Essays on State Interventions and Market Outcomes in Developing Economies

HARVARD UNIVERSITY
Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the Committee for the PhD in Business Economics have examined a dissertation entitled

Essays on State Interventions and Market Outcomes in Developing Economies

Presented by **Sagar Saxena**

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature ______________________
Ariel Pakes, Chair

Signature ______________________
Emily Breza

Signature ______________________
Edward Glaeser

Signature ______________________
Myrto Kalouptsidi

Signature ______________________
Gautam Rao

Date: April 24, 2023

Essays on State Interventions and Market Outcomes in Developing Economies

A dissertation presented

by

Sagar Saxena

to

The Department of Economics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Business Economics

Harvard University

Cambridge, Massachusetts

April 2023

Dissertation Advisor:
Professor Ariel Pakes

Author:
Sagar Saxena

Essays on State Interventions and Market Outcomes in Developing Economies

Abstract

This thesis consists of three essays which empirically analyze the equilibrium effects of various state interventions in markets in developing countries. The first essay studies the effects of a bundle of large-scale agricultural interventions in India, focusing on their distributional impact on farmers and consumers. The second essay quantifies the economic costs of protectionist industrial policy interventions in the utility-scale solar sector in India. The third essay examines the aggregate and distributional effects of different public healthcare provision strategies in India. Each essay develops and estimates a context-specific structural model of the relevant market, that is then used to evaluate counterfactuals and generate policy insights.

Table Of Contents

List of Tables

List of Figures

Acknowledgments

I am deeply grateful to my advisors, who have been instrumental in shaping my academic journey and supporting my research endeavors. Ariel Pakes and Myrto Kalouptsidi, my IO advisors, have been crucial in developing my skills as an empirical IO economist. Ariel's enthusiasm and genuine curiosity for my work, even when it ventured beyond his own domain, have not only shaped these essays but also laid the foundation for the kind of research I will undertake in the years to come. Myrto has been an extraordinary mentor and guide. Her critical feedback has significantly influenced the final work product of many of my papers, and over the last few months, she has been an incredible source of support and encouragement as I navigated the academic job market. Emily Breza and Gautam Rao, my development advisors, have been essential in fostering my pursuit of the policy questions that drive my research. Emily taught me the first graduate development economics class I took at Harvard, which introduced me to the research topics that have since fascinated me. Over the years, she has not only been supportive of my IO-style approach to development research but also kindly helped cultivate these interests. Gautam has been a very generous advisor and also an enthusiastic supporter of my research approach. I am very grateful for his unwavering willingness to share his time and expertise in both research and navigating the profession. Lastly, I would like to express my gratitude to Ed Glaeser, who has been a constant advocate for my research agenda. His generosity with his time and advice, as well as his ability to connect ideas across various subfields, has provided me with fresh perspectives and enriched my research experience.

This PhD journey would have been dull, lifeless, and a lot more tedious without the help of my friends in the department. Robbie Minton, Ashesh Rambachan, Giorgio Saponaro, and Roman Sigalov have been by my side, especially this past year, going through both highs and lows with me while patiently enduring my constant complaints about research setbacks and more. I am also thankful for friendships and shared experiences with Martin Aragoneses, Andrea Hamaui, François-Xavier Ladant, Daniel Ramos, Ricardo Rodríguez-Padilla, and Yin Wei Soon, who helped make the journey enjoyable through countless interactions, lunches, and after-hours activities. Lastly, I want to express my heartfelt gratitude to my friend and collaborator, Shresth Garg. Without his intellectual partnership, finishing this thesis would have been much more difficult and far less enjoyable.

I would also like to express my gratitude to those who shaped my academic journey prior to Harvard. Chetan Sanghvi, who hired me at NERA, my immediate supervisors Gabriella Monahova and Ilton Soares, and my undergraduate economics advisor Jeff Baldani all played crucial roles in nurturing my interest in economics and supporting my efforts to pursue a PhD program.

This PhD thesis represents the culmination of a series of milestones supported and enabled by my family. I thank my grandparents, uncles, and aunts who have enthusiastically checked in on me and encouraged me on this journey. Their support and interest in my work over the years has been a strong motivating force. I especially want to thank my parents, Sunil and Sarojini, who have made countless sacrifices in life to raise me and bring me to this stage of my career. I am also thankful to my sister Rashi and my brother-in-law Srikanth for their constant support and words of encouragement during stressful times. Finally, I am indebted to my partner LJ, who has been my go-to person for both emotional and intellectual support during the PhD. Her unconditional support and radiant positivity have been the guiding light that carried me through the most challenging and trying moments. I am incredibly lucky to have found her and to have her by my side as I embark on the next chapter of my career.

To my parents Sunil and Sarojini

Introduction

This thesis examines the effects of various state interventions on market outcomes in developing economies, with a focus on India. Through three distinct chapters, I analyze the distributional effects of large-scale agricultural programs, the costs of industrial policy interventions in the solar sector, and the equilibrium effects of public healthcare provision strategies. Each essay develops and estimates a context-specific structural model of the relevant market, that is then used to evaluate counterfactuals and generate policy insights.

In the first essay, co-authored with Shresth Garg, we investigate the distributional consequences of three government programs in Indian agricultural markets: fertilizer subsidies, procurement of crops at minimum support prices (MSP), and sale of subsidized grains to households. To examine their effects, we estimate a structural model of supply and demand with heterogeneous risk-averse producers, who choose a portfolio of crops and crop-specific inputs, and heterogeneous households who make consumption decisions. Using the estimated structural parameters, we solve for counterfactual equilibria in which these interventions are phased out. On the demand-side, we find these programs to be progressive. In their absence, consumption and expenditures of lower-income households would be affected more adversely. On the supply-side, we find these programs to be (weakly) regressive. Higher fertilizer prices, in the absence of subsidies, would be compensated by higher output prices so impact on farmer welfare would be minimal. Under no government-procurement at MSP, richer farmers would experience a greater welfare loss, while some of the poorest farmers would gain – a result driven partly by the inequitable implementation of the procurement program.

In the second essay, co-authored with Shresth Garg, we examine the economic costs of two types of industrial policy interventions aimed at supporting domestic producers: (1) import tariffs and (2) production subsidies. We focus on the Indian utility-scale solar sector, where the Indian government has employed both strategies to support the domestic solar module manufacturing industry. To fully account for their costs, we develop a structural model that captures the interdependence between the upstream solar panel/module industry and the downstream solar power plant industry. Through this empirical model, we quantify the short-run costs of expanding the size and share of domestic solar module manufacturing using tariffs and subsidies. Our results indicate that import tariffs lead to considerably higher welfare losses compared to production subsidies. Specifically, in our main counterfactual, we find that expanding domestic output using solely import tariffs results in a welfare loss that is 60 times higher than the welfare loss under production subsidies.

In the final essay, co-authored with Ljubica Ristovska, we examine the aggregate and distributional effects of two popular public healthcare provision strategies in developing countries: (1) directly providing care through tax-funded public hospitals and (2) subsidizing care for low-income patients at private hospitals. Our main results show that increasing government hospitals benefits all consumers along the income distribution, with larger welfare gains for lower-income patients. However, when patients with income below a specified threshold receive a subsidy for private hospital care, the gains for eligible patients are relatively modest compared to the indirect negative impact on non-eligible patients, as private hospitals raise prices in equilibrium. By comparing the cost-to-benefit ratios of both policies, policymakers can determine the optimal approach to improve health outcomes for targeted income groups.

Chapter 1

Distributional Effects of Indian Agricultural Interventions[1]

1.1 Introduction

Government programs that distort prices in agricultural markets, such as input subsidies and price supports, are ubiquitous.[2] A key objective of such programs is redistribution,[3] which leads one to ask: *how do these price interventions affect market participants along the income distribution?* Yet, assessing these distributional effects is difficult. These programs are typically executed at a large scale and their equilibrium impact may amplify or dampen any direct effects on market participants. For instance, by lowering costs, input subsidies may positively affect farm profits but the equilibrium increase in aggregate output may decrease output prices enough to hurt profits.[4]

---

[1]Co-authored with Shresth Garg

[2]Input subsidies lower the costs of farm inputs such as fertilizers and seeds, while price supports typically serve as price floors at which farmers may sell output to government agencies. All 54 countries studied in OECD (2022), including the 38 OECD countries, have programs which provide support to the agriculture sector. The study excludes African nations; for an overview of similar programs in Africa, see Holden (2019).

[3]See Acemoglu and Robinson (2001); OECD (2012).

[4]Large programs, in a variety of contexts, often generate equilibrium effects. Examples of papers studying such effects include Duggan and Morton (2006); Imbert and Papp (2015); Cunha, De Giorgi and Jayachandran

In this paper, we propose a structural model that enables us to examine large-scale agricultural interventions while accounting for spillover and equilibrium effects. We use our model to study the distributional effects of multiple price interventions along the agricultural supply chain in India. At the start of this supply chain, the government sells subsidized fertilizers to farmers. Next, upon harvest, the government buys a substantial share of the total output of key crops such as rice and wheat at prices known as *minimum support prices* (MSP); all other sales are made at market prices to private traders. Finally, the government sells subsidized foodgrains to households, subject to progressive income-based quotas, through the *public distribution system* (PDS). Jointly, these programs cost about 1.2% of India's GDP and impact nearly 800 million people.[5]

Our structural model closely follows the setup of the Indian agriculture sector. In our model, risk-averse farmers choose a portfolio of crops to plant and make crop-specific input allocations. Post-harvest, they sell output either to government agencies or to private traders. On the demand-side, households receive PDS entitlements from the government and make consumption decisions in the private market. In equilibrium, total PDS entitlements equal the sales made to government agencies and total household demand in the private market equals the sales made to private traders. We estimate model parameters by matching simulated moments with empirical moments from publicly-available farmer- and household-level microdata. Finally, using estimated parameters, we simulate counterfactuals in which we phase out (1) fertilizer subsidies, and (2) government procurement at MSP and PDS entitlements.

We find these interventions to be progressive on the demand-side, and (weakly) regressive on the supply-side; in their absence, lower-income consumers and higher-income producers are affected more adversely. By raising output (and procurement) and lowering prices, these programs greatly benefit lower-income consumers who we find to be more price

(2019); Egger, Haushofer, Miguel, Niehaus and Walker (2019); Rotemberg (2019); Breza and Kinnan (2021); Muralidharan, Niehaus and Sukhtankar (2022); Khanna (2022).

[5]See World Bank (2019).

elastic and more reliant on PDS entitlements. For producers, we find that direct gains from fertilizer subsidies are nearly offset by equilibrium changes in market prices, so any impact on farmer welfare is minimal. In contrast, large-farmer bias in government-procurement at MSP accords greater benefits of this program to wealthier farmers, thus making it regressive. In an additional counterfactual where within-region bias for larger farmers is eliminated, we find substantial gains for smaller farmers.

Below, we summarize the main sections of this paper in more detail.

Motivating evidence. In Section 1.2, we present a mix of causal and descriptive evidence that motivates our model. First, using a natural experiment wherein subsidies for non-urea fertilizers were partially phased out, we show that subsidies affect production decisions. We also provide suggestive evidence that minimum support prices (MSP) influence production decisions. Correspondingly, in our model, we allow farmers' planting decisions to be determined by these programs. Next, we provide descriptive evidence to show that a large share of farmers sell at prices well below the MSP, and that there are stark income and spatial inequities in sales made to government buyers – larger farmers and farmers located in some regions are more likely to sell to government buyers.[6] Importantly, MSP appears to have an impact on production decisions only when sales to government buyers are likely. We add these findings to our model by allowing the likelihood of encountering a government buyer to depend on crop, location, and farmer size; this likelihood, in turn, determines how MSP affects farming decisions.

Model. In Section 1.3, we develop a model of multiproduct producers with endogenous product and input choice.[7] Farmers differ by productivity, location, and wealth (proxied by farm size). Given fertilizer subsidies and minimum support prices, they choose a set of crops

[6]Regional differences may be due to two reasons. First, procurement is still heavily reliant on infrastructure set up in the 1960s when only a few states produced surplus rice and wheat that could be procured. Second, in recent years, some states have introduced independent procurement schemes which only benefit farmers located in those states. There's little systematic evidence to explain the bias in favor of larger farmers. Conversations with local researchers suggest that corruption and bribery may explain part of this bias.

[7]Similar in spirit to Wollmann (2018) which considers a setting with oligopolistic producers.

to plant and make crop-specific area and input allocations subject to a farm size constraint. Farmers make these choices to maximize a mean-variance payoff function mediated by a farmer-specific risk aversion parameter. Finally, farmers also pay a fixed cost for the set of crops they plant.

At the time of planting, farmers face both output and price risk. Output risk shows up in crop-specific production functions in the form of idiosyncratic output shocks which scale output in a Hicks-neutral sense. Consequently, higher input usage yields greater output variance – a force which leads risk-averse farmers to moderate demand for inputs.[8] Importantly, input subsidies help offset this force and promote greater input usage.

Price risk arises from uncertainty over the price offered by private buyers and the uncertainty in accessing government buyers. Private buyer offers are distributed around an average private market price for each crop; the realized offer depends on post-harvest realization of an output quality shock (e.g. dust and moisture content).[9] Upon harvest, farmers encounter government buyers with a likelihood that depends on farmer size, location, and crop while a private buyer is always accessible. If a government buyer is present, the farmer sells to the government buyer if the minimum support price (MSP) is greater than the private buyer offer; else sales are made to the private buyer. Thus, for farmers who are more likely to find a government buyer, MSP provides greater insurance against downside price risk.

We allow both risk and risk aversion to differ by farmer. At the time of planting, these differences induce different choices on both the extensive and intensive margins: farmers may choose different bundles of crops and, even for the same bundle, may allocate different

[8]Presence of output risk combined with a lack of risk-mitigating technologies is a known source of underinvestment in farm inputs. See Rosenzweig and Binswanger (1993); Rosenzweig and Wolpin (1993); Mobarak and Rosenzweig (2013); Karlan, Osei, Osei-Akoto and Udry (2014); Cole, Giné and Vickery (2017); Donovan (2021)

[9]In addition to output quality, other reasons such as transportation costs, storage costs, and intermediary market power/bargaining power may also explain the cross-sectional variance in prices in the private market but we do not model these. We assume that quality shocks are the only source of variance in private buyer offers; further, these only affect processing costs – lower quality crops have higher processing costs and therefore receive a lower private buyer offer. Finally, processed crops purchased by households in the private market do not differ in quality.

shares of their land to each crop. This risk channel, therefore, is an important determinant of how supply-side price interventions affect aggregate production and individual farmer welfare.

On the demand-side, households differ by income and entitlements from the public distribution system (PDS).[10] Quantity procured by the government is redistributed to households.[11] Residual demand, which depends on both income and PDS entitlements, is satisfied in the private market where households pay the average private market price, determined in equilibrium.

Estimation. We estimate the supply-side of our model in three steps. We rely primarily on publicly-available data from Cost of Cultivation Surveys (CCS) from 2008-2016, conducted each planting season by the Department of Agriculture in India. These include detailed information on prices, crop portfolio, and crop-specific input allocations for each farmer-season.

We begin by estimating parameters governing the distribution of price risk. Two challenges arise. First, private market prices below MSP are only observed if a government buyer is not found. Second, whether a government buyer is found is unobserved. While we can construct the likelihood of *selling* to government buyers – by crop, region, and farmer size – from administrative datasets,[12] this is not equal to the likelihood of *finding* a government buyer since farmers may sell to a private buyer if his offer is greater than MSP. We proceed with the help of a simulation-based estimator, described in detail in Section 1.4, which yields parameters that determine the likelihood of finding a government buyer and the distribution of private buyer offers. These allow us to assess the ex ante crop-, location- and farmer size-specific price risk faced by farmers.

[10]In addition to income, PDS entitlements may depend on household location. In our counterfactuals, we hold the targeting of the PDS system fixed.

[11]We treat these as in-kind transfers at zero cost to households.

[12]Data on the likelihood of selling to government buyers by farmer size, crop, and location are from the 77th round of the National Sample Survey (NSS) conducted in 2019.

Next, we estimate crop-specific production functions and the distribution of risk aversion which may depend on farmer size. Given a set of crops, these affect how farmers allocate land, labor, capital, and fertilizer to each crop in the set. Observed input choices and output are also influenced by unobserved farmer productivity (which we account for using farmer fixed effects) and the distribution of output shocks.[13] We estimate parameters using method of simulated moments: for each farmer, we solve the optimal portfolio choice problem for the observed set of crops planted and match simulated choices with observed moments that summarize crop-specific output, land share patterns, and input usage.

Finally, we estimate the fixed cost of planting. The fixed cost for a crop is independent of the level of area allocated to that crop and depends on whether the crop is a *new* crop for a farmer. If it was part of his portfolio in the previous year, this cost is discounted by a parameter we estimate. The choice of which set of crops to plant is akin to a discrete choice problem where the choice set is composed of sets of crops. Once distribution of prices, production function parameters, output risk, and risk-aversion are known, given some guess of fixed cost parameters, we can simulate choice probabilities for each set of crops. We match these with the probability of observing a given set of crops in the data to estimate fixed cost parameters.

Once supply-side parameters are known, we can compute, for each farmer, the optimal set of crops and crop-specific inputs for any given input and (distribution of) output prices. This allows us to trace out aggregate supply curves for the private market and the government stockpile ("PDS supply") as a function of private market prices.[14] To pin down equilibrium private market prices, we require aggregate demand curves for the private

---

[13]Note that standard production function estimation techniques fail given risk-averse farmers. These usually rely on a monotonicity assumption between productivity shocks and input demand. Productivity shocks, however, increase variance of output. For a risk-averse farmer, this yields a non-monotonic relationship between productivity and input demand; depending on the size of a positive shock, the farmer may choose to increase or decrease input demand.

[14]We abstract away from modeling how MSP is set. Motivated by data, we assume that MSP tracks prices in the private market. Specifically, we assume a level of MSP such that conditional on finding a government buyer, only 35% of farmers would sell to private buyers i.e. MSP is set at the 65th percentile of the private buyer offer distribution.

market. For PDS crops, rice and wheat, we estimate demand using household-level data, accounting for PDS entitlements and income.[15] This yields an aggregate private market demand curve for each level of government stockpile (or PDS entitlements). For non-PDS crops, we use demand elasticities from Deaton (1997).

Main results. In Section 1.5, we evaluate the distributional effects of fertilizer subsidies, government procurement at minimum support prices (MSP), and in-kind transfers through the public distribution system (PDS) using two counterfactuals which shut down these interventions.[16]

In the first counterfactual, we phase-out fertilizer subsidies. In the data, we approximate an average subsidy rate of 50% for all fertilizer products. As such, in our counterfactual, we double fertilizer prices and solve for a vector of private market prices which clear all markets. We find equilibrium output of all crops falls and private market prices go up ($\approx$ 5% for rice and wheat). For farmers, the impact of higher fertilizer prices is dampened by higher equilibrium private market prices. Therefore, we find a minimal impact on farmer welfare across the size distribution.

In the second counterfactual, we shut down government-procurement at MSP. Correspondingly, PDS entitlements of households also go to zero.[17] Now, farmers can only sell in the private market and households must satisfy all demand in the private market. Thus, both supply and demand in the private market go up. But farmers are also exposed to greater price risk now, especially those who were previously more likely to find government

[15]Data are from the 68th round of the National Sample Survey (2011). To deal with potential endogeneity of prices in our estimation, we construct Hausman-style price instruments (Hausman, Leonard and Zona (1994)) by computing state-level average prices excluding the district in which each household resides. These are valid instruments under the assumption of idiosyncratic district-level demand shifters but correlated state-level supply shifters (e.g. processing costs).

[16]While useful for evaluating these programs, these counterfactuals also help us understand the impact of some of the proposed reforms to these programs in public and political debates. Both fertilizer subsidy program and government procurement at MSP are highly contentious topics with many in favor of shrinking their scale and scope.

[17]This is an assumption. We can consider, for instance, a counterfactual where the government only shuts down procurement and offers consumption vouchers to households that can be redeemed in the private market.

buyers. For these farmers, MSP was a meaningful price floor that protected against low private buyer offers. We find that in the absence of procurement, price of rice goes up ($\approx 5\%$) while there is a negligible change in the private market price of wheat. This differential response is driven by a greater fall in the total output of rice as some farmers switch to other crops due to greater estimated variance in private buyer offers for rice. These switchers are also large (they were previously more likely to find government buyers) and therefore have a noticeable impact on the aggregate output of rice. Despite switching, we find larger farmers have significantly larger welfare losses. Some of the smallest farmers experience modest gains as they now receive higher private market prices for rice.[18]

On the demand-side, in both counterfactuals, lower income households are more adversely affected as prices go up and PDS entitlements fall. These effects are driven by higher estimated price elasticities and higher observed PDS entitlements for lower-income households. Using a Laspeyres index, we show that without fertilizer subsidies, the lowest income households pay 3%-4% more to consume the old bundle of rice, wheat, and a numeraire good. In contrast, without government procurement at MSP, households pay 15%-20% more, which highlights the value of in-kind transfers. The impact on highest income households, in both scenarios, is close to zero.

Three implications for policy emerge. First, the incidence of large-scale government programs may depend crucially on the equilibrium channel. In our setting, fertilizer subsidies lower input costs for farmers but ultimately benefit only low-income consumers through lower output prices.

Second, in settings with multiple programs, a joint evaluation may be necessary to understand potentially important interactions. For example, we find that fertilizer subsidies not only lower output prices but also help increase government procurement by raising aggregate output. This highlights important complementarities in the two programs for the objective of improving food security of lower-income households.

[18]This result is partly driven by our assumption of perfect passthrough of output prices to farmers. The presence of intermediary market power would dampen this feedback effect.

Finally, implementation matters. Unequal access to government procurement at minimum support prices directly disadvantages small farmers. But it also indirectly hurts small farmers as greater procurement raises PDS entitlements, lowers private market demand and therefore, lowers prices in the private market. A more equitable program could, for example, match government buyers with farmers at random. We test the equilibrium impact of this alternative by eliminating within-region bias for larger farmers and find positive gains for smaller farmers.[19]

Related literature. Our work relates to a growing literature that uses structural models to study the agriculture sector in developing countries (Costinot, Donaldson and Smith (2016); Sotelo (2020); Allen and Atkin (2022); Bergquist, Faber, Fally, Hoelzein, Miguel and Rodriguez-Clare (2022); Chatterjee (2022); Hsiao (2022)). We add to this literature by introducing a general simulation-based approach that integrates observational microdata on farmer- and household-level decisions in the estimation of structural models. Second, our paper relates to a large body of work that studies subsidies and transfers in agricultural markets, both on the supply-side (Duflo, Kremer and Robinson (2008, 2011); Karlan, Osei, Osei-Akoto and Udry (2014)) as well as the demand-side (Banerjee, Hanna, Kyle, Olken and Sumarto (2018, 2019); Gadenne (2020); Gadenne, Norris, Singhal and Sukhtankar (2022)). We contribute to this literature by jointly studying interventions that directly affect both producers and consumers. Finally, we provide new and timely empirical evidence on the impact of the largest agricultural interventions in India that are actively being discussed in public and political debates (Meenakshi and Banerji (2005); Krishnaswamy (2019); Gupta, Khera and Narayanan (2021); Chatterjee, Kapur, Sekhsaria and Subramanian (2022)).

[19]We do so by using the share of farmers in each region who found government buyers in the baseline as the uniform likelihood of finding government buyers in that region. This does not hold fixed the procurement in baseline since smaller farmers would be relatively more likely to find government buyers now. Finding a probability that holds procurement fixed is computationally non-trivial since it also affects farmers' production decisions.

1.2 Institutional Details: The Indian Agriculture Sector

The agriculture sector in India directly impacts the well-being and survival of over a billion people. On the supply side, nearly 300 million people rely on it for their livelihoods.[20] These agricultural households generally own small plots of farm land – average farm size in India is 2.8 acres compared to 445 acres in the United States (USDA, 2021) – and have lower incomes (proxied by consumption expenditures in Figure 1.1a) relative to non-agricultural households. On the consumption-side, the agriculture sector supports a population of 1.4 billion, over 200 million of which are undernourished.[21]

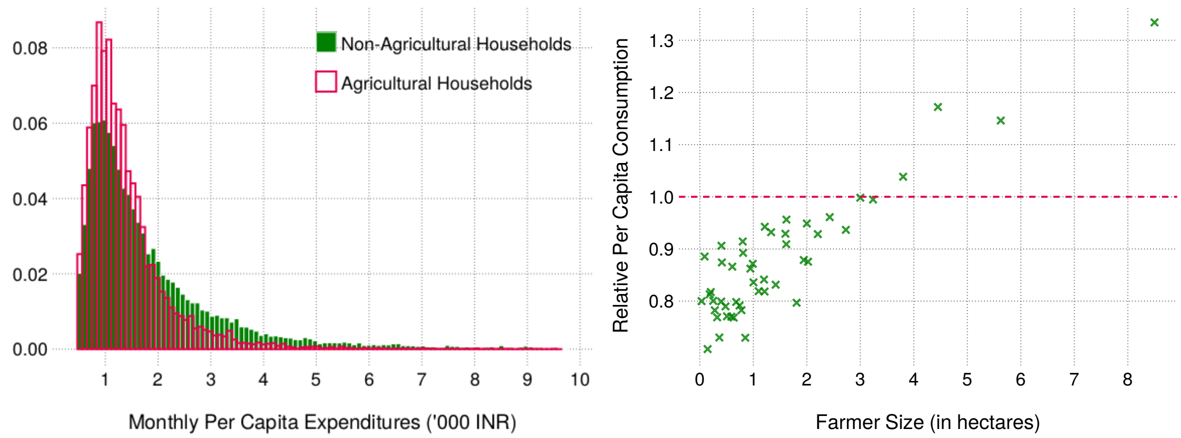
Against this backdrop, several government-sponsored programs exist to support agricultural households and bolster food security in the country.[22] In this paper, we focus on three of the largest and longest-running such programs. These include fertilizer subsidies for agricultural use, government procurement of staple crops at pre-announced minimum support prices (MSP), and the redistribution of these crops at highly subsidized rates to low income households through the public distribution system (PDS).

These interventions date back to at least the mid-1960s. Newly-independent India faced severe food shortages, exacerbated by two successive drought years, and relied heavily on imports and foreign food aid to feed its rapidly growing population. To encourage greater production of foodgrains, the government started supplying farmers with high-yield variety seeds and heavily-subsidized fertilizers. In addition, the government promised attractive purchase prices for staples such as rice and wheat. These policies marked the beginning of the Green Revolution of the 1960s in India, during which yields increased many-fold and

[20]Estimated as the weighted sum of number of family members in households which report farming as their principal source of income in the 68th round of the National Samply Survey (2013); excludes agricultural labor.

[21]See FAO, IFAD, UNICEF, WFP and WHO (2021). Despite tremendous gains in agricultural production in the last few decades, malnutrition remains an issue. In the 2021 Global Hunger Index, India ranks 101st out of 116 countries. Rankings depend on the prevalence of undernourishment, childhood wasting, childhood stunting, and child mortality.

[22]Other programs, in addition to those studied in this paper, include subsidized crop insurance under Pradhan Mantri Fasal Bima Yojana (PMFBY), minimum income support for small and marginal farmers under Pradhan Mantri Kisan Samman Nidhi Yojana (PM-Kisan Yojana) launched in 2018, pension scheme for small and marginal farmers under Pradhan Mantri Kisan Maan-Dhan Yojana (PM-KMY) launched in 2019 etc.



(a) *Per Capita Consumption By Household Type*

(b) *Consumption of Agricultural Households Relative to Median Household*

Figure 1.1: *Consumers and Producers in the Indian Agriculture Sector*

Notes. The left panel shows histograms of per-capita monthly expenditures of households who are identified as self-employed in agriculture against per-capita monthly expenditures of all other households in the 68th round of the NSS (2011-12). These values include home production valued at market prices. The plots only include households with reported per-capita expenditures greater than the 1st and lower than the 99th percentiles. In the right panel, we show median relative consumption of agricultural households binned by land size; relative consumption is defined as the ratio of per capita monthly expenditures and the median per capita expenditures in the data.

domestic production increased enough to allow India to become self-sufficient in food. Six decades after their introduction, these policies remain in place and make up a large share of the total budget of the central government – between 2010 and 2019, they amounted to 10% of annual government spending on average.[23]

Before discussing these interventions in more detail, we briefly describe the datasets used in our study.

1.2.1 Data

We bring together several publicly-available administrative datasets for our analysis.

To estimate the supply side of our model, we require detailed farmer-level micro-data on

[23]Central government spending on these programs went up in 2020-2021 due to COVID-19, and again in 2022 after the Russian invasion of Ukraine. In Figure A.1, we present these annual budget shares over time.

planting decisions including crop and input choice. We obtain these data from three rounds (2008-09 to 2016-17) of Cost of Cultivation Surveys (CCS) conducted by the Department of Agriculture in India. In each round, a sample of farmers is followed for all planting seasons for three consecutive years, and plot-level data on output and input usage are recorded. In particular, input data are recorded very well. For a given farmer-year-season, we observe not only input expenditures but also physical quantities (e.g. hours of labor/machinery) of all inputs, logged separately for each crop grown in the season. We provide additional details in Section A.2.1.

While CCS surveys include information on realized output prices, the identity of the buyer is unknown. These data are critical for understanding which farmers have access to government buyers and are able to avail MSP. To get at the identity of buyers, we rely on the 77th round of the National Sample Survey (NSS) conducted in 2019 which surveyed a nationally representative sample of agricultural households in India.[24] These survey data allow us to map farmer size, region, and crop to the likelihood of making sales to a government agency.

On the household-side, we leverage a nationally-representative consumer expenditure survey conducted from July 2011 to June 2012 as part of the 68th round of the NSS. Relevant variables include household size and income, and quantities and values of rice and wheat purchased. Household purchases of these crops are broken down by source, so for each household, we observe the share of consumption that comes from PDS shops.

In addition to the above datasets, we rely on two sources of aggregate agricultural data. These include the ICRISAT District Level Database (DLD) from 1966-2016, which provide annual district-level statistics on cropping patterns, fertilizer consumption, and output prices. We use these data in our reduced-form analyses of the impact of fertilizer subsidies and minimum support prices on production decisions. Second, we use an agricultural census of all farm holdings conducted in 2016 to construct a nationally representative farm size distribution by crop.

[24]The NSS is a nationally-representative repeated cross-sectional survey.

1.2.2 Fertilizer Subsidies

Prior to economic liberalization in India in 1991, the government controlled the prices of all fertilizer products in India. It set the price at which it procured fertilizers from fertilizer producers and importers, and it set the price at which fertilizer products were sold to farmers; the difference between these prices was borne by the taxpayer.[25] All fertilizer products continue to be subsidized, but over the years, the government has taken steps to progressively decontrol non-urea fertilizers, in 1991 and then again in 2010.[26] In contrast, the price of urea, the most popular fertilizer product in India, continues to be tightly controlled and set directly by the government.[27]

Do fertilizer subsidies affect production decisions? To study whether farmer behavior responds to these subsidies, we rely on a natural experiment. Starting 2010, subsidies for non-urea fertilizers, which are the only source of nutrients phosphorus (P) and potassium (K), were partially phased-out. In Figure 1.2a, we show that this resulted in a rapid increase in the price of fertilizer nutrients P and K, relative to nitrogen (N), as reported in the Cost of Cultivation Surveys.[28] Correspondingly, we find a decline in district-level consumption of nutrients P and K as shown in Figure 1.2b: this plot shows coefficients from a regression of (log) consumption on district fixed effects and year dummies (excluding 2009) estimated using ICRISAT District Level Database.

Next, to test how this partial phase-out of subsidies for nutrients P and K affected output, we construct a district-level measure of treatment intensity which captures the intensity

[25] Fertilizers were procured from producers under the Retention Price Scheme; producer prices were specific to production plant and based on plant-specific costs of production.

[26] Though non-urea fertilizers are decontrolled, non-urea fertilizer producers still receive production subsidies; however, the producers now have more control over the sale price of their products.

[27] While prices paid to producers are not publicly available, we can estimate subsidy rates based on prices paid by farmers, total consumption of fertilizers in the country, and the total fiscal costs of fertilizer subsidies. For example, in 2019, the government spent USD 232 per tonne of urea, and set the controlled price at USD 76 per tonne, which amounts to a subsidy of 75% on the price of urea.

[28] Urea only contains nitrogen (N), but non-urea fertilizers may also contain some nitrogen which might be why we see a small spike in the price of N as well.

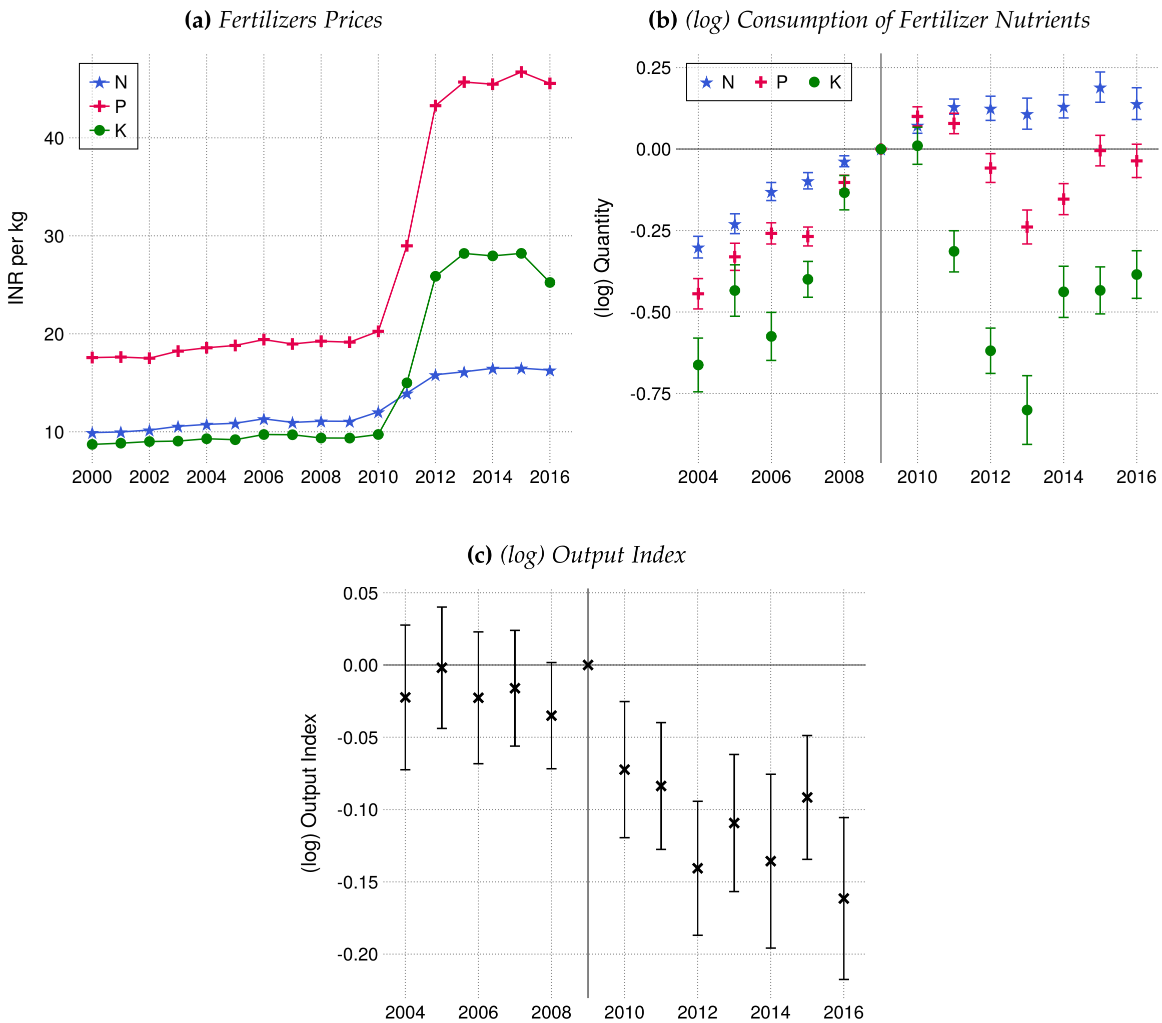


Figure 1.2: *Impact of Fertilizer Subsidies on Production Decisions and Output*

Notes. In the top-left panel, we plot (weighted) average reported prices of fertilizer nutrients N, P, and K in the Cost of Cultivation Surveys. In the top-right panel we show estimated coefficients from an event-study regression using district-level ICRISAT panel data. The dependent variable is (log) reported consumption of fertilizer nutrients (N, P, or K) at the district-level. The controls are year dummies (excluding 2009) and district fixed effects. In the bottom panel, we plot the estimated coefficients from a difference-in-differences specification with a continuous treatment variable using district-level ICRISAT panel data. Treatment intensity is defined as the per-unit area consumption of fertilizer nutrients P and K (aggregated using prices as weights) in the period 2004-2009, before prices of these nutrients increased sharply. The dependent variable is (log) output index at the district-level, constructed using output of all crops grown in that district aggregated using national median prices of those crops in the period 2004-2009. The controls are year and district fixed effects.

with which these nutrients were used in each district prior to 2010. We use this measure of usage intensity to run the following (continuous) difference-in-differences specification

$$\log Y_{dt} = \beta_0 + \sum_{k \neq 2009} \beta_k \log \text{Avg. Usage Intensity}_d \cdot \mathbb{1}\{k = t\} + \phi_d + \gamma_t + \epsilon_{dt},$$

where ϕ_d and γ_t are district and year fixed effects.[29] Our main outcome of interest is a district-level (price-weighted) output index, which captures the value of agricultural output in each year.[30]

The estimated coefficients, shown in Figure 1.2c, suggest that districts where nutrients P and K were used more intensively experienced a greater decline in output post-2010 when prices of these fertilizer nutrients increased sharply. We take findings from this natural experiment as strong evidence that fertilizer subsidies not only affect farmers' fertilizer usage decisions but also have an impact on final output.

1.2.3 Procurement at Minimum Support Prices (MSP) & Redistribution Through the Public Distribution System (PDS)

India has two main planting seasons for crops – kharif (monsoon) and rabi (winter). At the start of each planting season, the government announces minimum support prices (MSP); these are prices farmers can expect to receive at the time of harvest if sales are made to government agencies. These prices are based on government-administered surveys known as Cost of Cultivation Surveys (CCS) designed to estimate the average costs of growing various crops in the country.[31] While minimum support prices are announced for almost all major crops in India, only rice and wheat are subject to substantial procurement by

[29]We construct the (price-weighted) average usage intensity of P and K in district d as

$$\text{Avg. Usage Intensity}_d = \frac{1}{6} \sum_{t=2004}^{2009} \frac{r_P^F F_{Pdt} + r_K^F F_{Kdt}}{\text{Total Area Planted}_{dt}}$$

where F_{Pdt} and F_{Kdt} are quantities consumed of nutrients P and K, respectively, while prices r_P^F and r_K^F are national median prices of the nutrients in the period 2004-2010.

[30]Note that the prices used to construct the output index are national-level median crop prices in the period 2004-2009, and only serve as weights to combine output of different crops

[31]While these surveys inform minimum support prices, the prices are also subject to political considerations.

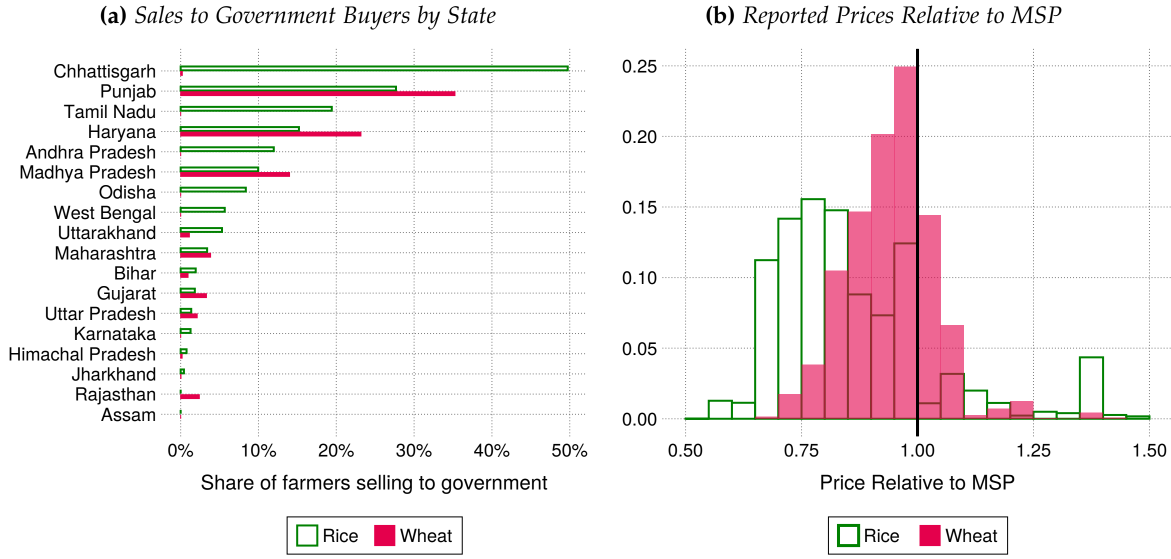


Figure 1.3: *Government-Procurement and Access to Minimum Support Prices (MSP)*

Notes. The left panel shows the share of rice (wheat) farmers in a state who made sales to government buyers. The right panel shows the distribution of prices received by farmers relative to the minimum support price (MSP) for that season. Source: 77th round of the National Sample Survey (NSS) conducted in 2019.

the central government at minimum support prices.[32] In addition, there is substantial geographic variation in how intensively government agencies procure these crops in a region. In Figure 1.3a, we plot, by state, the share of farmers growing rice and wheat that report selling their output to government buyers.

Importantly, minimum support prices are not a legal price floor. Upon harvest, when a farmer brings the output to a regional market, they may only encounter private traders who are free to make price offers below the MSP. As shown in Figure 1.3b, a large share of farmers report receiving prices below the MSP. At the same time, the likelihood of selling to government buyers and therefore availing MSP is strongly correlated with the size of a farmer. We show this with the help of NSS data where farmers report whether sales were made to government agencies. As shown in Figure 1.4a, we find that larger farmers –

[32] From 2011-2019, on average, the government procured over 30% of total annual output of rice and wheat in the country.

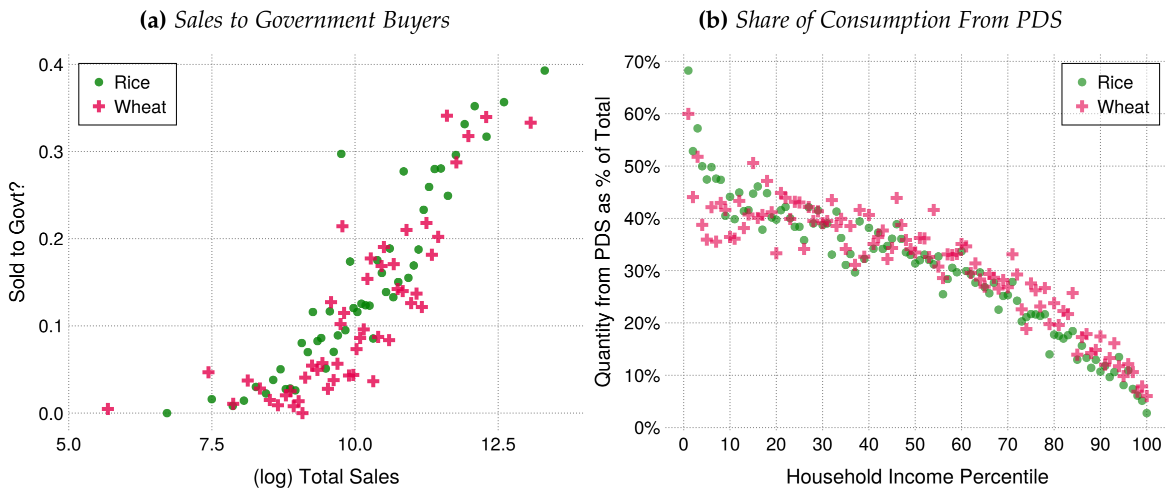


Figure 1.4: *Distributional Differences in the Impact of Government Interventions*

Notes. The left panel shows binned means of an indicator variable denoting whether sales were made to a government buyer against total sales made by the farmer, as reported in the 77th round of the NSS survey, conducted in 2019. The right panel shows binned means of the share of monthly consumption of rice and wheat obtained through PDS against total monthly expenses per capita as observed in the NSS Consumer Expenditure Survey, 2011.

proxied by total sales made – are more likely to sell their output to government buyers.[33] This relationship is robust to conditioning on farmer state.

The output procured by the government is fed into the public distribution system (PDS), which is a network of over half a million fair price or "ration" shops throughout the country where households can purchase staples rice and wheat at highly subsidized rates subject to income-based quotas.[34] Like fertilizer subsidies and government-procurement at MSP, the PDS has been in place since the 1960s, and is currently the largest food distribution program in the world (George and McKay 2019).[35]

The program assigns higher quotas to lower-income households. We confirm this in the data and also show, in Figure 1.4b, that lower-income households, proxied by total

[33]See Footnote 6 for a discussion of why these patterns emerge.

[34]Depending on the region, these shops may sell other commodities but rice and wheat are sold almost everywhere.

[35]About 70% of Indian households interact with the PDS (Gadenne, 2020); 800 million people receive subsidized grains through the system (World Bank, 2019).

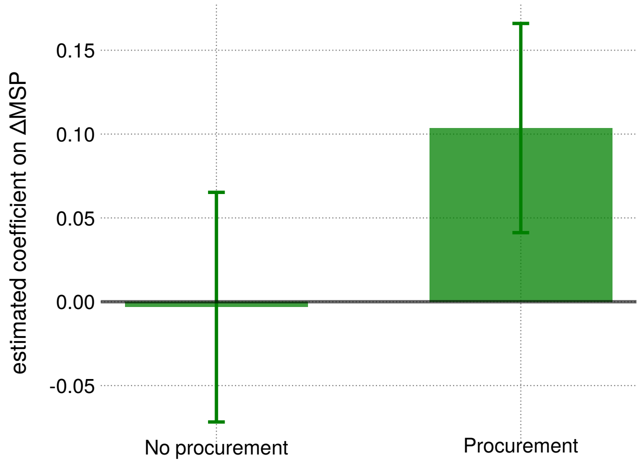


Figure 1.5: *Change in Share of Area Allocated to a Crop Responds to Minimum Support Prices If Government Buyers Active in the Region*

Notes. This figure plots estimated coefficients from a regression where the dependent variable is the change in share of area allocated to a given crop in a district in two consecutive years. The independent variable of interest is the change in (deflated) minimum support price for that crop expressed in hundreds of rupees. We interact this with an indicator variable for whether, in the previous year, the central government actively procured the given crop in the state in which the district is located. We also control for district $\times$ crop fixed effects. The data are from ICRISAT District Level Database (DLD) and reported at district $\times$ crop $\times$ year level. Only observations on rice and wheat are included in the regressions as these are the primary crops procured by central agencies. The plot shows the 95% confidence interval; standard errors are clustered at the district-level.

monthly expenditures, derive a larger share of their total consumption of rice and wheat from the PDS. This figure also highlights that PDS entitlements are inframarginal and that households across the income distribution rely on the private market for some share of their consumption.

Do minimum support prices affect production decisions? We provide evidence which suggests that farmers respond to higher MSP by increasing the share of area allocated to MSP crops, but only if the government actively procures in their state.

Specifically, let $X^1_{cs(d)t}$ be an indicator variable that equals one if in period $t-1$ the central government procured a nonzero quantity of crop c in state s of district d. Using the

ICRISAT District Level Database, we estimate the following regression

$$\Delta\text{Share Area}_{cdt} = \underbrace{\{a^0_{cd} + a^0_{MSP} \cdot \Delta MSP_{ct}\}}_{\text{no procurement}} \times (1 - X^1_{cs(d)t}) + \underbrace{\{a^1_{cd} + a^1_{MSP} \cdot \Delta MSP_{ct}\}}_{\text{procurement}} \times X^1_{cs(d)t} + u_{cdt}$$

where ΔShare Area$_{cdt}$ is the change in share of area allocated to crop c in district d relative to the previous year, while ΔMSP_{ct} is the change in minimum support price for the crop expressed in hundreds of rupees, deflated using a consumer price index. The intercepts a^0_{cd} and a^1_{cd} are crop $\times$ district fixed effects.

We plot coefficients a^0_{MSP} and a^1_{MSP} in Figure 1.5. Our estimates suggest that if the government procured a nonzero quantity of output of an MSP crop in a state, farmers in those states respond to a higher MSP for that crop by increasing the share of area allocated to it. Specifically, based on our estimate of a^1_{MSP}, increasing (deflated) MSP by 10 INR (mean = 15 INR) increases share of area allocated to that crop by 1 percentage point. In states with no government procurement in the previous year, change in MSP does not have a statistically significant effect on the change in cropping patterns.

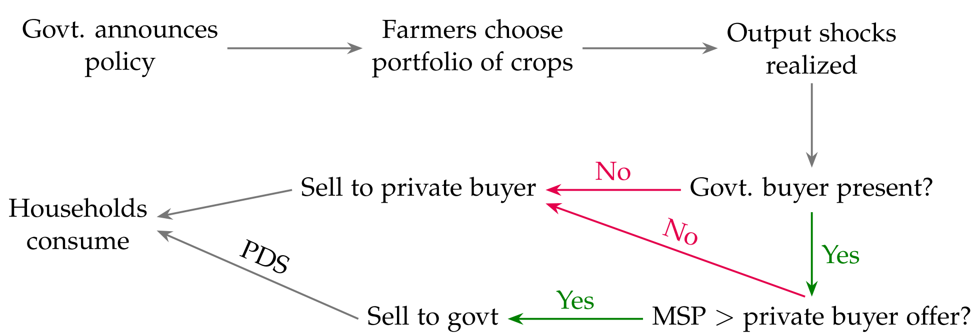
We incorporate these findings in our model, which we present in the next section.

1.3 Model

The structural model consists of a supply-side with farmers who make production decisions, and a demand-side with households who consume agricultural output.

1.3.1 Model Timeline

At the start of a planting season, the government announces fertilizer subsidies and crop-specific minimum support prices (MSP). Farmers observe these policy announcements and make planting decisions. After production decisions are made, idiosyncratic shocks are realized which affect output quantity as well as the price offer made by a private buyer. Farmers sell their output either to the government buyer or to the private buyer. Finally, households receive their PDS entitlements and make purchases from the private market.



Notes. This figure provides an overview of the model. Before planting decisions are made, the government announces fertilizer subsidies and minimum support prices. Farmers take these into account and make planting decisions. Upon harvest, output shocks are realized. Farmers bring their output to the market where a government buyer may be present. If the government buyer is present, the farmer sells his crop to the government buyer if MSP is greater than the price offered by the private buyer. Otherwise, sales are made to the private buyer. Quantity procured by the government is distributed to households through the public distribution system (PDS). Household satisfy residual demand in the private market.

Figure 1.6: *Model Timeline & Overview*

We summarize this in Figure 1.6.

1.3.2 Supply: Farmer's Problem

Planting decisions involve choices on both the extensive and intensive margins. Farmers choose which set of crops to plant and, for each crop in this set, they make crop-specific input allocations.

Farmer j in region r is endowed with a farm of total size A_j. In season t, he chooses a set s of crops to plant which maximize utility V_{jst}. This utility consists of two components: a mean-variance payoff U_{jst} and a fixed cost of planting κ_{jst}, and is expressed as

$$V_{jst} = U_{jst} - \kappa_{jst}, \quad \forall s \in \mathcal{S}_j$$

where $\mathcal{S}_j$ are all possible sets of crops farmer j can grow.

Farmers are risk averse with risk-aversion γ_j. Given a set of crops s, farmers choose how to allocate plots A_{jct} and inputs $\boldsymbol{X}_{jct}$ to each crop $c \in s$ in order to maximize the difference between expected total profits and risk-aversion weighted variance of total profits. The optimal allocation gives rise to the mean variance payoff U_{jst} defined as

$$U_{jst} = \max_{\{A_{jct}, \boldsymbol{X}_{jct}\}_{c\in s}} \mathbb{E}\left(\Pi_{jst}\right) - \gamma_j \operatorname{Var}\left(\Pi_{jst}\right)$$

$$\text{where } \Pi_{jst} = \sum_{c\in s} \pi_{jct}(A_{jct}, \boldsymbol{X}_{jct}) \text{ s.t. } \sum_{c\in s} A_{jct} \leq A_j \tag{1.1}$$

This optimization problem is only subject to an area constraint which requires that the sum of crop-specific area allocations is (weakly) less than the total land endowment A_j.[36]

Total profits Π_{jst} are the sum of crop-specific profits π_{jct} given by

$$\pi_{jct}(A_{jct}, \boldsymbol{X}_{jct}) = P_{jct} q_{jct}(A_{jct}, \boldsymbol{X}_{jct}) - \sum_{x\in \boldsymbol{X}_{jct}} w_t^x x$$

where P_{jct} is risky output price, q_{jct} is risky output, and w_t^x is the price of input x. One of the inputs is fertilizer F_{jct} and fertilizer subsidies directly affect the post-subsidy price of fertilizer w_t^F.

Next, we describe in detail (1) output, (2) output price, (3) risk aversion, and (4) fixed costs.

Risky Output

Output of crop c is given by $q_{jct}(A_{jct}, \boldsymbol{X}_{jct})$ which depends on plot area A_{jct} and inputs $\boldsymbol{X}_{jct}$. These inputs include labor L_{jct}, capital K_{jct}, and fertilizers F_{jct}. Further, output depends on a farmer-specific unobserved productivity term ω_j which is known at the time of planting. Output risk arises from an idiosyncratic output shock ε_{jct} that is realized at the time of

[36]We do not model credit constraints, which may also determine input choices, and assume that farmers can purchase any level of non-area inputs. In India, institutional credit for input purchases, through banks and cooperatives, might be readily available due to government support as agriculture is classified as a priority sector by the central bank, and banks and financial institutions are required to lend at least 18% of their credit to the agriculture sector. Further. all farmers are eligible for Kisan ("Farmer") Credit Cards that can be used to purchase inputs at low interest rates. In the 77th round of the NSS, we see that institutional lenders are responsible for 82% of agriculture loans, both by amount lent and number of loans extended. In Figure A.3 we show that the reported interest rates for farm loans are much lower than for consumption loans and these rates are similar across the farmer size distribution. Also see Karlan, Osei, Osei-Akoto and Udry (2014) which finds agricultural risk to be a more important determinant of production decisions than input credit constraints in northern Ghana.

harvest; only its mean and variance are known ex ante.[37]

We express this risky output as follows

$$
\begin{aligned}
q_{jct}(A_{jct}, \boldsymbol{X}_{jct}) &= q_c(A_{jct}, \boldsymbol{X}_{jct}) \exp\left\{\omega_j + \varepsilon_{jct}\right\} \\
&= A_{jct}^{\beta_{ac}} L_{jct}^{\beta_{lc}} (1 + K_{jct})^{\beta_{kc}} (1 + F_{jct})^{\beta_{fc}} \exp\left\{\omega_j + \varepsilon_{jct}\right\} \qquad (1.2)
\end{aligned}
$$

where $\boldsymbol{X}_{jct} = \{L_{jct}, K_{jct}, F_{jct}\}$. In estimation of crop-specific production functions $q_c(A_{jct}, \boldsymbol{X}_{jct})$, we also account for the impact of location and season using relevant fixed effects.

The unobserved productivity term, ω_j, captures the average fertility of farmer j's land as well as any technological know-how and ability. Importantly, it does not differ by crop. It does, however, affect input choices and therefore gives rise to the standard input endogeneity concern (Marschak and Andrews, 1944; Hoch, 1962; Griliches and Mairesse, 1995). We explain how we deal with it when we discuss estimation in the following section.

Risky Prices

Upon harvest, farmers bring their output to the market where they may or may not encounter a government buyer. A private buyer is always present. Government buyer offers to buy PDS crops, rice and wheat, at the pre-announced minimum support price (MSP). If a government buyer is found and MSP for a crop is greater than the price offered by the private buyer, farmer sells all output of that crop to the government; otherwise, the farmer accepts the private buyer's offer.[38]

What does the private buyer offer? The price offered by the private buyer depends on an idiosyncratic output *quality* shock that is realized post-harvest.[39] This output quality shock,

[37] Output quantity shocks are uncorrelated across farmers. Therefore, the model features no aggregate shocks. Aggregate shocks can be added to the model at the expense of significantly larger computation requirements. An important consequence of aggregate shocks would be the negative covariance between output and private market prices which might lower the total revenue risk faced by farmers (Allen and Atkin (2022)). In the absence of aggregate shocks, we may overestimate risk and therefore, underestimate risk aversion.

[38] Non-PDS crops are always sold to private buyers at the offered price.

[39] There are several alternative justifications for cross-sectional variance in private buyer offers. We discuss some of these in Footnote 9.

η_{jct}, captures factors such as dust and moisture content and only affects processing costs of the crop in the private market.[40] High quality crops have low processing costs and therefore receive higher private buyer offers. The expected price offer equals the equilibrium price in the private market, given by P_{ct}.[41] Now, we can express private buyer offers as

$$\widetilde{P}_{jct} = P_{ct} \cdot \exp\{\eta_{jct}\}$$

where $\eta_{jct} \sim N\left(-\frac{\sigma^2_{\eta c}}{2}, \sigma^2_{\eta c}\right)$. We assume that farmer j knows the distribution of private buyer offers. That is, he knows the distribution of quality shocks and the equilibrium private market prices for all crops at the time of planting.[42]

If a government buyer is present, the farmer only accepts private buyer offers if they are greater than MSP. Let $Z_{jct} = 1$ if farmer j encounters a govt buyer for crop c; $Z_{jct} = 0$ otherwise. Price received for crop c by farmer j is

$$P_{jct} = \mathbb{1}\left\{Z_{jct} = 1\right\} \max\left\{MSP_{ct}, \widetilde{P}_{jct}\right\} + \mathbb{1}\left\{Z_{jct} = 0\right\} \widetilde{P}_{jct} \tag{1.3}$$

Importantly, we assume that the farmer is uncertain about meeting a government buyer at the time of planting.[43] Motivated by data, the probability of finding a government buyer is a function of farmer size, crop, and location. Specifically,

$$\rho_{jct} = \Pr\big(Z_{jct} = 1\big) = \Phi(\alpha_{0rc} + \alpha_{1rc} \cdot \log A_j)$$

where $(\alpha_{0rc}, \alpha_{1rc})$ are crop- and region-specific coefficients, and A_j is the total area of farmer j. This probability, along with equation (1.3) and the distribution of private buyer offers,

[40] Processed crops that are purchased by households are homogeneous; there are no quality differences.

[41] Note that we do not model intermediary market power. The average price paid by private buyers to farmers equals the price households pay for private market purchases. Several studies document and analyze trader market power in agriculture in India (Meenakshi and Banerji (2005); Mitra, Mookherjee, Torero and Visaria (2018); Chatterjee (2022)) and elsewhere (Bergquist and Dinerstein (2020)). To check the robustness of our results, we plan to re-do our counterfactual analyses at different levels of calibrated passthrough.

[42] Knowing the equilibrium private market price requires solving a very complex problem. Alternatively, we can assume that farmers extrapolate equilibrium private market prices from the average prices in the previous year. This specification is easy to incorporate and we plan to add it as a robustness check.

[43] This can be relaxed and we can check the robustness of our results to this assumption; we plan to do this in the next iteration of estimation.

gives rise to a farmer size-, location-, and crop-specific mixture distribution which the farmer uses to compute mean-variance payoff given in equation (1.1). For non-zero ρ_{jct}, the mean of this distribution of prices is increasing in MSP.

Risk Aversion

Risk aversion is parameterized to be a function of total land holdings A_j (a proxy for wealth).

$$\ln \gamma_j = \gamma_0 + \gamma_A \ln A_j + \psi_j, \quad \psi_j \sim N(0, \sigma_\gamma^2).$$

where ψ_j is an idiosyncratic component of risk aversion.

Fixed Costs

Fixed costs depend on the set of crops planted and do not scale by area. These help us rationalize low crop diversification observed in the data: most farmers grow at most 3 crops in a season.

Let $s_{j,t-1}$ be the set of crops planted by farmer j in the same season but in the previous year. Fixed cost κ_{jst} of planting a set s of crops jointly is given by

$$\kappa_{jst} = \sum_{c \in s} \kappa_{jct} \text{ where } \kappa_{jct} = \begin{cases} \kappa_c & c \notin s_{j,t-1} \\ \lambda \cdot \kappa_c & c \in s_{j,t-1} \end{cases}$$

where κ_c is a constant crop-specific parameter, and λ is a discount on fixed costs for repeating crops. In our estimation, we allow λ to differ by staple crops (rice and wheat), and all other crops.

1.3.3 Demand

PDS crops: rice and wheat

Households differ by income and PDS entitlements.[44]

[44] In addition to household income, these entitlements may also depend on where the household is located and how easily it can access PDS shops. To hold targeting fixed, we calibrate the total share of government

Let q_{ch}^{PDS} denote the per-capita quantity of crop c received by household h through the PDS system.[45] Motivated by data, we assume this quantity is inframarginal to the total per-capita quantity of crop c consumed by the household, which is given by

$$q_{ch} = q_{ch}^{PVT} + q_{ch}^{PDS}$$

where q_{ch}^{PVT} is the per-capita quantity of crop c purchased in the private market. The total per-capita demand depends on the equilibrium price in the private market, P_c (time subscript t is suppressed). In addition, it also depends on per-capita income y_h. For tractability, we assume a log demand function given by

$$\log(1 + q_{ch}) = a_{cp} \log P_c + a_{cy} \log y_h + a_{cpy} \log P_c \cdot \log y_h + u_{ch} \tag{1.4}$$

which can be approximated using a utility function discussed in Section A.3.1. Importantly, this function is compatible with Engel's law (and our data) that higher income households spend a lower share of their income on food.

Non-PDS crops

For non-PDS crops, we consider an aggregate demand function given by

$$q_{ct} = \mu P_{ct}^{e_c} \quad \forall c \notin \{\text{rice, wheat}\} \tag{1.5}$$

where e_c is the price elasticity of demand for crop c.

1.3.4 Equilibrium

The total quantity procured by the government can be expressed as

$$Q_{ct}^{govt}(P_{ct}) = \sum_j \mathbb{E}\left[q_{jct}(P_{ct}) \cdot \mathbb{1}\left\{Z_{jct} = 1\right\} \cdot \mathbb{1}\left\{MSP_{ct} \geq P_{ct} \cdot \exp\{\eta_{jct}\}\right\}\right]$$

procurement that a household receives using the 68th round of the NSS which is a nationally-representative survey of households. We hold these calibrated PDS shares as fixed in our counterfactuals.

[45] We assume that the households do not pay for these entitlements. In reality, households may pay a small amount depending on their income level.

We assume that MSP is set to track the equilibrium price P_{ct} in the private market. In particular, motivated by our estimates, MSP is set at the 65th percentile of the private buyer offer distribution. This implies that conditional on finding a government buyer, 65% of farmers sell to the government. All other sales are made to private buyers. Total equilibrium quantity in the private market, therefore, is

$$Q_{ct}^{pvt}(P_{ct}) = \sum_j q_{jct}(P_{ct}) - Q_{ct}^{govt}(P_{ct})$$

Our notion of equilibrium is a vector of average private market prices which farmers and consumers take as given, and which clears all markets. More precisely, a static competitive equilibrium is a vector of private market prices, $\{P_{ct}\}_c$ such that

1. Government procurement equals sum of PDS entitlements received by households.

$$Q_{ct}^{govt}(P_{ct}) = \sum_h q_{cht}^{PDS} \quad \forall c$$

2. Total purchases by private buyers equals total private market demand for all crops.

$$Q_{ct}^{pvt}(P_{ct}) = \sum_h q_{cht}^{PVT}(y_h, P_{ct}, q_{cht}^{PDS}) \quad \forall c$$

3. Sum of government procurement and private buyer purchases equals total output.

$$Q_{ct}^{govt}(P_{ct}) + Q_{ct}^{pvt}(P_{ct}) = \sum_j q_{jct}(P_{ct})$$

1.4 Estimation

1.4.1 Supply

We estimate the supply-side of the model in three stages. First, we estimate the parameters governing the distribution of prices at the time of planting. Next, we estimate the production function and risk aversion parameters. Finally, we estimate the fixed costs. All stages rely on simulation-based estimators (Pakes, 1986; McFadden, 1989; Pakes and Pollard, 1989).

The Distribution of Output Prices

Three sets of parameters determine the distribution of prices at the time of planting. These are

1. the equilibrium (average) private market price, P_{ct} for all crops and years,
2. the variance of output quality shocks, $\sigma^2_{\eta c}$ for all crops, and
3. crop- and region-specific parameters governing the likelihood of finding a government buyer, $\{\alpha_{0rc}, \alpha_{1rc}\}$ the latter of which is the coefficient on farmer size.

Our analysis is complicated by the fact that private buyer offers are observed only if a government buyer is absent or if the offers are higher than the minimum support price (MSP).[46] To get an unbiased estimate of the mean and variance of the private buyer distribution, we need to condition on the presence of a government buyer. However, whether a government buyer is present is not known. Our data only includes information on realized sales.[47] Since farmers may choose to sell to private buyers even when government buyers are present, this measure is an imperfect proxy for the likelihood of finding government buyers.

We estimate these parameters as follows. Let $\theta_c = \{\{P_{ct}\}_t, \sigma^2_{\eta c}, \{\alpha_{0rc}, \alpha_{1rc}\}_r\}$ and θ^g_c be a guess of these parameters. For each θ^g_c, we can simulate whether a farmer found a government buyer given his location, size, and crop, for all farmers in the CCS data. We can also draw a private buyer offer given a guess of average private market price and the variance of output quality shocks. Simulated *realized* price is the private buyer offer if a government buyer is not found or if the private buyer offer is greater than the MSP. Otherwise, the simulated price equals the MSP for that crop. This generates a distribution of simulated prices that farmers receive.

[46] This is only an issue for PDS crops. For non-PDS crops, estimation is straightforward.

[47] In the 77th round of the NSS, agricultural households report whether sales were made to government agencies or to private buyers.

Table 1.1: *Standard Deviation of Output Quality Shocks Which Determine Private Buyer Offers*

	(1) σ_{η_c}	(2) 95% conf. interval
chickpea	0.124	[0.121, 0.126]
cotton	0.078	[0.076, 0.080]
finger millet	0.182	[0.177, 0.187]
groundnut	0.163	[0.159, 0.170]
maize	0.103	[0.101, 0.104]
mustard and rapeseed	0.082	[0.079, 0.084]
pearl millet	0.111	[0.109, 0.113]
pigeonpea	0.148	[0.143, 0.151]
rice	0.227	[0.225, 0.229]
sesamum	0.274	[0.263, 0.281]
sorghum	0.285	[0.282, 0.289]
sugarcane	0.133	[0.130, 0.136]
wheat	0.089	[0.088, 0.091]

Notes. This table shows the estimated standard deviation of output quality shocks, by crop, that determine private buyer offers. Column (2) is the 95% confidence interval estimated using bootstrap.

For each simulated distribution of prices, we compute the mean price by year, $\mathbb{E}[P_{jct}|\theta_c^g]$; recall that this does not necessarily equal the average private market price. We also compute the variance of this distribution. Finally, using the simulated data, we estimate the following probit model:

$$\Pr\Big(\text{Sold to government}_{jct} = 1|\theta_c^g\Big) = \Phi(\delta_{0rc}^g + \delta_{1rc}^g \cdot \log A_j) \tag{1.6}$$

which gives us region- and crop-specific coefficients δ_{0rc}^g and δ_{1rc}^g.

We construct empirical counterparts of these three sets of moments (mean, variance, and coefficients from probit model) using the Cost of Cultivation Surveys and the 77th round of the NSS. The former reports, in addition to farmer size and location, the realized output price for each crop planted. The latter (NSS) includes information on farmer size, location, and whether sales were made to government agencies or private buyers; this allows us to construct the auxiliary parameters δ_{0rc} and δ_{1rc} of equation (1.6) in the data.

For each crop, we estimate parameters θ_c by matching these empirical moments with the simulated moments.[48] We weigh the difference between empirical and simulated moments using inverse of the variance of empirical moments, which we estimate using bootstrap.

For non-PDS crops, identification is straightforward since the observed prices all come from the distribution of private buyer offers. For PDS crops, rice and wheat, identification is guaranteed if a positive mass of private buyer offers lies on both sides of MSP. Offers on the right of MSP guarantee that a non-trivial share of farmers sell to private buyers; thus, movements in the mean and variance of private buyer offers would shift the distribution of realized prices. Offers on the left ensure that if a government buyer is present, some sales would be made to government buyers. All else equal, a higher guess of the region-specific intercept α^g_{0rc} would uniformly (across farmer size) increase the share of farmers selling to government buyers in that region, and therefore yield a higher δ^g_{0rc}. Similarly, the slope coefficient α^g_{0rc} would directly affect the auxiliary coefficient δ^g_{1rc}. We present estimated parameters in Tables 1.1 and 1.2.[49]

Production Function and Risk Aversion

Production function parameters include crop-specific input elasticities for area (β_{ac}), labor (β_{lc}), capital (β_{kc}), and fertilizers (β_{fc}).[50] Additionally, we need to recover unobserved farmer productivities ω_j for all j. Finally, we also require the mean and variance of output quantity shocks ε_{jct} which enter the production function in equation (1.2). In logs, output of crop c

[48]Our approach is similar to an expectation–maximization (EM) algorithm (Dempster, Laird and Rubin, 1977) commonly employed to estimate parameters of mixture distributions. However, instead of maximizing a likelihood, we match moments. Our approach also relies on the literature on indirect inference (see Gourieroux, Monfort and Renault (1993))

[49]Not included in the interest of space: crop $\times$ year mean private market prices.

[50]Note that we treat fertilizers as a single composite input. As a robustness check, we plan to also estimate this production function using data on farmer-level consumption of fertilizer nutrients N, P, and K.

Table 1.2: *Parameters Governing the Likelihood of Finding a Government Buyer by State*

	(1) α_0 (Rice)	(2) α_1 (Rice)	(3) α_0 (Wheat)	(4) α_1 (Wheat)
Andhra Pradesh	-0.002	-0.019		
	[-0.004, -0.001]	[-0.036, -0.011]		
Bihar	-8.673	0.924		
	[-11.897, -8.105]	[0.893, 0.999]		
Chhattisgarh	-4.116	0.526		
	[-4.370, -3.351]	[0.480, 0.725]		
Gujarat	-6.323	0.580	-4.421	0.304
	[-7.812, -5.844]	[0.516, 0.738]	[-4.643, -4.113]	[0.268, 0.368]
Haryana	-3.921	0.373	-5.399	0.500
	[-5.469, -3.594]	[0.351, 0.468]	[-5.664, -5.311]	[0.482, 0.540]
Karnataka	-7.005	0.554		
	[-8.387, -6.578]	[0.496, 0.752]		
Madhya Pradesh	-7.311	0.665	-7.410	0.704
	[-8.731, -6.997]	[0.626, 0.786]	[-7.518, -7.259]	[0.676, 0.721]
Maharashtra	-5.719	0.462	-4.303	0.330
	[-5.946, -5.576]	[0.420, 0.665]	[-4.437, -4.116]	[0.263, 0.371]
Odisha	-8.063	0.745		
	[-11.297, -7.493]	[0.719, 0.832]		
Punjab	-3.110	0.307	-1.297	0.157
	[-4.298, -2.873]	[0.292, 0.344]	[-1.527, -1.181]	[0.148, 0.182]
Rajasthan			-5.223	0.373
			[-5.308, -5.121]	[0.357, 0.381]
Tamil Nadu	-4.038	0.417		
	[-4.571, -3.818]	[0.339, 0.441]		
Uttar Pradesh	-7.805	0.610	-7.586	0.649
	[-10.745, -7.321]	[0.596, 0.650]	[-7.676, -7.523]	[0.639, 0.659]
Uttarakhand	-12.485	1.171	-5.216	0.498
	[-15.250, -11.706]	[1.103, 1.376]	[-5.845, -4.411]	[0.439, 0.801]
West Bengal	-6.553	0.615		
	[-6.608, -6.500]	[0.595, 0.680]		

Notes. This table shows the estimated parameters governing the likelihood of finding a government buyer by state. Column (1) shows the intercept for rice. Column (2) shows the coefficient on (log) total farmer area for rice. Column (3) shows the intercept for wheat. Column (4) shows the coefficient on (log) total farmer area for wheat. Blank cells correspond to crop-states for which a negligible share ($<$ 1%) of farmers reported selling to government buyers. Confidence intervals are in square brackets below each point estimate and are estimated using bootstrap.

Table 1.3: *Risk Aversion Parameters*

	(1) estimate	(2) 95% conf. interval
Intercept, γ_0	-9.911	[-9.939, -9.895]
Coefficient on farmer size, γ_A	-0.118	[-0.125, -0.116]
Std. dev. of distribution, σ_γ	0.946	[0.929, 0.954]

Notes. This table shows the estimated parameters governing farmer risk aversion. Column (1) is the estimated parameters. Column (2) shows 95% confidence intervals estimated using bootstrap.

can be written as

$$\begin{aligned}\log q_{jct} &= \log q_c(A_{jct}, \boldsymbol{X}_{jct}) + \omega_j + \varepsilon_{jct} \\ &= \beta_{ac} \log A_{jct} + \beta_{lc} \log L_{jct} + \beta_{kc} \log\left(1 + K_{jct}\right) + \beta_{fc} \log\left(1 + F_{jct}\right) + \omega_j + \varepsilon_{jct} \qquad (1.7)\end{aligned}$$

The unobserved productivity term, ω_j, is constant across time and across crops.[51] This is in contrast to standard production function specifications which usually allow productivity to vary over time.[52] However, their estimation requires a monotonicity assumption between productivity and input demand which fails in our setting with risk-averse farmers as positive productivity draws increase the variance of output which may lead farmers to reduce input demand.[53] We cannot calibrate elasticities using cost shares either since that too relies on profit-maximizing choices.

Our approach involves jointly estimating production function and risk-aversion parameters using farmer's optimization problem in equation (1.1); risk aversion parameters include intercept γ_0, the coefficient on farmer size γ_A, and the variance of the mean-zero risk aversion draw ψ_j, denoted by σ_γ^2. We proceed as follows. Let $\theta_\beta = \{\beta_{ac}, \beta_{lc}, \beta_{kc}, \beta_{fc}\}_c$

[51]Since productivity does not differ by crop, selection into crops is not a concern for us.

[52]See Olley and Pakes (1996); Levinsohn and Petrin (2003); Ackerberg, Caves and Frazer (2015); Gandhi, Navarro and Rivers (2020)

[53]The monotonicity assumption allows researchers to construct control functions, using observed levels of intermediate inputs, which may account for unobserved productivity.

and $\theta_\gamma = \{\gamma_0, \gamma_A, \sigma_\gamma^2\}$. For each guess of parameters $(\theta_\beta^g, \theta_\gamma^g)$, where g denotes a candidate vector, we take the following sequence of steps.

1. Get $\xi_{jct}^g \equiv \omega_j^g + \varepsilon_{jct}^g$ by differencing out observed inputs from observed output using θ_β^g in equation (1.7); then regress ξ_{jct}^g on farmer fixed effects to get ω_j^g and ε_{jct}^g.

2. Compute mean and variance of output shocks ε_{jct}^g by crop.

3. Draw risk aversion γ_j for each farmer using θ_γ^g.

4. For the observed set of crops for each farmer-season, solve the portfolio choice problem in equation (1.1) using θ_β^g, ω_j^g, γ_j, mean and variance of output shocks for each crop, and the previously estimated parameters which govern the distribution of output prices.

The last step is computationally intensive; it gives us crop-specific input allocations of area, labor, capital, and fertilizers for the observed set of crops planted by each farmer in each season in the data. Since these input choices maximize the mean-variance utility for a given set of crops, they do not depend on fixed costs.

Using these simulated choices, we construct, by crop, first and second moments of simulated output, area, share of area conditional on planting, labor, capital, and fertilizer. These moments are sensitive to the guess of input elasticities and risk aversion parameters as both govern how farmers allocate inputs to different crops in a given set of crops.[54] We jointly identify these parameters by matching these simulated moments with their empirical analogs in the CCS data.[55] Estimated parameters are presented in Tables 1.3 and 1.4.

[54]For example, in Figures A.6 and A.7 we show how changing risk-aversion changes fertilizer usage and land share allocations.

[55]The differences are weighted by the inverse-variance weighted before summing.

Table 1.4: *Production Function Parameters*

	(1) land	(2) labor	(3) capital	(4) fertilizer
chickpea	0.455	0.437	0.259	0.059
	[0.444, 0.487]	[0.421, 0.471]	[0.240, 0.329]	[0.056, 0.069]
cotton	0.321	0.899	0.099	0.203
	[0.314, 0.334]	[0.898, 0.899]	[0.090, 0.136]	[0.192, 0.209]
finger millet	0.754	0.628	0.261	0.111
	[0.699, 0.857]	[0.610, 0.667]	[0.243, 0.333]	[0.100, 0.172]
groundnut	0.517	0.506	0.258	0.135
	[0.505, 0.549]	[0.492, 0.517]	[0.229, 0.302]	[0.132, 0.139]
maize	0.586	0.435	0.194	0.102
	[0.581, 0.592]	[0.428, 0.447]	[0.189, 0.208]	[0.098, 0.106]
mustard and rapeseed	0.581	0.290	0.183	0.078
	[0.571, 0.590]	[0.282, 0.309]	[0.178, 0.186]	[0.076, 0.080]
pearl millet	0.386	0.368	0.337	0.075
	[0.378, 0.396]	[0.359, 0.384]	[0.332, 0.358]	[0.071, 0.077]
pigeonpea	0.595	0.486	0.234	0.109
	[0.584, 0.611]	[0.478, 0.503]	[0.215, 0.250]	[0.104, 0.112]
rice	0.708	0.386	0.082	0.079
	[0.701, 0.741]	[0.380, 0.392]	[0.076, 0.086]	[0.074, 0.082]
sesamum	0.243	0.299	0.135	0.054
	[0.237, 0.265]	[0.289, 0.317]	[0.126, 0.140]	[0.046, 0.056]
sorghum	0.133	0.729	0.400	0.030
	[0.123, 0.146]	[0.722, 0.739]	[0.400, 0.400]	[0.023, 0.048]
sugarcane	0.551	0.665	0.126	0.091
	[0.527, 0.601]	[0.654, 0.675]	[0.118, 0.143]	[0.084, 0.104]
wheat	0.712	0.212	0.188	0.092
	[0.704, 0.747]	[0.207, 0.215]	[0.183, 0.192]	[0.088, 0.095]

Notes. This table shows the estimated production function parameters. Column (1) is the output elasticity of land. Column (2) is the output elasticity of labor. Column (3) is the output elasticity of capital. Column (4) is the output elasticity of fertilizer. Confidence intervals are in square brackets below each point estimate and are estimated using bootstrap.

Fixed costs

The fixed costs parameters to be estimated are crop-specific constants κ_c, and the discount parameter on repeated crops λ.[56] We denote these by $\theta_\kappa = \{\{\kappa_c\}_c, \lambda\}$.

Given a set $s \in \mathcal{S}_j$ of crops, the estimated parameters so far allow us to compute the mean-variance payoff U_{jst} for farmer j by solving the optimal portfolio choice problem. Farmer j computes this U_{jst} for all possible sets of crops he can plant, and then uses the fixed costs κ_{jst} to determine the set of crops which yield the highest utility $V_{jst} = U_{jst} - \kappa_{jst}$. This exercise is similar to a discrete choice problem where the choice set is a set of sets $\mathcal{S}_j$.

To estimate θ_κ, we proceed as follows. First, we build the set $\mathcal{S}_j$ for each farmer j. To do so, for each farmer, we take the set of crops planted by farmers in his state, and then take all combinations of up to length 3.[57] Then, we compute U_{jst} for all $s \in \mathcal{S}_j$ for each farmer in our data. This is a computationally intensive exercise but only needs to be done once.

Next, for each guess of parameters θ_κ^g, we find the set $s_j^*(\theta_\kappa^g)$ that maximizes utility V_{jst} for farmer j. Note that farmers are not forward-looking but their fixed costs depend on which crops they planted in the previous period; in our simulations, we take the set of crops in the previous period from the data, and then predict choices for this period given θ_κ^g. We take these simulated choices and compute simulated "market" shares for all sets $s \in \bigcup_j \mathcal{S}_j$. We also calculate, for each crop, the unconditional probability of being added to and dropped from a set between two consecutive periods.

We estimate θ_κ by matching these simulated market shares and switching probabilities with their empirical counterparts.[58] Moment conditions used in estimation are weighted by the inverse of the variance of empirical probabilities, estimated using bootstrap. We present the estimated parameters in Tables 1.5 and 1.6.

Our functional form assumption helps in the identification of the level of fixed costs. For

[56] We allow crop-specific constants to differ by season (monsoon or winter). The discount parameter is also allowed to be different for staples (rice and wheat) and all other crops.

[57] Almost all farmers in our data grow 3 or fewer crops in a season.

[58] In practice, we only match sets of crops with greater than 1% share in our data.

Table 1.5: *Crop-Specific Fixed Costs by Season*

	(1) season	(2) $\log(\kappa_c)$	(3) 95% conf. interval
chickpea	kharif (monsoon)	11.622	[10.235, 16.646]
chickpea	rabi (winter)	10.501	[9.985, 10.948]
cotton	kharif (monsoon)	11.356	[11.036, 11.641]
cotton	rabi (winter)	15.425	[12.664, 18.997]
finger millet	kharif (monsoon)	12.062	[10.670, 14.092]
finger millet	rabi (winter)	12.238	[10.075, 14.160]
groundnut	kharif (monsoon)	10.232	[9.906, 10.460]
groundnut	rabi (winter)	15.624	[13.281, 18.272]
maize	kharif (monsoon)	10.565	[10.353, 10.970]
maize	rabi (winter)	11.456	[10.455, 14.836]
mustard and rapeseed	rabi (winter)	9.843	[9.561, 11.401]
pearl millet	kharif (monsoon)	12.090	[8.683, 13.659]
pearl millet	rabi (winter)	10.373	[8.507, 12.854]
pigeonpea	kharif (monsoon)	10.315	[10.019, 10.543]
pigeonpea	rabi (winter)	15.869	[12.432, 20.763]
rice	kharif (monsoon)	10.605	[10.473, 10.981]
rice	rabi (winter)	11.328	[10.664, 12.296]
sesamum	kharif (monsoon)	8.808	[8.664, 8.987]
sesamum	rabi (winter)	10.660	[9.493, 11.542]
sorghum	kharif (monsoon)	10.196	[9.608, 12.015]
sorghum	rabi (winter)	11.353	[10.181, 12.932]
sugarcane	kharif (monsoon)	13.268	[12.064, 15.825]
sugarcane	rabi (winter)	13.029	[12.247, 13.968]
wheat	rabi (winter)	10.172	[9.700, 10.696]

Notes. This table shows the estimated crop-specific fixed cost constants κ_c by season. Farmers grow some crops in both seasons, while others are only grown in one season in our data. The last column reports the 95% confidence interval estimated via bootstrap.

example, consider the possible sets of crops with two crops c_1 and c_2: $\{c_1\}, \{c_2\}, \text{and}\{c_1, c_2\}$. If we increase the fixed cost associated with c_1 and c_2 by Δ, the relative attractiveness of $\{c_1\}\text{and}\{c_2\}$ will remain the same. However, $\{c_1, c_2\}$ will become relatively less attractive as costs go up by 2Δ and farmers will switch out of it. Thus, the share of farmers growing a given set of crops is informative about crop-specific constants κ_c. The discount on repeated crops, λ, is informed by the observed persistence of crop choice. We capture this persistence

Table 1.6: *Discount Parameters for Repeating Crops*

	(1) estimated λ	(2) 95% conf. interval
rice and wheat	-0.143	[-0.183, -0.091]
all other crops	-0.006	[-0.024, -0.001]

Notes. This table shows the estimated discount parameter on repeated crops. We estimate these separately for staple crops (rice and wheat), and all other crops. The last column reports the 95% confidence interval estimated via bootstrap.

by its counterpart, the probability of dropping a crop grown in the last period. Lower λ lowers the fixed cost of repeated crops, and therefore lowers the probability of dropping a crop.

Supply Model Fit

To assess model fit, we sample a set of farmers and solve their crop and input choice problem, keeping prices fixed. We compare these simulated choices with the data in Figures 1.7a and 1.7b.

1.4.2 Demand

PDS Crops: Rice and Wheat

For rice and wheat, we estimate the specification in equation (1.4) using household-level consumption data from the 68th round of the NSS, conducted in 2011-12. We proxy for household income using total monthly expenditures. All variables are measured at per capita level.

To address potential endogeneity of prices, we rely on Hausman, Leonard and Zona (1994) and instrument prices using average price in the state excluding own district. These are valid instruments under the assumption of idiosyncratic district-level demand shocks, which may enter the error term u_{ch} in equation (1.4), but correlated state-level supply shocks such as processing costs and/or transportation costs. Note that these are outside of our

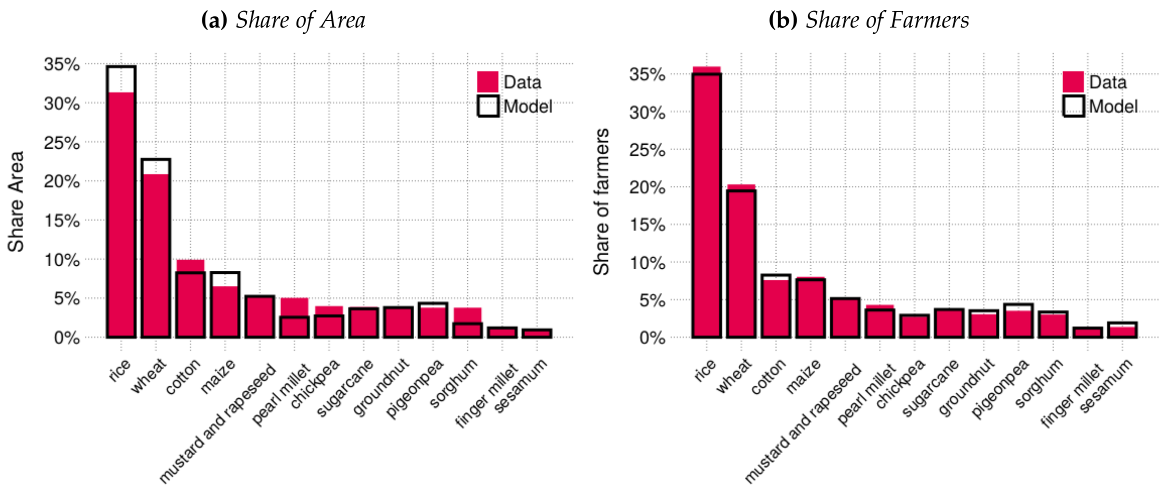


Figure 1.7: *Supply-Side Estimates: Comparing Model-Predictions with Data*

Notes. The left panel compares the share of area allocated to each crop as observed in the data and as predicted by the model. The right panel reports the same for the share of farmers growing a crop.

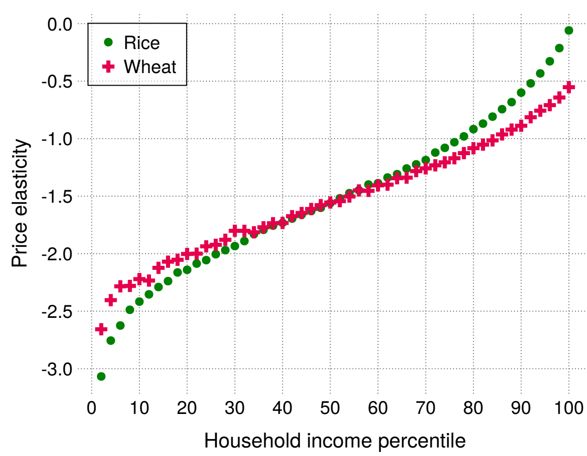


Figure 1.8: *Price elasticities of demand for rice and wheat by household income*

Notes. This figure shows simulated price elasticities using estimated demand parameters for households with different income levels – proxied using monthly household expenditures.

model and only used for the estimation of demand parameters; in our counterfactuals, households would face a single average private market price for each crop. We present the estimated parameters in Table 1.7. We also show the implied demand elasticities by

Table 1.7: *Estimated Demand Parameters for PDS Crops*

	$log\,(1+q)$
	(1)
(Rice) log price	-7.985***
	(0.268)
(Rice) log income per capita	-2.945***
	(0.105)
(Rice) log price × log income per capita	0.942***
	(0.034)
(Wheat) log price	-3.917***
	(0.150)
(Wheat) log income per capita	-1.055***
	(0.046)
(Wheat) log price × log income per capita	0.422***
	(0.016)
N	186,866
R^2	0.425

Notes. This table shows estimated parameters using the specification in equation (1.4) *for PDS crops, rice and wheat. Standard errors are reported in parentheses.*

household income in Figure 1.8.

Non-PDS Crops

We calibrate demand for non-PDS crops using estimates of price elasticities in Deaton (1997). These are given in Table 1.8.

1.5 Counterfactuals

In this section, we evaluate the distributional effects of fertilizer subsidies, government-procurement at minimum support prices (MSP), and redistribution of foodgrains through the public distribution system (PDS). We do so with the help of two counterfactuals in which we phase out these programs. These include: (1) no fertilizer subsidies, and (2) no

Table 1.8: *Calibrated Demand Elasticities for Non-PDS Crops*

	(1) elasticity
chickpea	-0.57
cotton	-0.85
finger millet	-3.29
groundnut	-0.28
maize	-3.29
mustard and rapeseed	-0.28
pearl millet	-0.45
pigeonpea	-0.57
sesamum	-0.28
sorghum	-0.45
sugarcane	-0.33

Notes. This table reports the calibrated demand elasticities for non-PDS crops.

government-procurement at minimum support prices; the latter also results in zero PDS entitlements for households.[59]

While these counterfactuals help us understand the effects of existing programs, they also help us study equilibrium effects of proposed reforms that aim to minimize government's role in the agriculture sector. These include proposals to end fertilizer subsidies (Gulati, 2014) as well as legislation to promote a greater role of private players and potentially smaller role of government buyers in output markets (Mashal, Schmall and Goldman, 2021).[60]

[59] In ongoing work, we also consider how to end government procurement without impacting household consumption through alternative programs such as consumption vouchers.

[60] In 2020, the Indian government attempted to pass bills which would have paved the way for greater private sector involvement in output markets (where farmers sell their harvest). But this attempt was met with a large-scale farmers' protest which lasted for a year, and ended with the repeal of these bills and a demand for a legal guarantee for MSP.

1.5.1 Solving for the baseline equilibrium

We begin by describing how we solve for the equilibrium in the prevailing regime of fertilizer subsidies, MSP procurement, and PDS entitlements using our estimated parameters. We compute this equilibrium for a sample of 20,000 farmers and all households in the 68th round of the NSS (weighted by sampling weights), that we hold fixed across counterfactuals.

Equilibrium consists of a vector of average private market prices, for the 13 crops we include in our sample, which clears all markets (see Section 1.3.4). We start with a guess of price vector, solve for optimal production and consumption decisions, and test if all equilibrium conditions hold. If not, we update our guess.[61]

On the supply-side, given this vector of prices, farmers choose which set of crops to plant and make crop-specific input allocations.[62]. We simulate whether sales are made to government buyers or to private traders based on the estimated likelihood of finding government buyers and the distribution of private buyer offers. This gives us the aggregate private market supply and the level of government stockpile of rice and wheat. We then redistribute the government stockpile to households in proportion to their observed entitlements.[63] Given the estimated demand function, and the guess of the price vector, we also know the total demand for each household. We can subtract their PDS entitlements from the total demand to get their private market purchases in the counterfactual. Summing across households gives the total private market demand for PDS crops. For non-PDS crops, private market demand is the predicted aggregate demand from equation (1.5).

---

[61] To update, we decrease prices for crops with excess private market supply and increase prices for crops with excess private market demand.

[62] MSP is set by the government taking into account cost of cultivation, and expected market prices. The announced MSP closely tracks the private market prices (see Figure A.2). We do not model this endogenous MSP setting process, but instead assume that the government announces an MSP based on the expected distribution of prices in the private market. On average, the announced MSP is at 59th percentile of the private market price distribution for wheat and 72nd percentile for paddy. We take the mid-point of the two, and assume that the announced MSP is at 65th percentile for wheat and paddy in the counterfactual.

[63] For each household, we compute the share of total entitlements received in the 68th NSS round. We hold these shares constant in each counterfactual simulation, and redistribute total quantities of rice and wheat procured by the government using these shares.

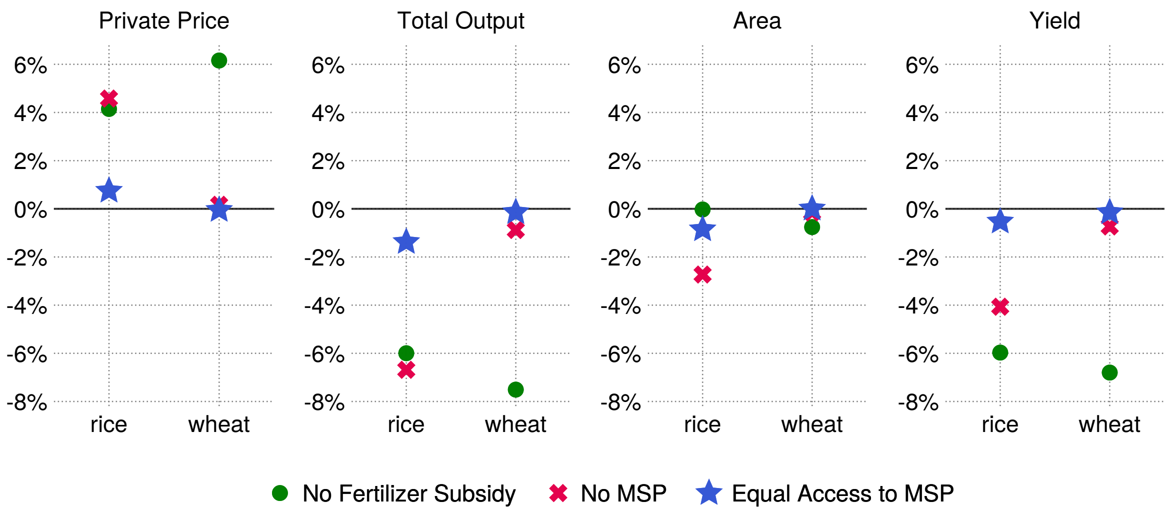


Figure 1.9: *Percent Changes in Key Variables Relative to Baseline*

Notes. The first panel shows percent change in equilibrium private market prices under different counterfactual policies relative to the baseline prices. The other three panels repeat this exercise for total output, total area, and average yield in the economy. For similar plots for all other crops in our data, see Figure A.4.

1.5.2 How do we measure distributional effects?

Before presenting results from our counterfactuals, we describe how we characterize the impact on farmer and consumer welfare along the income distribution.

For farmers, the net impact is captured by utility V_{jst} for each farmer j. We compute this utility for each farmer in each counterfactual and calculate changes relative to the baseline equilibrium described above. For example, when we consider the impact of ending government-procurement at MSP, negative changes in V_{jst} would imply that farmer j was relatively better off in the baseline. Then, we compute summary statistics of these changes grouped by income (or farmer size) bins and present them below.

Consumers, or households, differ along two dimensions – income and PDS entitlements; the latter is strongly correlated with the former. In our estimation, we find that lower-income households have relatively higher price elasticities. These households are also more reliant on the PDS for their consumption. Therefore, their consumption and their food expenditures are very sensitive to private market prices as well as the total size of the government stockpile.

While we present effects on consumption and expenditures along the income distribution on the demand-side, neither of these measures fully captures the impact of prices and in-kind transfers on consumers. To summarize this impact, we construct a Laspeyres index for each counterfactual, as described below.

Let the minimum food expenditures to consume a vector of quantities $\boldsymbol{q}^{TOT}$, given a vector of PDS entitlements $\boldsymbol{q}^{PDS}$, be given by

$$e(\boldsymbol{p}^{PVT}, \boldsymbol{q}^{TOT}; \boldsymbol{q}^{PDS}) = \left(\boldsymbol{q}^{TOT} - \boldsymbol{q}^{PDS}\right)' \boldsymbol{p}^{PVT}$$

where $\boldsymbol{p}^{PVT}$ is a vector of private market prices. In the baseline regime, household h consumes quantities $\boldsymbol{q}_{0,h}^{TOT}$ of crops rice and wheat, given by

$$\boldsymbol{q}_{0,h}^{TOT} = \boldsymbol{q}_{0,h}^{PVT} + \boldsymbol{q}_{0,h}^{PDS}$$

In addition, they consume quantity $c_{0,h}$ of the numeraire good, given by

$$c_{0,h} = y_h - e(\boldsymbol{p}_0^{PVT}, \boldsymbol{q}_{0,h}^{TOT}; \boldsymbol{q}_{0,h}^{PDS})$$

where y_h is the total monthly expenditures of household h. The modified Laspeyres index (MLI) under the two counterfactuals is given by

$$MLI_{\text{no fert subsidy},h} = \frac{c_{0,h} + e(\boldsymbol{p}_{\text{no fert subsidy}}^{PVT}, \boldsymbol{q}_{0,h}^{TOT}; \boldsymbol{q}_{\text{no fert subsidy},h}^{PDS})}{c_{0,h} + e(\boldsymbol{p}_0^{PVT}, \boldsymbol{q}_{0,h}^{TOT}; \boldsymbol{q}_{0,h}^{PDS})}$$
$$MLI_{\text{no msp},h} = \frac{c_{0,h} + e(\boldsymbol{p}_{\text{no msp}}^{PVT}, \boldsymbol{q}_{0,h}^{TOT}; \boldsymbol{0})}{c_{0,h} + e(\boldsymbol{p}_0^{PVT}, \boldsymbol{q}_{0,h}^{TOT}; \boldsymbol{q}_{0,h}^{PDS})}$$

In other words, this index captures the relative change in expenditures if the household were to continue to consume the baseline bundle of rice and wheat, and the numeraire good, in counterfactual regimes.[64]

Finally, in the main text below, we do not discuss the change in aggregate consumer surplus associated with non-PDS crops, but we present results in the appendix in Figure A.5.

[64]Since this excludes the impact of change in prices in other crops i.e. non-PDS crops, this is only an approximation to the actual relative change in expenditures.

1.5.3 The equilibrium without fertilizer subsidies

In a single-crop economy, the impact of removing fertilizer subsidies on downstream consumers is unambiguous. When fertilizer prices go up, fertilizer demand falls, individual and aggregate production falls, market price rises, and the consumption of downstream households falls. The impact on welfare of risk-neutral producers depends on the price elasticity of demand: if demand is inelastic, the decline in demand is low relative to the increase in price, so profits go up.

The conclusions in our setting with crop choice, on the extensive and intensive margins, and risk-averse farmers are more ambiguous. To understand equilibrium distributional effects, we solve for a new equilibrium without fertilizer subsidies. In our data, we approximate an average subsidy rate of 50% across all fertilizer products; as such, we double the price of fertilizer and solve for production and consumption decisions, accounting for government procurement, at different guesses of average private market prices. We stop when all equilibrium conditions are met.

Without fertilizer subsidies, aggregate output of all crops falls, driven by lower consumption of fertilizers, and private market prices go up (see Figure A.4 and **??** for all crops). For rice and wheat, prices go up by about 5%, output falls by about 7% (and so does yield) as shown in Figure 1.9. We estimate government savings to be approximately $4.35 billion or $70 per farmer.[65]

If we do not allow farmers to adjust input and crop choice, when fertilizer prices go up, farm profits would unambiguously fall. But since farmers are free to use less and produce less, in equilibrium, they receive higher prices in the private market. Our results suggest that these higher output prices are nearly enough to compensate for the higher per unit cost of fertilizers. As shown in Figure 1.11b, we find a minimal impact on farmer welfare in the absence of fertilizer subsidies.

On the demand-side, consumption falls. The fall in consumption of rice and wheat is greater for lower-income households who we estimate to be more price sensitive (see

[65]We use 2014 as our reference year and use the average exchange rate of 0.0164 USD = 1 INR.

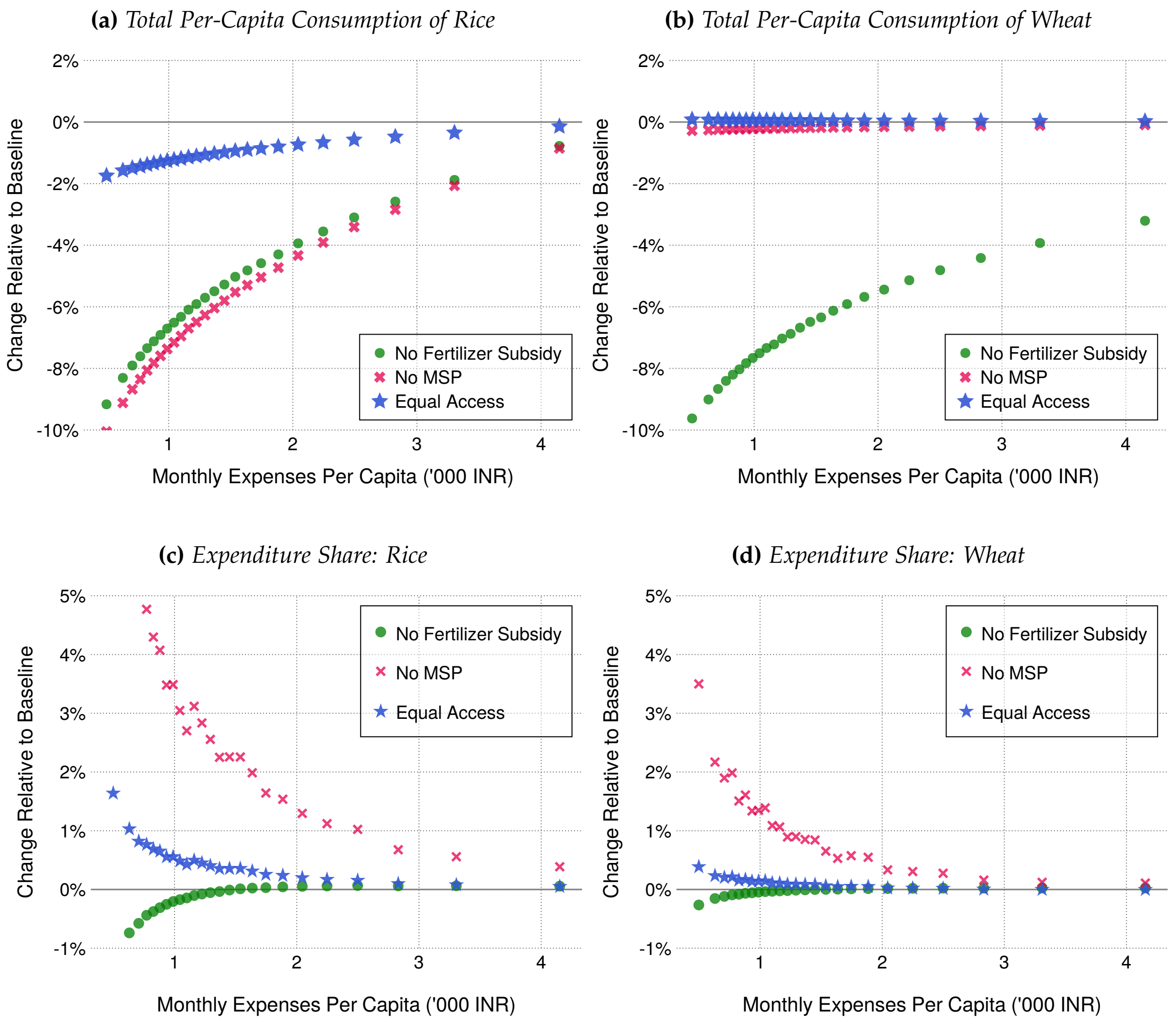


Figure 1.10: *Change in Household Consumption & Expenditures Relative to Baseline*

Notes. The top-left panel shows mean percent change in household consumption of rice relative to baseline, binned by total household expenses (proxy for household income) under different counterfactual policies. The top-right panel repeats the same for consumption of wheat. The bottom-left panel shows the change in expenditures on rice as a share of total household expenditures (in percentage points) relative to baseline. The bottom-right panel repeats the same for share of expenditures on wheat.

Figures 1.10a and 1.10b). Since total output falls and, therefore, government procurement is low, PDS entitlements go down. But this impact is small. Expenditures on rice and wheat, as a share of total expenditures, fall by 0.5 percentage points for the lowest-income households; this is just due to lower consumption of rice and wheat by these households.

Finally, we summarize the net impact on consumers in Figure 1.11a using the Laspeyres index described above. To consume the same bundle of rice, wheat, and numeraire good as in the baseline, lowest-income households must now spend 3%-4% more relative to baseline.

1.5.4 The equilibrium without government-procurement at MSP

Let us consider again the scenario with a single crop. For a risk-averse farmer, minimum support price (MSP), if available, increases the mean and reduces the variance of output price. In the absence of MSP, the farmer would face greater price risk; to reduce exposure to this risk, he would lower input usage and produce less. On the household side, removing government procurement would take PDS entitlements to zero.[66] Consequently, household demand in the private market would go up. Low supply and high demand would give rise to an equilibrium with higher prices. With zero PDS entitlements and higher private market prices, lower-income households would suffer more given their higher reliance on PDS entitlements and their greater price sensitivity.

With multiple crops, farmer response would depend on the relative impact on mean and variance of prices across crops. For example, while rice and wheat may both become less attractive, wheat may become more attractive relative to rice. This could result in more output for wheat when MSP for rice and wheat is taken away as farmers switch from producing rice to wheat. To understand equilibrium effects in our setting, we set the probability of finding a government buyer to zero, simulate farmer and household decisions, and solve for a new vector of equilibrium private market prices. We find that the private market price of rice goes up by about 5% and aggregate output falls by over 6%. We find a minimal impact on the private market price and output of wheat. These differences are due to the differential price risk of rice and wheat in the absence of government procurement at MSP – the estimated variance of private buyer offers for rice is much greater than that of wheat. In the absence of government-procurement at MSP, we estimate government savings

[66]This is an assumption of this counterfactual. We can also consider a scenario where the government stops procuring at MSP but continues to subsidize consumption of lower income households in the private market (e.g. through consumption vouchers).

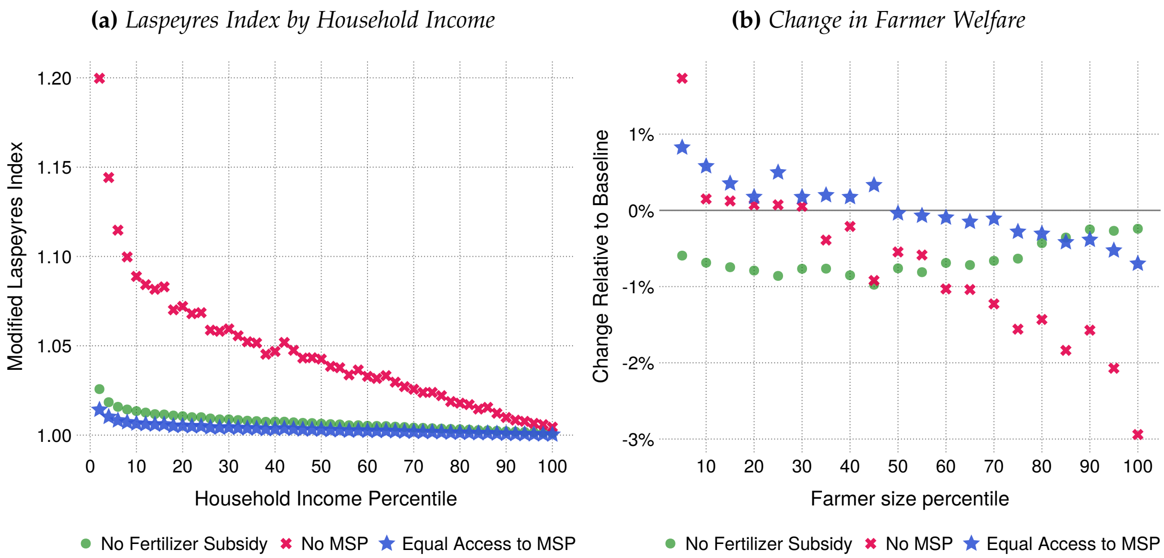


Figure 1.11: *Distributional Effects on Consumers and Producers*

Notes. The left panel shows a binscatter plot of Laspeyres index by household income under different policy regimes. The index accounts for changes in both prices and in-kind transfers. The right panel shows median percent change in farmer utilities, V_{jst}, relative to baseline, binned by farmer size under different counterfactual policies.

to be approximately \$8.5 billion or \$137 per farmer.

On the supply-side, we find that larger farmers experience a larger loss in welfare. This is because they were more likely to find a government buyer and avail MSP in the baseline. Some of the smallest farmers experience modest gains since they were less likely to sell to government buyers in the baseline and they now receive higher private market prices for rice, which is the preferred crop of small farmers.

On the demand-side, households must now satisfy all demand in the private market. As such, expenditures in the private market go up, as shown in Figures 1.10c and 1.10d. This increase in expenditures is greater for lower-income households since they derived a greater share of their total consumption of these crops from the public distribution system (PDS). For rice, not only do PDS entitlements go to zero, private market prices also go up which exacerbates the adverse effects of this counterfactual on lower-income households. Again, we summarize the impact using our Laspeyres index described above. As shown in

Figure 1.11a, the lowest-income households must now spend 15% to 20% more to consume the baseline bundle of rice, wheat, and a numeraire good.

1.5.5 What if there was no large-farmer bias in government procurement?

As an additional counterfactual, we consider the impact of a policy where the large-farmer bias in government-procurement at MSP is eliminated. To do so, we hold fixed the number of farmers in each state that the government procures from and randomly assign all farmers to government buyers. Note that this does not hold fixed the quantity of output procured by the government. Total procurement is expected to go down since government buyers would now match with smaller farmers with higher frequency than in the baseline. In fact, procurement of rice falls by about 17% and procurement of wheat falls by about 11%. As such, the government saves about $1.3 billion or $20 per farmer.

This alternative policy has minimum impact on private market prices and total output of rice and wheat, as shown in Figure 1.9, as well as all other crops shown in Figure A.4. Smaller farmers gain (not just the smallest) and the average gains are greater than the scenario where MSP procurement is phased out. Larger farmers are worse off but these losses are small. Importantly, on the demand-side, the impact on lower-income households is minimal. As shown in Figure 1.11a, under this counterfactual, lowest-income households only pay 1%-2% more to consume their baseline bundle of goods.

1.6 Conclusion

In this paper, we develop and estimate a structural model of the agriculture sector in India, accounting for the impact of various government-sponsored price interventions on production and consumption decisions. We estimate this model using observational microdata at the farmer and household level and run counterfactuals to characterize the distributional effects of these programs. On the demand-side, we find these interventions to be progressive – these accord greater benefits to lower-income households. In contrast, on the supply-side, we find these interventions to be (weakly) regressive due to inequities in implementation which favor wealthier farmers.

Chapter 2

Subsidies Trump Tariffs: Case of Utility-Scale Solar in India[1]

2.1 Introduction

Throughout history, economies have sought to protect domestic producers against foreign competitors. This protection may be in the form of import bans, import quotas, and import tariffs which restrict the supply of foreign goods. It could also be in the form of subsidies to domestic firms for capital investments, R&D, and production which help increase the supply of domestic goods. Regardless of their nature, these protectionist interventions are, in general, costly. Import restrictions may harm consumers of imported goods while subsidies to domestic firms require large upfront outlays. The objective of this paper is to measure and compare the welfare costs associated with two popular interventions – import tariffs and production subsidies.

There are many reasons why policymakers might intervene to protect domestic producers. The most popular theoretical justification is the infant-industry argument – that there exist external economies of scale or cross-industry spillovers which may be realized through the growth of the protected industry (Greenwald and Stiglitz, 2006; Bartelme, Costinot,

[1] Co-authored with Shresth Garg

Donaldson and Rodríguez-Clare, 2019). Reasons such as supply-chain fragility, recently highlighted by the COVID-19 pandemic, might also motivate policymakers to undertake efforts to onshore certain sectors of the economy and reduce reliance on foreign producers. Finally, a variety of political and strategic motives might also induce countries to favor domestic producers over foreign competitors (Grossman and Helpman, 1994).

Regardless of the underlying motivation, policymakers often set expansion targets for domestic industries. In this paper, with the help of a structural model of oligopolistic markets, we estimate the short-run costs associated with achieving such targets using either import tariffs on competitors or production subsidies for domestic producers. For tariffs, we compute the welfare losses suffered by downstream industries and final consumers, whereas for subsidies, we calculate the fiscal costs borne by the policymaker. At the same time, we net out any static benefits accorded to the direct beneficiaries, downstream industries and consumers, as well as to the government in the form of tariff revenues. When computing the costs of these two interventions, we take as given the objective of the policymaker to expand the domestic industry, and then compare the costs of achieving this goal using tariffs and subsidies.

We evaluate these interventions in the context of the Indian utility-scale solar sector. Over the past decade, the utility-scale solar sector in India has undergone a massive expansion[2] (see Figure 2.1a). The final output of this sector is solar power that is fed into transmission grids that supply electricity to agricultural, industrial, and residential consumers. The sector encompasses an upstream industry that supplies solar modules (or panels) and a downstream industry that develops solar power plants which generate electricity for end use. The expansion in the utility-scale solar sector has largely been fueled by imported solar modules. As such, the share of domestic manufacturers in the upstream sector has remained at very low levels (see Figure 2.4).

In recent years, the Indian government has taken two important steps to support the

[2]Installed capacity at the end of 2020 was 35 gigawatts compared to under 200 megawatts in 2010, which corresponds to a compound annual growth rate of approximately 70%.

(a) *Cumulative installed solar capacity* **(b)** *Solar module price per watt*

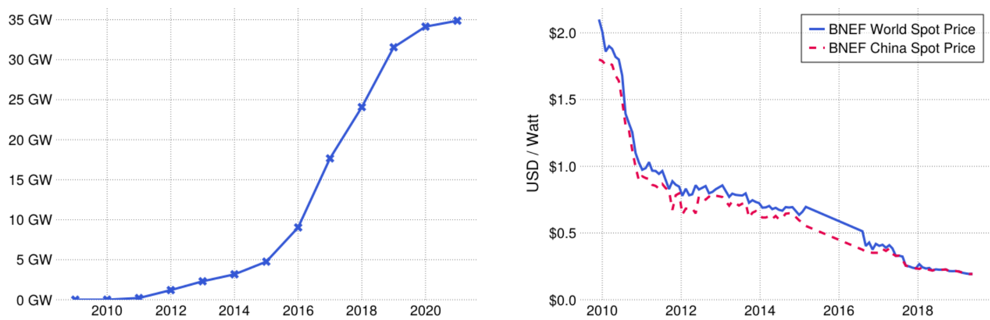


Figure 2.1: *Indian solar capacity and global solar module prices*

Notes. In the left panel, we plot the sum total of the capacities of all commissioned solar plants upto a given year as recorded in the projects database of our data provider, Bridge to India. In the right panel, we show the monthly average spot prices of multi crystalline silicon modules, expressed in price per watt. These values are obtained from the Bloomberg New Energy Finance (BNEF) Solar Spot Price Index.

domestic module manufacturing industry. First, starting August 2018, the government introduced a safeguard duty of 25% on solar cells and modules imported from China and Malaysia. From April 2022 onwards, this was converted into a basic customs duty of 40% on all imports of solar cells and modules. Second, in November 2020, the government introduced a production-linked incentive (PLI) scheme, which is essentially a per-unit subsidy on the production of solar modules[3]. Both policies are aimed at supporting the domestic solar module industry but the underlying motivation for engaging in these protectionist measures is unclear. That is, it is unclear whether the policies are in place to take advantage of cross-industry spillovers, for increasing manufacturing jobs, or for strategic reasons such as energy independence. As such, in this paper, we take the government's explicit objective of expanding domestic production at face value, and then calculate and compare the costs of achieving this objective using tariffs and subsidies.

The structure of the utility-scale solar sector in India makes it an attractive setting to study the relative costs of tariffs and subsidies. Solar panels produced by the upstream

[3]The subsidy to the solar sector under the PLI scheme is part of a broader push to provide production linked subsidies to a wide variety of manufacturing industries. As of 2022, subsidies totaling over $26 billion have been announced.

industry are primarily consumed in utility-scale solar projects[4]. Furthermore, the output of the downstream industry, solar power, is largely purchased by government-run power distribution companies. Thus, in our setting, we have a simple vertical structure which makes it straightforward to calculate all downstream welfare costs associated with policies affecting the upstream industry of solar panels. This is in contrast with other empirical settings where one might have to trace through a network of downstream industries to estimate total costs.

Our empirical approach involves specifying and estimating a structural model of the upstream and the downstream industries. We treat solar panels as homogeneous goods and model the upstream market as a Cournot oligopoly. Using observed market shares and estimated demand elasticities, we recover marginal costs of domestic and foreign module suppliers implied by the Cournot model. The downstream industry is organized as a multi-unit discriminatory price auction. These are procurement auctions run by government agencies to award long-run power purchase agreements to solar plant developers who build and operate solar plants. We use auction-level bid data to estimate an auction model which delivers the private costs of developing these solar plants. The auction model also allows us to estimate counterfactual winning bids, i.e. the price that the government agencies must pay to purchase solar power under different policy regimes.

With structural parameters in hand, we conduct our main counterfactual exercises. In our baseline regime of no intervention, we assume that downstream government agencies wish to auction off 1 gigawatt (GW) of solar capacity in a single auction. We simulate auction play and compute the equilibrium of the upstream and downstream markets under this baseline. Then, we consider the following two sets of objectives of the policymaker: (1) achieve a combined market size of 0.2 GW (or 0.4 GW) for domestic firms, and (2) achieve a combined market share of 20% (or 40%) for domestic firms. The first objective simulates a scenario where the policymaker wishes to expand the quantity of modules supplied by

[4]Given the small total size of the rooftop solar industry in India, we ignore any welfare losses experienced by rooftop solar projects. Instead, we treat utility-scale solar as the sole downstream consumer of solar panels in India.

domestic firms, while the second objective corresponds to a scenario where the policymaker wishes to reduce dependence on imported modules. For each objective, we compute the tariff and subsidy level that would allow domestic firms to achieve the desired market size or share. Then, we estimate the various welfare components associated with the required level of tariff and subsidy. These components include fiscal costs (or revenues), profits of domestic upstream suppliers, profits of downstream developers, and the consumer surplus of the power distribution companies who purchase solar power. For each objective and each policy instrument, we compare these welfare estimates with the estimates from the baseline regime of no intervention. Finally, to understand the role of market power in our welfare calculations, we treat the upstream market as a duopoly consisting of a foreign firm and a domestic firm, and repeat all of the above.

Our results suggest that subsidies are much more cost-effective than import tariffs and that the presence of market power makes subsidies all the more attractive. Expanding the output of the domestic solar module sector to 0.2 GW using just import tariffs results in a welfare decline of about 20 billion INR, while doing so using production subsidies reduces welfare by only about 0.3 billion INR. In relative terms, welfare loss under tariffs is 60 times the loss under subsidies. If we treat the upstream market as a duopoly consisting of a representative foreign firm and a representative domestic firm, we find that subsidies are welfare enhancing, while import tariffs continue to impose severe welfare costs relative to baseline.

It is important to highlight that the analysis in this paper has several important limitations. First, even for temporary interventions, the associated costs and benefits might take years, if not decades, to be realized. One needs long-run data which capture the full trajectory of all relevant sectors to calculate total costs and benefits (Head, 1994; Irwin, 2000b; Hansen, Jensen and Madsen, 2003; Harris, Keay and Lewis, 2015). While we capture the short-run components of welfare we abstract away from long-run or dynamic considerations such as learning by doing, declining barriers to entry, or reputation (Schmalensee, 1982) which provide additional incentives for offering industrial protection. Second, it is difficult to fully capture the spillover effects. Subsidies benefit not only the targeted industry but also the full network of upstream and downstream industries. Similarly, tariffs may benefit input suppliers of the protected domestic industry and at the same time, hurt the downstream buyers. In addition, an expansion in the targeted sector might impose costs on other sectors a la Dixit and Grossman (1986). We capture the effect on downstream firms and consumers, but do not incorporate effects upstream or in adjacent industries. Third, there is no easy way to compute the value of strategic benefits gained from reducing reliance on foreign producers. Industrial protection measures which might appear to be very costly in a myopic sense might help improve the international bargaining power of a country resulting in long-run political and economic benefits. Finally, despite our focus on solar energy, we do not consider environmental costs or benefits associated with changes in the size of this sector.

Related literature. Our work adds to the large literature evaluating the impact of industrial policy interventions. The set of papers closest to our work focus on individual sectors and use structural modeling techniques to estimate the costs and benefits of tariffs and subsidies. These include Irwin (2000a) and Irwin (2000b) which study the impact of tariffs on the US pig iron industry and the US tin industry, respectively. Baldwin and Krugman (1988a) measure the extent and welfare impact of European subsidies to Airbus. In a similar vein, Baldwin and Krugman (1988b) study the effect of the tacit "buy domestic" policy of Japanese firms on the Japanese RAM industry. They find that while the protection was crucial for the development of the domestic industry, the net welfare effects were negative. Head (1994) looks at the impact of steel rail tariffs on welfare and finds a small positive effect. Kalouptsidi (2018) estimates a model of the global shipbuilding industry to measure Chinese subsidies to the industry, and finds that these subsidies reallocated production to China, with no significant gains in consumer surplus. Barwick, Kalouptsidi and Zahur (2021) evaluate a suite of policies aimed at developing the Chinese shipbuilding industry and find production subsidies to be the most cost-effective means of achieving output targets.

There's also a large literature that evaluates the effects of industrial policy using reduced form methods. This work tries to address the endogeneity of typical policy interventions, exploiting natural experiments for identification. Juhász (2018) looks at the French cotton spinning industry which was protected from British competition due to Napoleon blockades. Temporary protection increased adoption of mechanized cotton spinning technology, and the effect persisted even after the protection was lifted. Protection also increased industrial output in the long run. Lane (2021) studies the Korean industrial policy directed towards heavy chemical and industry (HCI). Similar to Juhász (2018), the policy increased the size of domestic industry which persisted even after the policy was withdrawn. A string of papers also analyze place-based subsidies given to firms in Italy and the UK, and find largely positive effects on employment, investment, and innovation. These include Bronzini and Piselli (2016); Cerqua and Pellegrini (2014, 2017); Pellegrini and Muccigrosso (2017); Criscuolo, Martin, Overman and Van Reenen (2019).

We also contribute to the literature measuring the welfare costs of trade policy. Following the US-China trade war of 2018, there has been a renewed interest in measuring costs of trade policy. Fajgelbaum, Goldberg, Kennedy and Khandelwal (2020) estimate that even after accounting for domestic producer surplus and government tariff revenue, the trade war cost US $7.2 billion in the short run. Amiti, Redding and Weinstein (2019) also estimate a similar number. Cavallo, Gopinath, Neiman and Tang (2021) find that tariff costs were largely borne by US importers and absorbed by retailers. Thus, while US consumers did not suffer welfare losses, retail margins fell. Flaaen, Hortaçsu and Tintelnot (2020) zoom in on the US washing machine industry and find a complete pass-through of 2018 import tariffs to US consumers. They find that these tariffs increased consumer costs by $1.5B annually compared to $82 million collected in tariff revenue. Other recent papers include Irwin (2019) which estimates a high pass through of early twentieth century tariffs on imported sugar to prices faced by consumers. Finally, a large body of literature measures welfare effects of import tariffs by linking it to domestic firm productivity (Amiti and Konings, 2007; Topalova and Khandelwal, 2011; Edmond, Midrigan and Xu, 2015), quality of domestic products (Amiti and Khandelwal, 2013), and availability of new products (Goldberg, Khandelwal, Pavcnik and Topalova, 2010).

The remainder of this paper is organized as follows. In Section 2.2, we provide an overview of the solar sector in India and provide key institutional details relevant for our structural model. Section 2.3 lays out our data sources and presents preliminary evidence on the impact of module prices on module demand and solar power rates. We describe our structural model in Section 2.4, and present our estimation strategy and estimates in Section 2.5. In Section 2.6, we discuss results from the aforementioned counterfactual scenarios computed using the estimated parameters. Finally, we conclude in Section 2.7.

2.2 The Indian utility-scale solar sector

Since 2010, India has embarked on an ambitious plan to rapidly develop and deploy utility-scale solar plants in the country (see Figure 2.1a). Less than a decade ago, the total installed solar photovoltaic (PV) capacity in India was under 200 megawatts (MW). At present, the total installed capacity is over 35 gigawatts (GW) with another 52 GW in pipeline. Two major factors have contributed to this massive expansion. First, rapid advancements in solar photovoltaic (PV) technology have reduced dramatically the costs of solar cells and modules (see Figure 2.1b). Second, India has made a concerted effort to expand utility-scale solar capacity. As part of its National Action Plan for Climate Change (NAPCC), in 2010, the Government of India launched the National Solar Mission (NSM) which set an initial target of deploying 20 GW of solar capacity by 2022. Halfway through the decade, this was revised upwards to a target of 100 GW of solar capacity by 2022, 60 GW of which is to be deployed through utility-scale solar projects[5].

The term *utility-scale* is used to indicate the scale of each solar plant ($>$ 1 MW) and the intended end-use — solar power generated through these plants is fed into the electricity transmission grid operated by various state-run power distribution companies. Electricity from other sources such as hydroelectric, coal, wind etc. is also fed into the same grid and

[5]Solar power capacity can be achieved through utility-scale solar projects or through rooftop solar projects. Under the new target of the NSM, 40 GW of solar capacity is to be deployed via rooftop solar projects.

then supplied to agricultural, industrial, and residential consumers. For the purposes of this paper, we treat power distribution companies as the end users of solar power. As explained in Section 2.4, we model them as price sensitive "consumers" and compute consumer surplus statistics with respect to their utility function. The bulk of our analysis focuses on two vertically-linked industries which are crucial to the utility-scale solar sector. In the downstream industry, solar plant developers build and operate solar plants, while in the upstream industry, firms manufacture solar modules which are the main input for solar power plants. In the remainder of this section, we describe these two industries in more detail.

2.2.1 The downstream industry: solar plants

The downstream industry comprises of solar power plants, or *projects*, which generate electricity that is eventually sold to power distribution companies. To incentivize large upfront investments in the construction of these plants, power distribution companies sign long-term power purchase agreements (PPA), usually 25 years, which guarantee long-run revenues for the solar plant developers. These power purchase agreements can be bilaterally negotiated, or, as in most cases, be awarded through an auction process. In these auctions, participants bid on the rate at which they would sell electricity for the duration of the PPA. The government can then award PPAs to the lowest-cost providers.

Over the years, India has experimented with multiple auction formats to award these power purchase agreements. In early years, it relied on single-stage sealed bid auctions. It has also experimented with auctions where the price of electricity is nominally fixed and firms instead bid on capital subsidy they require from the government to build these solar plants. However, in recent years, the most frequently used auction format has been a two-stage multi-unit English auction. We describe this auction game in detail below.

Suppose the government wants to develop a solar plant of capacity Q_a (say, 1 GW). As such, it will broadcast a call for applications, formally known as a Request for Selection (RFS). Interested developers submit a sealed *initial* bid containing a quantity bid (say, 200

MW) and a price bid (say, INR 4/kWh). In this example, the bidder is proposing to erect a solar plant of capacity 200 MW and sell the electricity generated by it at a rate of INR 4 per kilowatt-hour. Based on some basic financial and techno-commercial criteria[6], as well as the initial sealed bids, a subset of initial respondents are invited to participate in the English auction[7].

This English auction is conducted online. Starting at the initial sealed bid, bidders are allowed to adjust their price bids downwards while holding quantity bids fixed. At all times, the bids (but not the identities) of all other participants are visible to everyone. The auction ends when no player adjusts their bids for a pre-specified duration of time (say, 8 minutes). Capacity allocations are made in the order of increasing price bids until the initial target Q_a is met.

In Figure 2.2, we plot the (weighted) winning bids in these auctions over time. The dramatic decline in these winning bids is reminiscent of the fall in solar module prices over the same period and hints at the potential significance of this input for the solar plant industry. We formally test for the effect of solar module prices on auction bids in **??** and find that it has a large and significant impact on these bids. In the next subsection, we discuss the characteristics of the solar module industry.

2.2.2 The upstream industry: solar modules

A solar module, also known as a panel, is a collection of solar photovoltaic (PV) cells which convert sunlight to electricity. These panels are marketed in terms of watts (W) per piece and serve as the main input for the downstream solar power plants. For instance, a 100 MW solar plant would require 400,000 pieces of 250 W solar modules. At a conservative price of \$100 a piece, that equals an investment of over \$40 million in solar modules alone.

---

[6]This is just to ensure that the bidder would be able to build and operate a plant of the proposed size.

[7]In our data, we do not observe the initial price bid nor do we see the full set of initial respondents. For each auction, we only observe the set of participants invited to the second-stage of the auction process and their final price and quantity bids. As such, in our model and estimation, we disregard the first-stage selection process.

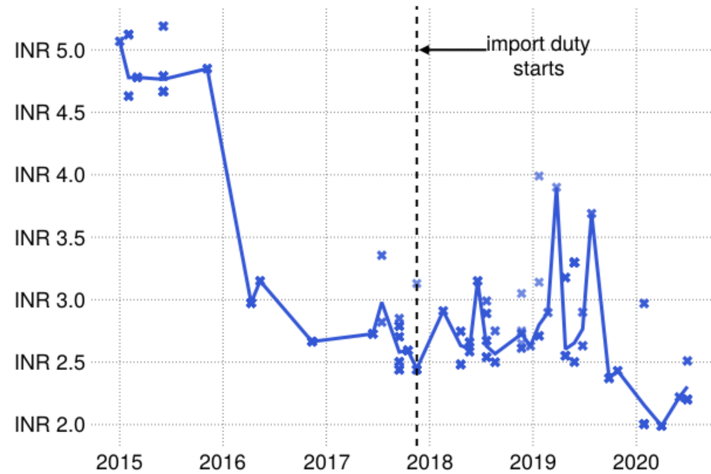


Figure 2.2: *(Weighted) Winning Bids*

Notes. In the above figure, each dot represents a distinct English auction. The height of each dot is the quantity-weighted winning price bid in that auction. The solid line shows these weighted winning prices at a monthly level.

Despite the large demand generated for these modules by the downstream industry, domestic solar module manufacturing has failed to take off. Globally, and in India, solar modules from China dominate this industry. In the first half of the past decade, the market share of Chinese solar modules in the utility-scale solar sector in India was approximately 100%. While there has been an uptick in domestic manufacturing in recent years, Chinese solar modules still command a majority share in the market. Industry experts point to several reasons behind China's relative dominance in the industry including availability of cheap credit, free land, manufacturing subsidies by the Chinese government, and the presence of an "ecosystem" that makes it easier to procure raw materials such as cells, wafers, and poly silicon.

Over the years, the Indian government has also tried to support the domestic module manufacturing industry through two avenues — Domestic Content Requirement (DCR) auctions and Modified Special Incentive Package (MSIP) Scheme. The former is a class of auctions where the winners must procure their solar modules from domestic manufacturers, while the latter is a set of investment incentives to support manufacturing industries[8]. The

[8]Conversations with industry experts suggest that take-up of the MSIP scheme, announced in 2012, has

(a) *Import value (billion USD)* **(b)** *Share of total imports*

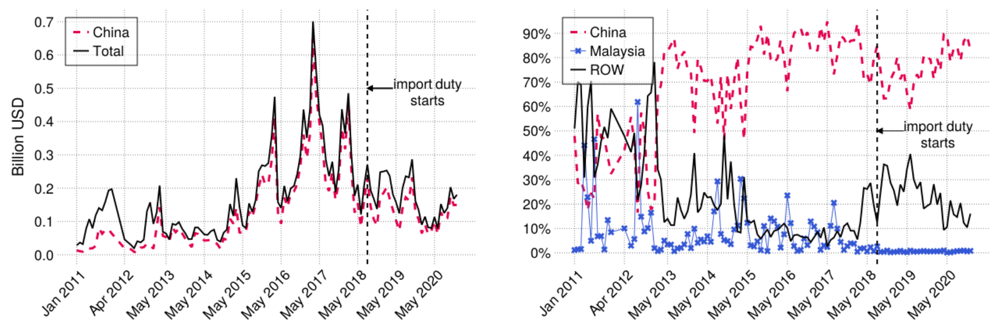


Figure 2.3: *Impact of safeguard duties against China and Malaysia*

Notes. This figure plots monthly imports of products categorized under HS code 854140 into India as recorded under UN Comtrade Database. In the right panel, share of imports are calculated using value of imports recorded in US dollars; ROW refers to value of all imports excluding China and Malaysia.

impact of these policies on domestic solar module production is unclear and not investigated in this paper. Instead, we focus on two big steps taken by the Indian government in recent years to protect the domestic solar module industry. These include (1) import tariffs and (2) production subsidies.

The Indian government first introduced safeguard tariffs against Chinese and Malaysian modules in August 2018. The initial import duty was set at 25% for one year, and then reduced by 5 percentage points every six months until July 2020. At present, the import duty is around 15% but India has already announced plans to levy a customs duty of 40% on solar modules and 25% on solar cells from April 2022 onwards. The new tariffs will apply uniformly to imports from all countries. We show the impact of the initial safeguard tariffs in Figure 2.3. Tariffs affected the composition of imports into the country. The value of imports from Malaysia essentially dropped to zero and there was a small dip in the imports from China; while imports from the rest of the world picked up. However, imports from China continued to make up a large share of total imports into India.

In 2020, the Indian government also announced plans to subsidize manufacturing in the

been very low.

domestic solar module industry. As part of a broader push to boost manufacturing in 13 key sectors, the government has pledged $25 billion under Production Linked Incentive (PLI) Schemes. Of this, approximately $600 million has been earmarked to incentivize production in the solar PV module industry. The PLI scheme is quantity-based i.e. firms will receive subsidy per unit of module produced. The subsidy would be offered to pre-approved plants for a period of five years, and depending on module efficiency and temperature, can range from INR 2.25 to INR 3.75 per watt. Using BNEF world spot prices in early 2020 as a benchmark (see Figure 2.1b), this corresponds to a subsidy rate of 15-25%.

2.3 Data and exploratory evidence

2.3.1 Data

We rely on three main sources of data for our estimation. These include data on (1) government-run auctions, (2) solar plants/projects, and (3) imports of solar modules.

The *auctions* dataset contains auction-level data on the universe of solar auctions held in India. We obtained this dataset from a market research firm, Bridge to India, which aggregates these data from various official and private sources. For each auction, we observe its characteristics such as total open capacity, whether DCR or not, offtaker i.e. the power distribution company that would commit to purchase electricity, and various dates associated with the auction such as announcement date, bid submission date, and results date. Each auction is also linked to detailed bid-level data which includes the price and quantity bids of all bidders, along with the associated outcome of their bid.

We restrict the auctions dataset to those that were held as two-stage multi-unit discriminatory price English auctions[9]. This left us with 63 auctions with over 100 unique bidders. At first glance, this may appear as a highly competitive market but out of these 100+ developers, around 70 have participated in only one auction, and just 20 firms make

[9]Some auctions involved simultaneous bids on multiple tenders which were not disaggregated by our data provider. As such, we observe the same bidder submitting multiple bids for the same RFS. We exclude these auctions too.

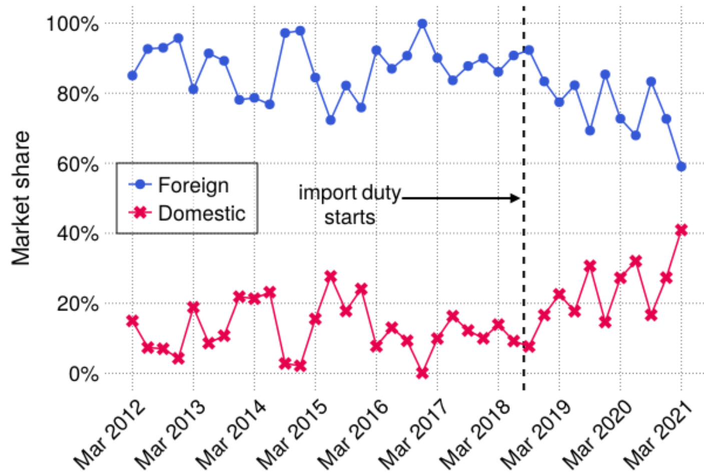


Figure 2.4: *(Smoothed) Quarterly Market Shares*

Notes. This figure uses data on project-level module suppliers to plot market shares of foreign and domestic module manufacturers. For each project, the total project capacity is evenly split over the three months prior to its commissioning and assigned to its supplier. Aggregating over all projects yields a smooth monthly series on sales by each module manufacturer. The above figure plots quarterly market shares derived from this smooth monthly series.

up close to 85% of total capacity won in auctions.

The winning quantity-weighted bids for these auctions are shown in Figure 2.2. This chart excludes single-stage sealed bid auctions held before 2016 which delivered winning bids as high as INR 12 per kWh. The dramatic decline in winning bids evident in the data is strongly correlated with the fall in module prices over the same period. This relationship between module prices and winning bids is the main motivation behind this paper.

The *projects* database is also compiled and provided to us by Bridge to India. It includes solar project-level data on the status of all solar projects in India, including those developed without a formal auction process. The key variables of interest for us are the (1) commissioning date of a project and (2) the identity of the solar module supplier. Using data on 1,970 projects totaling 45 GW in solar capacity for which the identity of the solar module supplier is available, we construct supplier-level market shares in the upstream industry.

It takes about 12-18 months to build a solar plant, and our conversations with industry experts suggest that solar modules are one of the last items to be deployed in a solar plant. We use this feature of the industry to construct a smooth time series for firm-level market shares. For each project, we divide the total project capacity equally over the three months prior to its commissioning. Aggregating over all projects supplied by a given supplier yields a smooth monthly series on sales by each module manufacturer. We aggregate these monthly series by quarter, and present the market shares of foreign and domestic firms in Figure 2.4. Domestic market share is stable at around 10% before the import duty starts, and gradually increases after the duty starts.

Finally, the *imports* database records transaction-level data between 2014 and 2020 on imports of products categorized under HS code 8541[10]. We manually clean text fields describing the product being imported to construct a monthly price per watt series for imported modules. This involved identifying the peak-wattage of each product being imported (e.g. 250 W) and the number of panels being imported, and then dividing total value of the shipment by the total imported watts. As a robustness check, we plot our constructed measure of imported module prices against the spot prices recorded under the BNEF Price Index in Figure 2.5.

2.3.2 Exploratory evidence

Policy interventions such as import tariffs and production subsidies would change the price of solar modules faced by downstream solar plant developers. If the price bids in downstream auctions are very sensitive to solar module prices, these interventions would also affect the solar electricity prices as well as the demand for solar power in the country. While our structural model will capture these relationships, in this section, we provide suggestive evidence for reduced-form relationships between (1) module prices and auction bids, and (2) module prices and demand for modules.

To test for the sensitivity of auction bids to module prices, we regress firm-level price bids on the prevailing import price per watt of modules in the month in which the bids were submitted. We also include a dummy variable to capture the post-import tariffs regime

[10]HS Code 8541 is defined as "Diodes, transistors and similar semiconductor devices; photosensitive semiconductor devices, incl. photovoltaic cells whether or not assembled in modules or made up into panels (excluding photovoltaic generators); light emitting diodes; mounted piezoelectric crystals; parts thereof".

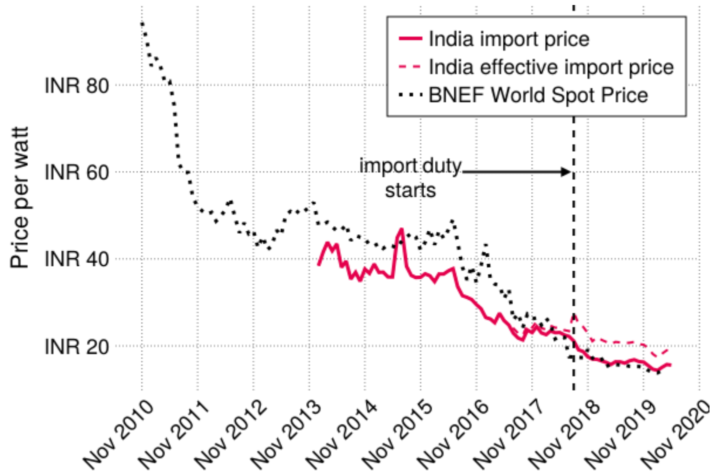


Figure 2.5: *Monthly imported module price per watt*

Notes. This figure plots monthly module prices obtained from transaction-level imports data. The solid red line is the pre-tariff price per watt of imported solar modules, while the dashed red line is the effective *price inclusive of duty that is faced by Indian importers. In dotted black, we show the BNEF world spot price index for reference. The correlation coefficient between our manually constructed price series in solid blue and the BNEF price index is 0.96.*

Table 2.1: *Effect of module prices on auction bids*

	(1) All bids	(2) All bids	(3) Winning bids
Module import price	0.10***	0.15***	0.14***
	(0.01)	(0.01)	(0.01)
Post-August 2018		0.82***	0.96***
		(0.21)	(0.19)
R^2	0.71	0.79	0.71
N	312	312	152

Notes. This table contains the results from regressing price bids on imported module prices per watt before import tariffs are applied. Post-August 2018 is a dummy indicating the period in which solar modules were subject to import tariffs. Standard errors are clustered at the auction-level and given in parentheses.

starting August 2018. In all specifications in Table 2.1, the coefficient on module price remains positive and significant, highlighting a strong pass-through of module prices to equilibrium bids. Our estimate from column (3) suggests that a 1 INR increase in module price is associated with a 0.14 INR increase in price per kilowatt-hour (kWh) of a winning bid.

While this number may seem small, recall that power purchase agreements (PPA) are signed for 25 years. We can use a back-of-the-envelope calculation to gauge the economic significance of the point estimate. For a 100 MW plant, an increase of 0.14 INR per kWh corresponds to an additional payment of INR 360 million in net present value terms[11]. If there is no substitution to domestic module suppliers, the government raises INR 100 million in import duties. Thus, for each rupee raised by the government, the power distribution companies pay 3.6 times more in electricity prices.

Next, we investigate the sensitivity of module demand to module prices. For this, we rely on two data sources. First, we look at monthly imports into India as recorded in the imports data. Second, we use the smooth monthly module consumption series constructed using projects data. For each data source, we calculate quantity as total watts imported (or deployed as in the projects data) in a given month. Our price variable of interest is the post-tariff *effective* import price per watt of modules derived from the imports data. We instrument this price using spot prices for polysilicon, which is a key input used in the production of solar modules. Our estimates from a Two-Stage Least Squares (2SLS) procedure suggest a highly elastic demand for solar modules. Both quantity variables outlined above deliver an elasticity between -2.2 and -3. We present these estimates and the associated first-stage estimates in Table 2.2.

In our setting, lower module prices lead to higher demand for solar modules through an increase in auction activity by various state and central agencies including power distribution companies. When module prices fall, price bids in solar auctions also fall. This lowers the cost of solar energy to government-run power distribution companies who may respond by organizing more auctions or auctions with larger open capacities, thus leading to higher demand for solar modules. Indeed, in our data, auctions with much larger open capacities were organized in the latter half of the past decade when module prices were much lower relative to the first half.

[11]The additional payment is calculated as $\sum_{\tau=1}^{25}(1+r)^{1-\tau} \times c.u.f \times 0.14 \times 100 \times 10^3 \times 365 \times 24 = 362,981,086$. In the expression above, we set $r = 5\%$ and $c.u.f. = 0.2$, which is a constant that captures the capacity utilization rate of solar plants.

Table 2.2: *Effect of module prices on demand for solar modules*

	(1) First Stage	(2) 2SLS (Imports)	(3) 2SLS (Projects)
Polysilicon USD/kg	0.631***		
	(0.030)		
India effective import price		-2.972***	-2.221***
		(0.472)	(0.534)
Observations	75	75	75
R^2	0.627	0.473	0.389
F stat.	434	40	17

Notes. This table shows the price elasticity of demand obtained from a Two-Stage Least Squares (2SLS) regression of quantity on effective import price of modules, using price of polysilicon as an instrument. Column (2) uses quantity of imports into India as the dependent variable, while column (3) uses the (smooth) monthly solar module consumption derived from the projects data.

2.4 Model

In this section we outline the structural model. We use this model to compute the market equilibrium and the associated welfare statistics under different policy regimes.

2.4.1 Upstream Industry Model

Firms in the upstream industry manufacture and supply solar modules. There are two kinds of producers, home and foreign. Marginal cost of production, for firm j in period t is given by c_{jt}. The *effective* market price per watt (inclusive of import duties) for a solar module in time period t is P_t, and total demand from downstream developers is denoted by Q_t. Government can impose ad-valorem tariffs τ_h and τ_f on home and foreign producers, where a subsidy is treated as a negative tariff. Module manufacturers play a Cournot game, maximizing their static profits while taking demand and actions of other manufacturers as given. The objective function of a manufacturer j facing a tariff of τ_j is given by

$$\max_{q_{jt}} \pi_{jt}(q_{jt}) = \max_{q_{jt}} (1 - \tau_{jt}) P_t(Q_t) q_{jt} - c_{jt} q_{jt}. \tag{2.1}$$

Furthermore, we assume that the demand function is isoelastic with constant price elasticity α. That is, downstream demand for solar modules has the form

$$Q_t = \mu P_t^{\alpha}$$

We interpret this elasticity as the composite elasticity of two structural objects — the sensitivity of power distribution companies' demand for solar power to winning solar bids, and the sensitivity of winning bids to the price of solar modules. Formally,

$$\alpha = \frac{\partial Q_t}{\partial b_t} \cdot \frac{\partial b_t}{\partial P_t}$$

where b_t is the weighted winning bid from a solar auction. Using our model for the downstream industry, we can compute $\frac{\partial b_t}{\partial P_t}$. Then, using an estimate for α, the first term, $\frac{\partial Q_t}{\partial b_t}$ can be recovered and used to calculate consumer surplus statistics.

2.4.2 Downstream Industry Model

Firms in the downstream industry participate in auctions and develop utility-scale solar plants. We develop a static model of downstream industry and abstract away from dynamic considerations[12].

In the auction game, the auctioneer announces an auction a in time period t, with an open capacity Q_{at}. There are two stages to the game after the announcement. In the first stage firms decide whether to participate or not. They also decide the size of the solar plant that they want to build, subject to an upper bound of Q_{at}. We call this the participation stage. In the second stage the set of entering firms take part in a reverse English auction[13].

[12]There are multiple dynamic aspects of the industry that we abstract away from. Firms when bidding may take into consideration the effect of winning an auction on construction costs for future auctions due to either learning by doing (which reduces costs) or capacity constraints (which increases costs). Further, the payoff from winning an auction in our model is the NPV of expected profits flow from the project. In reality operating projects over a long horizon will add additional complexities.

[13]Firms that choose to participate in the first stage may not be selected to take part in the second stage depending on their initial tariff bid. Further, the initial tariff bid constrains the second stage bids. As we do not observe the firms which are not selected and the first stage bids, we do not model that part.

Stage 1: Participation Stage

In the first stage firms decide whether to enter or not. Entry entails an entry cost κ. Before entry, firms observe a vector of state variables $\boldsymbol{s_{iat}} = \{\boldsymbol{s_{at}}, \boldsymbol{\eta_{iat}}\}$, where $\boldsymbol{s_{at}}$ is a common state vector observed by all the firms in the industry and η_{iat} is private information. $\boldsymbol{s_{at}}$ consists of input prices, import duties and production subsidies, and the announced auction capacity Q_{at}. $\boldsymbol{\eta_{iat}} = \{\eta^0_{iat}, \eta^1_{iat}\}$ are the idiosyncratic entry cost shocks. Firm i receives an entry cost shock of η^1_{iat} if it enters and η^0_{iat} if it stays out.

In addition to the entry decision the firm also decides the quantity bid q_{iat}. Quantity is chosen to maximize expected profits from entry. The participation stage problem can be written as,

$$\max_{e_{iat} \in \{1,0\}} (1 - e_{iat}) \cdot \eta^0_{iat} + e_{iat} \cdot g(q^*_{iat}, \boldsymbol{s}_{iat})$$

where $e_{iat} \in \{0, 1\}$ is the participation decision and $g(q^*_{iat}, \boldsymbol{s}_{iat})$ is the benefit from entry, net of costs. It is defined as

$$\begin{aligned} g(q^*_{iat}, \boldsymbol{s}_{iat}) &= \max_{q_{iat}} \; g(q_{iat}, \boldsymbol{s}_{iat}) \\ &= \max_{q_{iat}} \; \mathbb{E}\, \pi_{iat}(q_{iat}, \boldsymbol{s}_{at}) - \kappa + \eta^1_{iat}. \end{aligned}$$

$\mathbb{E}\, \pi_{iat}(q_{iat}, \boldsymbol{s}_{at})$ is the expected profits for firm i from auction a conditional on entry with a quantity bid q_{iat}.

Stage 2: Auction Stage

The auction is organized as a reverse English auction. The auction takes place online where the participants can see the current tariff bids and the quantity bids of all the other players.

The tariff bid b_{iat} is the price at which the firm, if selected, will supply electricity to the grid for the next 25 years. Thus, winning the auction is associated with a revenue stream over the next 25 years. We compute the payoff from winning 1 watt capacity in the auction by calculating the net present value of the revenue stream as follows,

$$NPV(b_{iat}) = \sum_{\tau=1}^{25} (1+r)^{1-\tau} \times b_{iat} \times c.u.f. \times 24 \times 365 \tag{2.2}$$

where r is the interest rate and $c.u.f$ is the capacity utilization factor which adjusts for the fact that a solar plant does not generate electricity all hours of the day. We set r to be 0.1 and the c.u.f. to be 0.2. Given the monotonic relationship between b_{iat} and $NPV(b_{iat})$, we can redefine the auction where participants submit NPV_{iat} bids. A lower NPV_{iat} bid means that the firm requires a lower revenue stream and therefore is more attractive to the auctioneer.

We impose a button press auction model on our data. A button press auction starts at a high bid level with all the participants willing to supply solar energy at the high bid. The bid level falls over time and participants can opt out at any point of time. Participants will opt out when the current bid is not enough to cover the cost of construction. The auction ends when the set of remaining participants are just sufficient to build the desired capacity Q_{at}. Under the model it is a dominant strategy for a firm to drop out when the benefit from staying in is equal to the cost of building the plant.

Proposition 1. *In a button press auction model, it is a dominant strategy for a firm to exit when the announced NPV is equal to the cost of constructing the plant.*

We parameterize the *per unit* cost function for a solar plant to depend on prevailing module prices as well as a private firm-specific unobserved cost shock, ϵ_{iat}. We assume that the private cost shock is only realized for auction participants. The cost per-watt is given by

$$C_{iat} = \beta \text{Module Price Per Watt}_t + \epsilon_{iat},\ \epsilon_{iat} \sim \mathcal{N}(0, \sigma^2). \tag{2.3}$$

Combining Equation (2.2) and Equation (2.3), profit from a solar plant of capacity q_{iat} is,

$$\pi_{iat}(q_{iat}, b_{iat}; \boldsymbol{s_{at}}, \beta, \sigma) = q_{iat} \cdot (NPV_{iat} - C_{iat})\,.$$

where β, σ are the second stage parameters to be estimated.

2.4.3 Equilibrium and Timing

The timing and equilibrium of the model are as follows. At the start of time period t, the government announces the set of tariffs on foreign and domestic upstream firms.

Upstream firm j takes the tariff it faces τ_{jt}, the demand elasticity α, and actions of other producers $-j$ as given and chooses quantity q_{jt} to maximize profits. Total quantity determines the price P_t of solar modules. After observing the market price, the auctioneer, announces an auction with capacity Q_{at}. In equilibrium, $Q_{at} = \sum_j q_{jt}$.

In the downstream market, developers observe prices P_t and quantity Q_{at} and make entry and quantity decisions. The set of developers that enter draw cost shocks from $\mathcal{N}(0, \sigma^2)$ and participate in the auction. As entry of other developers and cost-shock draws are unknown at the time of entry, developers form expectations on realized profits. The entry and quantity decisions are optimal given the expectations and the expectations are correct in equilibrium.

2.5 Estimation & Results

2.5.1 Upstream Industry Model

Given our Cournot assumption for the upstream industry, recovering firm-level marginal costs is straightforward. The upstream firms take the demand function and the associated price elasticity of demand, α, as given and choose quantity to maximize their profits. The FOC for firm j is given by,

$$(1 - \tau_{jt}) P_t \left(1 + \frac{\mathrm{ms}_{jt}}{\alpha}\right) = c_{jt},$$

where ms_{jt} is the market share for firm j and c_{jt} is the constant marginal cost of production for firm j. We take price elasticity of module demand, α, from column (3) of Table 2.2. Market shares are calculated using the projects database. With α and ms_{jt}, the FOC allows us to back out marginal costs c_{jt} for all firms j.

For our baseline and counterfactual exercises, we take market shares in the fourth quarter of 2019 as representative and estimate marginal costs for firms active in this quarter. Our

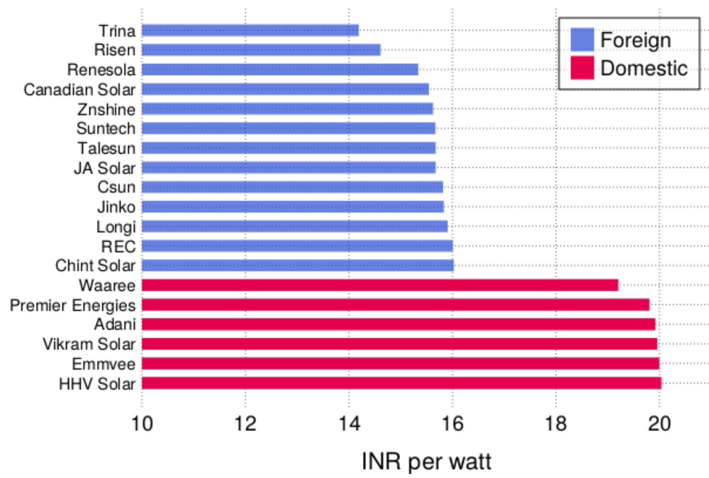


Figure 2.6: *Estimated Marginal Costs*

Notes. The above figure plots the estimated marginal costs of solar module producers, using market shares from 2019 Q4.

estimates of marginal costs can be found in Figure 2.6.

2.5.2 Downstream Industry Model

We estimate the downstream parameters in three steps. In the first step we estimate the parameters governing the cost function (β, σ), in the second step we estimate the expected profits from participation, and in the final step we estimate the entry cost parameters.

Step 1: Cost Parameters

We use the observed auction bids to back out the underlying cost parameters, β and σ. β is the coefficient on module price in the cost specification and σ is the standard deviation of firm-level idiosyncratic cost shocks.

Given Proposition 1, the observed NPV bid of a losing firm will be equal to its cost per watt of building the solar plant. However, we cannot directly estimate the cost function (2.3) using losing bids as it is not a random sample. In particular, firms with a higher draw of cost shock (ϵ) are more likely to end up in the set of losing firms.

We use order statistics to take the selection problem into account during estimation.

Consider a guess of parameters (β^g, σ^g). For firm i in auction a the cost shock under the parameter guess is given by $\epsilon^g = NPV(b_{iat}) - \beta^g \text{Module Price}_t$. Let the firm i's bid be at rank j out of n total bids in the auction. Then the cost shock realization will also follow the same order, and the likelihood will be equal to the likelihood of j^{th} order statistic out of n draws from $F \sim \mathcal{N}(0, (\sigma^g)^2)$. This likelihood is given by,

$$n f(\epsilon) \binom{n-1}{k-1} F(\epsilon)^{j-1} (1 - F(\epsilon))^{n-j}.$$

Therefore, for a guess of parameters (β^g, σ^g), we can compute the likelihood of the data. We search over the parameter values to maximize the likelihood. The estimated parameters and bootstrapped standard errors are given in Table 2.3.

Step 2: Expected Profits

Next, we estimate the expected profits from entry, $\mathbb{E}\, \pi_{iat}(q_{iat}, \boldsymbol{s}_{at})$. The profits depend on the set of firms that choose to participate and the cost shocks of all the participants. As analytical characterization of $\mathbb{E}\, \pi_{iat}(q_{iat}, \boldsymbol{s}_{at})$ is hard we rely on a simulation estimator. To simulate auction play we need firm policy functions associated with the decision to enter and the amount of quantity bid. We estimate these policy functions empirically using observed entry and quantity decisions. To estimate the entry probability function we use a Probit model regressing the observed entry choice on the vector of state variables and year dummies. We estimate the quantity policy function using a linear model, regressing observed quantity bids on the vector of state variables and year dummies. We limit the predicted quantity to the total auction capacity Q_{at}.

Across simulations, the variation in outcomes will be driven by variation in set of firms entering (due to different draws of η_{iat}) and variation in cost shock realizations. We run a simulation as follows. We draw the set of firms from the set $\{\mathcal{N} \backslash i\}$ using our estimated entry probability functions. Firm i enters with probability 1. Next, we use our estimated quantity policy function to draw quantity bids for all the firms. Finally, we draw the cost shocks from the distribution $\mathcal{N}(0, \hat{\sigma}^2)$. The cost shocks, along with $\hat{\beta}$, determine the cost to

Table 2.3: *Downstream cost function parameters*

	(1) Estimate	(2) Standard error	(3) 95% Confidence interval
Constant	-18.81	7.85	[-34.99, -6.40]
Module (effective) import price	2.80	0.30	[2.32, 3.44]
Std. dev. of cost shock, σ	13.49	2.48	[8.66, 17.89]
Entry cost, κ	5.63	1.02	[3.65, 7.58]

Notes. This table presents the estimates of parameters governing the cost structure of the downstream industry. Module (effective) import price is the price of modules, inclusive of duties, faced by downstream firms. σ^2 is the variance of the firm-specific unobserved cost shock that enters the cost function of downstream developers. κ is the entry cost paid by downstream developers to enter solar auctions. Standard errors are calculated using 500 bootstraps in which we draw auctions with replacement.

each firm for building capacity q_{iat}. In line with our button press auction model assumption, the lowest cost bidders are allocated capacity until open capacity Q_{at} is filled, and the price paid is the lowest bid (cost) amongst the set of losers. If firm i is amongst the set of winners then the profits are the difference between its cost realization and the highest winning bid. If not, firm i earns a profit of 0. We average across 2000 auction plays to compute $\mathbb{E}\,\pi_{iat}(q_{iat}, \boldsymbol{s}_{at})$.

Step 3: Entry Costs

In the final step we estimate the auction entry cost κ. We assume that the entry cost shock parameters $\{\eta^1_{iat}, \eta^0_{iat}\}$ are drawn from a T1EV distribution. Then the ex-ante probability of entry for firm i in auction a is,

$$\Pr(e_{iat} = 1 | s_{at}) = \frac{\exp(\mathbb{E}\,\pi_{iat}(q_{iat}, \boldsymbol{s}_{at}) - \kappa)}{1 + \exp(\mathbb{E}\,\pi_{iat}(q_{iat}, \boldsymbol{s}_{at}) - \kappa)}. \tag{2.4}$$

The empirical entry policy function, $\Pr(e_{iat} = 1 | s_{at})$, is estimated while computing expected profits, $\mathbb{E}\,\pi_{iat}(q_{iat}, \boldsymbol{s}_{at})$, as discussed under Step 2. With these two components in hand, we can recover κ as the value that minimizes the difference between LHS and RHS of 2.4. We estimate κ to be 5.63 as shown in Table 2.3.

2.5.3 Estimating Demand for Solar Power

The downstream industry generates electricity that is purchased by power distribution companies. In essence, this demand for solar power by power distribution companies is what generates the demand for solar modules upstream. If equilibrium solar power prices, determined by auctions, are high, power distribution companies may demand less solar power, and hence fewer solar modules would be purchased. We use this relationship to recover the implicit price sensitivity of power distribution companies to the equilibrium solar power prices.

Specifically, we first calculate how winning bids respond to module prices. We start by taking observed module prices at the end of 2019 as given and then run the downstream auction model to get the downstream solar power prices i.e. the winning NPV bid. Next, we increase the upstream module price by 1% and rerun the downstream model to compute the impact on downstream prices. This allows us to compute the elasticity of downstream bids with respect to upstream module prices. We denote this by ε_{bid}.

Now, the price elasticity of power distribution companies with respect to winning bids, denoted by ε_{power}, can be recovered as follows

$$\varepsilon_{power} = \frac{\alpha}{\varepsilon_{bid}}$$

where α is the price elasticity of modules with respect to module prices. We use this derived elasticity, ε_{power}, to compute changes in consumer surplus under different policy regimes.

2.6 Counterfactuals

The welfare effects of either of the two policy interventions – tariffs or subsidies – can be decomposed into four components. First, these interventions generate *revenues for the government*. Tariffs raise government revenues while subsidies generate negative revenues since the government makes payments to domestic producers. Second, both of these interventions increase the *profitability of the upstream sector*, by direct transfers in the case of subsidies and through protection from foreign competitors in the case of tariffs. Third, the policies affect *profitability of downstream developers*. Subsidies reduce input prices faced by developers which increases their profits while tariffs raise prices which reduce profits. Finally, the *consumer surplus* of power distribution companies goes down in the case of tariffs, while it goes up in the case of subsidies.

Our baseline scenario is one in which the auctioneer conducts an auction with total open capacity of 1 GW with both tariffs and subsidies set to zero[14]. We compare the baseline equilibrium with two sets of counterfactuals. In the first set of counterfactuals, the policymaker aims to expand the *size* of the domestic solar module sector. Specifically, we consider two scenarios in which the total output of the domestic module producers is expanded to 0.2 GW and 0.4 GW using appropriate levels of tariffs and subsidies. In the second set of counterfactuals, the policymaker aims to expand the *market share* of the domestic sector module sector to 20% and 40%. For each counterfactual, we find the level of tariffs and subsidies required to achieve the objective, and compute the equilibrium at that level of tariff and subsidy. The auction size, originally set at 1 GW, will change under tariffs and subsidies depending on price responses and elasticities. Then we compute the four welfare components under tariffs and subsidies, and compare them with the welfare components under the baseline.

Next we discuss how we compute each of the four components of welfare.

For each level of tariff and subsidy, we need to compute (a) the equilibrium market price of solar modules implied by the upstream model, (b) the set of upstream firms that will continue to supply positive quantities of solar modules, (c) the equilibrium capacity offered in the downstream auction, and (d) the equilibrium winning bid in the downstream auction. This equilibrium quantities will allow us to compute all four components of welfare described above.

We first focus on the upstream industry. At the observed import tariff in 2019 Q4, all

[14]The results will be the same qualitatively if instead of a 1 GW auction we have an auction with higher capacity or have multiple auctions.

module producers find it profitable to supply positive quantities of modules. But at different levels of tariffs and subsidies, some foreign and domestic firms may want to stop producing. For our counterfactuals, to arrive at the set of firms that would continue to supply modules, we follow an iterative procedure. For a given value of tariff or subsidy, we rank firms in the order of increasing *effective* marginal costs, net of tariff/subsidy level. Next, we compute the equilibrium price implied by the Cournot model where only the lowest effective marginal cost producer produces. At this equilibrium price, if it is profitable for the single firm to produce, we add the firm with the second lowest effective marginal cost and repeat the exercise making sure that all operating firms have non-negative profits. We keep on adding firms, going from lowest to highest effective marginal cost, until adding one more firm makes it unprofitable to produce for at least one firm. This process gives us the new equilibrium market price of solar modules as well as the set of firms which can profitably produce at the given level of subsidies and tariffs. All remaining firms are assigned a market share of zero.

Next, we estimate how the auction quantity changes under different levels of tariffs and subsidies. In our baseline scenario, auction capacity is fixed at 1 GW and we know the prevailing module price. Under counterfactual values of tariffs and subsidies, we can recover the market price of solar modules implied by the Cournot model as described above. Using the percent change in module prices, along with the module demand elasticity α, we can compute the percent change in auction capacity offered. This yields the new equilibrium capacity auctioned off.

The estimated counterfactual auction capacity and the new equilibrium price of solar modules can be used to compute the profits of all the domestic solar module producers. Government revenues are also straightforward to compute. Under tariffs, for every unit of output sold by foreign firms, the government receives τP_t in tariff revenue. Under subsidies, for every unit of output sold by domestic firms, the government pays τP_t under production subsidies.

To compute the equilibrium winning bid in the downstream auction, we use the simu-

Table 2.4: *Counterfactual Results*

	1 ΔGovt. Rev.	2 ΔUpstream Profits	3 ΔAuction Profits	4 ΔCS	5 ΔTotal
A. Upstream Domestic Size at 0.2 GW					
With import tariff	0.92	0.19	-2.94	-18.28	-20.11
With production subsidy	-0.81	0.15	0.04	0.29	-0.34
B. Upstream Domestic Size at 0.4 GW					
With import tariff	0.00	0.47	-4.01	-23.30	-26.83
With production subsidy	-1.79	0.32	0.09	0.56	-0.81
C. Upstream Domestic Share at 20%					
With import tariff	1.71	0.00	-1.44	-9.67	-9.40
With production subsidy	-0.82	0.15	0.06	0.30	-0.32
D. Upstream Domestic Share at 40%					
With import tariff	1.12	0.14	-2.89	-17.39	-19.02
With production subsidy	-1.86	0.34	0.11	0.62	-0.79

Notes. This table reports changes in aggregate statistics under different counterfactual scenarios relative to a baseline scenario in which tariffs and subsidies are set to zero. In the baseline, the auctioneer conducts an auction of total open capacity 1 GW. In the absence of import tariffs, the market share of domestic firms is zero as their marginal costs are too high to operate with non-negative profits. In panels A through D, we present the change in aggregate statistics for different potential objectives of the policymaker. For each objective, we estimate the level of tariff or subsidy it would take to achieve it, and then compute the equilibrium at that level of tariff or subsidy using our estimated structural parameters. Note that auction profits are just average sum of profits of all winners computed using our simulation estimator; these exclude auction entry costs.

lation estimator outlined in Section 2.5. The size of the downstream auction and module input price are set from the upstream equilibrium. All other variables are set at the baseline level. Within a particular simulation we use the entry and quantity policy function to draw the set of firms that will participate in the auction and their quantity choices. The firms then draw cost shocks from the estimated distribution. In line with the auction model, capacity is sequentially allocated to the lowest cost bidders until auction capacity is fulfilled. Auction price is the cost of the lowest cost loser. We run the simulation many times, sum over profits of all winners in each auction and then average to get our estimate of developers' profits.

The final welfare component is the consumer surplus. This is estimated using the elasticity ε_{power} discussed in Section 2.5. Assuming a constant elasticity demand function, we compute change in consumer surplus, when the winning bid goes from b_0 to b_1 as

$$\Delta CS = \frac{b_1 Q_1 - b_0 Q_0}{1 + \varepsilon_{power}}.$$

where we take b_0 and b_1 to be the average winning bid in our simulated auctions.

Results. We present results from our main set of counterfactuals in Table 2.4. In each of the panels A through D, we take the indicated objective of the policymaker as given, and calculate the level of import tariff and production subsidy it would take to achieve it. At these implied levels of tariff and subsidy, we compute the four components of welfare described above. Under all objectives, we find both tariffs and subsidies to be welfare-reducing, but find welfare loss under tariffs to be much larger than the welfare loss under subsidies. To achieve a total combined output of 0.2 GW by domestic module producers using just import tariffs leads to a decline of about 20 billion INR, while doing so using production subsidy only lowers welfare by about 0.3 billion INR. In relative terms, welfare loss under tariffs is 60 times the loss under subsidies. Most of this decline in welfare comes through the decline in consumer surplus of the "auctioneer" or the downstream purchasers of solar energy. The results are qualitatively similar when we consider scenarios where the policymaker wishes to expand the *market share* of domestic firms.

The drastic difference in the loss in welfare under tariffs and subsidies might also be driven by the presence of market power in the upstream sector, which we explicitly account for in our structural model. Economic theory suggests that subsidies can correct for the presence of market power distortions. To further investigate the role of this channel, we consider a baseline scenario where the upstream sector is a duopoly consisting of a representative foreign firm and a representative domestic firm. We aggregate the market shares accordingly and recover the marginal costs of these two firms, as implied by the Cournot model. Finally, we run the same counterfactuals as above and present the results in Table 2.5. We continue to find large welfare losses under tariffs but estimate welfare gains under production subsidies. The results suggest that appropriately accounting for market power may be crucial when comparing outcomes from different policy interventions.

Table 2.5: *Counterfactual Results Under a Duopoly*

	1 ΔGovt. Rev.	2 ΔUpstream Profits	3 ΔAuction Profits	4 ΔCS	5 ΔTotal
A. Upstream Domestic Size at 0.2 GW in a Duopoly					
With import tariff	2.81	0.71	-3.65	-8.80	-8.93
With production subsidy	-0.73	0.29	0.91	2.58	3.05
B. Upstream Domestic Size at 20% in a Duopoly					
With import tariff	2.77	0.28	-1.97	-4.18	-3.09
With production subsidy	-1.18	0.48	1.06	2.93	3.29

Notes. This table reports changes in aggregate statistics under different counterfactual scenarios starting with a hypothetical baseline scenario in which the upstream sector is a duopoly and in which tariffs and subsidies are set to zero. In this hypothetical duopoly, there is one representative foreign firm and one representative domestic firm. We aggregate their market shares accordingly and re-derive marginal costs implied by the Cournot model. As in our main counterfactuals, in the baseline, the auctioneer conducts an auction of total open capacity 1 GW. In panels A and B, we present the change in aggregate statistics for different potential objectives of the policymaker. For each objective, we estimate the level of tariff or subsidy it would take to achieve it, and then compute the equilibrium at that level of tariff or subsidy using our estimated structural parameters. Note that auction profits are just average sum of profits of all winners computed using our simulation estimator; these exclude auction entry costs.

2.7 Conclusion

A number of reasons including supply chain fragility, strategic importance, and infant industry arguments may motivate policymakers to undertake steps to protect domestic producers against foreign competitors. Standard theory models recommend production subsidies as more cost-efficient than tariffs on foreign firms, yet tariffs remain very popular around the world and are implicitly favored over subsidies by WTO policies. Using the Indian utility-scale solar sector as our empirical setting, in this paper, we quantify just how much worse tariffs can be relative to production subsidies.

To answer our question, we develop and estimate a structural model of two vertically-linked industries in the Indian utility-scale solar sector. In the upstream industry, firms produce solar modules or panels, while in the downstream industry, firms purchase these solar modules and develop large utility-scale solar power plants. In recent years, the Indian government has taken steps to protect the domestic upstream industry through import tariffs on solar modules as well as production subsidies for domestic producers of solar modules.

Our structural model allows us to compare the welfare costs of these two interventions in the upstream industry on the entire sector, and our findings show that tariffs are significantly more welfare-reducing than subsidies.

Chapter 3

Build Public Hospitals or Subsidize Care at Private hospitals? An Equilibrium Analysis of Public Healthcare Strategies[1]

3.1 Introduction

What are the equilibrium effects of popular large-scale public healthcare provision strategies in developing countries? Two popular strategies include directly providing care through tax-funded public hospitals, which offer services at low prices but often suffer from congestion and inadequate resources, and subsidizing care for low-income patients at private hospitals. Both policies can have direct and indirect effects by altering equilibrium behavior of both patients and providers, and they may also have important distributional effects by differentially impacting patients from different income groups. In this paper, we develop and estimate a structural model that allows us to characterize the net effects of these

[1] Co-authored with Ljubica Ristovska

interventions at each point along the income distribution of patients.

We focus on India where the government has made significant investments in developing an extensive network of low-cost public hospitals. As an alternative approach, India has also recently introduced a government-run insurance scheme that subsidizes healthcare access for low-income individuals at private hospitals. To compute and contrast the equilibrium effects of these interventions, we zoom in on a particular ailment, diabetes, which is a growing epidemic in India and a leading cause of hospitalization.

To analyze the equilibrium effects of the two public healthcare provision strategies, we develop a static structural model with a demand-side and a supply-side. The demand-side consists of patients in need of hospitalization, differentiated by income levels, choosing between public and private hospitals. In making this choice, patients take into account both price and non-price characteristics of each type of hospital in the market. The supply-side consists of hospitals setting prices. Public hospitals use marginal-cost pricing, while private hospitals maximize profits given demand and marginal costs; the latter are allowed to depend on hospital characteristics. The model is tailored to a specific disease category, in our case, diabetes, but can be generalized to other diseases as well.

We estimate this model using observed choices of diabetes patients who sought hospitalization, as reported in the 75th round of the National Sample Survey (2017-18). Our estimated demand-side parameters are reasonable – patient utility is declining in prices, but the sensitivity to price is lower for patients with higher incomes. We also find that patients prefer private hospitals, with a stronger preference for private hospitals for higher-income patients. Finally, we find that patient utility is increasing in the number of beds per unit population i.e. patients prefer less congestion. The estimated supply-side parameters are also reasonable. Ceteris paribus, the marginal cost of providing care at private hospitals is slightly higher than doing so at public hospitals. Furthermore, the marginal cost of providing care is increasing in congestion or declining in the number of beds per unit population.

Main results. With structural parameters in hand, we simulate counterfactuals to understand the equilibrium effects of different public healthcare provision strategies. Specifically, we consider two main counterfactual scenarios: (i) an increase in the number of beds at government hospitals, and (ii) a subsidy to patients, with income below a given threshold, for receiving treatment at private hospitals. We run the counterfactuals for two distinct markets: a low-density market (Jharkhand) and a high-density market (Delhi), where *density* is the number of government hospital beds per unit population.

In the first counterfactual, where we increase the number of government hospital beds, our findings indicate that all consumers along the income distribution are better off in both markets. Lowering congestion reduces the marginal cost of providing care at government hospitals, which in turn leads to lower prices and increased utility for patients. In equilibrium, the price of care at private hospitals also falls as the increase in the relative attractiveness of government hospitals puts competitive pressure on private hospitals. The welfare gains are larger for lower-income patients as we estimate them to have a higher marginal utility of income, and they now benefit from lower prices at both government and private hospitals. Furthermore, these gains are higher in the low-density market than in the high-density market for all patients – a result due to the fact that the low-density market had both higher congestion and higher prices in the baseline scenario.

In the second counterfactual, where we examine a policy in which patients with income below a specified threshold receive a subsidy on the price of care at private hospitals, our results reveal that eligible patients indeed experience welfare gains. However, the gains are relatively modest compared to the indirect negative impact on non-eligible patients, who now face higher prices at private hospitals due to the equilibrium response of these hospitals. Our estimates of utility parameters suggest that higher-income patients prefer private hospitals, so higher equilibrium prices at private hospitals lead to welfare losses that scale up with income. Overall, our findings highlight that this policy can create winners and losers, and policymakers should carefully consider the potential distributional effects of such an intervention.

To arrive at an optimal policy, we first need to account for the costs of implementing the first policy, i.e. expanding the number of beds at government hospitals. Although we do not estimate these, a policymaker can still use our results to arrive at an optimal policy. As an illustration, suppose the policymaker wants to improve health outcomes for patients with income below the 10th percentile of the income distribution. Under each policy, we can compute the gain in welfare for patients in this income group. In the second counterfactual, we can also compute the total cost of the policy, which equals subsidy expenses minus the change in welfare for all patients not in the target income group. This gives us the cost to benefit ratio of the second policy. The policymaker would prefer to expand the number of beds at government hospitals as long as the cost of doing so (net of the impact on non-targeted patients) yields a cost to benefit ratio that is lower than the cost to benefit ratio of the second policy.

Related literature. This study contributes to three main strands of literature. First, it relates to a number of papers comparing public and private healthcare provision (Villa and Kane, 2013; Basu, Andrews, Kishore, Panjabi and Stuckler, 2012; Frakes, Gruber and Justicz, 2020; Chan, Card and Taylor, 2022). This literature focuses on the differences in the quality and efficiency of public vs. private healthcare delivery systems. Our paper extends this literature by empirically examining the equilibrium effects of improving access to either government or private hospitals through government interventions.

Second, our work adds to a large body of literature focused on improving our understanding of healthcare systems in developing countries. This literature has examined healthcare access and quality, specifically in the context of low-income countries (Banerjee, Deaton and Duflo, 2004; Das and Hammer, 2014; Mills, 2014; Das, Holla, Mohpal and Muralidharan, 2016b; Das, Chowdhury, Hussam and Banerjee, 2016a; Banerjee, Finkelstein, Hanna, Olken, Ornaghi and Sumarto, 2021; Das and Do, 2023). Our paper complements this literature by providing a framework to investigate the equilibrium effects of different healthcare provision strategies in a context where financial and physical access to healthcare is still limited. Through the lens of this framework, we generate insights that can inform the design of public healthcare provision strategies in similar contexts.

Lastly, this paper connects to the literature on structural models of patient choice and provider behavior in healthcare markets (Ho, 2009; Gaynor, Ho and Town, 2015; Ho and Lee, 2017). We develop a structural model that incorporates insights from this literature, but is tailored to our specific context involving non-profit-maximizing public hospitals in a developing-country setting.

The rest of the paper is organized as follows. In Section 3.2, we provide background on the healthcare system in India. In Section 3.3, we describe our data and present summary statistics. In Sections 3.4 and 3.5, we present our model and estimation strategy. In Section 3.6, we discuss results from counterfactual simulations, and in Section 3.7, we conclude.

3.2 Public Healthcare Provision Strategies in India

India's healthcare system has undergone considerable transformation over the past few decades. In the early years following independence, governmental efforts were primarily directed toward addressing public health challenges such as communicable diseases, malnutrition, and maternal and child health. In order to meet these objectives, a comprehensive network of government clinics and hospitals was established throughout the nation, which continues to form the backbone of the Indian healthcare system. These tax-funded public hospitals provide affordable healthcare to the Indian population, particularly catering to the needs of low-income households. However, government hospitals grapple with numerous challenges, including underfunding, inadequate infrastructure, shortage of medical personnel, and long waiting times.

In parallel, the past few decades have seen the emergence of a large number of private hospitals in India. These institutions typically offer advanced diagnostic and treatment facilities and reduced waiting times, although at considerably higher prices than government hospitals. Consequently, private hospitals have gained popularity among the burgeoning middle and upper classes, who can afford to pay for their services out-of-pocket. While private hospitals could potentially alleviate some of the pressure faced by government hospitals, their services remain unaffordable for a majority of the population, leaving public hospitals as the primary healthcare providers for most Indian citizens.

In recent years, the Indian government has taken several measures aimed at enhancing healthcare access and affordability. Given the overcrowded and often inadequate conditions in government hospitals, one proposed solution involves providing subsidized care to specific segments of the population at private hospitals. An important policy intervention in this regard is the National Health Protection Scheme (NHPS), also known as Ayushman Bharat, launched in 2018. This initiative offers health insurance coverage of Rs. 500,000 (USD 6,000) per family per year for secondary and tertiary care hospitalization. Under the scheme, eligible households can access treatments at pre-approved private hospitals, with the government directly reimbursing the hospital for the treatment while the patient pays only a minimal co-payment. The program aims to cover 10% of India's population and is slated for full implementation by 2022.

Given its scale, the NHPS has the potential to significantly alter health outcomes for both eligible and non-eligible populations. But the empirical research on the equilibrium effects of large healthcare interventions like the NHPS is limited. In addition, for effective policy making, it is also important to understand how these effects compare with the alternative of providing care directly through tax-funded government hospitals. This paper takes the first step towards filling this gap in the literature by analyzing the impacts of two public healthcare provision strategies that resemble the two interventions discussed herein, namely directly providing care and improving access to private hospitals.

3.2.1 Diabetes in India

In order to empirically examine the two alternative healthcare provision strategies outlined earlier, our analysis focuses on a single ailment – diabetes. According to the World Health Organization (WHO), India is host to approximately 77 million individuals with diabetes, making it the second-largest diabetic population globally. There is mounting apprehension that India is in the throes of a diabetes epidemic[2], with projections indicating that the number of diabetics may escalate to 134 million by 2045 (Pradeepa and Mohan (2021)). As the prevalence of diabetes continues to increase, it is expected to put additional pressure on India's healthcare infrastructure.

Moreover, lower-income populations are increasingly becoming vulnerable to diabetes, owing to the rising prevalence of obesity and sedentary lifestyles. As the cost of treatment can often be prohibitive, public interventions may be essential to guarantee access to quality healthcare for all. Consequently, thinking about effective public healthcare strategies in the context of diabetes may be crucial to address this growing health challenge.

3.3 Data

Our primary data source for the analysis in this paper is the 75th round of the National Sample Survey (NSS) conducted between July 2017 and June 2018. The NSS is a nationally representative survey of households in India, and in this specific round, households were asked about their health expenditures and health outcomes.

Importantly, each member of every sampled household was asked whether they had been hospitalized within the previous 12 months. If the answer was yes, they were requested to provide additional details about the hospitalization, including the type of hospital (government or private), the number of days spent in the hospital, and the total cost of the hospitalization. The data also reports household size and consumption expenditures of each household, which allows us to construct a proxy for income per capita of the household.

For the purposes of our study, we focus on households that reported a hospitalization for a patient with diabetes. We use the observed choices of these patients (i.e. whether they chose a government or a private hospital) to estimate key parameters of the model outlined later in this paper. We also use their reported costs of hospitalization to construct a measure for the price of receiving treatment for diabetes at government and private hospitals in their

[2]See NDTV (2015).

Table 3.1: *Summary Statistics of Market-Level Variables*

	(1) N	(2) Mean	(3) Std. Dev.	(4) 25th Percentile	(5) Median	(6) 75th Percentile
Price at public hospitals	20	4.0	2.6	1.9	3.0	5.6
Price at private hospitals	20	17.7	4.7	15.5	18.0	20.5
Market share of private hospitals	20	0.6	0.2	0.4	0.7	0.7
Beds / 1000 at public hospitals	20	0.9	0.7	0.4	0.6	1.1
Beds / 1000 at private hospitals	20	1.0	0.7	0.5	0.9	1.4

Notes: *This table presents summary statistics for the 20 markets (states) we include in our sample. Price is reported in thousands of Indian rupees, and is computed from the raw data as the median cost of hospitalization for a patient with diabetes.*

market (state).

One limitation of our data is the lack of precise information about the specific hospital where each patient sought treatment; we only know whether the patient received treatment at a government or private hospital. Therefore, when we aggregate prices to the market level, we assume that the price of receiving treatment at a government (or private) hospital is the same across all government (or private) hospitals in the market. Relatedly, this data constraint also forces us to assume that there is a single government hospital and a single private hospital in each market.

Our secondary dataset is derived from Kapoor, Sriram, Joshi, Nandi and Laxminarayan (2020), which compiled the number of hospitals and hospital beds by hospital type and state in India. We employ this data to calculate the number of beds per capita in each state.

By combining these datasets, we generate a comprehensive overview of market-level variables, the summary statistics of which are presented in Table 3.1. It is important to note that the cost of receiving treatment at government hospitals is significantly lower than at private hospitals. However, private hospitals still maintain a substantial market share, with an average of around 60%. Additionally, the number of beds per capita is marginally higher at private hospitals, suggesting greater congestion at government hospitals. In the subsequent section, we will incorporate these characteristics into our model to better understand their impact on patient choices.

3.4 Model

We present a static model with a demand-side and a supply-side. On the demand-side, households in need of hospitalization choose where to seek care. On the supply-side, hospitals set prices – public hospitals use marginal-cost pricing while private hospitals set prices to maximize profits. The model is for a specific disease category (in our case, diabetes), but can be generalized to any other disease or multiple diseases.

3.4.1 Demand for public vs. private hospitals

Patient i is located in market m and is choosing where to seek treatment. She has two choices: (1) public hospital and (2) private hospital, denoted by GOV and PVT, respectively.[3] Her utility from choosing choice $j \in \{GOV, PVT\}$ is given by

$$u_{ijm} = \beta_{ip} \cdot p_{jm} + \beta_{i,pvt} \cdot \textit{whether private}_j + \beta_x \cdot \boldsymbol{x}_{jm} + \varepsilon_{ijm} \tag{3.1}$$

where p_{jm} is the price of receiving treatment at hospital j in market m, $\textit{whether private}_j$ is an indicator variable which denotes whether j is a private hospital, $\boldsymbol{x}_{jm}$ is a vector of observed non-price characteristics of hospital j, and ε_{ijm} is unobserved idiosyncratic taste for hospital j which is iid distributed according to a Type 1 extreme value distribution.

Patient-specific coefficients on price and preference for private hospitals, denoted by β_{ip} and $\beta_{i,pvt}$, are parameterized to be functions of patient income, y_i, so that

$$\beta_{ip} = \beta_p + \beta_{py} \cdot y_i$$
$$\beta_{i,pvt} = \beta_{pvt} + \beta_{pvt,y} \cdot y_i$$

Thus, given patient income and observed price and non-price hospital characteristics, the probability that patient i chooses hospital j can be expressed as

$$s_{ijm} = s_{jm}\left(p_{jm}, \boldsymbol{x}_{jm}, y_i; \boldsymbol{\beta}\right)$$

[3]The model does not feature an outside option; all patients must seek care at either a public or a private hospital.

where

$$\boldsymbol{\beta} = \{\beta_p, \beta_{py}, \beta_{pvt}, \beta_{pvt,y}, \beta_x\}$$

Finally, given the T1EV assumption, choice probability s_{ijm} can be expressed as

$$s_{ijm} = \frac{\exp\{\beta_{ip} \cdot p_{jm} + \beta_{i,pvt} \cdot whether\ private_j + \beta_x \cdot \boldsymbol{x}_{jm}\}}{\sum_{k \in \{GOV,PVT\}} \exp\{\beta_{ip} \cdot p_{km} + \beta_{i,pvt} \cdot whether\ private_k + \beta_x \cdot \boldsymbol{x}_{km}\}} \tag{3.2}$$

While hospital characteristics $\boldsymbol{x}_{jm}$ are determined exogenously, hospital prices are chosen by hospitals. In the following section, we describe how these prices are set.

3.4.2 Supply-side

Let q_{jm} be the number of patients in market m who choose hospital j, given by

$$q_{jm} = \sum_{i \in N_m} s_{ijm}$$

where N_m is the total number of hospitalizations in market m.

Profits of hospital j in market m are given by

$$\pi_{jm} = (P_{jm} - c_{jm})q_{jm}$$

where c_{jm} is the constant marginal cost of providing care at hospital j in market m.

Public hospitals set prices to break even i.e.

$$P_{GOV,m} = c_{GOV,m} \tag{3.3}$$

Private hospitals set prices to maximize profits such that

$$P_{PVT,m} = \underset{\widetilde{P}_{PVT,m}}{\arg\max}\, \pi_{PVT,m}(\widetilde{P}_{PVT,m}) \tag{3.4}$$

The static equilibrium, therefore, in each market is the set of prices $(P_{GOV,m}, P_{PVT,m})$ which satisfy conditions equation (3.3) and equation (3.4).

The optimal price at both types of hospitals crucially depends on their marginal costs. We parameterize these marginal costs to be a function of hospital type and hospital characteristics. Specifically, we assume that

$$\log c_{jm} = \gamma_{0j} + \gamma_x \cdot \boldsymbol{x}_{jm} + \nu_{jm} \tag{3.5}$$

where $\nu_{jm} \sim N\left(0, \sigma_j^2\right)$. We denote the set of marginal cost parameters with $\boldsymbol{\gamma}$, where

$$\boldsymbol{\gamma} = \{\gamma_{0,GOV}, \gamma_{0,PVT}, \gamma_x, \sigma_{GOV}, \sigma_{PVT}\}$$

3.4.3 Limitations of the model

Our model has several limitations, primarily due to data constraints. First, we assume that the choice set only consists of two choices: public vs. private, and ignore all heterogeneity within public and private hospitals. In other words, we assume the market consists of a single government hospital and a single profit-maximizing private hospital. This is because in the data, we do not observe exactly which hospital each patient went to. Patients only report whether they went to a public hospital or a private hospital.

Second, we only know the total number of public and private hospitals in each state[4], and do not know the precise location of each hospital. Consequently, one crucial hospital characteristic missing from our utility specification is distance to the hospital. In future work, we plan to collect more detailed and comprehensive data on hospital location, and include some notion of distance to the hospital from the center of each district. This additional information will allow us to better understand the role of accessibility in determining patient choices, as well as provide a more accurate representation of the competitive landscape in the local healthcare market.

3.5 Estimation Strategy

In this section we describe how we estimate the parameters governing patient utility, $\boldsymbol{\beta}$, as well as the parameters governing hospital marginal costs, $\boldsymbol{\gamma}$.

[4]Our data also gives the total number of beds in public and private hospitals in each state. We use this to construct the number of beds per 1000 people, and include it in the utility specification in equation (A.1).

Let $\widehat{\theta} = \left\{\widehat{\boldsymbol{\beta}}, \widehat{\gamma}\right\}$ be a candidate vector for the parameters we are interested in estimating. From equation equation (3.2), for each guess of $\widehat{\boldsymbol{\beta}}$, we can construct the probability that patient i chooses hospital j in market m. In the data, we observe the choice each patient makes. Thus, at the candidate vector of parameters, we can compute the likelihood of the data, denoted by $\mathcal{L}(\widehat{\theta})$.

At the same time, given our assumption on the supply-side of the model, we can use the observed prices and the guess of utility parameters, $\widehat{\boldsymbol{\beta}}$, to compute the implied marginal costs for each hospital in each market. Let the vector of implied marginal costs be denoted by $c(\widehat{\boldsymbol{\beta}})$. Now, given our guess for marginal cost parameters, $\widehat{\boldsymbol{\gamma}}$, we can also construct this same vector of marginal costs, and denote it by $c(\widehat{\boldsymbol{\gamma}})$. The distance between these two vectors of marginal costs is denoted by $\Gamma(\widehat{\theta})$, and defined as

$$\Gamma(\widehat{\theta}) = \|c(\widehat{\boldsymbol{\beta}}) - c(\widehat{\boldsymbol{\gamma}})\|$$

The goal of our estimation strategy is to find the vector of utility and marginal cost parameters, θ, which maximizes the likelihood of the data, $\mathcal{L}(\theta)$, while minimizing the distance between the implied and predicted marginal costs, $\Gamma(\theta)$. That is, we want to find θ such that

$$\theta = \arg\min_{\widehat{\theta}} \left\{-\lambda \cdot \mathcal{L}(\widehat{\theta}) + (1-\lambda) \cdot \Gamma(\widehat{\theta})\right\}$$

where λ is a parameter that controls the relative weight we give to the likelihood of the data and the distance between the implied and predicted marginal costs.

We present our estimated parameters in Table 3.2. As expected, patients dislike price, but this sensitivity to price declines with income as our coefficient on the interaction between price and income is positive.[5] We also find a positive preference for private hospitals, which is stronger for higher income patients. Finally, we find that patients dislike congestion, or in other words, prefer hospitals with more beds per thousand people in the market. To compare the fit of our utility parameters with the observed data, we compare the predicted choice

[5] Income is proxied using per capita household expenditures, expressed in thousands of rupees.

Table 3.2: *Parameters governing patient utility and hospital marginal costs*

		(1) Estimate
A. Patient utility	price	-0.257
	price × income	0.026
	private	1.486
	private × income	0.370
	beds per ’000 persons	0.566
B. Marginal cost	constant (government)	1.590
	constant (private)	1.700
	log beds per ’000 persons (government)	-0.254
	log beds per ’000 persons (private)	-1.434
	cost shock variance (government)	0.018
	cost shock variance (private)	0.700

Notes. This table presents the estimated structural parameters of the model. The top panel presents the parameters governing patient utility. The bottom panel presents the parameters governing hospital marginal costs. Patient utility depends on hospital price and whether the hospital is a private hospital, both interacted with household income. Price is expressed in thousands of rupees, and household income is proxied using per capita household expenditures, also expressed in thousands of rupees. Additionally, patient utility depends on how many beds are available per thousand people in the market, or the inverse of congestion. *Hospital marginal costs depend on a hospital type-specific intercept, as well as number of beds per thousand people. Additionally, hospital marginal costs are also subject to a hospital type-specific shock, whose estimated variance is reported above.*

probabilities for private hospitals in each market with the observed choice probabilities. We present these in Figure 3.1.

On the supply-side, we find that the intercept of the (log) marginal cost function is higher for private hospitals, suggesting that ceteris paribus, government hospitals can provide care at a lower cost. The coefficient on number of beds per thousand people is negative, suggesting that relaxing capacity constraints could lower the cost of providing care. Finally, we find that the variance of the hospital type-specific shock is larger for private hospitals, which hints at greater unobserved heterogeneity in the cost of care at private hospitals.

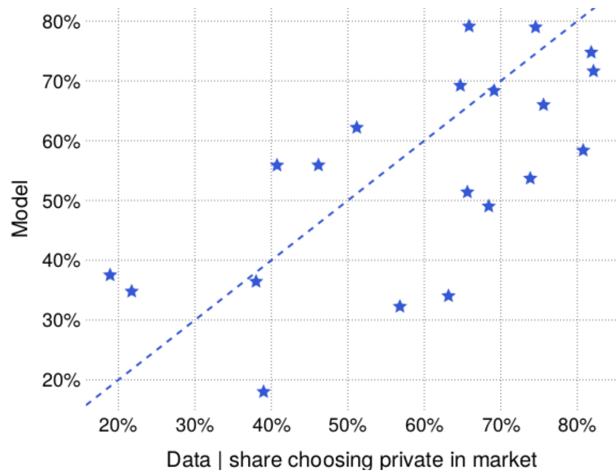


Figure 3.1: *Model fit: observed vs. predicted market share of private hospitals*

Notes: *This figure shows the observed share of private hospitals in each market, as well as the predicted share of private hospitals in each market, given the estimated parameters of the model. The dashed line represents the 45 degree line.*

3.6 Counterfactuals

In this section, we use the estimated structural parameters to simulate the effects of different public healthcare provision strategies on equilibrium market and patient outcomes. We consider two main counterfactual scenarios: (i) an increase in the number of beds at government hospitals, and (ii) a subsidy to patients for receiving treatment at private hospitals. To evaluate the impact of these policy interventions, we run the counterfactuals for two distinct markets: (1) a low density market, and (2) a high density market. The density here refers to the number of beds at government hospitals per 1000 people, which can be seen as a measure of the relative availability of government beds in the market or the inverse of *congestion*.

For the low density market, we take the state of Jharkhand, which has only 0.32 beds per 1000 people at government hospitals. In contrast, for the high density market, we take the city of Delhi, which has 1.45 beds per 1000 people at government hospitals. For each market, we consider 10,000 patients in need of hospitalization, with incomes drawn from a nationally-representative income distribution. Our analysis focuses on the effects of these policy interventions in these two markets, considering both direct and indirect impacts on patient welfare.

3.6.1 Greater Number of Beds at Government Hospitals

In this counterfactual, we increase the density by 0.2 in both the low and high-density markets. Given our estimates of marginal cost parameters, lowering congestion lowers the marginal cost of providing care at government hospitals. Since government hospitals use marginal-cost pricing, all reductions in marginal costs are completely passed through to patients. As a result, the utility associated with government hospitals increases due to both lower prices and reduced congestion, which directly enters the utility function.

In equilibrium, the price of care at private hospitals also falls, as the increase in the relative attractiveness of government hospitals puts competitive pressure on private hospitals. To quantify the impact of this policy on consumers, we compute the associated compensating variation (CV) for each patient in the market. We present the estimated impact on individual patients in Figure 3.2. We first analyze the impact without accounting for the equilibrium response of private hospitals (referred to as "partial") and then consider the impact accounting for this equilibrium response.

Our findings indicate that, in both markets, all consumers along the income distribution are better off. The associated change in consumer surplus is higher for lower-income consumers, as they have a higher marginal utility of income. Accounting for the equilibrium impact on the price at private hospitals leads to higher CV for all patients, with a substantial increase for higher-income patients, as we estimate them to have a stronger preference for private hospitals. This analysis shows that both the direct and indirect benefits of the policy are substantial, and not accounting for the equilibrium impact of this intervention would lead to an underestimation of the welfare gains.

Lastly, we find that the CV is higher in the low-density market than in the high-density market for all patients. This is because the relative increase in expected utility is higher in the low-density market, which had higher congestion and higher prices at both government

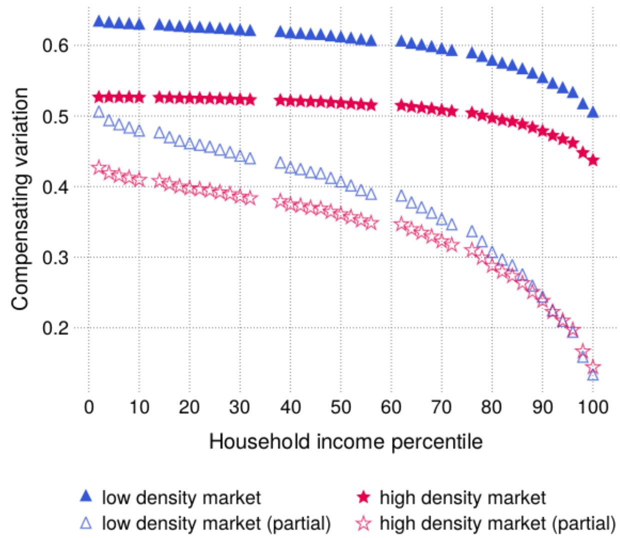


Figure 3.2: *Compensating variation by household income under additional government hospitals*

Notes: *This figure shows the compensating variation for patients across different household income levels under the counterfactual scenario of additional government hospital beds. The horizontal axis represents the income quantiles, while the vertical axis shows the compensating variation. The results are presented for both partial and equilibrium responses; the latter accounts for the impact of the policy intervention on private hospital prices. The figure also distinguishes between the impact on patients in low- and high-density markets.*

and private hospitals in the baseline scenario.

It is important to note that our analysis presented above focuses on the benefits associated with increasing the number of beds at government hospitals, and does not account for the fixed costs associated with this expansion. In practice, adding more beds at government hospitals would involve costs such as infrastructure investments, personnel hiring and training, and ongoing maintenance. Later, we discuss how to use our estimates of just gross welfare gains presented here to think about optimal policy interventions.

3.6.2 Subsidy to Patients for Receiving Treatment at Private Hospitals

In this counterfactual, we examine a policy in which patients with income below a specified threshold receive a subsidy on the price of care at private hospitals. This approach differs from the actual implementation of Ayushman Bharat, where eligible patients can access pre-approved treatments up to a certain amount at designated private hospitals. Although

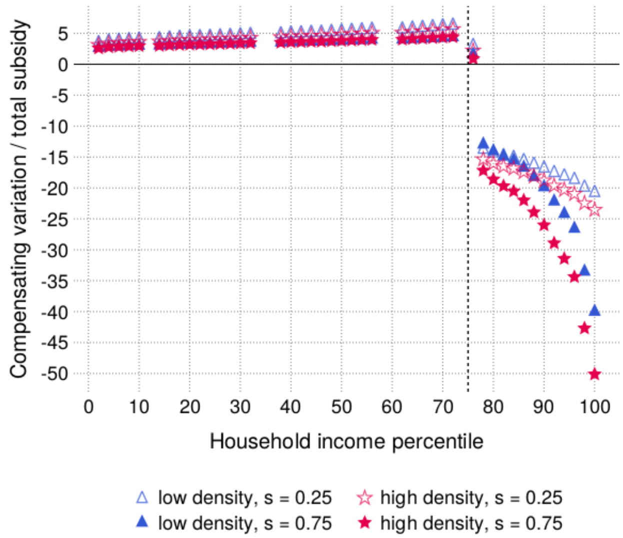


Figure 3.3: *Compensating variation per unit subsidy expenditures by houshold income*

Notes: *This figure shows the compensating variation per unit of subsidy expenditures for eligible and non-eligible patients across different household income levels. The horizontal axis displays the income quantiles, while the vertical axis represents the compensating variation per unit of subsidy expenditures. The results are presented for two subsidy levels, 25% and 75%, with the income threshold fixed at the 75th percentile of the income distribution. The figure also breaks down the results by market type – low- vs. high-density.*

not identical, the policy we investigate is a viable alternative, and we intend to extend our analysis to the actual implementation of Ayushman Bharat in future research.

Let s denote the subsidy rate and q represent the quantile of the income distribution that defines the income threshold for eligibility. Consequently, if the price of care at private hospitals is p, eligible patients pay $p(1-s)$, taking into account the subsidy. The immediate effect of this policy is that eligible patients can access treatments at private hospitals at a reduced price. However, private hospitals may respond to this policy by raising their prices, resulting in the government subsidy being partially absorbed by these hospitals. At the same time, this equilibrium response of private hospitals may adversely affect non-eligible patients, as they now face higher prices at private hospitals.

To assess the impact of this policy, we simulate the counterfactual for various values of s and q. In Figure 3.3, we present the results for two subsidy levels, 25% and 75%, while maintaining the income threshold at the 75th percentile of the income distribution.

The vertical axis represents the compensating variation per unit of subsidy expenditures, enabling a comparison between the two subsidy levels in terms of "bang for the buck" of each policy.

Among eligible patients, we observe that both subsidy levels yield increased welfare, with benefits rising along with patient income, as higher-income patients exhibit a stronger preference for private hospitals. However, these gains are relatively modest compared to the indirect negative impact on non-eligible patients, which is more pronounced with patient income. As mentioned earlier, this is because non-eligible patients now face higher prices at private hospitals due to the equilibrium response of these hospitals.

Interestingly, the benefits to eligible patients are greater under the 25% subsidy, while the losses to non-eligible patients are also less severe at this subsidy level. This finding suggests that a lower subsidy rate may be more effective in balancing the positive and negative effects of the policy across different income groups. Lastly, when comparing the low and high-density markets, we discover that the benefits are diminished and the losses are more substantial in the high-density market.

3.6.3 Optimal Policy

In this subsection, we discuss the net effects of the two policy interventions and make progress towards providing a recommendation for the optimal policy, assuming that the policymaker aims to maximize the benefits to patients per unit of government expenditures.

As demonstrated above, the first counterfactual, which involves increasing the number of government hospital beds, yields substantial benefits to consumers across the income distribution. However, we lack information on the costs associated with this policy. Nonetheless, we can use our results to bound the maximum amount the government should be willing to spend on this policy.

Suppose that the income threshold for private subsidy eligibility is set at the 25th percentile, i.e., $q = 0.25$. In the first counterfactual, we can compute the total benefits, b_1, for patients below the 25th percentile, and the benefits to all other patients as l_1. For a

Table 3.3: *Results from counterfactual simulations*

			(1) Δ private price	(2) Δ private share	(3) CV net of subsidy expenses
Build government hospitals	low density market		-3%	-0.02	5,685
	high density market		-3%	-0.02	4,836
Subsidize private hospitals	low density market	$s = 0.25$, $q = 0.25$	7%	0.01	-4,867
		$s = 0.25$, $q = 0.50$	13%	0.02	-9,483
		$s = 0.25$, $q = 0.75$	21%	0.03	-14,319
		$s = 0.50$, $q = 0.25$	17%	0.00	-9,390
		$s = 0.50$, $q = 0.50$	38%	0.01	-19,867
		$s = 0.50$, $q = 0.75$	61%	0.03	-29,761
		$s = 0.75$, $q = 0.25$	30%	-0.03	-7,395
		$s = 0.75$, $q = 0.50$	204%	-0.22	-49,269
		$s = 0.75$, $q = 0.75$	229%	-0.08	-53,951
	high density market	$s = 0.25$, $q = 0.25$	7%	0.01	-3,558
		$s = 0.25$, $q = 0.50$	14%	0.01	-7,057
		$s = 0.25$, $q = 0.75$	23%	0.01	-10,866
		$s = 0.50$, $q = 0.25$	18%	0.00	-7,166
		$s = 0.50$, $q = 0.50$	39%	0.00	-14,658
		$s = 0.50$, $q = 0.75$	64%	0.01	-22,225
		$s = 0.75$, $q = 0.25$	34%	-0.02	-7,771
		$s = 0.75$, $q = 0.50$	169%	-0.17	-35,132
		$s = 0.75$, $q = 0.75$	230%	-0.09	-41,228

Notes: *This table presents the results from various counterfactual policy simulations, comparing the net effects of increasing the number of government hospital beds and providing subsidies for private care. The table reports percent change in price at private hospitals, the change in market share of private hospitals (in percentage points), as well as total compensating variation net of subsidy expenditures. The results are broken down by market type.*

given subsidy rate, s, in the second counterfactual, we can compute the total benefits, b_2, for patients below the 25th percentile, as well as the total costs to the government, e_2, and the total benefits to patients above the 25th percentile, l_2. Let the cost to benefits ratio under the second counterfactual be $(e_2 - l_2)/b_2$. Now, by comparing the two policies, we can determine the maximum amount the government should be willing to spend on the first policy as $l_1 + b_1 * ((e_2 - l_2)/b_2)$. In other words, the government should be willing to invest in the first policy up to the point where the costs per unit benefits from it are equal to the costs per unit benefits of the second policy.

This cost-benefit analysis is a work in progress. We present preliminary statistics computed using various counterfactual policies the government could implement in Table 3.3.

3.7 Conclusion

In this paper, we examine the effects of various public healthcare provision strategies on market equilibrium and patient outcomes. In doing so, we analyze two main counterfactual scenarios: (i) an increase in the number of beds at government hospitals, and (ii) a subsidy to patients for receiving treatment at private hospitals. We focused our analysis on two distinct markets, a low-density market represented by the state of Jharkhand and a high-density market represented by the city of Delhi.

Our findings indicate that increasing the number of beds at government hospitals leads to substantial benefits for consumers across the income distribution. On the other hand, providing a subsidy for treatment at private hospitals yields mixed results. While eligible patients experience increased welfare, non-eligible patients face negative impacts due to the equilibrium response of private hospitals, which raises prices in response to the subsidy.

References

ACEMOGLU, D. and ROBINSON, J. A. (2001). Inefficient Redistribution. *American Political Science Review*, **95** (3), 649–661.

ACKERBERG, D. A., CAVES, K. and FRAZER, G. (2015). Identification Properties of Recent Production Function Estimators. *Econometrica*, **83** (6), 2411–2451.

ALLEN, T. and ATKIN, D. (2022). Volatility and the Gains From Trade. *Econometrica*, **90** (5), 2053–2092.

AMITI, M. and KHANDELWAL, A. K. (2013). Import Competition and Quality Upgrading. *Review of Economics and Statistics*, **95** (2), 476–490.

— and KONINGS, J. (2007). Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia. *The American Economic Review*, **97** (5), 1611–1638.

—, REDDING, S. J. and WEINSTEIN, D. E. (2019). The Impact of the 2018 Tariffs on Prices and Welfare. *Journal of Economic Perspectives*, **33** (4), 187–210.

BALDWIN, R. and KRUGMAN, P. (1988a). Industrial Policy and International Competition in Wide-Bodied Jet Aircraft. In *Trade Policy Issues and Empirical Analysis*, University of Chicago Press, pp. 45–78.

— and — (1988b). Market Access and International Competition: A Simulation Study of 16K Random Access Memories. In *Empirical Methods for International Trade*, 1st edn., Cambridge, MA: MIT Press, pp. 171–197.

BANERJEE, A., DEATON, A. and DUFLO, E. (2004). Health Care Delivery in Rural Rajasthan. *Economic and Political Weekly*, **39** (9), 944–949.

—, FINKELSTEIN, A., HANNA, R., OLKEN, B. A., ORNAGHI, A. and SUMARTO, S. (2021). The Challenges of Universal Health Insurance in Developing Countries: Experimental Evidence from Indonesia's National Health Insurance. *American Economic Review*, **111** (9), 3035–3063.

—, HANNA, R., KYLE, J., OLKEN, B. A. and SUMARTO, S. (2018). Tangible Information and Citizen Empowerment: Identification Cards and Food Subsidy Programs in Indonesia. *Journal of Political Economy*, **126** (2), 451–491.

—, —, —, — and — (2019). Private Outsourcing and Competition: Subsidized Food Distribution in Indonesia. *Journal of Political Economy*, **127** (1), 101–137.

Bartelme, D., Costinot, A., Donaldson, D. and Rodríguez-Clare, A. (2019). *The Textbook Case for Industrial Policy: Theory Meets Data*. Working Paper 26193, National Bureau of Economic Research.

Barwick, P. J., Kalouptsidi, M. and Zahur, N. B. (2021). Industrial Policy Implementation: Empirical Evidence from China's Shipbuilding Industry.

Basu, S., Andrews, J., Kishore, S., Panjabi, R. and Stuckler, D. (2012). Comparative Performance of Private and Public Healthcare Systems in Low- and Middle-Income Countries: A Systematic Review. *PLOS Medicine*, **9** (6), e1001244.

Bergquist, L. F. and Dinerstein, M. (2020). Competition and Entry in Agricultural Markets: Experimental Evidence from Kenya. *American Economic Review*, **110** (12), 3705–3747.

—, Faber, B., Fally, T., Hoelzein, M., Miguel, E. and Rodriguez-Clare, A. (2022). Scaling Agricultural Policy Interventions.

Breza, E. and Kinnan, C. (2021). Measuring the Equilibrium Impacts of Credit: Evidence from the Indian Microfinance Crisis. *The Quarterly Journal of Economics*, **136** (3), 1447–1497.

Bronzini, R. and Piselli, P. (2016). The impact of R&D subsidies on firm innovation. *Research Policy*, **45** (2), 442–457.

Cavallo, A., Gopinath, G., Neiman, B. and Tang, J. (2021). Tariff Pass-Through at the Border and at the Store: Evidence from US Trade Policy. *American Economic Review: Insights*, **3** (1), 19–34.

Cerqua, A. and Pellegrini, G. (2014). Do subsidies to private capital boost firms' growth? A multiple regression discontinuity design approach. *Journal of Public Economics*, **109**, 114–126.

— and — (2017). Industrial policy evaluation in the presence of spillovers. *Small Business Economics*, **49** (3), 671–686.

Chan, D., Card, D. and Taylor, L. J. (2022). Is There a Va Advantage? Evidence from Dually Eligible Veterans.

Chatterjee, S. (2022). Market Power and Spatial Competition in Rural India.

—, Kapur, D., Sekhsaria, P. and Subramanian, A. (2022). New Facts, Constitutional Vision. (36), 10.

Cole, S., Giné, X. and Vickery, J. (2017). How Does Risk Management Influence Production Decisions? Evidence from a Field Experiment. *The Review of Financial Studies*, **30** (6), 1935–1970.

Costinot, A., Donaldson, D. and Smith, C. (2016). Evolving Comparative Advantage and the Impact of Climate Change in Agricultural Markets: Evidence from 1.7 Million Fields around the World. *Journal of Political Economy*, **124** (1), 205–248.

Criscuolo, C., Martin, R., Overman, H. G. and Van Reenen, J. (2019). Some Causal Effects of an Industrial Policy. *American Economic Review*, **109** (1), 48–85.

Cunha, J. M., De Giorgi, G. and Jayachandran, S. (2019). The Price Effects of Cash Versus In-Kind Transfers. *The Review of Economic Studies*, **86** (1), 240–281.

Das, J., Chowdhury, A., Hussam, R. and Banerjee, A. V. (2016a). The impact of training informal health care providers in India: A randomized controlled trial. *Science*, **354** (6308), aaf7384.

— and Do, Q.-T. (2023). The Prices in the Crises: What We Are Learning from 20 Years of Health Insurance in Low- and Middle-Income Countries. *Journal of Economic Perspectives*, **37** (2), 123–152.

— and Hammer, J. (2014). Quality of Primary Care in Low-Income Countries: Facts and Economics. *Annual Review of Economics*, **6** (1), 525–553.

—, Holla, A., Mohpal, A. and Muralidharan, K. (2016b). Quality and Accountability in Health Care Delivery: Audit-Study Evidence from Primary Care in India. *American Economic Review*, **106** (12), 3765–3799.

Deaton, A. (1997). *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. The World Bank.

Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data Via the *EM* Algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, **39** (1), 1–22.

Dixit, A. K. and Grossman, G. M. (1986). Targeted Export Promotion With Several Oligopolistic Industries. *Journal of International Economics*, **21** (3–4), 233–249.

Donovan, K. (2021). The Equilibrium Impact of Agricultural Risk on Intermediate Inputs and Aggregate Productivity. *The Review of Economic Studies*, **88** (5), 2275–2307.

Duflo, E., Kremer, M. and Robinson, J. (2008). How High Are Rates of Return to Fertilizer? Evidence from Field Experiments in Kenya. *American Economic Review*, **98** (2), 482–488.

—, — and — (2011). Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya. *American Economic Review*, **101** (6), 2350–2390.

Duggan, M. and Morton, F. M. S. (2006). The Distortionary Effects of Government Procurement: Evidence from Medicaid Prescription Drug Purchasing. *The Quarterly Journal of Economics*, **121** (1), 1–30.

Edmond, C., Midrigan, V. and Xu, D. Y. (2015). Competition, Markups, and the Gains from International Trade. *American Economic Review*, **105** (10), 3183–3221.

Egger, D., Haushofer, J., Miguel, E., Niehaus, P. and Walker, M. (2019). *General Equilibrium Effects of Cash Transfers: Experimental Evidence from Kenya*. Tech. Rep. w26600, National Bureau of Economic Research, Cambridge, MA.

Fajgelbaum, P. D., Goldberg, P. K., Kennedy, P. J. and Khandelwal, A. K. (2020). The Return to Protectionism. *The Quarterly Journal of Economics*, **135** (1), 1–55.

FAO, IFAD, UNICEF, WFP and WHO (2021). *The State of Food Security and Nutrition in the World 2021. Transforming Food Systems for Food Security, Improved Nutrition and Affordable Healthy Diets for All.* Rome: FAO.

Flaaen, A., Hortaçsu, A. and Tintelnot, F. (2020). The Production Relocation and Price Effects of US Trade Policy: The Case of Washing Machines. *American Economic Review*, **110** (7), 2103–2127.

Frakes, M., Gruber, J. and Justicz, T. (2020). *Public and Private Options in Practice: The Military Health System*. Tech. Rep. w28256, National Bureau of Economic Research, Cambridge, MA.

Gadenne, L. (2020). Can Rationing Increase Welfare? Theory and an Application to India's Ration Shop System. *American Economic Journal: Economic Policy*, **12** (4), 144–177.

—, Norris, S., Singhal, M. and Sukhtankar, S. (2022). In-Kind Transfers as Insurance.

Gandhi, A., Navarro, S. and Rivers, D. A. (2020). On the Identification of Gross Output Production Functions. *Journal of Political Economy*, **128** (8), 2973–3016.

Gaynor, M., Ho, K. and Town, R. J. (2015). The Industrial Organization of Health-Care Markets. *Journal of Economic Literature*, **53** (2), 235–284.

Goldberg, P. K., Khandelwal, A. K., Pavcnik, N. and Topalova, P. (2010). Imported Intermediate Inputs And Domestic Product Growth: Evidence From India. *The Quarterly Journal of Economics*, **125** (4), 1727–1767.

Gourieroux, C., Monfort, A. and Renault, E. (1993). Indirect Inference. *Journal of Applied Econometrics*, **8**, S85–S118.

Greenwald, B. and Stiglitz, J. E. (2006). Helping Infant Economies Grow: Foundations of Trade Policies for Developing Countries. *American Economic Review*, **96** (2), 141–146.

Griliches, Z. and Mairesse, J. (1995). *Production Functions: The Search for Identification*. NBER Working Paper w5067, National Bureau of Economic Research, Cambridge, MA.

Grossman, G. M. and Helpman, E. (1994). Protection for Sale. *The American Economic Review*, **84** (4), 833–850.

Gulati, A. (2014). A fertile mess.

Gupta, P., Khera, R. and Narayanan, S. (2021). Minimum Support Prices in India: Distilling the Facts.

Hansen, J. D., Jensen, C. and Madsen, E. S. (2003). The establishment of the danish windmill industry—Was it worthwhile? *Review of World Economics*, **139** (2), 324–347.

Harris, R., Keay, I. and Lewis, F. (2015). Protecting infant industries: Canadian manufacturing and the national policy, 1870–1913. *Explorations in Economic History*, **56**, 15–31.

Hausman, J., Leonard, G. and Zona, J. D. (1994). Competitive Analysis with Differenciated Products. *Annales d'Économie et de Statistique*, (34), 159–180.

Head, K. (1994). Infant industry protection in the steel rail industry. *Journal of International Economics*, **37** (3-4), 141–165.

Ho, K. (2009). Insurer-Provider Networks in the Medical Care Market. *American Economic Review*, **99** (1), 393–430.

— and Lee, R. S. (2017). Insurer Competition in Health Care Markets. *Econometrica*, **85** (2), 379–417.

Hoch, I. (1962). Estimation of Production Function Parameters Combining Time-Series and Cross-Section Data. *Econometrica*, **30** (1), 34–53.

Holden, S. T. (2019). Economics of Farm Input Subsidies in Africa. *Annual Review of Resource Economics*, **11** (1), 501–522.

Hsiao, A. (2022). Coordination and Commitment in International Climate Action: Evidence from Palm Oil.

Imbert, C. and Papp, J. (2015). Labor Market Effects of Social Programs: Evidence from India's Employment Guarantee. *American Economic Journal: Applied Economics*, **7** (2), 233–263.

Irwin, D. A. (2000a). *Could the U.S. Iron Industry Have Survived Free Trade After the Civil War?* Working Paper 7640, National Bureau of Economic Research.

— (2000b). Did Late-Nineteenth-Century U.S. Tariffs Promote Infant Industries? Evidence from the Tinplate Industry. *The Journal of Economic History*, **60** (2), 335–360.

— (2019). Tariff Incidence: Evidence From U.S. Sugar Duties, 1890–1914. *National Tax Journal*, **72** (3), 599–616.

Juhász, R. (2018). Temporary Protection and Technology Adoption: Evidence from the Napoleonic Blockade. *American Economic Review*, **108** (11), 3339–3376.

Kalouptsidi, M. (2018). Detection and Impact of Industrial Subsidies: The Case of Chinese Shipbuilding. *The Review of Economic Studies*, **85** (2), 1111–1158.

Kapoor, G., Sriram, A., Joshi, J., Nandi, A. and Laxminarayan, R. (2020). COVID-19 in India : State-wise estimates of current hospital beds, intensive care unit (ICU) beds and ventilators.

Karlan, D., Osei, R., Osei-Akoto, I. and Udry, C. (2014). Agricultural Decisions After Relaxing Credit And Risk Constraints. *The Quarterly Journal of Economics*, **129** (2), 597–652.

Khanna, G. (2022). Large-Scale Education Reform in General Equilibrium: Regression Discontinuity Evidence from India.

Krishnaswamy, N. (2019). At What Price? Price Supports, Agricultural Productivity, and Misallocation.

Lane, N. (2021). Manufacturing Revolutions: Industrial Policy and Industrialization in South Korea.

Levinsohn, J. and Petrin, A. (2003). Estimating Production Functions Using Inputs to Control for Unobservables. *The Review of Economic Studies*, **70** (2), 317–341.

Marschak, J. and Andrews, W. H. (1944). Random Simultaneous Equations and the Theory of Production. *Econometrica*, **12** (3/4), 143–205.

Mashal, M., Schmall, E. and Goldman, R. (2021). What Prompted the Farm Protests in India? *The New York Times*.

McFadden, D. (1989). A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration. *Econometrica*, **57** (5), 995–1026.

Meenakshi, J. and Banerji, A. (2005). The unsupportable support price: An analysis of collusion and government intervention in paddy auction markets in North India. *Journal of Development Economics*, **76** (2), 377–403.

Mills, A. (2014). Health Care Systems in Low- and Middle-Income Countries. *New England Journal of Medicine*, **370** (6), 552–557.

Mitra, S., Mookherjee, D., Torero, M. and Visaria, S. (2018). Asymmetric Information and Middleman Margins: An Experiment with Indian Potato Farmers. *The Review of Economics and Statistics*, **100** (1), 1–13.

Mobarak, A. M. and Rosenzweig, M. R. (2013). Informal Risk Sharing, Index Insurance, and Risk Taking in Developing Countries. *The American Economic Review*, **103** (3), 375–380.

Muralidharan, K., Niehaus, P. and Sukhtankar, S. (2022). General Equilibrium Effects of (Improving) Public Employment Programs: Experimental Evidence from India.

NDTV (2015). Diabetes: The Epidemic That Indians Created. https://food.ndtv.com/health/diabetes-the-epidemic-that-indians-created-771755.

OECD (2012). *Agricultural Policies for Poverty Reduction*. OECD Publishing.

OECD (2022). *Reforming Agricultural Policies for Climate Change Mitigation*, OECD.

Olley, G. S. and Pakes, A. (1996). The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica*, **64** (6), 1263–1297.

Pakes, A. (1986). Patents as Options: Some Estimates of the Value of Holding European Patent Stocks. *Econometrica*, **54** (4), 755–784.

— and Pollard, D. (1989). Simulation and the Asymptotics of Optimization Estimators. *Econometrica*, **57** (5), 1027–1057.

Pellegrini, G. and Muccigrosso, T. (2017). Do subsidized new firms survive longer? Evidence from a counterfactual approach. *Regional Studies*, **51** (10), 1483–1493.

Pradeepa, R. and Mohan, V. (2021). Epidemiology of type 2 diabetes in India. *Indian Journal of Ophthalmology*, **69** (11), 2932–2938.

Rosenzweig, M. R. and Binswanger, H. P. (1993). Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments. *The Economic Journal*, **103** (416), 56–78.

— and Wolpin, K. I. (1993). Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-Income Countries: Investments in Bullocks in India. *Journal of Political Economy*, **101** (2), 223–244.

Rotemberg, M. (2019). Equilibrium Effects of Firm Subsidies. *American Economic Review*, **109** (10), 3475–3513.

Schmalensee, R. (1982). Product Differentiation Advantages of Pioneering Brands. *The American Economic Review*, **72** (3), 349–365.

Sotelo, S. (2020). Domestic Trade Frictions and Agriculture. *Journal of Political Economy*, **128** (7), 2690–2738.

Topalova, P. and Khandelwal, A. (2011). Trade Liberalization and Firm Productivity: The Case of India. *Review of Economics and Statistics*, **93** (3), 995–1009.

USDA (2021). *Farms and Land in Farms: 2021 Summary*. National Agricultural Statistics Service.

Villa, S. and Kane, N. (2013). Assessing the Impact of Privatizing Public Hospitals in Three American States: Implications for Universal Health Coverage. *Value in Health*, **16** (1), S24–S33.

Wollmann, T. G. (2018). Trucks without Bailouts: Equilibrium Product Characteristics for Commercial Vehicles. *American Economic Review*, **108** (6), 1364–1406.

World Bank (2019). Schemes To Systems | The Public Distribution System: Anatomy of India's Food Subsidy Reforms. https://www.worldbank.org/en/news/feature/2019/02/21/schemes-to-systems-public-distribution-system.

Appendix A

Appendix to Chapter 1

A.1 Additional Figures

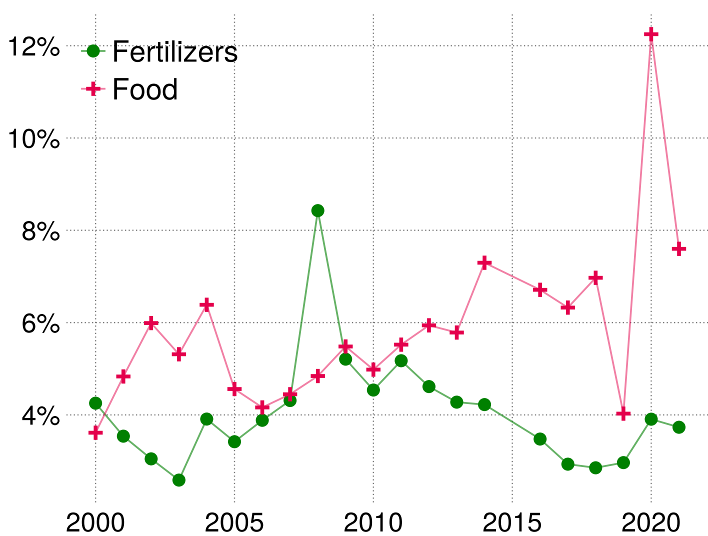


Figure A.1: *Program Costs As a Share of Total Government Spending*

Source. (Revised) budget estimates of the Government of India

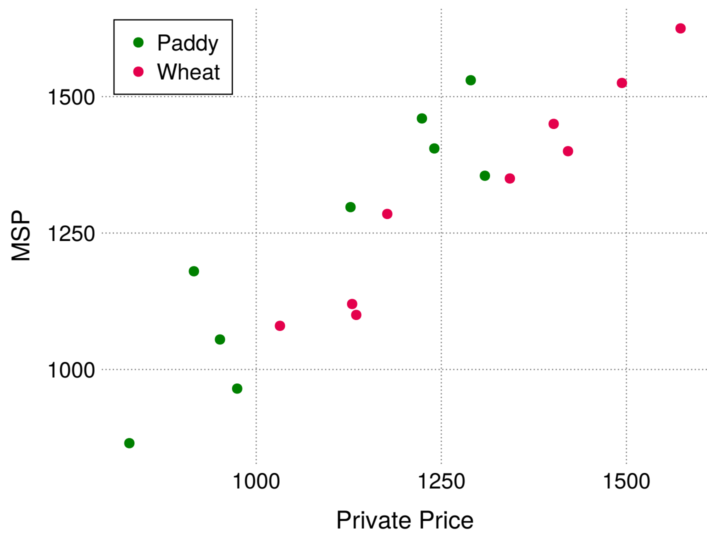


Figure A.2: *MSP Relative to Mean Private Market Price*

Notes. This figure plots the government announced MSP for rice and wheat on the private market mean prices recovered from our estimation.

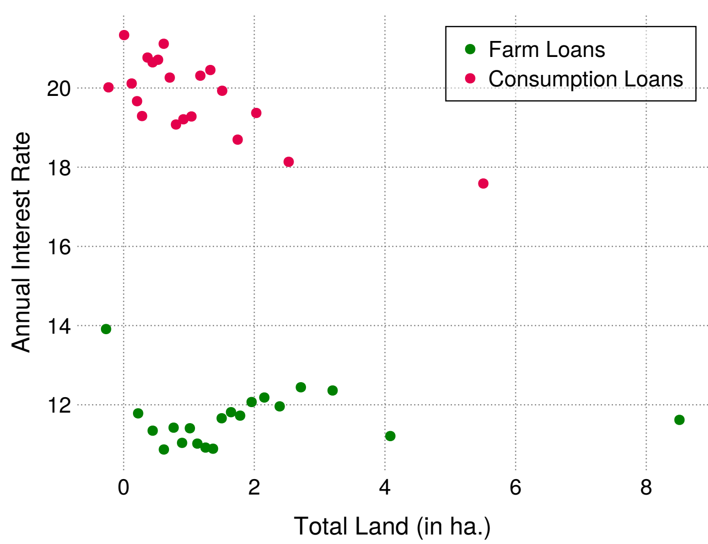


Figure A.3: *Annual Interest Paid for Farm and Consumption Loans*

Notes. The figure plots the average interest rate paid by farmers for farm and consumption loans on total land holdings (in ha.) of the farmer. The data are from the 77th round of the NSS (2019).

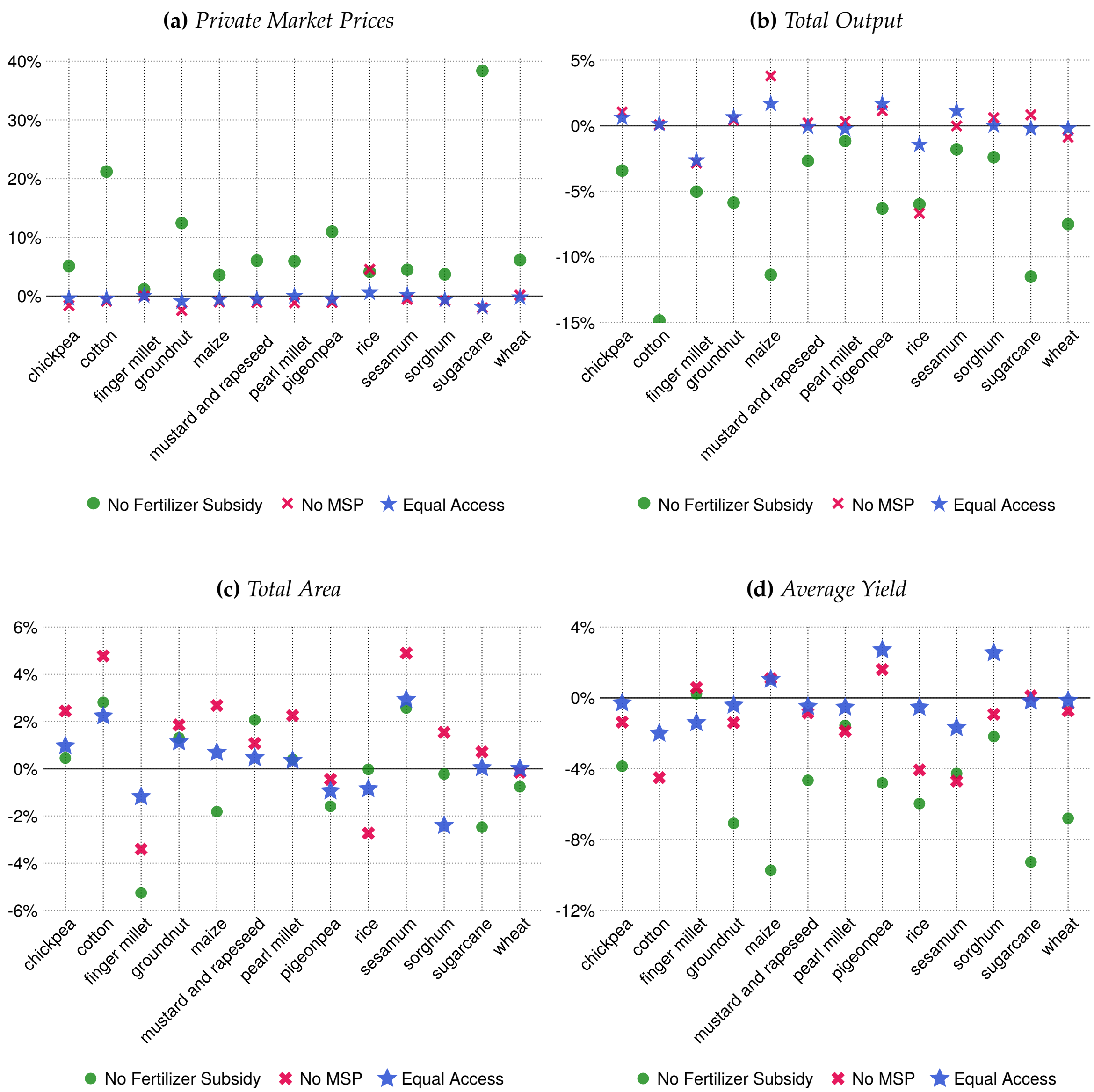


Figure A.4: *Counterfactuals: Percent Change Relative to Baseline, By Crop*

Notes. These plots show the relative change in key aggregate statistics in the various counterfactuals for each crop in our sample.

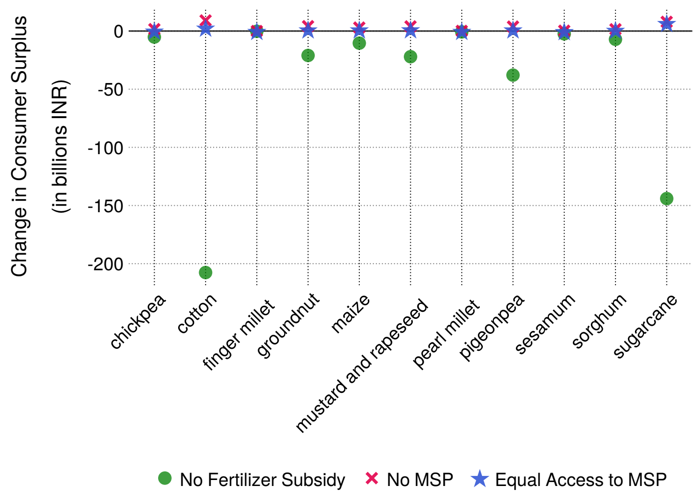


Figure A.5: *Non-PDS Crops: Change in Consumer Surplus*

Notes. This figure shows the change in consumer surplus in different counterfactuals relative to baseline for non-PDS crops. Change in consumer surplus is defined as the area under the demand curve in equation (1.5) *between the baseline price and the counterfactual price.*

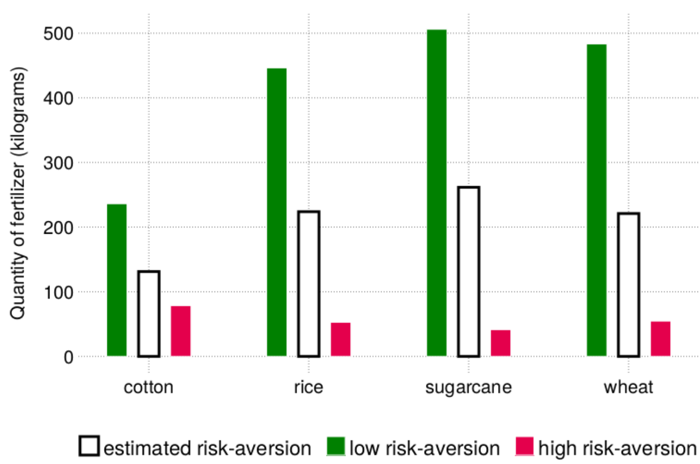


Figure A.6: *Impact of Risk Aversion on Fertilizer Usage*

Notes. This figure shows how predicted average fertilizer usage by crop would differ if risk-aversion were set to a very low or very high level, relative to the model-predicted level of risk-aversion. For all crops, average fertilizer usage falls as risk aversion goes up.

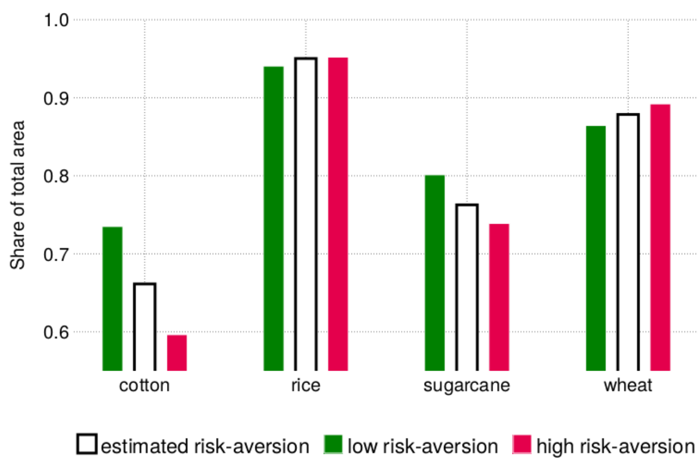


Figure A.7: *Impact of Risk Aversion on Crop Area Allocation*

Notes. This figure shows how predicted average area allocated to each crop would differ if risk-aversion were set to a very low or very high level, relative to the model-predicted level of risk-aversion. For staple crops such as rice and wheat, conditional on growing these crops, average area allocated goes up as risk aversion goes up. The converse is true for cash crops such as cotton and sugarcane.

A.2 Data Appendix

A.2.1 Cost of Cultivation Surveys (CCS)

Table A.1: *Descriptive Statistics for Cost of Cultivation Surveys*

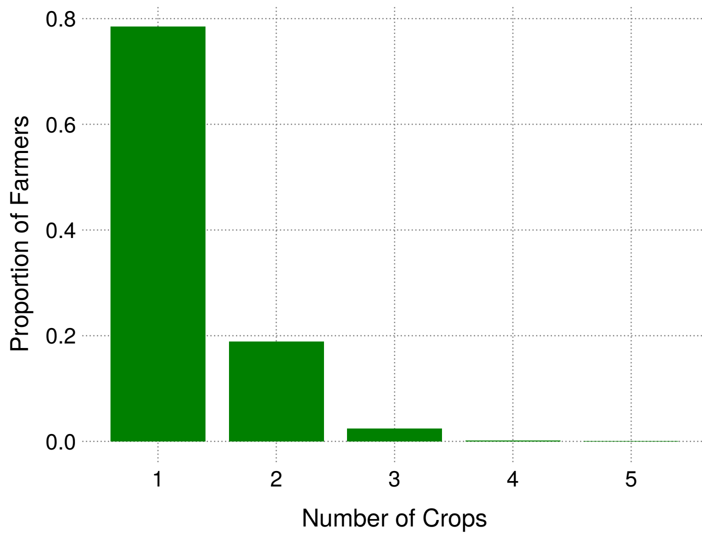
	(1) area share	(2) observation share	(3) fertilizer use / ha.	(4) labor use / ha.	(5) capital use / ha.
rice	29.70	34.84	150.59	829.96	11.34
wheat	19.76	20.24	156.74	380.14	12.63
cotton	10.08	7.54	184.87	915.09	16.09
maize	6.83	8.16	123.78	514.46	9.12
pearl millet	5.57	4.47	47.41	354.74	9.12
mustard and rapeseed	5.01	4.98	113.13	428.85	10.99
chickpea	4.37	2.98	46.03	299.57	13.91
pigeonpea	4.34	3.69	64.95	474.17	17.51
groundnut	4.20	3.29	93.57	634.47	12.72
sorghum	4.08	3.15	61.65	380.01	8.79
sugarcane	3.82	3.84	377.01	1688.76	10.21
sesamum	1.31	1.56	47.67	371.63	6.36
finger millet	0.93	1.27	88.10	767.92	7.93

Notes. This table shows some descriptive statistics from the Cost of Cultivation Survey, after resampling to match the agricultural census. Column (1) is the share of land allocated to different crops. Column (2) is the share of observations for different crops. Column (3), (4), and (5) are average fertilizer, labor, and capital per hectare for different crops. (3) is recorded in kgs per hectare, while (4) and (5) are hours of use per hectare.

Resampling CCS The government runs the Cost of Cultivation Surveys to get an unbiased estimate of the average cost of growing different crops in the country for farmers of different sizes. The sampling strategy makes the survey unrepresentative due to two reasons. First, within each primary survey unit (PSU - typically a village) the government will sample 2 farmers from each quintile of farm size distribution. Second, PSUs are sampled in proportion to area under cultivation instead of number of farmers in the PSU.

To get a representative sample at the national level, we reweight CCS using the 2016 agricultural census. Agricultural census gives us the proportion of farmers in each size-group $\times$ crop bin. For example, the proportion of farmers that have marginal land holdings ($<$ 0.5 ha.) and grow paddy. We reweight our sample to match this distribution as follows.

Figure A.8: *Share of farmers growing a given number of crops*



Notes. This figure plots the proportion of farmers that grow different number of crops within the same season. The data is from (resampled) Cost of Cultivation Surveys.

Let G denote a group defined by size-category and crop. Let $P_{\text{ag census}}(G)$ be the probability of the group in agricultural census and $P_{\text{ccs}}(G)$ be the probability of the group in Cost of Cultivation Surveys. The probability in CCS is computed as the proportion of G at the farmer-season-crop level, i.e., the share of farmer-season-crops that belongs to G. We assign a new weight for each farmer-season-crop observation in CCS as,

$$\text{weight} = \frac{P_{\text{ag census}}(G)}{P_{\text{ccs}}(G)}.$$

To compute the farmer weights, we take the mean over all season-crops for the farmer. We resample farmers according to these weights.

A.3 Model Appendix

A.3.1 Household Demand for PDS Crops

Given income y_h, household h chooses private market quantities of rice and wheat, denoted by $q^{PVT}_{\text{rice},\,ht}$ and $q^{PVT}_{\text{wheat},\,ht}$ to maximize consumption utility given by

$$\begin{aligned} \max_{q^{PVT}_{\text{rice},\,ht},\, q^{PVT}_{\text{wheat},\,ht}} \mathcal{U}_{ht} = & \left(1 + q^{PVT}_{\text{rice},\,ht} + q^{PDS}_{\text{rice},\,ht}\right)^{\delta_{\text{rice},\,h}} + \left(1 + q^{PVT}_{\text{wheat},\,ht} + q^{PDS}_{\text{wheat},\,ht}\right)^{\delta_{\text{wheat},\,h}} \\ & + \delta_{yh} \left(y_h - P_{\text{rice},\,t} \cdot q^{PVT}_{\text{rice},\,ht} - P_{\text{wheat},\,t} \cdot q^{PVT}_{\text{wheat},\,ht}\right) \end{aligned} \tag{A.1}$$

where $P_{\text{rice},\,t}$ and $P_{\text{wheat},\,t}$ are the equilibrium private market prices of rice and wheat.

Differentiating equation (A.1) with respect to private market quantity for crop c gives

$$\delta_{ch} \left(1 + q^{PVT}_{cht} + q^{PDS}_{cht}\right)^{\delta_{ch}-1} = \delta_{yh} P_{ct}$$

Taking logs and re-arranging gives

$$\begin{aligned} \log\left(1 + q^{PVT}_{cht} + q^{PDS}_{cht}\right) &= \log\left(1 + q_{cht}\right) \\ &= \frac{\log \delta_{ch} - \log \delta_{yh}}{1 - \delta_{ch}} - \frac{\log P_{ct}}{1 - \delta_{ch}} \end{aligned}$$

where q_{cht} is the total consumption of crop c.

Consider the approximation

$$\begin{aligned} \frac{\log \delta_{ch} - \log \delta_{yh}}{1 - \delta_{ch}} &\approx \alpha_{cy} \log y_h \\ \frac{1}{1 - \delta_{ch}} &\approx -(\alpha_{cp} + \alpha_{cpy} \log y_h) \end{aligned}$$

Plugging it back in gives

$$\log\left(1 + q_{cht}\right) = a_{cp} \log P_{ct} + a_{cy} \log y_h + a_{cpy} \log P_{ct} \cdot \log y_h$$

which is the specification proposed in equation (1.4).