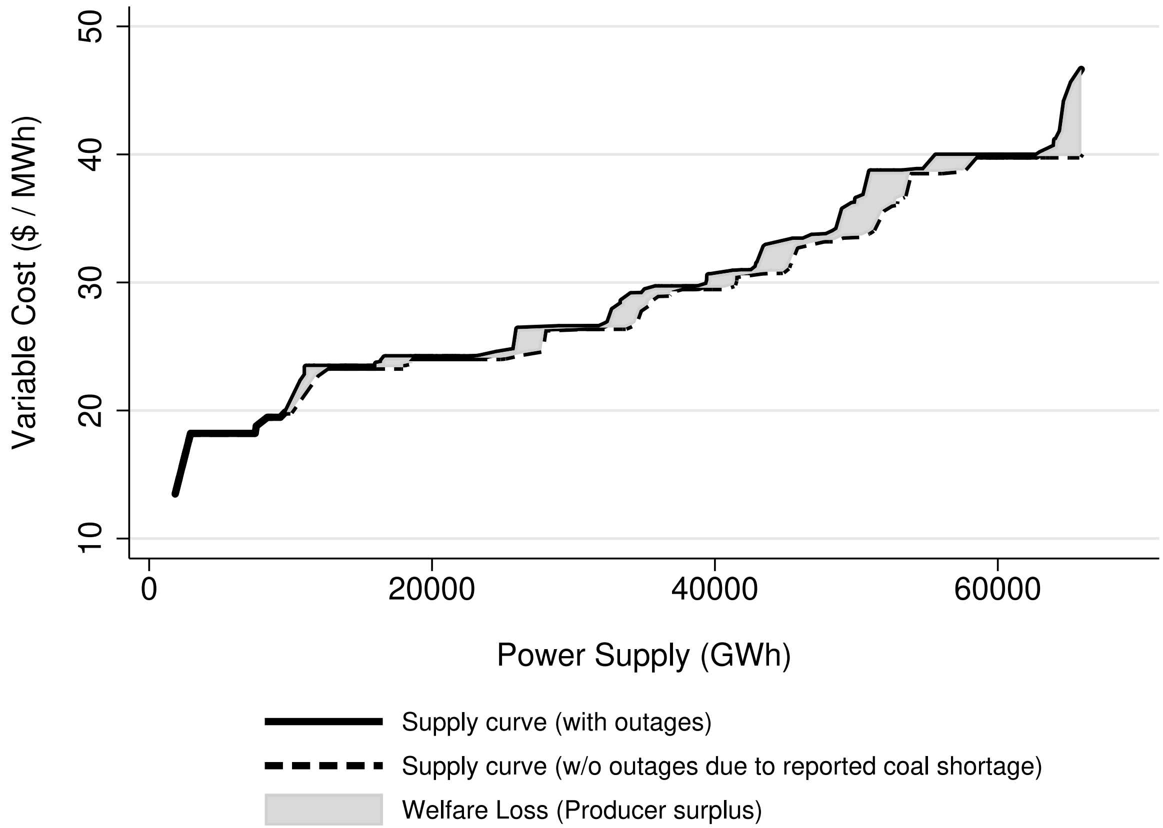
Notes: The figure shows the actual and simulated least-cost dispatch supply curves for the month of October 2014. The area in gray represents the out-of-merit generation costs under least-cost dispatch and totals 51.61 Million US Dollars.

Figure D.3: Loss in producer surplus due to uncertain fuel supply - November 2014



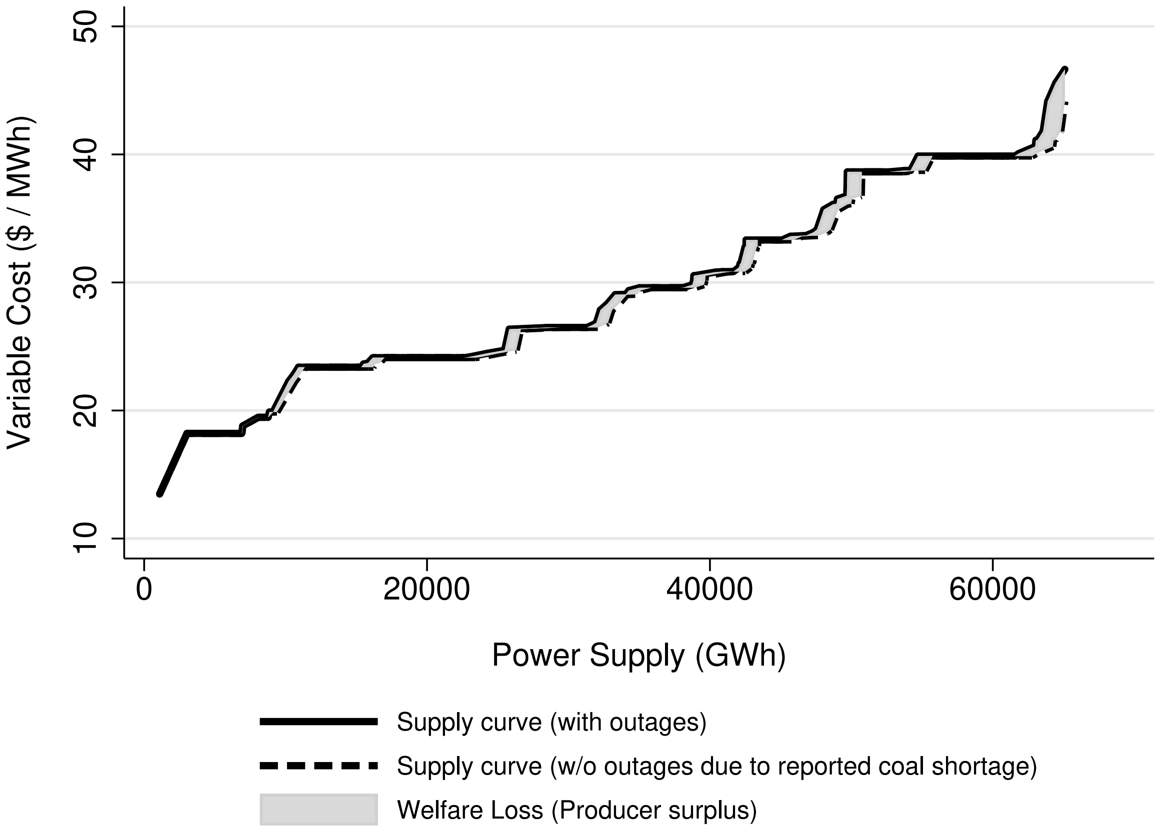
Notes: The figure shows the actual and simulated least-cost dispatch supply curves for the month of November 2014. The area in gray represents the out-of-merit generation costs under least-cost dispatch and totals 45.35 Million US Dollars.

Figure D.4: Loss in producer surplus due to uncertain fuel supply - December 2014



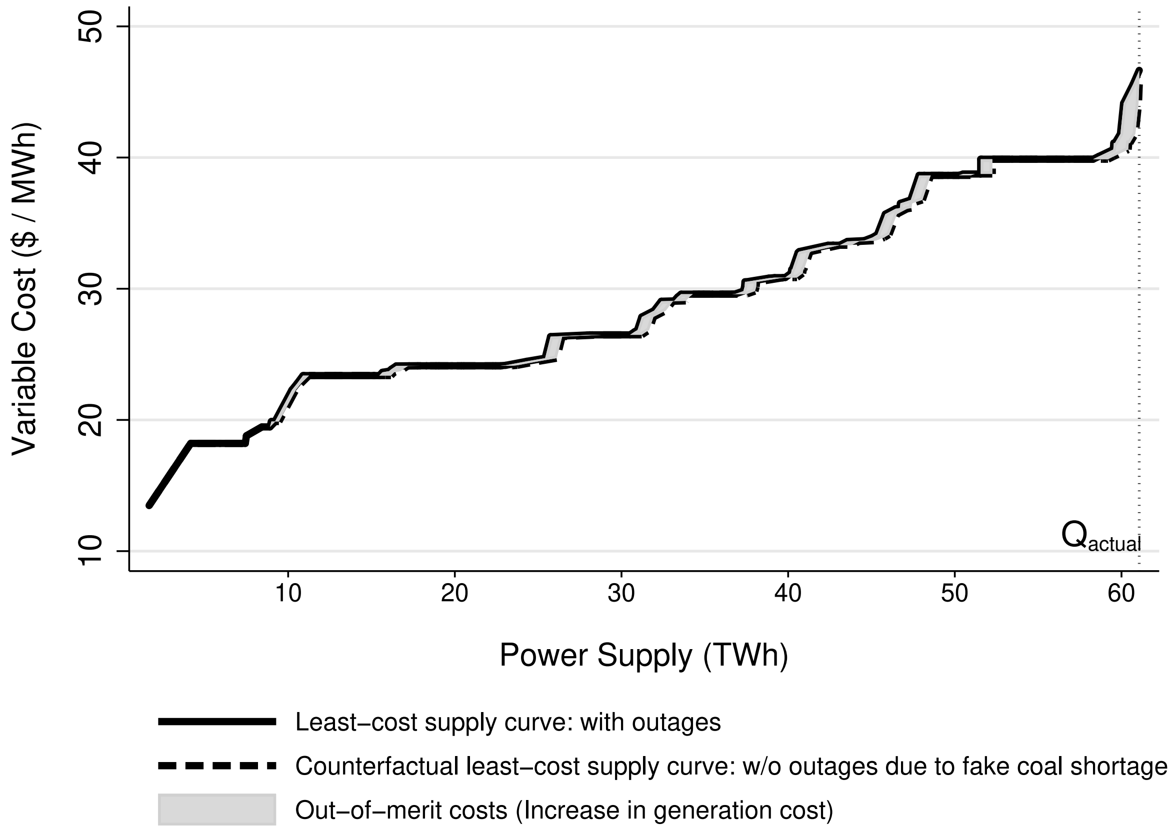
Notes: The figure shows the actual and simulated least-cost dispatch supply curves for the month of December 2014. The area in gray represents the out-of-merit generation costs under least-cost dispatch and totals 48.43 Million US Dollars.

Figure D.5: Loss in producer surplus due to uncertain fuel supply - January 2015



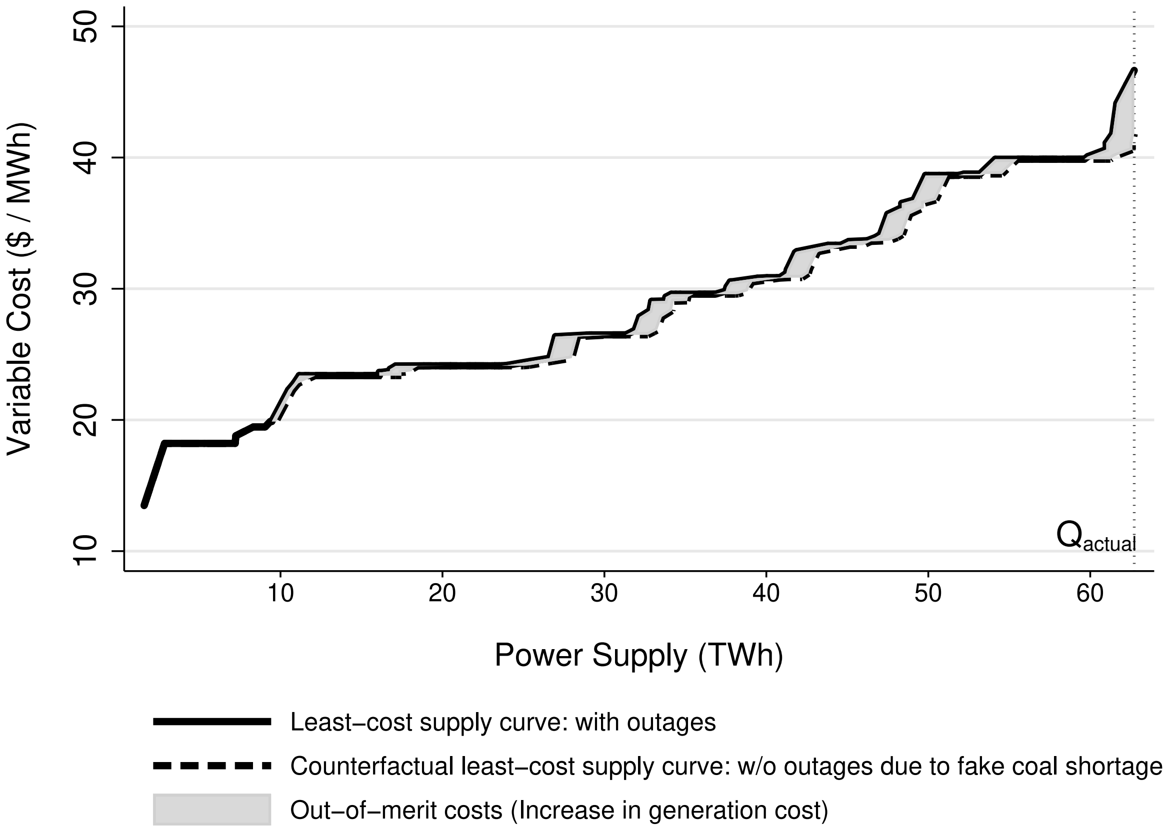
Notes: The figure shows the actual and simulated least-cost dispatch supply curves for the month of January 2015. The area in gray represents the out-of-merit generation costs under least-cost dispatch and totals 23.54 Million US Dollars.

Figure D.6: Loss in producer surplus due to uncertain fuel supply - February 2015



Notes: The figure shows the actual and simulated least-cost dispatch supply curves for the month of February 2015. The area in gray represents the out-of-merit generation costs under least-cost dispatch and totals 17.79 Million US Dollars.

Figure D.7: Loss in producer surplus due to uncertain fuel supply - March 2015



Notes: The figure shows the actual and simulated least-cost dispatch supply curves for the month of March 2015. The area in gray represents the out-of-merit generation costs under least-cost dispatch and totals 31.60 Million US Dollars.

REFERENCES

[1] Allcott, Hunt, Allan Collard-Wexler, and Stephen D. O'Connell. 2016. "How do electricity shortages affect industry? Evidence from India." American Economic Review 106, no. 3 (2016): 587-624.

[2] Abeberese, Ama Baafra. 2017. "Electricity cost and firm performance: Evidence from India." Review of Economics and Statistics 99, no. 5 (2017): 839-852.

[3] Baker, S. R., Bloom, N., and Davis, S. J. 2016. "Measuring Economic Policy Uncertainty." The Quarterly Journal of Economics, 131(4), 1593–1636.

[4] Ryan, Nicholas. 2017. "The Competitive Effects of Transmission Infrastructure in the Indian Electricity Market." Technical report. National Bureau of Economic Research Working Paper No. 23106.

[5] Central Electricity Authority (CEA). 2014 - 2018. "Load Generation Balance Report. New Delhi: Central Electricity Authority".

[6] Central Electricity Authority (CEA). 2014 - 2018. "Coal Statistics. New Delhi: Central Electricity Authority".

[7] Chan, Hei Sing (Ron), Maureen L. Cropper, and Kabir Malik. 2014. "Why Are Power Plants in India Less Efficient Than Power Plants in the United States?" American Economic Review, 104 (5): 586-90

[8] Cicala, Steve 2015. "When Does Regulation Distort Costs? Lessons from Fuel Procurement in US Electricity Generation." American Economic Review, 105(1), 411–444.

[9] Dorsey, Jackson. 2019. "Waiting on the Courts: Effects of Policy Uncertainty on Pollution and Investment." *forthcoming*

[10] Fuchs-Schündeln, Nicola. 2008. "The Response of Household Saving to the Large Shock of German Reunification." American Economic Review 98 (2008), 1798–1828.

[11] Fuchs-Schündeln, Nicola, and Matthias Schündeln. 2005. "Precautionary Savings and Self-Selection: Evidence from the German Reunification 'Experiment'.", Quarterly Journal of Economics 120 (2005), 1085–1120.

[12] Francesco Giavazzi and Michael McMahon. 2012. "Policy Uncertainty and Household Savings." The Review of Economics and Statistics 2012 94:2, 517-531

[13] Kira R. Fabrizio. 2013. "The Effect of Regulatory Uncertainty on Investment: Evidence from Renewable Energy Generation." The Journal of Law, Economics, and Organization, Volume 29, Issue 4

[14] Mahadevan, Meera. "The Price of Power: Costs of Political Corruption in Indian Electricity." *working paper*, 2019

[15] Ministry of Coal, India reports: various years

[16] Ministry of Power, India reports: various years

[17] McRae, Shaun. 2015. "Infrastructure quality and the subsidy trap." American Economic Review 105, no. 1 (2015): 35-66.