

Market Power and Spatial Competition in Rural India^{*}

Shoumitro Chatterjee[†]

January 2019

[Find the most updated version here](#)

Abstract

In this paper, I argue that market power of intermediaries plays an important role in contributing to low incomes of farmers in India. I study the role of spatial competition between intermediaries in determining the prices that farmers receive in India by focusing on a law that restricts farmers to selling their goods to intermediaries in their own state. I show that the discontinuities in market power generated by the law translate into discontinuities in prices. Increasing spatial competition by one standard deviation causes prices received by farmers to increase by 6.4%. To shed light on spatial and aggregate implications, I propose and estimate a quantitative spatial model of bargaining and trade. Using this structural model, I estimate that the removal of the interstate trade restriction in India would increase competition between intermediaries substantially, thereby increasing the prices farmers receive and their output. Estimates suggest that average farmer prices and output would increase by at least 11% and 7% respectively. The value of the national crop output would therefore increase by at least 18%.

JEL Classifications: D43, F12, L13, L81, O13, Q13, R12

^{*}I am deeply indebted to my advisor Esteban Rossi-Hansberg for his continued guidance and support. Special thanks to Bo Honoré, Kala Krishna, Eduardo Morales, Richard Rogerson, and Tom Vogl for their suggestions and encouragement. I have hugely benefited and learned from endless conversations with Skand Goel, Seher Gupta, Ajay Jakhar, Devesh Kapur, Mekhala Krishnamurthy, Rohit Lamba, David Nagy, Anders Nielsen, Sutirtha Roy, Gaurav Sabharwal, Sanjay Raj Singh, and Dean Spears. I am extremely grateful to Arvind Subramanian as the intellectual roots of this project lie in the work I did while interning at his office in the Ministry of Finance in India. I gratefully acknowledge financial support from the International Economics Section and the Princeton Institute for International and Regional Studies at Princeton University. I would also like to acknowledge that most of the work reported on in this paper was performed at the TIGRESS high performance computer center at Princeton University. All errors are mine.

[†]Department of Economics, The Pennsylvania State University, 605 Kern Building, University Park, PA 16801, USA. Email: sc20@psu.edu. Website: <https://www.personal.psu.edu/sc20>.

1 Introduction

Small farmers in low-income countries are among the poorest people in the world. Their revenues may be low partly because they are unproductive but also because they may receive low prices for what they produce. One reason farmers receive low prices could be the monopsony power of intermediaries or “middlemen,” who are the main buyers of farmers’ output in developing countries. A source of their market power is the limited ability of farmers to arbitrage between different intermediaries due to high costs of transportation or policy regulations. This channel of spatial competition between intermediaries, i.e. the ability of farmers to conduct spatial arbitrage, may be an important determinant of the income and welfare of farmers. In this paper, I use microdata from India on locations of intermediary markets, prices, and data on agriculture production to study the importance of spatial competition between intermediary markets for farmer incomes and production.

Key to my approach are the Agriculture Produce and Marketing Committee (APMC) Acts of Indian states that restrict farmers to selling their output to government-licensed intermediaries in government-regulated markets of their own state. The analysis has two parts. In the first part of this paper, I establish that spatial competition between intermediaries is an important determinant of the prices that farmers receive. There are two challenges in analyzing the relationship between spatial competition and prices: first, credibly measuring spatial competition and second, the endogeneity in the location of intermediaries. Besides its intrinsic interest, the Indian context helps me address both of these challenges. Monopsony rights of the APMC markets and their fixed spatial location allow me to measure competition between markets and relate it to price realizations. More importantly, the restriction on farmers’ crossing state boundaries ensures a discontinuity at the border in the competition faced by markets. This allows me to establish a causal relationship between competition and prices.

In particular, I define spatial competition as the number of markets in the neighborhood of each market weighted by the inverse of their distance.¹ I form market pairs by matching markets that are in close proximity to each other but lie on different sides of a state border. Because farmers cannot sell across state borders, the competition faced by markets changes discontinuously at the border. However, other factors that influence prices, like geography and local demand, are likely to be similar

¹This is a standard measure of market access in the trade literature. e.g. [Harris \(1954\)](#), [Donaldson and Hornbeck \(2016\)](#), and [Allen and Atkin \(2016\)](#).

and vary continuously across the border. Therefore, I recover a causal effect of increased competition by comparing the difference in prices that farmers receive in each market pair to the difference in competition faced by them. My estimates suggest an average increase of 6.4% in farmer prices due to one standard deviation increase in spatial competition.

In the above context, a policy that would increase spatial competition is the removal of interstate trade restriction on farmers. The effects of such a policy change cannot be directly inferred from the data, and evaluating them requires a structural model. Specifically, a reduced-form exercise will fail to capture at least three crucial elements of the change. First, removal of trade restrictions will not only affect the markets on the border, but will also cause ripple effects in the markets located in the interior of states as farmers eliminate arbitrage between markets. The magnitude of these effects will depend on transportation costs, the actual location of markets, local demand conditions, and cropping patterns. Second, a change in the prices received by farmers reflects only a part of the benefit. Higher prices in turn incentivize farmers to reoptimize the use of intermediate inputs and hence alter production. Third, changes in supply can lead to changes in retail prices of crops, which would feed back into the prices received by farmers.

In the second part of the paper, I develop a spatial model of trade in agricultural markets in India to estimate gains from removing interstate trade restrictions. The model uses the economic geography and flexibly captures the determinants of spatial competition. In the framework, post-harvest, farmers optimally choose a market to sell their output. At the market, they Nash-bargain with the associated intermediary,² who in turn sells the purchased goods in the local retail market. When farmers bargain with an intermediary at a market, they alternatively consider transporting their goods to another market nearby and selling at a higher price. Geography and policy create spatial variation in the farmers' outside options, and therefore, spatial variation in the degree of market power that intermediaries can exert.

I recover a structural relationship between the prices that the farmer receives at a particular market site, the prices in the neighboring markets, and the local retail price. This relationship depends on the Nash-bargaining parameter, geographic location of the markets, and the transport

²In practice, there may be 15-20 intermediaries in a market, but empirical ([Banerji and Meenakshi, 2004](#); [Meenakshi and Banerji, 2005](#)) and ethonographic ([Krishnamurthy, 2011](#)) works suggest collusion among intermediaries in these markets. Therefore, in the absence of data on the number of intermediaries within a market, I assume that each market is associated with a representative intermediary.

costs. I estimate the model parameters using a method of simulated moments procedure to match the overall spatial distribution of farmer prices. This procedure helps me address issues stemming from the simultaneity in the determination of prices at each location.

Next, I use the estimated model to conduct policy counterfactuals in which I remove the restriction on farmers' selling across state borders. This change increases spatial competition between markets. I quantify the associated benefits to farmers under four different scenarios. In the first exercise, I keep all decisions of the farmer fixed except the threat of the farmers' selling at a different market. Specifically, farmers go to the same market as before but while bargaining with the intermediary, can threaten to sell in a market outside the state.³ This increases farmer prices by about 3% on average. For the farmers with above-median gain, the average increase in prices is about 7%. The maximum increase in farmer prices is 20%. Increased competition between markets close to state borders changes the outside option of the farmers trading in the markets adjacent to them. This mechanism ripples across space, and as a result even farmers in the interior of states experience an increase in prices. However, due to costly transportation, the price increases are larger in markets that are closer to state borders.

In the second exercise, I additionally allow farmers to optimize their choice of the market where they trade. This increases farmer prices by about 11% on average. For the farmers with above-median gain, the average increase in prices is about 21%. The key channel that generates the large increase in prices over the previous exercise is the geographical location of farmers relative to the location of markets. Consider farmers near state borders who are closer to markets of a neighboring state but may be far from markets in their own state. These farmers incur large transportation costs when border crossing restrictions are in place. With no restriction on crossing state boundaries, these farmers benefit not only from greater spatial competition but also from gaining access to closer markets, which reduces their transportation costs.

In a third exercise, I also allow farmers to reoptimize the use of intermediate inputs in response to getting higher prices for their output and estimate the associated change in production. I find that the crop output of farmers increases by about 9% on average. For the farmers with above-median gain, the average increase in crop output is about 15%. My estimates suggest this increases the

³In the model, this changes the threat point of the farmer in the Nash-bargaining problem but not his market choice.

national crop output by about 8% and the value of the national crop output, at farmer prices, by about 18%.

Since increased crop output can depress retail prices and feed back into farmer prices, in a fourth exercise, I allow the retail prices to adjust as a result of the change in supply. I estimate the change in prices and production in the long-run equilibrium of the model. Although a small fraction of farmers lose because of the downward adjustment in retail prices, the losses are small with respect to the overall gains. Farmer prices increase by about 9% on average. For the farmers with above-median change, the average increase in prices is about 19%. Among the farmers who lose, the average decline in prices is about 10%. The reason why the average increase in prices is similar to the earlier exercises, even after accounting for losses, is the following. Upon removal of border crossing restrictions, farmers from regions with lower initial prices transport their output and sell in regions with higher initial prices. Thus, regions with higher initial prices experience an increased supply and hence a reduction in retail prices. The opposite happens in regions with low initial prices, and retail prices adjust upwards. Compared to the scenario with no adjustment of retail prices, farmers in the high initial retail price regions lose on average. The effect on farmers in regions with low initial prices is ambiguous. In the scenario with no adjustment of retail prices, these farmers incurred transport costs to get access to better markets. When I allow retail prices to adjust, they can get higher prices in their local regions and save the transport cost. Therefore, they could potentially gain more compared to the scenario with no retail price adjustment.

The quantitative spatial model also allows me to show that spatial competition interacts in important ways with other development policies. The market access of farmers can be increased either by lowering transport costs or by letting them sell in markets in other states. Using the model, I show that there is a large interaction effect of the two policies. Farmers gain substantially more from the simultaneous reduction of transport costs and the elimination of border restrictions to trade, than the sum of the gains from enacting each of the two policy changes alone. Moreover, the additional gains occur in locations with greater market power of intermediaries, where reductions in transport costs on their own would not have a large effect.

The quantitative magnitudes from the counterfactual exercises should matter for policymakers. First, they provide an estimate of the gains of a potential reform in India that the central government has been pushing the states toward—creating a National Agriculture Market (NAM) that lets farmers

trade in any market in the country. Second, an understanding of the spatial distribution of market power provides insight on policy design that is likely to yield better results for raising farmer incomes (and concomitant poverty alleviation) in specific regions. For example, farmers in more competitive regions are likely to gain more from investments such as road construction. On the other hand, policymakers need to explicitly target competition between intermediaries in regions where they exert considerable market power. Third, while subsidies in the input and output markets have conventionally been the focus of policies meant to improve farmer incomes in developing countries, this paper highlights the importance of physical output market sites and policies that restrict market access for farmer incomes. Subsidies not only place a fiscal burden on the exchequer but also distort market prices, whereas gains from implementing policies that increase competition⁴ can essentially be achieved with little financial cost.

In summary, in this paper I make three contributions. First, I establish that spatial competition between intermediaries is an important channel for determining prices that farmers get. Second, I show that policies that increase spatial competition can have large effects on farmers' welfare which manifest through two channels; (a) a ripple effect of one market affecting the other; and (b) alteration of the production decisions of farmers. Third, I show that spatial competition can have quantitatively large interaction effects with other development interventions such as road construction.

The work in this paper relates to a number of different strands of literature. Most relevant is the work that considers the distribution of gains from trade in the presence of variable markups and intermediation. Although there is recent work that examines gains from trade in the presence of variable markups, the focus has been on producers rather than intermediaries with market power (see, for example, [Melitz and Ottaviano, 2008](#); [Edmond et al., 2015](#); [De Loecker et al., 2016](#)). In this paper, I introduce the dimensions of intermediation and spatially varying market power, and estimate gains from changes in the market power of intermediaries. In that respect, my work is complementary to [Atkin and Donaldson \(2015\)](#), who show that the market power of intermediaries is an important determinant of actual prices paid by consumers. However, the goal in [Atkin and Donaldson \(2015\)](#) is to separately identify trade costs from market power. I contribute by establishing

⁴Many other policies can increase spatial competition, such as creating new markets or allowing farmers to directly sell to supermarkets. Quantifying gains from those policy changes will require additional information, such as estimates of the fixed cost of entry of new markets or the location and capacities of existing supermarkets. In the absence of such data, I choose to relax a realistic policy regulation, which allows me to study the same channel without having to worry about other margins.

the micro-foundations of a distinctive channel—that of bargaining with spatially varying threat points—through which intermediaries in remote locations may enjoy market power. This is essential to conducting the policy counterfactuals.

Although the core mechanism in this paper is not surprising and has been conjectured by others (e.g. [Dillon and Dambro, 2016](#)), the theoretical literature on intermediation and trade ([Antras and Costinot, 2011](#); [Bardhan et al., 2013](#); [Krishna and Sheveleva, 2017](#)) does not explicitly model it. I contribute to this literature by building a spatial model of bargaining and trade with intermediation where intermediaries enjoy market power because it is costly for the farmers in remote locations to access alternative buyers.

This paper is also related to a growing body of quantitative economic geography literature that estimates welfare gains from integration. For example, [Donaldson \(2015b\)](#) and [Donaldson and Hornbeck \(2016\)](#) estimate welfare gains as a result of reduction in transport costs. [Allen and Atkin \(2016\)](#) explore the effect of trade on the link between integration and volatility of returns for Indian farmers. While this literature has largely focused on perfectly competitive environments often building on the seminal work of [Eaton and Kortum \(2002\)](#), I show that gains from integration or reduction in trade costs are heterogeneous in space and the magnitudes relate to the spatial variation in market power of intermediaries.

The existing empirical literature that measures market power of intermediaries in agricultural markets (surveyed in [Dillon and Dambro, 2016](#)) provides mixed evidence, with some papers estimating sizable market power of intermediaries ([Bergquist, 2017](#); [Casaburi and Reed, 2016](#)) and others finding that it is not so much ([Fafchamps et al., 2006](#)). This body of work has focused on testing competitiveness of particular agricultural markets without regard to their spatial locations. Papers either estimate pass-through rates at markets (e.g. [Bergquist, 2017](#); [Casaburi and Reed, 2016](#)), directly measure margins of traders ([Fafchamps and Minten, 2002](#)), or account for the entry and exit of intermediaries ([Fafchamps et al., 2005](#)). This paper shows that the interaction of economic geography with spatial competition can generate spatially varying market power for intermediaries. It, therefore, provides a possible rationale for isolated studies finding different estimates of the market power of intermediaries.

There are three related papers on agricultural markets in India. [Banerji and Meenakshi \(2004\)](#) and [Meenakshi and Banerji \(2005\)](#) analyze transactions-level data from markets in North India to

identify collusion among traders. [Mitra et al. \(2017\)](#) estimate high trader margins in the state of West Bengal. They conclude that their results are inconsistent with long-term contracts between farmers and traders but consistent with a model of ex-post bargaining. While these papers provide useful insights about the working of these markets, I further contribute to this literature by using microdata to estimate the spatial distribution of the market power of intermediaries in most of India.

Finally, since the seminal work of [McCallum \(1995\)](#), an extensive literature has documented significant costs of crossing national borders (e.g. [McCallum, 1995](#); [Anderson and van Wincoop, 2003](#); [Engel and Rogers, 1996](#)). I show that significant border costs exist even within a large developing country. This paper is complementary to [Cosar et al. \(2015\)](#), that studies spatial competition between wind turbine manufacturers and the impact of the Denmark-Germany border on their production costs.

The rest of the paper is organized as follows. In the next section, I provide the institutional background of agricultural trade and production in India. Section [3](#) outlines data sources. Section [4](#) presents reduced-form evidence for the main mechanism, i.e. spatial competition causing an increase in prices that farmers get. In section [5](#), I develop the quantitative framework that I will use to conduct policy counterfactuals. Section [6](#) discusses the structural estimation of the model. In section [7](#), I study the effect of increased competition due to removal of inter-state trade restrictions on farmer incomes and production. Section [8](#) discusses robustness and sensitivity of results. Finally, in section [9](#), I offer concluding remarks.

2 Agricultural Trade in India

In this section, I provide a brief description of agricultural trade in India, the current institutional setting, and a brief history of the regulated agricultural markets.

2.1 Empirical Context

Agriculture is an important sector of the Indian economy: in 2011, 54.6% of the total workforce was employed in agriculture, and the sector comprised 18.52% of India's total GVA.⁵ Eight non-perishable

⁵Agriculture Statistics at a Glance 2016, Ministry of Agriculture & Farmer Welfare, Government of India.

crops—rice, wheat, maize, sorghum, barley, finger millet, pearl millet, and soybean—account for 70% of India’s gross cropped area. In this paper, I will restrict attention to these crops.

There are two cropping seasons each year: fall or *kharif* (July to November during the south-west monsoon) and spring or *rabi* (November to March). Some regions also have a third summer crop between March and June. Rice, sorghum, maize, millet, and soybean are primarily grown in the *kharif* season, while wheat and barley are grown in the *rabi* season.

The median Indian farming household operates a small farm of 1.5 hectares and cultivates it with the help of family or village labor. Its net annual income⁶ is approximately USD 365.⁷ Usually, the farmers keep a small fraction of the final crop output for personal consumption and sell the rest to licensed intermediaries in a government-regulated marketplace (known as a *mandi*).

Thus, the institutional setting is comprised of three economic markets: (a) a market for intermediate inputs for farming; (b) the regulated market where farmers sell their output to government-licensed intermediaries; and (c) the retail markets where the intermediaries sell. The focus of my analysis is the regulated market and the transaction between the farmer and the intermediary.

Present-day trade in agricultural commodities in India is regulated by two key laws, the Essential Commodities Act (1955) of the central government, which prohibits hoarding, and the autonomous state-level APMC Acts. The APMC Acts mandate that after harvest the *first* sale and purchase of agricultural commodities produced in the state must be carried out in government-designated marketplaces, and buyers of agricultural output must obtain a license from the marketing committee of the marketplace. The laws de facto prohibit farmers from selling in marketplaces of a neighboring state (Chand, 2012). Thus, these Acts restrict the set of buyers of farmers’ output to intermediaries within the state. Thereby these laws also reduce competition between marketplaces across state borders and provide a source of variation that can be exploited to estimate gains from increasing competition in the presence of spatially varying market power.

According to officials in the central and state governments,⁸ the enforcement of the border restriction is not hard. Typically, whenever farmers try to cross state borders, intermediaries protest with local law enforcement agencies who in turn punish a farmer to set an example. This does not

⁶Includes personal consumption valued at market prices.

⁷Economic Survey of India, Ministry of Finance, Government of India, 2016

⁸While interning at the Office of the Chief Economic Advisor of India in 2015–16, I interviewed senior bureaucrats at the Ministry of Finance, the Ministry of Agriculture, and in three state governments (Punjab, Karnataka and Madhya Pradesh) to understand the functioning of agricultural markets and the implementation of the laws.

preclude all leakages but it is well understood in public policy debates that this restriction hurts farmers. Moreover, the central government has been trying to push the state governments to reform these laws.⁹ However, these efforts have had little success.¹⁰

As such, the regulated markets are the most important institutions for determining farmer incomes in India. The laws, as discussed, create licensed intermediaries who have monopsony rights over the farmers' output in the state. An intermediary has a license to a particular marketplace. Although multiple licenses are given per market, only a few are active and cartelization among intermediaries is common. Incumbent intermediaries also prevent new entrants (Chand, 2012). Once an intermediary has bought crops from farmers, he moves them up the supply chain and sells them to a large miller or in the retail markets. There are no legal restrictions on where the intermediaries can sell.

Given the importance of the regulated markets and the APMC Acts, I present a brief history of them next.

2.2 The *Mandi* Institution: A Brief History

The idea of the regulated agricultural marketplace in India has its genesis in a 1928 report by the British colonial government's Royal Commission on Agriculture in India. The chairman of this commission, Lord Linlithgow, had presided over a similar committee in 1922 that was supposed to investigate and comment on the "prevailing legal chaos" that raged around British agricultural markets. Four years later, in 1926, Linlithgow brought with him many of his unimplemented ideas derived from the British markets to his new job in India (Harris, 1984; Krishnamurthy, 2011). In his comprehensive 1928 report, he declared that the countrywide establishment of regulated markets "would confer an immense boon on the cultivating classes of India" (Royal Commission on Agriculture in India 1928).

This idea was entrenched in the post-colonial policy for agricultural development and regulated markets were seen as key to helping farmers realize a reasonable price in an environment where private trade was underdeveloped and controlled by mercantile power. Therefore, when the constitution of

⁹See, for example, "Will a national common market for farm products curb price rise?," in Livemint, July 10, 2014 ([Click here](#)), and "Agriculture reform: Government takes first step for a national farm market," in The India Express, July 3, 2015 ([Click here](#)), which describe how the current cross-border restrictions hurt farmers and the efforts of the central government to ameliorate them.

¹⁰For details, see "Farm Market Reforms: Not in NAME alone" in The Indian Express, June 30th, 2016 ([Click here](#)).

independent India gave legislative powers regarding agriculture to the states, each state of India adopted its own APMC Act in the 1960s.

Under the aegis of the APMC Acts, new markets were created, especially during the Green Revolution that started in the mid-1970s, when agricultural production started increasing exponentially. Between 1976 and 1991, agricultural production grew by 74%, and there was a concomitant 76% increase in the number of markets. The construction of *mandis* has plummeted since the late 1990s. Since 2001, very few new markets have been constructed (Chand, 2012).¹¹

2.3 Geographical Placement of Markets

Most of these market were constructed between 1960 and 1985, during the peak of the green revolution in India, when grain production increased exponentially. According to the senior bureaucrats I interviewed in the state and central ministries of agriculture, historically, more markets were created in regions that had a greater yield or output per capita. In Appendix B, I verify that controlling for size, cropped area, and population of districts in 1970–1985, the more productive districts have more markets today.

3 Data

Geospatial data on the location of markets and prices of commodities sold in them is necessary to assess the importance of spatial competition between intermediary traders in price determination. To capture the demand side, this should be matched geographically to local retail markets and must capture the local distribution of the population and retail prices. Further, credible quantification of production losses requires fine geospatial data on land productivity, land use, rainfall, and crop choice. Because such a data set is not publicly available, I have assembled a rich and novel microdata set for India covering the decade of 2005–2014 by combining data from various sources.

¹¹In my data set, I do not observe the year of construction or age of a market. As such, I will assume that in my sample period the number of markets is constant. Chand (2012) reports the aggregate growth rate of markets after 2001 to be 0.7%, and hence, this will not be a major concern for me.

Intermediary Markets: My main data set is comprised of monthly data on the modal price¹² of eight major non-perishable commodities, specifically rice, wheat, maize, sorghum, barley, pearl millet, finger millet, and soybean, sold in any regulated rural agricultural market, along with the village names of the market. This is the price farmers get when they sell in these markets. I obtained this from the Ministry of Agriculture in India. I use Google Maps API to geocode the location of these markets.

I combine this with data on retail prices and production. These data are available only for administrative districts of India. There are 455 districts in the sample I consider, so they can substantially capture the spatial heterogeneity.

Retail Prices: I use monthly data on retail prices at the district level from the National Sample Survey (NSS) Schedule 3.01(R) - Rural Price Collection Survey (RPC) survey of the Central Statistical Organization of India. These data are collected from a fixed set of 603 villages spread across India and are available for the years 2005–2011, except 2008.

Production and Yields: To understand the cropping patterns of a district, I use data from the National Sample Survey (NSS) Schedule 33 - Situation Assessment Survey of Agricultural Households (Round 70) 2013. This is a large survey of rural agricultural households conducted twice a year, once in each cropping season, in 4529 villages covering 35000 households. Because the survey is representative at the district level, it provides me with a good estimate of the local cropping patterns.

In addition, I use estimates of the price elasticity of demand from [Deaton \(1997\)](#) and of local crop yields at the district level for the years 2005–2014 provided by the Ministry of Agriculture of India.

I further match the above data on prices, production, and consumption with finer data on land use, land elevation, and distribution of population. In particular, I obtained gridded data on land use from Princeton University’s Geospatial Information Systems library, gridded data on land elevation produced by NASA’s Shuttle Radar Topography Mission (SRTM),¹³ gridded data on rainfall from

¹²These are producer prices,–i.e. the price that a farmer receives. The Indian government also has a Minimum Support Price (MSP) program and announces the procurement prices for 26 crops prior to the start of an agricultural season. Although legally, the government promises to buy farmers’ output at the announced MSP whenever market price falls below it, reality is far removed. The MSP is only implemented for rice and wheat, and most procurement occurs in two states, Punjab and Haryana. Even in these states, the government actually relies on the intermediaries to procure the crops ([Chatterjee and Kapur, 2017](#)). Appendix Figures [J.1–J.10](#) plot the distribution of log prices alongside the MSPs for the crops. There is little evidence of price heaping just at or above the MSPs, as well as substantial mass below MSPs, suggesting any biases from excluding MSPs from the model is limited.

¹³[National Aeronautics And Space Administration And The National Geospatial Intelligence Agency, 2000](#)

Willmott and Matsuura (2001), and geocoded data on village population from the Census of India 2001. I compute district level monthly rainfall by taking an inverse-distance weighted average of all the grid points within the boundary of any district.

Sample: For my analysis I exclude the mountainous states of India because of the difficulty in measuring distances. I exclude the state of Bihar and all North Eastern states because they do not have an APMC Act, and hence there is no data on market-level prices. I also exclude some other territories and islands where agriculture is not practiced on any substantial scale. Therefore, my sample includes 15 states in mainland India covering 89% of the total cropped area and accounting for 90% of total production. The states are shown in Figure A.1. In my data set, this results in a total of 2978 primary grain markets.

Table 1 presents summary statistics on prices at the market and retail levels. The average variance of farmer prices across markets within crop-month is 0.029. The variance in retail prices across regions, conditional on the crops' being produced in that region, is lower than the variance in farmer prices. The average variance in retail prices across districts within state-crop-month is 0.014, whereas the average variances in farmer prices across markets within state-crop-month is 0.022.

4 Does more spatial competition increase farmers' prices?

The analysis in this section is focused on the regulated marketplace where the transaction between the farmer and the government-licensed intermediary occurs.¹⁴ There are two potential forms of local competition among buyers (i.e. the licensed intermediaries in markets): between and within market sites. In this paper, I will focus on between-market competition because current empirical evidence suggests that intermediaries within a market collude.¹⁵ This section provides two forms of reduced-form evidence that suggest: (a) local competition between market sites increases prices that farmers get for their output (hereinafter farmer prices); and (b) state borders attenuate competitive forces, i.e. the presence of more markets in the neighboring state does not influence farmer prices.

The hypothesized mechanism influencing farmer prices, which will be explicit in the quantitative model, hinges on a farmer's access to alternative buyers while negotiating with a buyer at a given market site. More markets in the vicinity increase the set of alternatives available to the farmer,

¹⁴See Section 2.1 for a discussion of the overall economic environment of agriculture trade in India.

¹⁵See, for example, Banerji and Meenakshi (2004) and Meenakshi and Banerji (2005).

increasing the competition faced by the present buyer. Consequently, the farmer is likely to be offered a better price. Therefore, greater competition is likely when alternative markets are closer and there are more of them. Moreover, since the APMC laws prevent farmers from crossing state borders, more markets in a neighboring state should not have an influence on farmer prices. I test for these implications below.

4.1 Effect of within- and out-of-state competition

To explore the association between local competition and farmer prices, I regress farmer prices at a particular market site on a measure of local market density (or local competition faced by a market). Similar to the market access measure in [Harris \(1954\)](#) and [Donaldson and Hornbeck \(2016\)](#), I construct a local competition measure by taking a weighted sum of other markets near a particular market site but in the same state. The weights are the inverse of distances of the neighboring markets to the origin market.¹⁶ For any market m ,

$$\text{comp}_m = \sum_{j \in \mathcal{M} \setminus \{m\}} \left\{ \frac{1}{\text{distance}_{mj}} \right\} \mathbb{1} \{ \text{state of } m = \text{state of } j \} \quad (1)$$

\mathcal{M} is the set of all markets in the country. As competition (in any market m) is driven by a farmer's ease of access to alternative markets, the comp_m measure assigns a greater weight to a closer market. I also create an analogous competition measure for a market site from the markets not in the same state. Under the null hypothesis, this measure should have no association with prices received by farmers at any market site.

$$\text{comp}'_m = \sum_{j \in \mathcal{M} \setminus \{m\}} \left\{ \frac{1}{\text{distance}_{mj}} \right\} \mathbb{1} \{ \text{state of } m \neq \text{state of } j \} \quad (2)$$

My main specification takes the following form:

$$\log p_{cndst}^f = \beta_0 + \beta_1 \text{comp}_m + \beta_2 \text{comp}'_m + \mathbf{X}'_{cndt} \beta_3 + \gamma_t + \gamma_c + \gamma_s + \epsilon_{cndt}, \quad (3)$$

¹⁶Donaldson and Hornbeck (2013) compute a weighted average of the total trade in a market. I do not have data on trade and therefore do not compute the trade-weighted measure. However, the number of proximate competitors, as a measure of competition, has also been used by Machiavello and Morjaria (2016).

Here, p_{cmdst}^f is the price that a farmer receives at market site m located in district d , state s , for crop c , at time (month-year) t . All price observations are at market-crop-month level. \mathbf{X} includes controls for district specific time (year) varying controls like crop yields, crop area, population, and district-crop specific rainfall shocks. γ_t is a month-year fixed effect and controls for crop and district invariant unobservables such as macroeconomic shocks, large scale droughts, etc. γ_c controls for crop-specific factors such as crop-specific price levels and national tastes for particular crops. γ_s controls for state-specific and crop-time invariant effects such as state-level policies, relative incomes of states, and local tastes. Since cropping decisions and other shocks are likely to be spatially correlated, I compute robust standard errors clustered at the district level.

Variation in comp_m comes from geographical differences in the placement of markets. Because the local crop production to population ratio drives this variation, I control for local yields, local cropped area, and population in the regression (3). β_1 is identified from within-state variation in comp_m under the assumption that comp_m is uncorrelated with the residuals. As very little *mandi* construction occurred in my sample period, and I do not have data on the date of construction of markets, comp_m does not vary over time. Therefore, threats to identification can come only from spatially varying factors, and I have tried to control for as many of them as possible. I have not included district fixed effects because there are very few markets within a district, and therefore, I do not have enough within-district variation in spatial competition to identify its effect on prices.

Within-market competition is an omitted variable in all my regressions and can potentially confound my estimates if markets in local areas with a greater market density also have greater within-market competition. In the absence of any data on the number of buyers within a market or their transactions, I cannot control for it directly. However, within-market competition is unlikely to be important because existing evidence points towards collusion among intermediaries within agricultural markets, not only in India (Banerji and Meenakshi, 2004; Meenakshi and Banerji, 2005) but also in other parts of the developing world (Bergquist, 2017). In the Indian context, incumbent intermediaries have also actively tried to prevent entry of new traders (Chand, 2012).

Table 2 reports the regression results. Column 1 reports results from the base specification given in equation (3). The coefficient on competition, β_1 , is significant and equals 0.0163. In Column 2, I also include crop-time fixed effects, which control for monthly world price shocks, and state-year fixed effects, which control for state-specific income levels and policy changes that vary over time.

This does not change the estimate of β_1 much, which now equals 0.0146 and is still significant. The coefficient on out-state competition β_2 is close to zero and not significant in both columns.

The key message is that greater local competition increases farmer prices. A one standard deviation increase in competition increases prices by 2.7%. Moreover, the competition from markets in other states has no impact on prices. To get a sense of the overall magnitude of gains, note that if we removed all border restrictions to trade, the median increase in competition would be 1.6 standard deviations. This would increase prices in half the markets in India by at least 4.4%.

4.2 Causal Effects of Local Competition on Prices

Although the previous regression design shows that local competition increases farmer prices and borders reduce out-of-state market competition, one can still be concerned about other forms of unobserved heterogeneity. Therefore, to test this idea more concretely, I implement a border discontinuity design with market pairs. I match all markets, which are less than x kilometers apart but lie on different sides of a state boundary. Here, $x = 25, 30$, and 35 . I regress the difference in prices for each market pair on the difference in their competition measures.

The basic idea is that other determinants of prices like demand, productivity via soil quality, and rainfall vary continuously across a state boundary. Indian state boundaries were primarily determined along linguistic or cultural lines rather than geography (Guha, 2007). Thus, the only determinant of prices that changes discontinuously across state borders is local competition because farmers are not allowed to sell their output in the markets of other states.¹⁷ Therefore, along a state boundary, the correlation between the difference in prices and the difference in market density between market pairs should help us identify causal effects of local competition on prices.

In the previous design, since I relied on within-state variation for identification, I could not fully address concerns about unobserved heterogeneity. The advantage of this design is that I can difference out unobserved factors other than competition that affect prices by choosing market pairs very close to each other.

¹⁷Indian languages change very gradually over distance. Therefore, it is not the case that farmers in one state will not be able to communicate with farmers or intermediaries in the neighboring state.

I start with a levels specification as in (3) and then take a difference between all market pairs $(m - m')$ that are within the selected bandwidth but lie in different states.

$$\begin{aligned}\log p_{cmdt}^f &= \beta_0 + \beta_1 \text{comp}_m + \gamma_t + \gamma_c + \gamma_s + \epsilon_{cmdt} \\ \Delta \log p_{cmdt}^f &= \beta_1 (\Delta \text{comp}_m) + \gamma_{ss'} + \tilde{\epsilon}_{cmt}\end{aligned}\tag{4}$$

β_1 is the causal effect of competition on prices. $\gamma_{ss'}$ is a border fixed effect, specific to two states sharing a common border. This specification controls for unobservables like differences in state-specific policies, law enforcement, and tax regimes that do not vary along the common border but might potentially differ across state pairs. Because the market pairs are very close to each other, shocks can be spatially correlated. I report (in square brackets) standard errors corrected for spatial correlation.

Following Conley (1999), I allow for spatial dependence of an unknown form within a 100 km surface from the mid-point of the line segment joining the market pair. To be more precise, the variance-covariance matrix is estimated as a weighted sum of cross-products of the OLS moments ($E[x_i \epsilon_i] = 0$). I use two kernels for determining the weights: (a) the uniform kernel, which puts equal weight on all observation pairs within 100 km and 0 otherwise; and (b) the Bartlett kernel, which gradually decreases the weight the further the observations are. For comparison, I also report robust standard errors in parentheses.¹⁸

I report the results in Table 3. The preferred estimates are reported in column (2) which has a bandwidth of 30 km. The marginal effect of competition on prices, β_1 , equals 0.035. Choosing bandwidths lower than 30 km reduces the number of market pairs and then I do not have enough power to precisely estimate the effect of competition on prices. As reported in Column 1 of Table 3, where the bandwidth is 25 km, β_1 is equal to 0.025 but is not precisely estimated. Also, as can be seen from Column 3 of the same table, the estimate of β_1 settles down to about 0.035 as we increase the bandwidth further.¹⁹

Like before, there is a positive effect of local competition on farmer prices. However, the effect size is almost double compared to the estimate of 0.016 from the levels specification reported in Table 2.

¹⁸I am restricted to computing cluster robust standard errors for the levels specification (3) because of the computational burden of computing Conley (1999) standard errors.

¹⁹Note that my choice of bandwidths is much smaller than similar border discontinuity designs as in Dell (2010).

A farmer selling in a market that faces the 75th percentile of competition compared to one that faces the 25th percentile of competition gets a 3.5% higher price on average. A one standard deviation increase in competition would lead to a 6.4% increase in price on average. Appendix C presents placebo tests to show that there are no discontinuities at state borders in other observable factors like potential crop productivity, land elevation, and distance to the nearest district headquarter that might influence farmer prices.

Another way of estimating this causal effect is to use a levels specification like (3) to mimic the border discontinuity design. β_1 could be biased in (3) if it is correlated with other omitted factors that also affect prices. However, if we limit our sample to markets very close to all state borders, then these omitted factors are likely to affect comp'_m in a manner similar to their effect on comp_m . Therefore, $\beta_1 - \beta_2$, should help us identify the “true” competition effect as β_2 purges β_1 from any contamination.

I report these results in Columns 3 and 4 of Table 2, where we indeed find that the estimated effect of 0.032 is very close to the estimated effect of 0.035 in the border discontinuity design.

To summarize, in this section I have provided evidence for two aspects of spatial competition between intermediary markets. First, local competition, determined via density of markets, matters for farmer prices. Second, border restrictions have real effects in that they nullify the competitive forces from markets in neighboring states. As such, the quantitative effects of increased competition on prices, agricultural output from an induced change in use of intermediate inputs, and other welfare measures can be large if border restrictions on trade are removed. With farmers’ being able to freely cross state boundaries, not only does competition increase among markets at the border but also potentially throughout the country via a ripple effect—since a price increase in one market influences the prices in adjacent markets and so on—and therefore affecting prices even in the interior markets.

Quantification of such ripple effects needs the aid of a spatial model of trade that uses the location of markets and the economic geography of the country and allows farmers to remove price arbitrages in space. In the next section, I set out such a model. In the baseline, I limit the choice set of farmers to be the markets only in their own state and expand this to be any market in the country in the counterfactual.

5 Model

In this section, I develop a simple model of trade in agricultural markets in India that will flexibly capture the forces of spatial competition between markets and aid us in quantifying the income and productivity effects of border restrictions on trade. The basic structure of the model is as follows. The economy consists of many geographic regions. Within each region there are two types of agents: farmers and intermediaries. Each intermediary is identified with a market. The location of markets is determined exogenously by the government. Each farmer first chooses input levels. After realization of output, he chooses the market where he wants to sell his output. Transporting goods is costly. The price that a farmer receives at a given market is determined by Nash bargaining between the farmer and the intermediary after the farmer has arrived at the market with the output. Intermediaries sell all of the purchased output in the retail market at an exogenously given price. Figure 1 shows the timing of events.

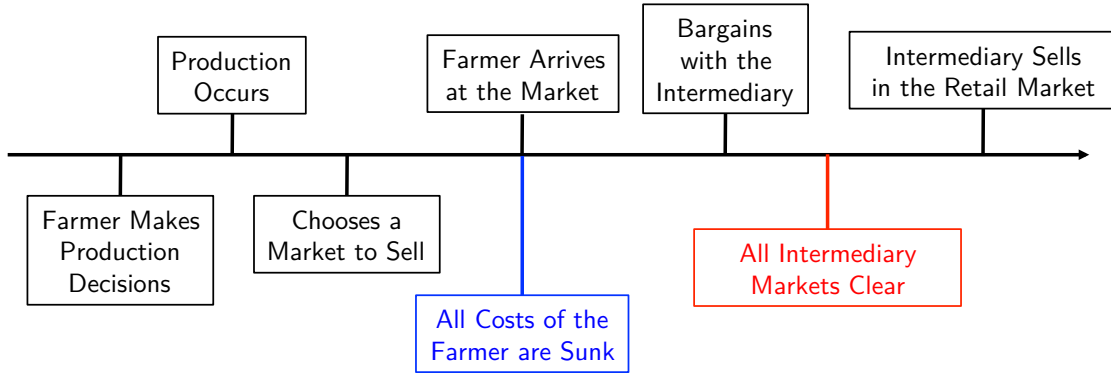


Figure 1: Timing of Events

I provide formal details of the model next.

5.1 Setup

5.1.1 Geography

There are S regions in the economy. Regions in the model are analogous to the states of India. For now, and motivated by the regulatory structure in India that prohibits farmers from selling their output across different states, there will be no interactions between regions. In what follows, I will focus on a generic region and abstract from indexing the region.

The region has two types of agents: F farmers and M intermediaries. \mathcal{F} and \mathcal{M} are the sets of all farms and intermediaries in the region, respectively. Each farmer f is identified by a given plot of land, or farm, and each intermediary m is identified by a specific market. Farms and markets are distributed in space, so I will also refer to f and m as locations. Geography matters because it is costly to move goods across locations. I model these costs as iceberg trade costs and denote by $\tau_{xy} (> 1)$ the amount of goods that must be moved from location x in order for 1 unit to arrive at location y .

5.1.2 Farmers

There is a single crop in the economy.²⁰ Land and labor available to a given farm are fixed and denoted by h_f and l_f respectively.²¹

A farmer purchases $\{x_f^k\}_k$ units of K intermediate inputs and his output is given by:

$$y_f = \tilde{A}_f \left(h_f^\gamma l_f^\nu \prod_{k=1}^K (x_f^k)^{\alpha_k} \right), \quad \sum_{k=1}^K \alpha_k + \nu + \gamma = 1; \alpha_k > 0 \forall k, \nu, \gamma > 0, \quad (5)$$

\tilde{A}_f is the total factor productivity for farm f . The market for inputs is perfectly competitive. w^k , the price of the k th input, is the same for all farmers in the region and is exogenously given. Farmers are small compared to the input, output, and the retail markets and therefore cannot influence prices directly.

5.1.3 Market Choice

Upon harvest, farmers optimally choose the market where they want to sell. Because farmers are subject to iceberg trade costs, the quantity of output actually reaching any market m from farm f is

$$q_{fm} = \frac{y_f}{\tau_{fm}}. \quad (6)$$

²⁰I will incorporate multiple crops later but will continue to treat crop choice as exogenous.

²¹This is the farm of the median Indian farmer who owns about 1.5 hectares of land (Economic Survey of India, Ministry of Finance, Government of India, 2016) and mostly uses family or village labor for cultivation. To the extent that large farmers in India may hire migrant labor, my estimates of gains in income and productivity should be thought of as lower bounds that ignore adjustment on the labor market.

The price that farmers would get in market m is denoted by $p^f(m)$. Thus, farmers choose market m to sell their output such that their income $\frac{p^f(m)y_f}{\tau_{fm}}$ is maximized.

5.1.4 Price Determination and Intermediaries

Once a farmer f reaches a market m with his goods, all his costs are sunk. He bargains with the intermediary associated with market m over the price for a unit of his output. The outcomes are determined via Nash bargaining. The outside option of the farmer, $\underline{p}(m)$, is the value that he would receive by going to the best alternative market. The best alternative is given by

$$\underline{p}(m) = \max_{k \in \mathcal{M} - \{m\}} \left\{ \frac{p^f(k)}{\tau_{mk}} \right\}. \quad (7)$$

The outside option for the intermediary is zero. If an intermediary m purchases q units of output from a farmer, he will sell this on the local retail market and receive a price p_m^r for each unit of output. A local retail market is the Indian administrative district in which the market is located. Each market is small compared to the local retail market and therefore the intermediary takes the retail price as given.²²

If δ denotes the bargaining weight for the farmer, the Nash bargaining outcome is the solution to:

$$\max_{\lambda} \left(\lambda - \underline{p}(m) q_m \right)^{\delta} (p_m^r q_m - \lambda)^{1-\delta} \quad (8)$$

Here, λ is the farmer's income. Since $p_m^r q_m$ is the total surplus, $p_m^r q_m - \lambda$ is the income of the intermediary trader from this transaction. We should note that all farmers trading at a particular market m receive the same price because what matters is the distance of the next best alternative from a particular market and not the distance to that farmer's farm. Therefore, I omitted the subscript f in the bargaining problem.

²²For the main analysis, I read local retail prices from the data. In Section 7.4, I will endogenize retail prices and let them adjust in response to changes in local supply.

5.1.5 Equilibrium

To solve for the spatial equilibrium, I employ the Nash-in-Nash (à la [Horn and Wolinsky, 1988](#)) solution concept—price negotiated at a market will be the Nash bargaining solution given that farmer-trader pairs in all other markets reach an agreement. The equilibrium for a region S is defined as follows:

Definition 1. Given parameters $\{\alpha_k\}_{k=1}^K, \nu, \gamma$ and δ , endowments $\{h_f, l_f, \tilde{A}_f\}_{f \in \mathcal{F}}$, transportation cost function τ , and retail and intermediate input prices $\{p_m^r\}_{m \in \mathcal{M}}$ and $\{w^k\}_{k=1}^K$ respectively, an equilibrium is a set of farmer prices $p^f(\cdot)$, intermediate input choices $\{x_f^k\}_{k \in \mathcal{K}, f \in \mathcal{F}}$ and the optimal market choice at each farm, denoted by $\{\mu(f)\}_{f \in \mathcal{F}}$, such that:

1. Farmers choose intermediate inputs to maximize their profits, to solve:

$$\max_{\{x_f^k\}_k} p^f(\mu(f)) \tilde{A}_f \left(h_f^\gamma l_f^\nu \prod_{k=1}^K (x_f^k)^{\alpha_k} \right) - \sum_{k \in \mathcal{K}} w^k x_f^k, \quad (9)$$

2. Farmers optimally choose a market to sell their output, according to:

$$\mu(f) = \arg \max_{m \in \mathcal{M}} p^f(m) q_{fm} = \arg \max_{m \in \mathcal{M}} \left\{ \frac{p^f(m) y_f}{\tau_{fm}} \right\}, \quad (10)$$

3. Farmers and intermediaries in any market Nash bargain over the total value of the farmer's output assuming that bargaining in all other markets have reached an agreement.

5.2 Solving the Model

We can now solve the model backwards from the timeline in figure 1. Conditional on the retail prices that intermediaries get, we first solve the bargaining problem between the farmers and the intermediaries.

5.2.1 Solution of the Bargaining Problem

The solution of the bargaining problem (8) gives us:

$$p^f(m) = (1 - \delta) \underline{p}(m) + \delta p_m^r \quad (11)$$

where, $p^f(m) = \frac{\lambda}{q_m}$.

We will have one such equation for every market and all of them are interrelated through the outside option, $p(m)$. Equilibrium farmer prices is then just the fixed point for the system of M equations in M endogenous variables defined in (11).²³

Theorem 1. *Given retail prices $\{p_m^r\}_{m \in \mathcal{M}}$, there exists a unique fixed point $p^{f*}(\cdot) : \mathcal{M} \rightarrow \mathbb{R}_+$, that solves the system of equations given in (11).*

The proof can be found in Appendix D. Intuitively, I can show that $p^f(\cdot)$ is a contraction as $\delta \in (0, 1)$ and $\tau \geq 1$, and therefore there is a unique solution.

Once we know the equilibrium farmer price in any market $m \in \mathcal{M}$, denoted by $p^{f*}(m)$, the optimal market choice is given by (10). Subsequently, the input choices would be the result of a standard profit maximization problem (9), given the production function and intermediate input prices.

Lemma 1. *Given retail prices, removing restrictions on farmers from trading across regions will improve the prices that they receive.*

The formal proof of Lemma 1 is relegated to Appendix E; the main intuition is as follows. In the system of equations (11) that determine equilibrium prices, the only variable that changes in the counterfactual is the set of outside available options to the farmer. In particular, \mathcal{M} will include all markets in India as opposed to markets within the region (or that state). As the local retail prices are held fixed and we are taking the maximum over all outside options, the prices they receive in equilibrium cannot be strictly lower than before.

Lemma 1 guarantees that as long as retail prices do not change, no farmer will be worse off as a result of removal of trade restrictions. When farmers are allowed to trade across states, the degree of spatial competition between intermediary markets on both sides of the border increases resulting in an increase in farmer prices in markets on both sides of the border. The only way farmers can lose in this setup is if due to an increase in supply, the retail prices fall enough to lower the total surplus and thus the prices farmers receive in markets.

²³In the quantitative section, I will assume the problem to be independent across crops and time periods. Therefore, the problem with all crops and multiple periods involves finding a fixed point of a system of $M \times C \times T$ equations in $M \times C \times T$ variables, where T is the number of time periods.

5.3 Demand in Retail Markets

To complete the model, I define the demand side here. Each region above is further divided into subregions, indexed by d . Each subregion corresponds to an Indian administrative district and as discussed in Section 5.1.4 is the local retail market. The aggregate demand function for a subregion is given by

$$Q^d = Q(p_d^r). \quad (12)$$

Let η be the price elasticity of demand. All intermediaries m located in subregion d sell in the same subregion and hence, $p_m^r = p_d^r \forall m \in d$.²⁴ Finally, for retail markets to clear, total quantity demanded in a subregion d must equal total quantity sold in that subregion

$$\sum_{f \in \{\mu^{-1}(m) | m \in d\}} y_f = Q^d. \quad (13)$$

6 Estimating the Model

In order to take the model to data, I need to expand it to include multiple crops and time periods. Crops and time periods are assumed to be independent and therefore, in the quantitative section we will have $M \times C \times T$ equations of the form (11), where C is the number of crops and T is the number of time periods. To quantify the gains from increased spatial competition due to removal of restrictions on farmers' trading across state borders, I structurally estimate the model presented in the previous section. First, I parametrize τ and then estimate δ and the parameters in the trade cost function jointly to match the spatial distribution of farmer prices. I conduct a sensitivity analysis to understand what features of the data inform the parameters. I set the production function parameters to match to land and labor shares in my data and I take demand elasticities from the literature.

6.1 Estimating τ and δ

I parametrize trade costs as a function of distance, and I estimate the parameters using data on the location of markets. I define the trade cost between two markets m and k , for crop c , and at time t

²⁴This is an assumption. Equivalently, retail markets can be assumed to be perfectly integrated and then retail prices would differ across regions only due to transport costs. In the latter case, the intermediary will be indifferent between selling in different retail markets. Table 1 provides support for this by showing low variance of retail prices across different regions that produce a certain crop.

as:

$$\tau_{mkct} = \begin{cases} 1 & m = k \\ 1 + A \cdot d_{mk} + \epsilon_{mct} & m \neq k \end{cases}, \quad (14)$$

$$\log \epsilon_{mct} \sim N(0, \sigma),$$

where, d_{mk} is the geodesic distance²⁵ between locations m and k and A is a scale parameter. The shock term, ϵ_{mct} , represents origin market-crop specific costs like broken roads, availability of a truck, or a strike among intermediaries, which are not observable to the econometrician but are known to the farmers. τ_{mkct} is set to infinity if markets m and k lie in different states to incorporate trade restrictions.

If I denoted crops by c and time period by t , then the key equilibrium equation of the model (11) becomes

$$p^f(m, c, t) = (1 - \delta) \max_{k \in \mathcal{M} - \{m\}} \left\{ \frac{p^f(k, c, t)}{1 + A \cdot d_{mk} + \epsilon_{mct}} \right\} + \delta p_{mct}^r, \quad (15)$$

From equation (15), we can see that prices in the different markets are interrelated. Therefore, the marginal likelihood function of price at one market is not independent of prices in the other markets. Hence, estimation cannot be carried out via a standard maximum likelihood or a method of moments procedure. To address this, I use a method of simulated moments procedure to estimate (δ, σ, A) . In particular, I choose (δ, σ, A) to minimize the distance between moments of the data and their simulated counterparts.

Formally, the method of simulated moments chooses parameters $\Theta = (\delta, \sigma, A)$ to minimize the distance between simulated and data moments, solving:

$$\hat{\Theta} = \arg \min \left(\Psi^d - \Psi^s(\Theta) \right)' W \left(\Psi^d - \Psi^s(\Theta) \right),$$

where W is the optimal weighting matrix—the inverse of the variance-covariance matrix of the moments; Ψ^d are the data moments; and $\Psi^s(\Theta)$ are the simulated moments. I compute robust standard errors clustered at the state-season level and use numerical methods to find the required gradients.

²⁵The shortest distance along the earth's surface.

6.1.1 Choice of Moments and Estimation

To identify moments that will inform the estimation, I choose an auxiliary regression model that mimics the equilibrium equation (15) as closely as possible.²⁶

$$\ln p_{mct}^f = \tilde{\beta}_0 + \tilde{\beta}_1 \left(\max_{k \in \mathcal{M} - \{m\}} \ln p_{kt}^f \right) + \tilde{\beta}_2 \log d_{mk^*} + \tilde{\beta}_3 \ln p_{mct}^r + \gamma_c + \xi_{mct}, \quad (16)$$

Here, p_{mct}^f and p_{mct}^r are the prices farmers get and the retail prices in market m for crop c at time t respectively. d_{mk^*} is the distance from market m to market k^* , which has the highest price in the neighborhood, and γ_c is a crop fixed effect. I match five moments from the auxiliary regression: $\tilde{\beta}_0, \tilde{\beta}_1, \tilde{\beta}_2, \tilde{\beta}_3$, and the Mean Squared Error of the auxiliary regression. In addition, I also match the distribution of farmer prices using the mean and the variance of farmer prices, and the fraction of data in different parts of the distribution. Let us denote the k -th percentile of the distribution of farmer prices in the actual data by pc_k , and simulated farmer prices by \tilde{p}_{mct}^f . Then the moment condition is defined as:

$$\psi(\Theta) = \mathbb{1} \left[\tilde{p}_{mct}^f \in [pc_k, pc_{k+\zeta}] \right] - \zeta,$$

where ζ is the chosen bin size. I use four bins: 5th–25th, 25th–50th, 50th–75th and the 75th–95th percentiles.

While in principle all parameters are jointly estimated to match the distribution of prices, some moments are more important for identifying some parameters. Intuitively, the Nash-bargaining parameter, δ , is identified from the correlation between retail and farmer prices, $\tilde{\beta}_3$, and that between farmer price in a market and the maximum price in the neighborhood, $\tilde{\beta}_1$. Given δ , the scale parameter on distance, A , adjusts to match the level of farmer prices and is therefore identified of, $\tilde{\beta}_3 - \tilde{\beta}_1$, and the mean of farmer prices. Finally, the variance of the shock term, σ , is identified from the residual variance in prices captured by the MSE of the auxiliary regression and the distribution of prices. However, because the shock term enters inside the max function in the denominator, it is the smallest shocks that matter for identification. As the log normal distribution is bounded below

²⁶Following [Gourieroux et al. \(1993\)](#), identification of the auxiliary regression is not necessary for consistent estimation of the structural parameters. It just needs to capture key moments of the data that can inform identification of the structural parameters. One reason not to use only the border regression (4) as an auxiliary model is that it does not capture the relationship between farmer prices, distances to other markets, and retail prices. These moments are important for identification of the model parameters.

by zero, this parameter is not estimated very precisely and this is reflected by the standard errors. Finally, $\tilde{\beta}_0$ reflects the scale in the auxiliary regression and therefore informs all parameters in the model.

To carry out the estimation, I set up the data in the following manner. First, because retail prices are only available at the level of an Indian administrative district, I assign the same retail price to all markets in a given district-time period. Second, data on retail prices are available only between 2005 and 2011, barring 2008²⁷; therefore I focus on the data in these years for estimation. Third, to reduce the computational burden, I use the median prices at the crop-market-season level for estimation.²⁸ There is little variation in retail prices of a district or farmer prices at a market within a season (see Table 1) and thus there is no consequential loss of information. However, it helps me to create a balanced panel as retail prices in certain months are erroneously reported or are missing. Fourth, I exclude the states for a crop where it has never been produced. To infer this, I use the information in the crop calendar and production statistics reported by the Ministry of Agriculture of India in its annual publication, Agriculture Statistics at a Glance, for the years in my sample. I provide the details in Appendix I.3. This leaves me with data on 8 crops in 2978 markets for 15 time periods.

Table 4 reports the estimated parameters with robust standard errors clustered at the state-time level. The bargaining power of the farmer, δ , is estimated to be 0.69. The estimated parameters of the trade costs correspond to a deterministic 16% transport cost over a distance of 100 km. In Section 8, I also consider other functional forms for trade cost such as a power function, a quadratic, and the exponential function. In general, the data admit any functional form for trade costs that is concave in distance.

6.1.2 Sensitivity of Parameters

In this subsection, I compute four different statistics to understand the sensitivity of the parameters to different aspects of the data. First, following Andrews et al. (2017), I compute the semi-elasticity of the probability limit of each estimator with respect to the moments scaled by their standard deviations. Second and third, following Honore et al. (2017), I compute the elasticity of each estimator with respect to the diagonal elements of the weighting matrix and the elasticity of the variance

²⁷2008 was a drought year. Therefore, I am not worried about sample selection. If anything, the data would have been outliers.

²⁸I use three crop seasons—fall, spring, and summer.

co-variance matrix of the estimator with respect to the variance co-variance matrix of the moments. Fourth, following [Cooper and Liu \(2016\)](#), I compute the elasticity of the moments with respect to the parameters. All the estimates are reported in [Tables 5a & 5b](#) and the results are also shown in [Figures 2–5](#).

The first statistic helps us understand to which moments the estimators are most sensitive. In particular, if a moment is not exactly equal to 0 but equal to some positive number ρ , i.e. $E(\psi(x_i, \theta_0)) = \rho$, then I compute the derivative of the probability limit of $\hat{\theta}$ with respect to ρ . I scale the moments by their respective standard deviations. The results are presented in [Figure 2](#). A general takeaway is that the coefficients $\tilde{\beta}_1, \tilde{\beta}_2$ and $\tilde{\beta}_3$ aid in identification of all the parameters.

The second statistic helps us ask what would happen to the estimate of a parameter if I gave a little more weight to a moment, holding all other weights in the weighting matrix fixed. This is calculated as the derivative of the probability limit of the estimator with respect to the kk th element of the weighting matrix. The results in [Table 5b](#) (and [Figure 3](#)) are qualitatively very similar to the previous statistic. An additional insight here is that σ is sensitive to the weight on almost all moments and thus the estimate of σ is very sensitive to the choice of the weighting matrix.

The third statistic corresponds to the thought experiment: How much precision would we lose if we had a different data set in which a certain moment is subject to a little additional noise? This differs critically from the previous two statistics in that we allow the weighting matrix to change optimally as more noise is added. In my case, the results presented in [Figure 4](#) provide useful insurance against noise in the retail prices. Because data on retail prices is aggregated at the level of the district, it is measured with noise. Panel 1 of [Figure 4](#) tells us that δ is not much affected by such noise in the retail prices. In fact, none of the parameters are.

The fourth statistic is a direct measure of sensitivity of the moments with respect to model parameters. In contrast to the other statistics, this does not take into account optimal weighting at all and asks the opposite question: If data were generated from the same model with different parameter values, which moments are likely to change? It is again comforting to note that changing δ does not change $\tilde{\beta}_3$, the coefficient on log retail prices in the auxiliary regression, as much. What stands out in the results presented in [Figure 5](#) is that the bottom of the price distribution is very sensitive to δ . This makes sense because the lowest farmer prices are likely to occur in very isolated markets such that distances to other markets are very large. This would make the simulated prices

in the bottom tail very sensitive to the scale parameter on distance, shocks to transport costs, and retail prices. This has little implication for estimation per se, given that the optimal weight matrix will ensure that only relevant information is extracted from this moment.

6.1.3 Goodness of Fit

To assess how well the model fits the data within the sample, I regress log of actual prices to log of prices generated from the model. A linear regression without an intercept term has a coefficient of 1.065 with an R-square of 0.59. The fit is shown in Figure 6. Figure 7 shows a density plot of model-generated prices and actual prices and Table 6 presents different moments of the price distribution. We can notice that the model does a good job of capturing the distribution of prices. In Figure 7, the humps in the density of actual data are because of different crop specific means. This is not exactly replicated in the model generated data as I do not have any crop-specific parameters in the model. However, I take it to be a success of the model that I match the distribution of prices well even without incorporating crop-specific parameters.

6.2 Calibrating the production function and demand elasticities

To calibrate $\gamma + \nu$, the share of land and labor, I use the first order conditions from the farmers' profit maximization problem. Because of fixed factors, I need data on input costs and farm profits, which I obtain from the Situation Assessment Survey of Agricultural Households in India 2012–13.²⁹ I find that $\gamma + \nu = 0.57$. In section (8), I present results for other values of $\gamma + \nu$.

I take estimates of the price elasticity of demand, η , in India from Deaton (1997). For rice $\eta = -1.19$, wheat $\eta = -1.29$, and other crops $\eta = -3.31$.

7 Counterfactual Analysis

The model I have laid out allows me to compute changes in farmer prices, production, and revenues as a result of more competition between intermediary markets. In this section, I use the removal of the restriction on farmers' selling in markets of other states as a convenient experiment to increase spatial competition between intermediary markets. The key mechanism is the following. Consider

²⁹For details, see Appendix F.1.

two markets on the Madhya Pradesh-Maharashtra border marked in Figure 8. The farmer living close to the market in Madhya Pradesh gets low prices because competition is low in his local region as the next market is quite far away. Further, this farmer cannot cross the border and sell in Maharashtra. When we remove restrictions on farmers’ selling across state borders, these two markets start competing for the farmer’s output. This increases prices in both these markets. The farmer, wherever he chooses to sell, is better off. Moreover, these two markets are the “threat points” to other markets nearby. As the prices increase in the border markets, they also increase in other markets and this ripple effect increases prices even in the interior of the states. This is the direct benefit to farmers via an increase in prices. In response to an increase in output prices, farmers also adjust their use of intermediate inputs, and their incomes improve further as a result of increased production. Finally, as a result of changes in production and changes in the market-site where farmers choose to sell, retail prices may adjust and feed back into the prices farmers get.

In the model, this removal of trade restrictions on farmers is equivalent to changing the trade cost between two locations in different states from infinity to what it would be under equation (14). To study the role of the different channels mentioned in the previous paragraph, I conduct four exercises relaxing one constraint at a time. In the first exercise, I do not allow retail prices to adjust. I keep all decisions of the farmer fixed except their threat to sell at a market outside the state (Section 7.1). In addition to allowing farmers to offer a different threat, in the second exercise, I also allow them to choose a different market to trade in (Section 7.2). In a third exercise, I let farmers also adjust their intermediate input choices to study the change in their crop output (Section 7.3). In the final exercise, I let retail prices at the district level adjust in response to changes in supply to understand medium-run implications for farmer prices and output (Section 7.4).

I also use the model to understand the implications of reductions in transport costs and examine the quantitative magnitudes of the interaction of such public investment with trade liberalization for farmers (Section 7.5).

Note that all these exercises require data on crop choices of farmers. For this, I use data from the Situation Assessment Survey of Agricultural Households 2012–13, and I match them with data on retail prices from 2011–12. These are the two closest years for which I have detailed data on crop

choice and retail prices.³⁰ Since regional cropping patterns in India have been stable over the last decade, this is not a major concern for interpreting my results.

7.1 Changing Farmers' Threat Point

In this exercise, I use my model and estimated parameters to study what happens to farmer prices when we remove the restriction on farmers' selling in markets of other states but keep all other decisions fixed. We can think of this as a hypothetical exercise where farmers go to the same market where they went when restrictions were in place. However, they can threaten intermediaries with going to a market in another state when bargaining with them.

The changes in farmer prices at the markets are presented in Figures 9 and 10. Figure 9 plots the relationship between the change in prices in each market in the counterfactual scenario and the change in competition for each market, where competition is defined as in the reduced-form regressions using (1). The bottom of Figure 9 plots the empirical distribution of the change in competition as a result of removal of the interstate trade restriction. There is a monotonic relationship between price increase and increase in competition. The median increase in competition is 2.9, which means that half of all markets experience at least a 2.8% increase in price.

Figure 10 plots the estimated relationship between gain in price at a market and the distance to the nearest border. The bottom of the figure plots the empirical distribution of the location of markets with respect to state borders. We can see that the biggest gains are in markets closest to state borders and the gains decline the farther the markets are. However, due to the ripple effect, even markets in the interior of some states realize an increase in competition and therefore an increase in price.

To understand what this means for farmers living in different locations, I divide India into grid cells of 5 arc minute by 5 arc minute.³¹ I use data on land use to exclude cells that are uncultivated like deserts, large water bodies, mines, marshes etc. The remaining cells are hypothetical farms in India for which I will compute changes in prices and production as a result of increased spatial

³⁰The Central Statistical Organization of India should soon release data on retail prices from 2012–13, at which point I will update this analysis.

³¹This corresponds to a cell having land area of approximately of 10 km by 10 km. Although, this is much larger than a typical farm, it is not a major concern because I consider a production function which is homogeneous of degree 1, and cropping patterns in India are similar within 10 km distances. Limitations to computing power does not allow for consideration of finer grid cells.

competition. In order to determine the price the farm gets, I also need to know what crops this farm grows. I use data from the Situation Assessment Survey of Agricultural Households 2012–13 to estimate the district level land shares for each crop in each crop season, i.e. the crop choice of the district. I then assign the crop choice of the district to each cell that lies in a majority in that district. I compute the price received by a farm as the Laspeyres price index. I use yield estimates from the Situation Assessment Survey of Agricultural Households to estimate the quantities produced on each farm.

The results for the fall season are presented in Figure 11; Figure 12 presents the results for the spring season. The average increase in farmer prices is 2.74% and 2.65% in the fall and spring seasons respectively. The average increase in prices for the top 25% of farmers with the largest gains is 7.36% in the fall season and 5.84% in the spring season. The maximum increase in prices is 9% and 13% in the fall and the spring season respectively. These increases in farmer prices are purely due to an increase in spatial competition because we do not allow farmers to change any other decision.

7.2 Allow Farmers To Sell At A Different Market

In the second exercise, in addition to providing a greater threat to intermediaries while bargaining, I also allow farmers to choose a different market to trade in. Their new market choice could be another market in the same state or a neighboring state. As we will see, the largest increases in farmer prices will come from changing market choice.

Figures 13 and 14 present the changes in prices that farmers get as a results of increased competition because farmers can sell in markets of other states. We can see that farmers are significantly better off in many regions of India. The price increases are generally larger near state borders and lower in the interior of the states. The average increase in prices is 12.6% in the fall and 10.7% in the spring. The distribution of price increases is, however, skewed. The top 50% of the farmers who experience an increase in prices get 23.37% and 20.04% higher prices on an average in the fall and the spring. Consistent with Lemma 1, no farmer loses.

In the fall season, we notice bigger gains in central India (in particular, in the state of Madhya Pradesh). To explore the mechanism further, I will zoom into the Madhya Pradesh and Maharashtra border.³² This region mainly grows soybean in the fall. Figures 15 and 16 plot the log of soybean

³²Refer to Figure A.1 for state names.

prices that farmers get in this region before and after the policy change respectively. We can see that price levels before the policy change are higher in Maharashtra than in Madhya Pradesh. This is so because near the state border, the density of markets is higher in Maharashtra.

After the policy change, markets on the border compete and the price in them increases. Through a ripple effect of one market affecting the other, prices increase even in the interior of the states. As can be seen in Figure 16, prices equalize on both sides of the state border. Of course, because prices were lower in Madhya Pradesh to begin with, the percentage increase there is higher (see Figure 17).

Prices increase by nearly 60% in Madhya Pradesh near its border with Maharashtra for the following reason. Farmers in Madhya Pradesh close to the border are generally far away from markets in their own state. Therefore, with border restrictions, they pay a substantial transport costs to trade in these markets. When states open up to trade, these farmers get access to markets in Maharashtra, which are much closer. This saves farmers' transport costs, leading to even larger gains.

7.3 Letting Farmers Reoptimize the Use of Intermediate Inputs

The Economic Survey of India 2015–16 documents a huge variation in the use of intermediate inputs. Farmers in more productive regions are usually more input-intensive. I highlight a possible reason behind this pattern in this exercise. I allow farmers to change their use of intermediate inputs in response to the change in prices we observe in the counterfactual exercises of the previous section. From the model, we can show that for farm f , the new output (y'_f) to the old output (y_f) ratio is a function of the ratio of the new to old prices. In particular³³

$$\frac{y'_f}{y_f} = \left(\frac{p^{f'}}{p^f} \right)^{\frac{1-(\gamma+\nu)}{\gamma+\nu}},$$

here γ and ν are the share of land and labor in production, and primes represent the new output and prices. Since $\gamma + \nu < 1$, the percentage increase in production is smaller than percentage increases in prices. The results presented in Figures 18 and 19 use the same crop choices as before. To measure quantity produced at each farm, I compute the Laspeyres quantity index. The average increase

³³For a formal proof see Appendix F.

in production is 9.1% and 7.75% in the fall and the spring respectively. The average increase in production for the top 50% of the farmers is 16.9% and 14.46% in the fall and the spring respectively.

Farmers in less competitive regions use fewer intermediate inputs as a result of low output prices. Therefore, the benefit to farmers from the increase in competition comes from two sources—an increase in prices and an increase in quantities. This leads to an increase in the total revenue of farmers. The aggregate increase in the value of production is 17.3% in the fall season and 20.9% in the spring season.

7.4 Incorporating a Demand Side Response

The gains described above could be overestimates because an increase in production could depress retail prices, which would feed back into farmer prices. To understand this channel better, I use estimates of demand elasticities from [Deaton \(1997\)](#) and an iterative algorithm³⁴ to find the new equilibrium after border restrictions to trade are removed.

Figures 20 and 21 present the changes in prices incorporating adjustment of retail prices for the fall and the spring respectively.³⁵ Although there are regions in the interior of some states where farmers lose, most farmers gain. The average increase in farmer prices is 9.62% in the fall season and 9.56% in the spring season. The average increase in prices for the farmers with above-median changes is about 20%. For the farmers who lose, the average decline in prices is about 10%.

The average price increases here are only marginally lower than in the exercise when we kept retail prices fixed despite a fall in prices of some farmers. The main reason is the following. When I remove the restriction on farmers to trade across state borders, supply, on average, increases in regions with high initial retail prices. The supply increases because these regions attract more farmers, who in turn, produce even more as they get paid more. This adjusts the retail prices downwards, and farmers from such regions lose on average. Since supply reduces in regions with low initial retail price, as farmers sell in other regions, there is an upward adjustment of retail prices on average. In equilibrium, the net effect on farmers from these regions is ambiguous. Earlier, when retail prices were held fixed, in response to removal of border restrictions these farmers incurred transport costs to get access to higher prices (Section 7.2). Now, these farmers can get higher prices in their local

³⁴See Appendix G for details.

³⁵Compare to Figures 13 and 14, that present changes in farmer prices when we do not allow retail prices to adjust.

neighborhood. These prices are not as high as in the counterfactual before, but the farmers save their transport costs. Hence, in principle, some farmers can gain more as a result of the adjustment of retail prices.

The results from the exercises above are summarized in Table 7. I have shown that increased competition due to the removal of trade restrictions substantially increases the prices that farmers receive. The average farmer gets about a 10% higher price for his output. The gains are skewed as the average increase in price for the top 50% of farmers is about 19%. Approximately 6% of the increase is purely due to competition, while a majority of it comes because farmers also reoptimize their market choice and save on transport costs. Increase in price is only part of the gain that farmers get as this is reinforced by an increase in production as farmers adjust their use of intermediate inputs. As a result, the total value of production increases by about 19.90%. Finally, I also allay concerns that an increase in output might reduce retail prices so much that farmers in most regions end up losing.

7.5 Effect of a Reduction in Transport Costs

The exercises that I have presented heretofore reflect the direct gains from competition via increased prices for farmers. Spatial competition can also interact with other development policies in interesting ways, and I can use my model to shed some light on this aspect. Poor roads and lack of access to good vehicles increase transportation costs in India, and they are especially high for farmers (Kapur and Krishnamurthy, 2014). However, over the last 15 years, the central government has invested around \$40 billion in the construction of rural roads (Chatterjee and Kapur, 2017). Between 2001 and 2015, approximately 185,000 habitations have been connected by all-weather roads. The aim in this section is to show that farmers can realize substantially greater benefits of such public investments in the presence of greater competition among intermediaries.

I consider reducing A by 40%, and for illustrative purposes, I present results only for the fall season. Figure 22 presents the increase in farmer prices due to reduction in transport costs in the fall season in the presence of border restrictions. Figure 23 presents the additional increase in farmer prices due to more spatial competition in the absence of border restrictions to trade. I computed this by calculating the increase in farmer prices due to simultaneously reducing the transport costs and removing the border restrictions, and then subtracting the increase in prices just due to the

reduction in transport costs (computed in Figure 22) and the increase just due to the elimination of border restrictions to trade (computed in Figure 13). Farmer prices increase by an additional 8% on average, because of increased spatial competition.

We can draw two lessons. First, a reduction in transport costs, even in the presence of border restrictions, leads to increases in farmer prices in the interior of states as farmers can travel farther, and this changes their threat point in the bargaining problem. Second, the additional gains from the increase in spatial competition occur in those regions where transport cost reduction was not effective in improving farmer prices, e.g., in the states of Odisha, Madhya Pradesh, and Rajasthan.³⁶ Therefore, transport infrastructure policies can be thought of as complements to policies that increase competition.

8 Robustness and Sensitivity of Quantitative Results

In this section, I argue that my results are robust to two potential misspecifications in the setup above. First, it could be possible that farmers in villages do not sell in markets but to village intermediaries who aggregate from all farmers in a village and in turn sell to market intermediaries. Second, there are costs associated with activities of the intermediaries that are missing in the model. Additionally, I show that my quantitative estimates of increases in farmer prices and output are not very sensitive to the way I assign crop choices to farm-cells, the choice of $\gamma + \nu$, and the functional form of trade cost.

It could be possible that farmers in some regions do not travel to the markets themselves but sell in villages to village intermediaries who in turn sell in the markets to market intermediaries. In this case, all the quantitative results remain unchanged if we allow for a constant pass-through from the village intermediary to the actual farmer. In particular, the village intermediaries perform the same role that farmers play in the current model and pass a share of their earnings to the farmers. Therefore, p^f is now the price that the village intermediaries receive as they engage in Nash-bargaining (defined in 8) with the market intermediaries. Let us further suppose that these village intermediaries pass a fixed share to the actual farmers, κ^f . I can allow this pass-through rate, κ^f , from the village intermediaries to the farmers to vary by each farm. However, we have to assume

³⁶For state names, see Figure A.1

that it does not change in the counterfactual scenarios. Since all my results are computed in changes, they remain the same.

In my model, the intermediaries buy from farmers and sell in retail markets and have no associated costs. In the presence of costs to intermediaries, the actual prices that intermediaries get must be lower than the retail prices measured in my data. δ is insensitive to noise in retail prices following the arguments I presented in Section 6.1.2. Therefore, to match the level of farmer prices, A would be lower than is estimated. This would make trade costs lower and the competition channel more important. Thus my results can be viewed as a lower bound on the actual gains.

To show that my results are not sensitive to the allocation of crop choice to farm-cells, in Figures H.1 and H.2 I present the counterfactual increases in prices that farmers receive assuming that farms grow the main crop (by crop sown area) of the district. The results look very similar to Figures 13 and 14, because in the data, the main crop of the district occupies on average 80% of the cropped area of the district. The mean increase in farmer prices is about 11%, and the average increase in prices for the 50% of farmers with the largest gains is about 20%.

In Table 8, I present estimates of increases in production and revenues for higher values of $\gamma + \nu$. The increase in output is decreasing in $\gamma + \nu$. For $\gamma + \nu$ between 0.57 and 0.63, the average increase in output of farmers is between 6% and 9%. The average increase in output for 50% of the farmers with the largest gains is between 12% and 19%.

As far as different functional forms for trade cost is concerned, in Table 9, I present results if trade cost is assumed, following Redding and Sturm (2008), to be $\tau_{mk} = (1 + A \cdot d_{mk})^{\frac{1}{3}} + \epsilon_{mct}$. Since the trade costs in this case increases at a slower rate with distance than with the linear trade costs, the associated price increases are larger. The average increase in farmer prices is about 20%, and the average increase for the 50% of farmers with the largest gains is about 35%.

The data seem to be rejecting functional forms that are convex in distance. I estimated a quadratic function for trade costs: $\tau_{mk} = 1 + A \cdot d_{mk} + B \cdot d_{mk}^2 + \epsilon_{mct}$. The estimation procedure resulted in $B = 0$, rejecting this functional form in favor of the linear functional form. Similarly, the estimation procedure results in $A = 0$ for an exponential functional form $\tau_{mk} = \exp(A \cdot d_{mk}) + \epsilon_{mct}$.

9 Concluding Remarks

The median annual farmer income in India is USD 365 ([Ministry of Finance, Government of India 2016](#)). In this paper, using unique data on the location of intermediary markets and farmer prices, I have shown that spatial competition between intermediaries is an important determinant of the prices that farmers in India get for what they produce. High transport costs and policies that limit the ability of farmers to arbitrage between different intermediaries cause the market power of intermediaries to vary in space. Farmers who live in regions where there is more competition between intermediaries receive higher prices for their output. A one standard deviation increase in competition increases farmer prices by 6.4%.

I also show that increasing competition in one region spreads through the rural economy via a ripple effect. In particular, I consider the effect of reforming the laws of Indian states that prohibit farmers from selling in markets of other states. I show that such a policy not only increases incomes of farmers who live close to state borders but also of those living in the interior. I also show that increased competition further increases farmer incomes because of the increased output of farmers as they optimize the use of intermediate inputs. I find that the average increase in prices and output for the farmers with above-median gains is about 21% and 15% respectively. Further, the value of national crop output increases by about 18%. Moreover, the results indicate that isolated studies in agricultural markets can indeed find varying estimates of market power of intermediaries because they are partly driven by spatial competition.

Finally, I document that there are large complementarities between market integration and spatial competition. Adding to the large literature on how reduction in trade costs affects market integration and gains from trade (summarized in [Donaldson, 2015a](#)), I show that welfare benefits of market integration and trade cost reduction are substantially more in presence of greater competition.

Although I have focused on one institutional feature in one particular developing country, I believe that these findings have broader importance. Similar institutions and policies exist elsewhere. Monopsony powers of agricultural marketing boards or intermediaries is a common feature across developing countries (e.g. see [Bates, 1981](#)). Policies that restrict trade through intermediaries exist even outside agriculture. For example, until 2004 medium-sized firms in China were required by law to export via government approved intermediaries ([Bai et al., 2017](#)). The underlying mechanism of

spatial competition will apply to all such environments and future work could study its importance in other settings.

References

- Allen, Treb and David Atkin**, “Volatility and the Gains from Trade,” *NBER Working Paper*, 2016, (22276).
- Anderson, James E. and Eric van Wincoop**, “Gravity with Gravitas: A Solution to the Border Puzzle,” *The American Economic Review*, 2003, *93* (1), 170–192.
- Andrews, Isaiah, Matthew Gentzkow, and Jesse M. Shapiro**, “Measuring the Sensitivity of Parameter Estimates to Estimation Moments,” *The Quarterly Journal of Economics*, 2017, *132* (4), 1553–1592.
- Antras, Pol and Arnaud Costinot**, “Intermediated Trade,” *The Quarterly Journal of Economics*, August 2011, *126* (3), 1319–1374.
- Atkin, David and Dave Donaldson**, “Who’s Getting Globalized? The Size and Implications of Intra-national Trade Costs,” *NBER Working Paper*, 2015, (21439).
- Bai, Xue, Kala Krishna, and Hong Ma**, “How You Export Matters: Export Mode, Learning and Productivity in China,” *Journal of International Economics*, 2017, *104*, 122–137.
- Banerji, A. and J. V. Meenakshi**, “Buyer Collusion and Efficiency of Government Intervention in Wheat Markets in Northern India: An Asymmetric Structural Auctions Analysis,” *American Journal of Agricultural Economics*, 2004, *86* (1), 236–253.
- Bardhan, Pranab, Dilip Mookherjee, and Masatoshi Tsumagari**, “Middlemen Margins and Globalization,” *American Economic Journal: Microeconomics*, 2013, *5* (4), 81–119.
- Bates, Robert H**, *Markets and States in Tropical Africa: The Political Basis of Agricultural Policies*, Berkeley: University of California Press, 1981.
- Bergquist, Lauren Falcao**, “Pass-through, Competition, and Entry in Agricultural Markets: Experimental Evidence from Kenya,” *UC Berkeley Mimeograph*, 2017.
- Casaburi, Lorenzo and Tristan Reed**, “Competition and Interlinkages in Agricultural Markets: An Experimental Approach,” *Working Paper*, 2016.

- Chand, Ramesh**, “Development Policies and Agricultural Markets,” *Economic and Political Weekly*, 2012, 47 (52).
- Chatterjee, Shoumitro and Devesh Kapur**, “Six Puzzles in Indian Agriculture,” *India Policy Forum* 2016–17, 2017, 13.
- Conley, Timothy G**, “GMM Estimation with Cross Sectional Dependence,” *Journal of Econometrics*, 1999, 92 (1), 1–45.
- Cooper, Russell and Huacong Liu**, “Money or Grit? Determinants of Mismatch by Race and Gender,” *NBER Working Paper*, 2016, (22734).
- Cosar, A Kerem, Paul Grieco, and Felix Tintelnot**, “Borders, Geography, and Oligopoly: Evidence from the Wind Turbine Industry,” *Review of Economics and Statistics*, 2015, 97 (3), 623–637.
- Deaton, Angus**, *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*, World Bank Publications, 1997.
- Dell, Melissa**, “The Persistent Effects of Peru’s Mining Mita,” *Econometrica*, 2010, 78 (6), 1863–1903.
- Dillon, B. and C. Dambro**, “How Competitive are Food Crop Markets in Sub-Saharan Africa? A Review of the Evidence,” *Working Paper, Evans School of Public Policy and Governance, University of Washington*, 2016.
- Donaldson, Dave**, “The Gains from Market Integration,” *The Annual Review of Economics*, 2015, 7 (1), 619–647.
- , “Railroads of the Raj: Estimating the Impact of the Transportation Infrastructure,” *American Economic Review*, *forthcoming*, 2015.
- **and Richard Hornbeck**, “Railroads and American Economic Growth: A Market Access Approach,” *The Quarterly Journal of Economics*, 2016, 131 (2), 799–858.
- Eaton, Jonathan and Samuel Kortum**, “Technology, Geography, and Trade,” *Econometrica*, 2002, 70 (5), 1741–1779.

- Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu**, “Competition, Markups, and The Gains from International Trade,” *The American Economic Review*, 2015, 105 (10), 3183–3221.
- Engel, Charles and John H. Rogers**, “How Wide Is the Border?,” *The American Economic Review*, 1996, 86 (5), 1112–1125.
- Fafchamps, M, E Gabre-Madhin et al.**, “Agricultural markets in Benin and Malawi,” *African Journal of Agricultural and Resource Economics*, 2006, 1 (1), 67–94.
- Fafchamps, Marcel and Bart Minten**, “Returns to social network capital among traders,” *Oxford Economic Papers*, April 2002, 54 (2), 173–206.
- , **Eleni Gabre-Madhin, and Bart Minten**, “Increasing returns and market efficiency in agricultural trade,” *Journal of Development Economics*, 2005, 78 (2), 406–442.
- Gourieroux, Christian, Alain Monfort, and Eric Renault**, “Indirect Inference,” *Journal of Applied Econometrics*, 1993, 8 (S1).
- Great Britain and Royal Commission on Agriculture in India**, *Report of the Royal Commission on Agriculture in India.*, London: H.M. Stationery Office, 1928.
- Guha, Ramachandra**, *India after Gandhi: The History of The World’s Largest Democracy*, New Delhi: Harper Collins, 2007.
- Harris, Barbara**, *State and market: State Intervention in Agricultural Exchange in a Dry Region of Tamil Nadu, South India*, New Delhi: Concept Pub. Co., 1984.
- Harris, Chauncy D.**, “The Market as a Factor in the Localization of Industry in the United States,” *Annals of the Association of American Geographers*, 1954, 44 (4), 315–348.
- Honore, Bo E., Thomas Jorgensen, and Aureo de Paula**, “Sense of Sensitivity,” *Working Paper*, 2017.
- Horn, Henrick and Asher Wolinsky**, “Bilateral Monopolies and Incentives for Merger,” *The RAND Journal of Economics*, 1988, pp. 408–419.

- Kapur, Devesh and Mekhala Krishnamurthy**, “Market Towns and the Dynamics of India’s Rural and Urban Transformations,” *Working Paper, International Growth Center*, 2014, (F-35007-INC-1).
- Krishna, Kala and Yelena Sheveleva**, “Wheat or Strawberries? Intermediated Trade with Limited Contracting,” *American Economic Journal: Microeconomics*, August 2017, 9 (3), 28–62.
- Krishnamurthy, M.**, “Harda Mandi: Experiencing Change in an Agricultural Market in Central India (1980-2010).” Ph.D. thesis, UCL (University College London) October 2011.
- Loecker, Jan De, Pinelopi K Goldberg, Amit K Khandelwal, and Nina Pavcnik**, “Prices, Markups, and Trade Reform,” *Econometrica*, 2016, 84 (2), 445–510.
- McCallum, John**, “National Borders Matter: Canada-US Regional Trade Patterns,” *The American Economic Review*, 1995, 85 (3), 615–623.
- Meenakshi, J.V. and A. Banerji**, “The Unsupportable Support Price: An Analysis of Collusion and Government Intervention In paddy Auction Markets in North India,” *Journal of Development Economics*, 2005, 76 (2), 377 – 403.
- Melitz, Marc J. and Gianmarco Ottaviano**, “Market Size, Trade, and Productivity,” *The Review of Economic Studies*, 2008, 75 (1), 295–316.
- Ministry of Finance, Government of India**, “Economic Survey of India 2015–16,” 2016.
- Mitra, Sandip, Dilip Mookherjee, Maximo Torero, and Sujata Visaria**, “Asymmetric Information and Middleman Margins: An Experiment with Indian Potato Farmers,” *The Review of Economics and Statistics*, forthcoming, 2017.
- National Aeronautics And Space Administration And The National Geospatial Intelligence Agency**, “Shuttle Radar Topography Mission 30 Arc Second Finished Data,” 2000.
- Redding, Stephen J and Daniel M Sturm**, “The Costs of Remoteness: Evidence from German Division and Reunification,” *The American Economic Review*, 2008, 98 (5), 1766–1797.
- Willmott, C. J. and K. Matsuura**, “Terrestrial Air Temperature and Precipitation: Monthly and Annual Time Series (1950 - 2014),” 2001.

Tables and Figures

TABLE 1: SUMMARY STATISTICS

Log Farmer Prices (Rs. per kg)	
Mean	
Barley	2.24
Finger Millet	2.21
Maize	2.15
Paddy	2.25
Pearl Millet	2.17
Sorghum	2.33
Soybean	2.92
Wheat	2.46
Variance Across Markets, Within crop x month	0.029
Variance Across Markets, Within state x crop x month	0.022
Variance Across Months, Within market x crop x agricultural season	0.003
Log Retail Prices (Rs. per kg)	
Mean	
Barley	2.55
Finger Millet	2.29
Maize	2.18
Paddy	2.11
Pearl Millet	2.20
Sorghum	2.34
Soybean	3.33
Wheat	2.49
Variance across districts, within crop x month	0.053
Variance across districts, within crop x month (Conditional on production)	0.026
Variance across districts, within state x crop x month	0.029
Variance across districts, within state x crop x month (Conditional on production)	0.014
Variance Across Months, Within district x crop x agricultural season	0.002
Standard Deviation of comp (as defined in equation 1)	1.85

TABLE 2: FARMER PRICES AND LOCAL COMPETITION

Dependent Variable:	log price			
	All Markets		Border Markets within 30 km	
	(1)	(2)	(3)	(4)
comp (β_1)	0.0163 (0.0043)***	0.0146 (0.0035)***	0.0322 (0.0154)***	0.0292 (0.0132)***
comp' (β_2)	-0.0005 (0.0012)	-0.0007 (0.0009)	-0.0057 (0.0047)	-0.0056 (0.0041)
$\beta_1 - \beta_2$	0.0169 (0.0041)***	0.0154 (0.0034)***	0.0381 (0.0144)***	0.0348 (0.0123)***
Observations	211963	211954	15838	15751
R-Squared	0.60	0.68	0.57	0.68
State, Crop FE	✓	✓	✓	✓
Month-Year FE	✓		✓	
State-Year, Crop-Month-Year FE		✓		✓

Notes: Ordinary least squares. Each observation is a crop-market-month-year. All regressions include controls for district-year specific crop-yields, crop area, crop-specific rainfall shocks, and local population. Robust standard errors clustered at the district level reported in parenthesis. Crops in sample are barley, finger millet, maize, paddy, pearl millet, sorghum, soybean and wheat.

*p < 0.1, ** p < 0.05, *** p < 0.01

TABLE 3: BORDER DISCONTINUITY REGRESSIONS

Dependent Variable:	$\Delta \log \text{ price}$		
	Markets <25 km apart (1)	Markets <30 km apart (2)	Markets <35 km apart (3)
$\Delta \text{ comp } (\beta_1)$	0.025	0.035	0.036
Conley (Uniform)	[0.015]	[0.016]**	[0.012]**
Conley (Bartlett)	[0.015]	[0.017]**	[0.012]**
Robust	(0.011)***	(0.013)***	(0.009)***
Obs	3538	5292	9018
Mkt Pairs	108	173	284
State-Pair FE	✓	✓	✓

Notes: Observations are market-pair-crop-month. Column (a) reports regressions with markets pairs on state borders and (b) reports placebo regressions with market pairs not on state borders. Brackets contain Conley (1999) standard errors that allow for spatial and serial correlation of the error terms. Arbitrary spatial correlation in error terms is allowed within 100 km from the geographic midpoint of the market pairs. Serial correlation is allowed over 6 months, equivalent to one crop season in India. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

TABLE 4: STRUCTURAL ESTIMATION RESULTS

PARAMETER	POINT ESTIMATE	CLUSTER ROBUST SE
δ	0.69	0.0025
σ	0.00039	0.0048
A	0.0016	0.0007
J-Statistic	30.00	

Notes: Standard Errors Clustered at State-Time Level

TABLE 5A: SENSITIVITY MEASURES

	Andrews, Gentzkow & Shapiro (2017)			Cooper & Liu (2016)		
	δ	σ	A	δ	σ	A
β_0	-1.57	2.12e+05	114.87	327.31	4.42e-4	-4.85
Max P(β_1)	0.23	-32110	-16.77	-122.77	4.30e-4	2.20
Distance (β_2)	0.08	-10652	-5.91	72.16	1.31e-4	-1.04
Retail Price (β_3)	0.17	-21988	-12.64	29.90	7.20e-5	-0.41
Mean Sq Error	1.75e-2	-2924.4	-1.27	0.07	-1.98e-5	-0.01
Mean	1.55e-5	-33.02	-1.17e-3	-15.97	-1.96e-6	0.14
Variance	-6.24e-6	0.83	4.55e-4	-6.91	-9.97e-7	-0.011
5 th -25 th P'ctile	-2.20e-3	-334.66	0.15	-210.49	-0.03	7.19
25 th -50 th P'ctile	7.89e-2	-1478.7	-0.57	-2.60	1.14e-3	-0.004
50 th -75 th P'ctile	1.2e-2	-1872.5	-0.90	-40.07	1.83e-3	1.24
75 th -90 th P'ctile	8.79e-3	-1379.2	-0.64	-9.97	-1.38e-4	0.11

TABLE 5B: SENSITIVITY MEASURES (HONORÉ, JØRGENSEN & DE PAULA 2017)

Elasticity of	Estimator wrt Diagonal of			Variance of Estimator wrt		
	the Weighting Matrix			Variance of Moments		
	δ	σ	A	δ	σ	A
β_0	37.09	-1.08E+06	-2426.7	0.45	1768.7	151.65
Max P(β_1)	-50.30	1.66E+06	3447.3	129.09	1230	236.05
Distance (β_2)	-8.33	2.42E+05	545.19	91.23	7.14	7.02
Retail Price (β_3)	-20.12	5.90E+05	1321.8	7.46	2.54	0.94
Mean Sq Error	-1.63e-3	40.28	0.10	1.37	1.34	0.41
Mean	-4.74	1.59E+05	328.67	1.59	40.85	4.71
Variance	-2.34	93086	177.29	0.075	27.41	1.80
5 th -25 th P'ctile	-0.45	12957	29.627	0.56	0.20	0.007
25 th -50 th P'ctile	-3.62	17806	169.4	97.46	13.80	20.11
50 th -75 th P'ctile	-0.44	-7490.6	12.92	43.80	41.19	13.38
75 th -90 th P'ctile	-0.36	12010	25.10	15.76	14.04	4.82

TABLE 6: GOODNESS OF FIT

Moment of Farmer Prices	Data	Model
Mean	2.18	2.31
Variance	0.13	0.14
25 th Percentile	1.90	2.04
Median	2.19	2.25
75 th Percentile	2.39	2.54

TABLE 7: SUMMARY OF COUNTERFACTUAL EXERCISES

Change in Farmer Prices (in %)	Median		Mean		Mean For the Top 50%		Mean For the Top 25%	
	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring
#1: Change the treat point	1.75	2.50	2.74	2.65	4.97	4.50	7.36	5.84
#2: #1+Market Choice	8.35	5.82	12.60	10.72	23.37	20.04	32.67	29.69
#3: #2+Adjust Retail Price	5.98	6.21	9.62	9.56	20.32	19.49	31.70	28.74

TABLE 8: SENSITIVITY TO $\gamma + \upsilon$

$\gamma + \upsilon$	Mean Increase in Output (%)			
	All Farmers		Top 50%	
	Fall	Spring	Fall	Spring
0.57	9.10	7.75	16.10	14.46
0.60	7.91	6.73	14.68	12.55
0.63	7.00	5.96	12.99	11.11

TABLE 9: SENSITIVITY TO FUNCTIONAL FORM FOR TRADE COST

Functional Form	δ	σ	A	Mean Increase in Prices (%)			
				All Farmers		Top 50%	
				Fall	Spring	Fall	Spring
1+A.distance	0.69	0.00039	0.0016	12.6	10.72	23.37	20.04
(1+A.distance) ^{0.33}	0.69	0.00058	0.0057	21.5	21.75	37.04	33.36

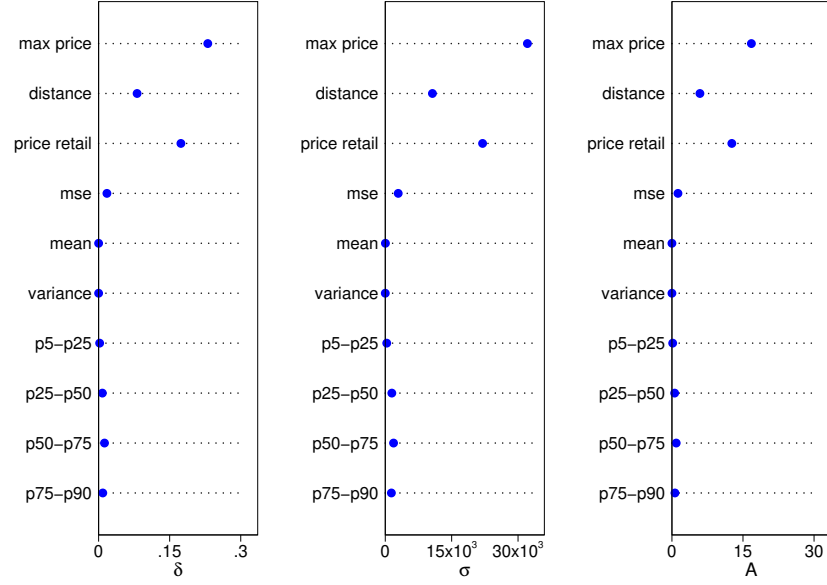


Figure 2: Semi-Elasticity of the Prob Limit of the Estimator wrt Standardized Moments.

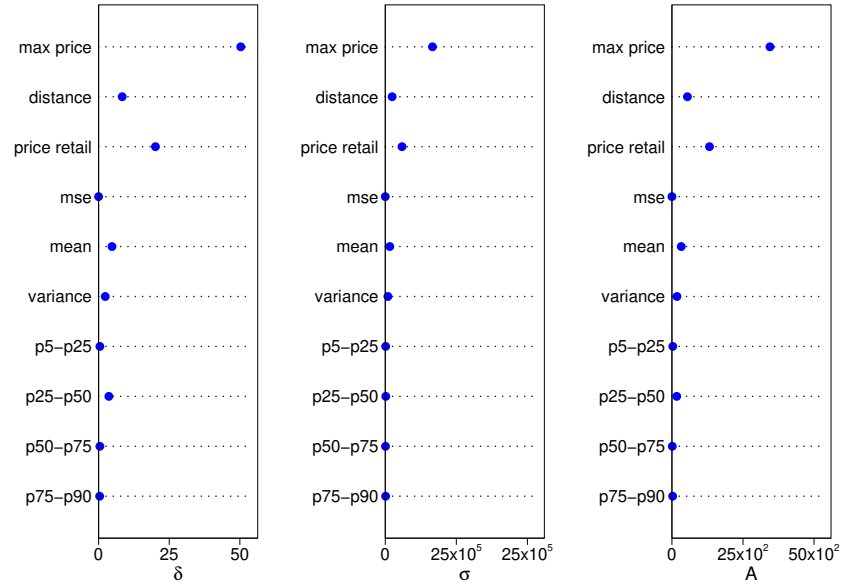


Figure 3: Elasticity of the Estimator wrt the Weight on a Moment.

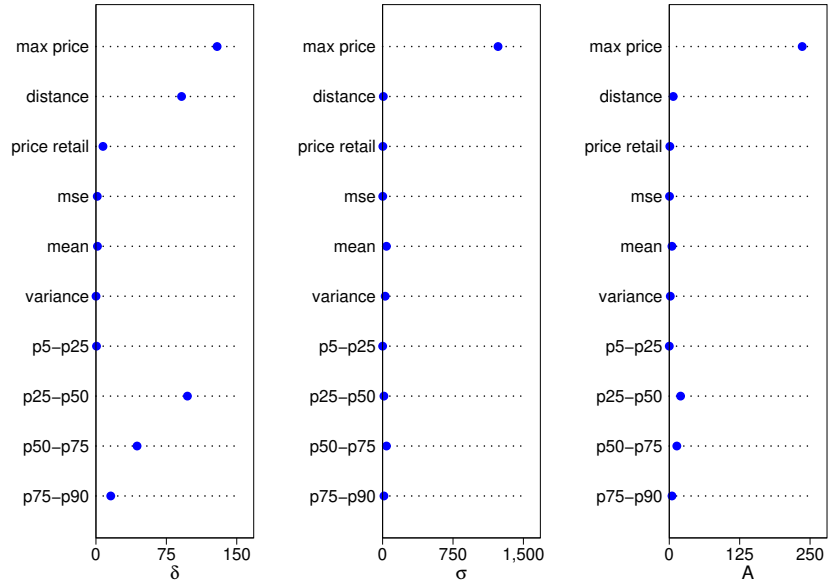


Figure 4: Elasticity of the Variance of Estimator wrt Variance of Moments.

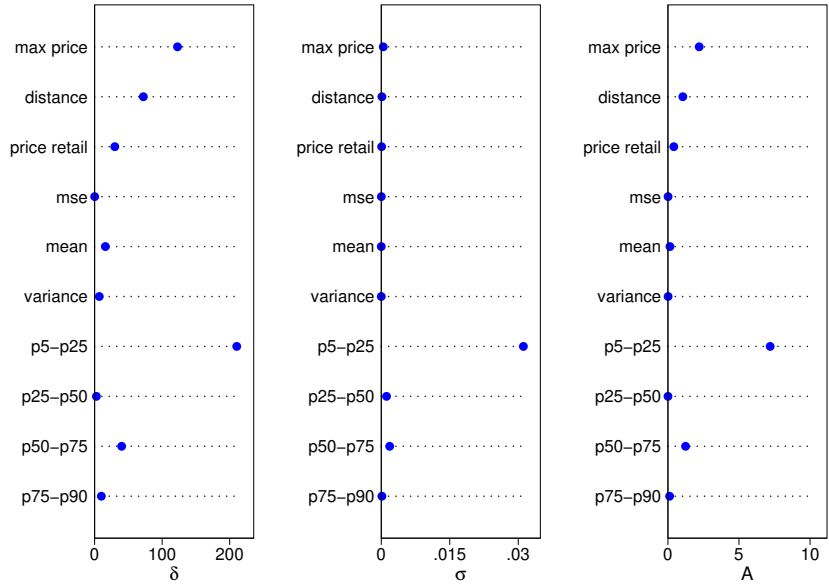


Figure 5: Elasticity of the Moment wrt the Parameter.

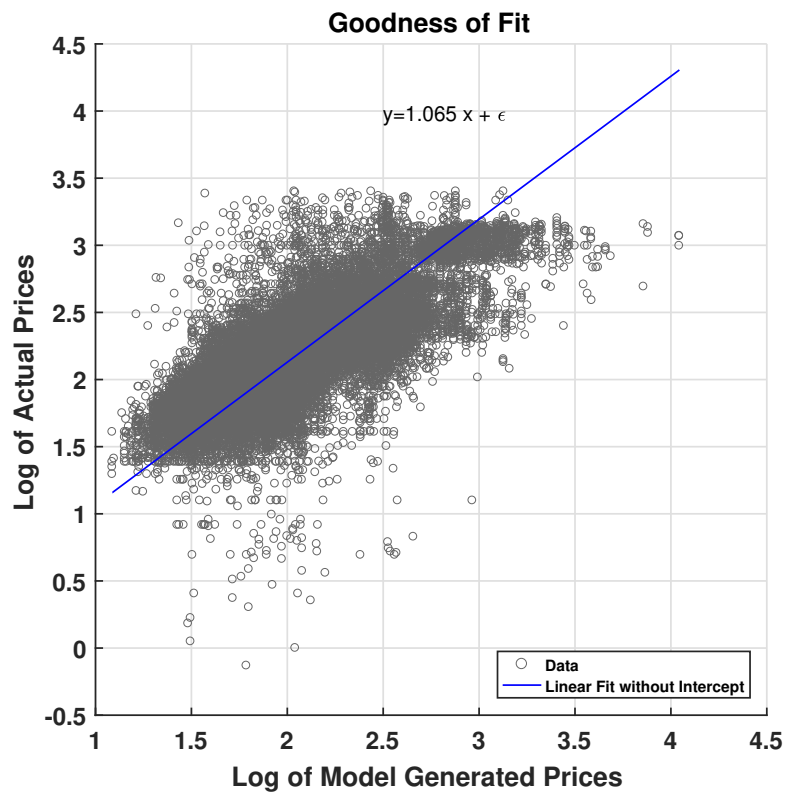


Figure 6: Linear Fit: Data vs. Model.

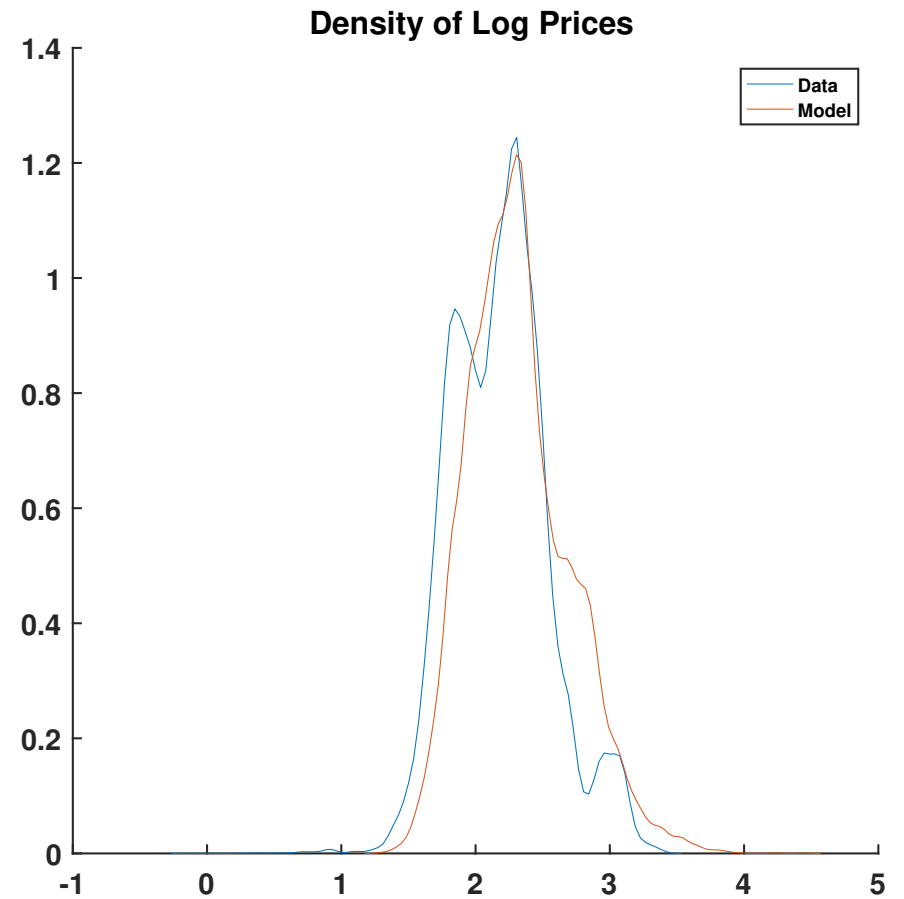


Figure 7: Empirical Density of Prices: Data vs. Model.

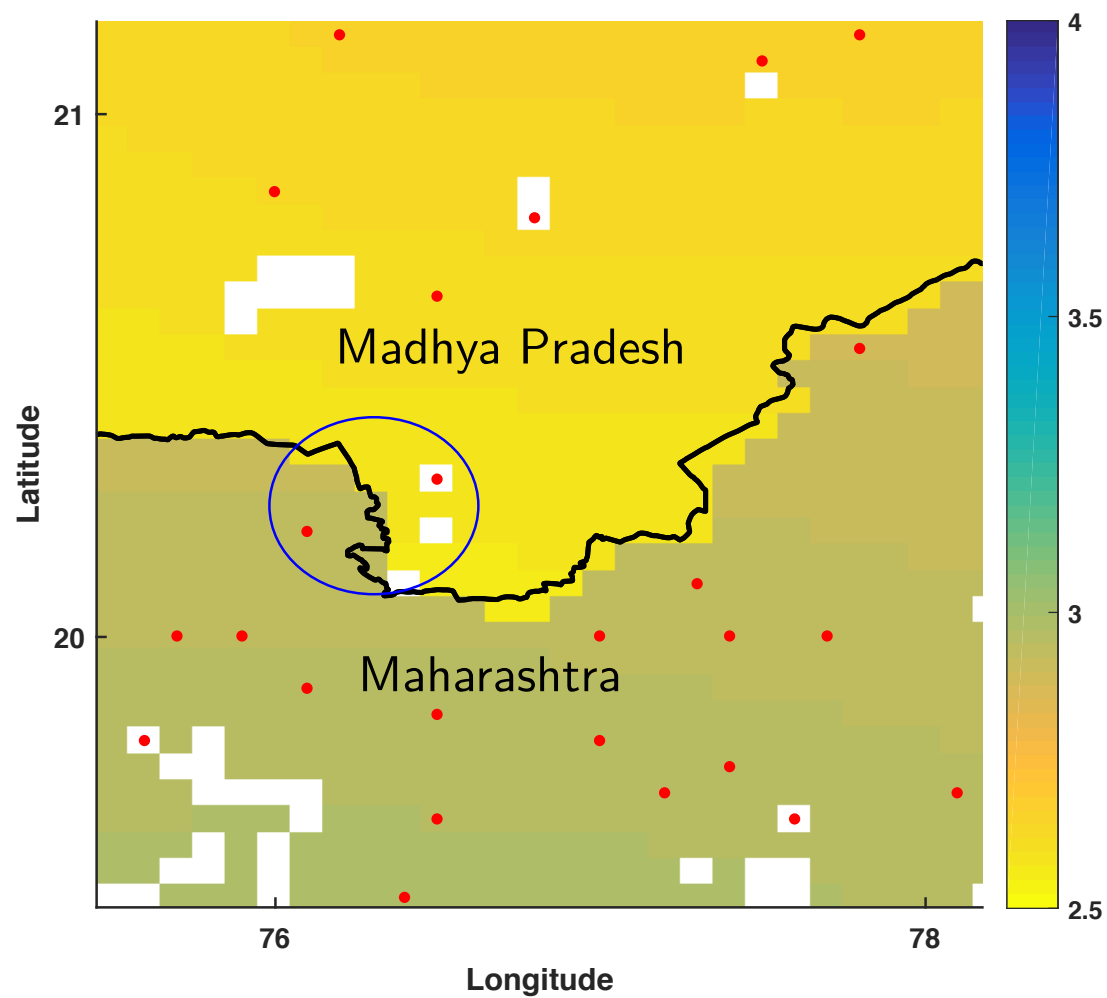


Figure 8: Illustrative Example.

Notes: The red dots are primary grain markets. The border markets in the example are encircled in blue. The figure also plots the log of soybean prices (Rs. per kilo) received by farmers.

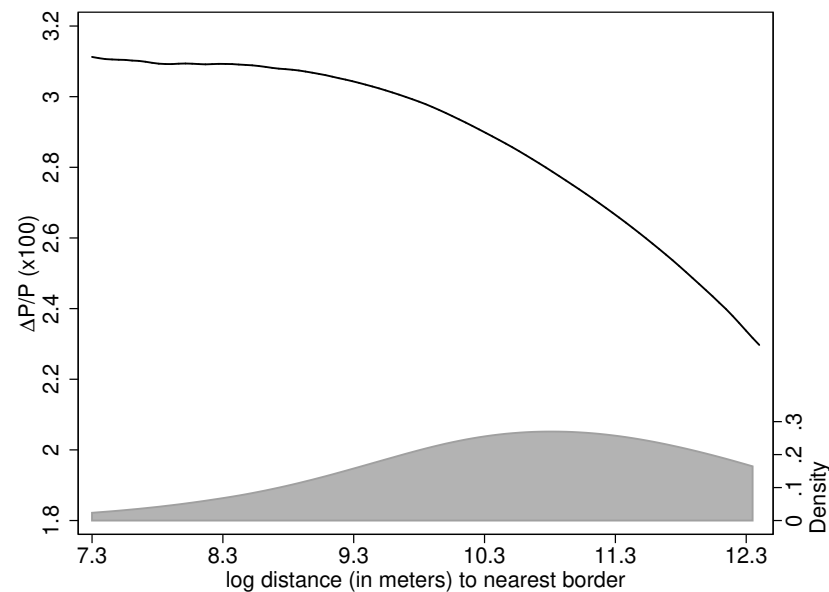
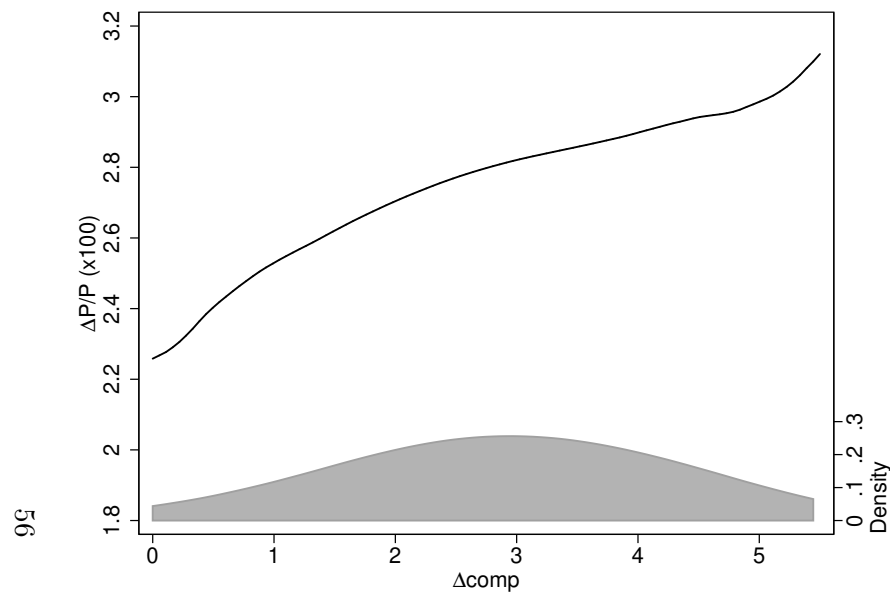


Figure 9: Change in Price at Each Market vs. Change in Competition. Figure 10: Change in Price at Each Market vs. Distance to Nearest Border.

Notes: The solid line in Figure 9 is the estimated relationship between change in price at a market and the change in competition, $\Delta comp$, due to removal of the interstate trade restriction. Competition is measured as in the reduced-form regressions (defined in equation 1). The solid line in Figure 10 is the estimated relationship between change in price at a market and the distance from the market to the nearest state border. Estimated relationships using local linear regressions of bandwidth 1. Domain is 1st to 99th Percentile. Bottom of both figures plot the empirical distribution of the x-axis variable.

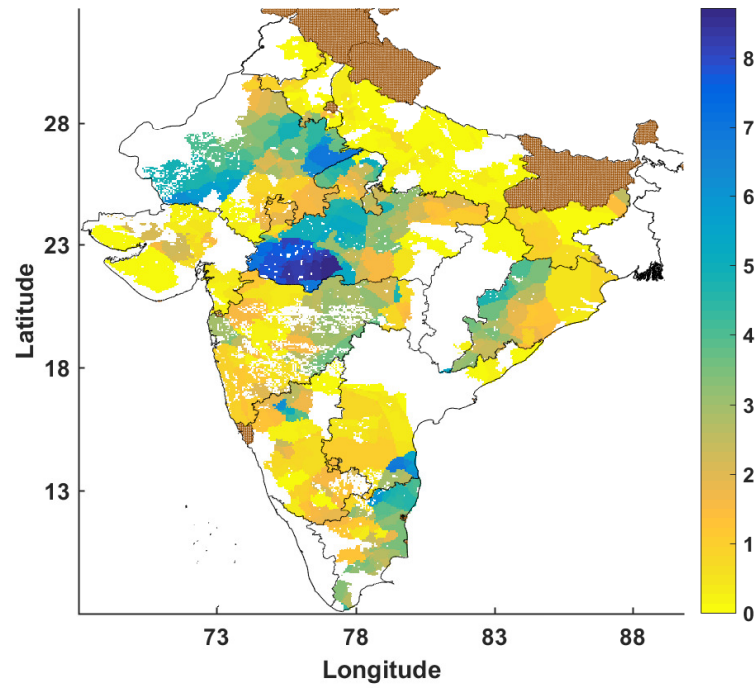


Figure 11: % Increase in Prices in the Fall Holding Market Choice of Farmers Fixed.

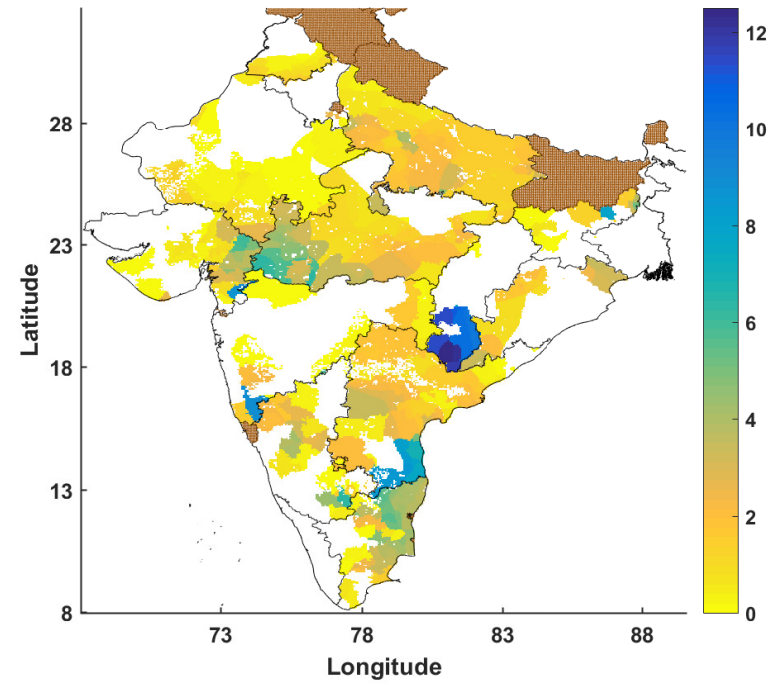


Figure 12: % Increase in Prices in the Spring Holding Market Choice of Farmers Fixed.

Note: Brown color represents no data and white represents zero gains.

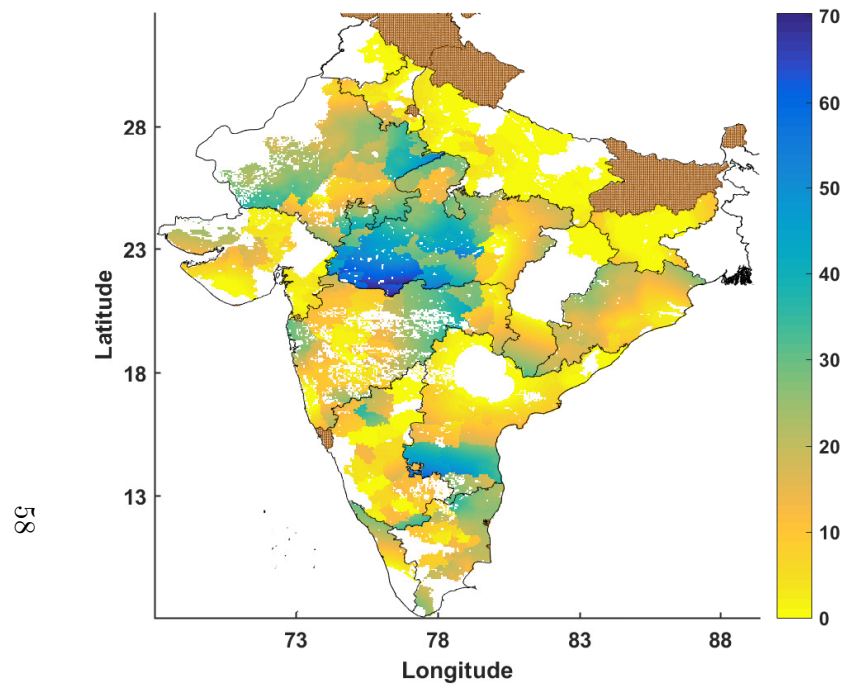


Figure 13: % Increase in Prices allowing for Optimal Market Choice—Fall Season.

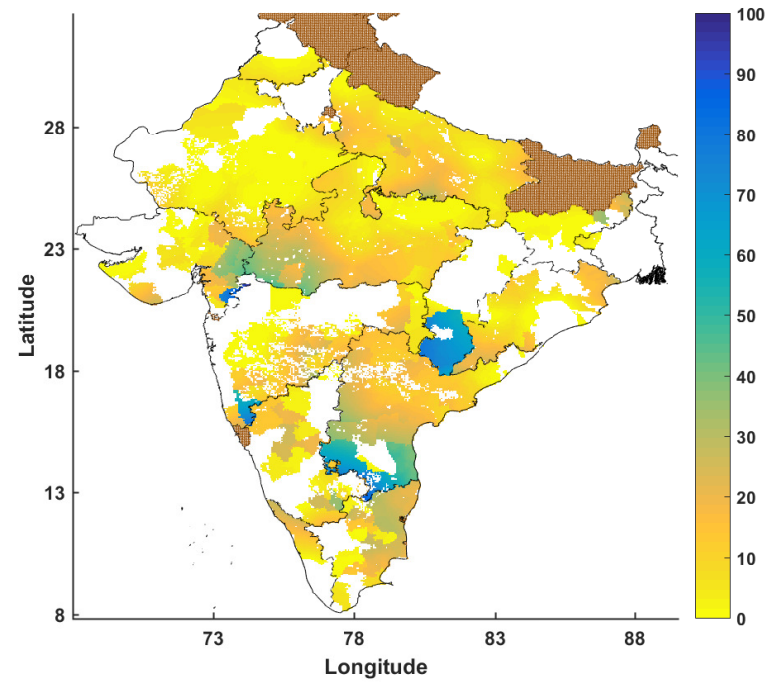


Figure 14: %Increases in Prices allowing for Optimal Market Choice—Spring Season.

Notes: Brown color represents no data and white represents zero gains.

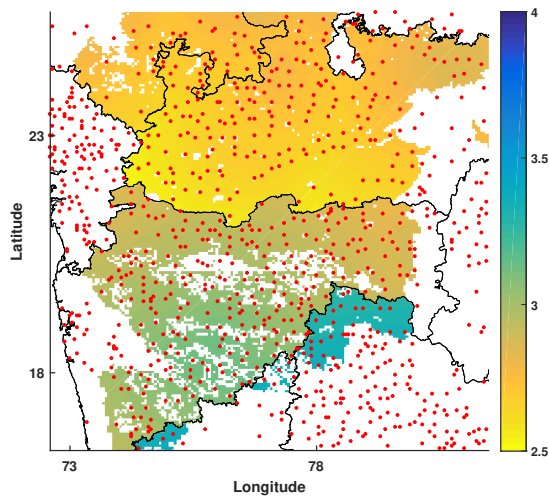


Figure 15: Log of Soybean Prices (Rs. per kilo) Before Policy change.

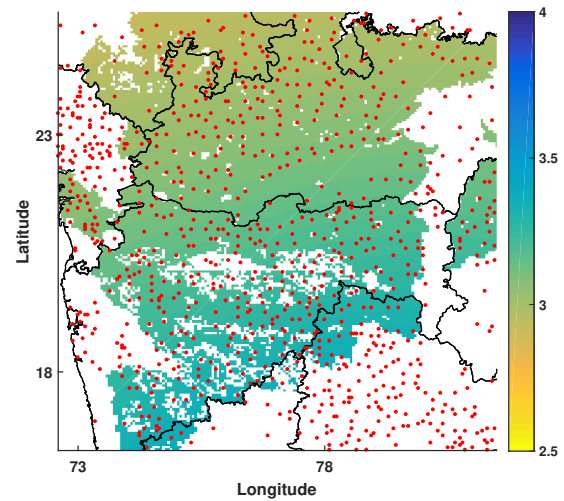


Figure 16: Log of Soybean Prices (Rs. per kilo) After Policy change.

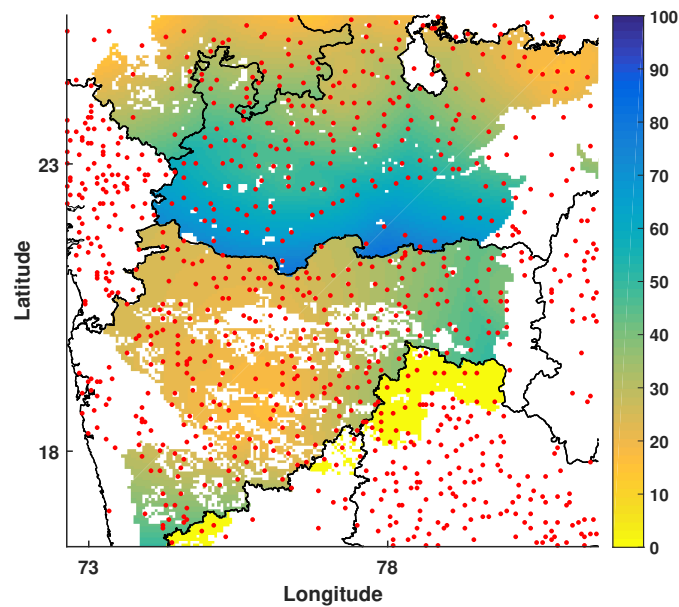


Figure 17: %Change in Soybean Prices in Maharashtra and Madhya Pradesh.

Note: The red dots are primary grain markets.

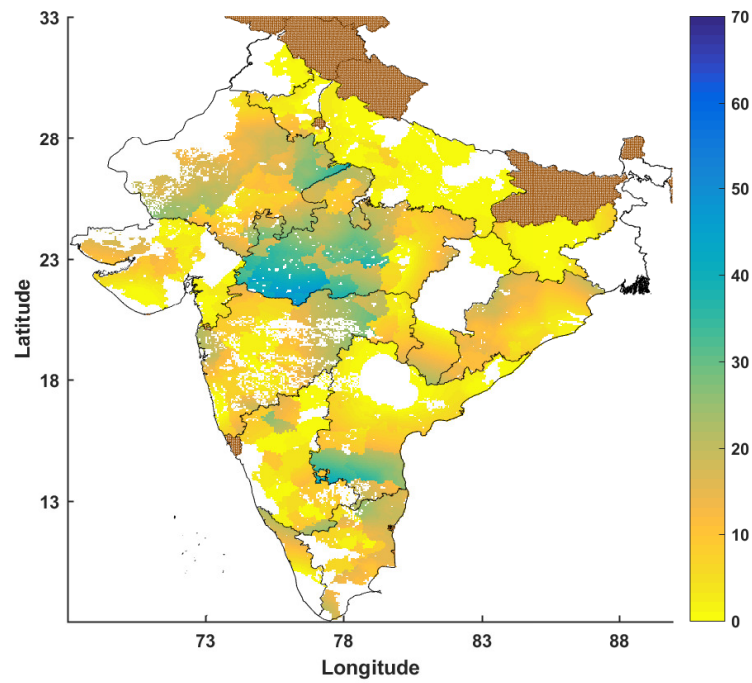


Figure 18: % Increase in Production in the Fall Season.

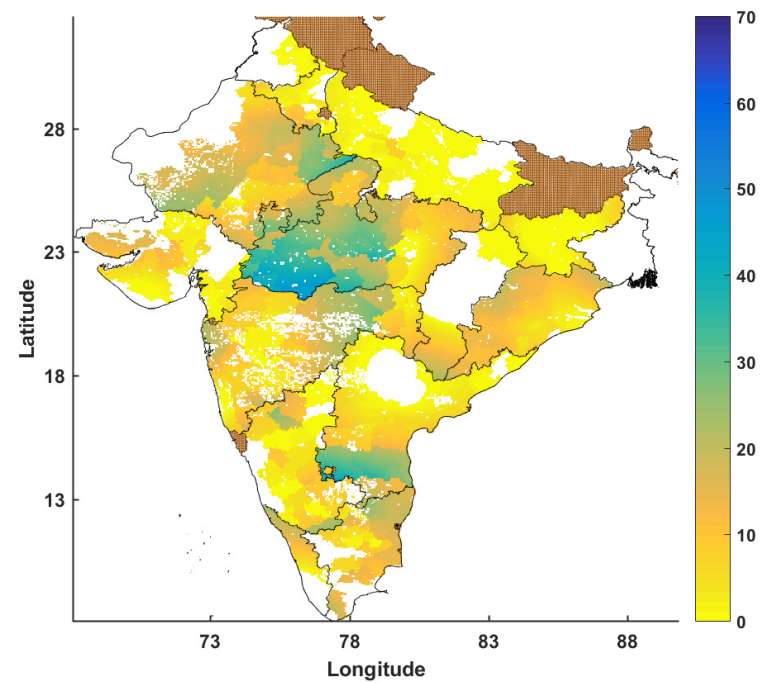


Figure 19: % Increase in Production in the Spring Season.

Notes: Brown color represents no data and white represents zero gains.

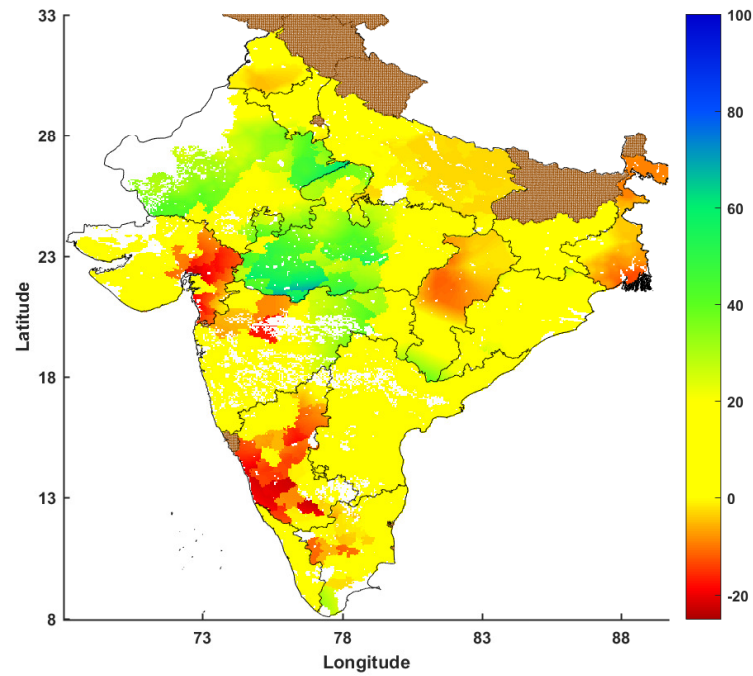


Figure 20: % Increase in Prices in the Fall Season
Incorporating a Demand Side Response.

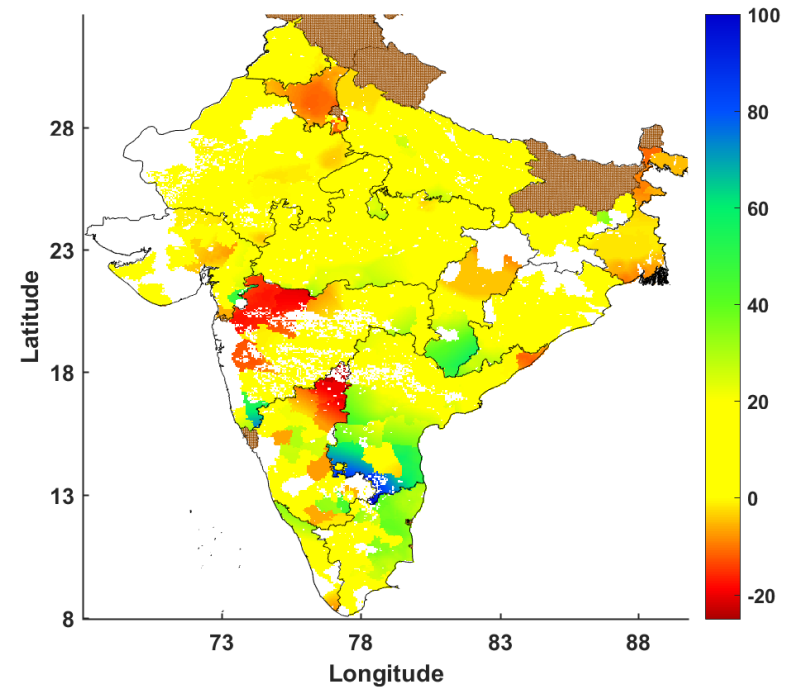


Figure 21: % Increase in Prices in the Spring Season
Incorporating a Demand Side Response.

Notes: Brown color represents no data and white represents zero gains.

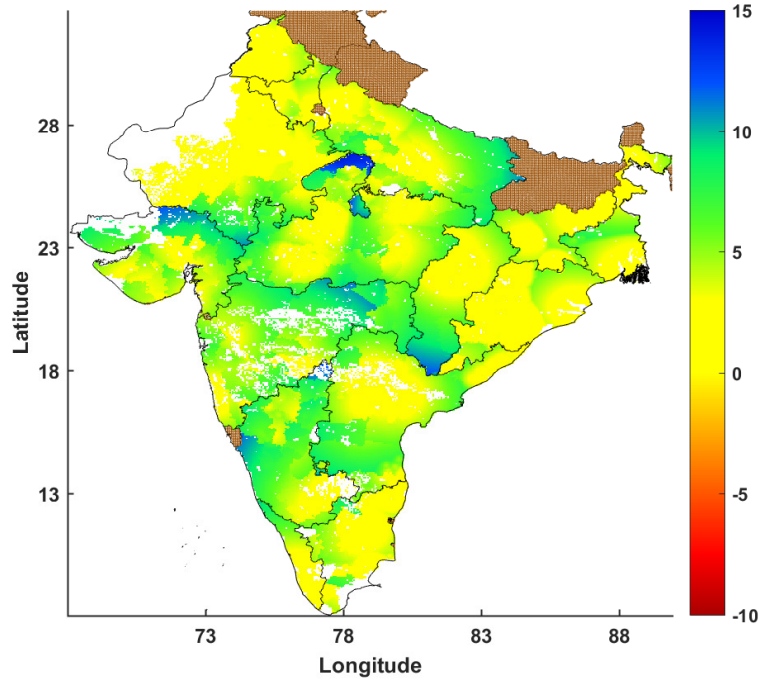


Figure 22: % Increase in Prices due to Reduction in Trade Costs with Border Restrictions.

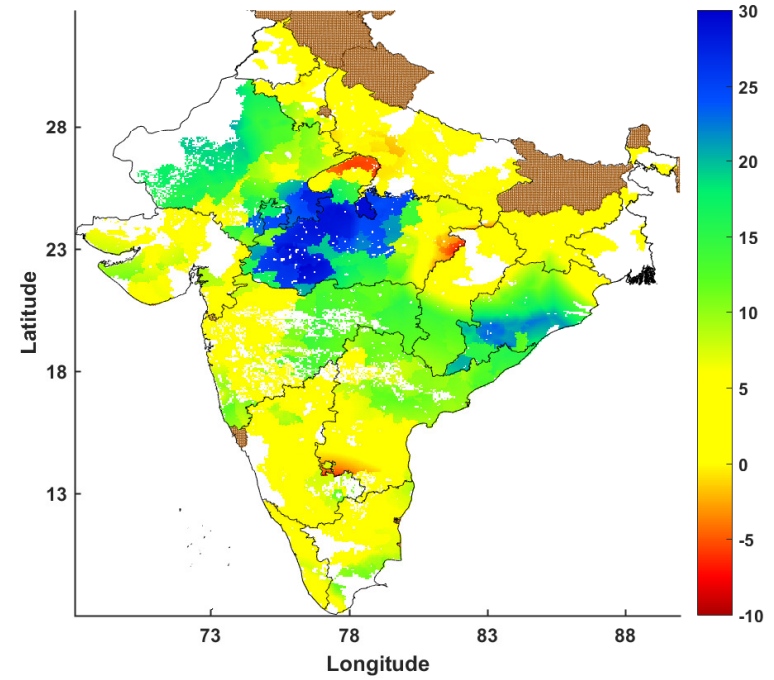


Figure 23: Additional Changes in Prices (in %) because of a Reduction in Trade Costs and Elimination of Border Restrictions.

Notes: Brown color represents no data and white represents zero gains. Estimates are for the fall season. To compute the changes in the Figure (23), I first compute changes in prices due to reduction in trade cost and elimination of border restrictions. From this I subtract the gains in prices just due to reduction in trade costs (Figure 22) and just due to elimination of border restrictions to trade (Figure 13).

Appendix

A States in Sample



Figure A.1: States in Sample

B Construction of Markets

In this section, I present conditional correlations between factors that drove geographical placement of markets. I regress the number of markets in a district as observed in my main data set on average cereal output, cereal yield, district area, cropped area, population between 1970-1985, and state fixed effects. This time period was during the peak of the green revolution when most new market

construction occurred. District level data on production and population for this period come from the Village Dynamics in South Asia (VDSA) “meso” data set. We can see from Table B.1 that conditional on district size, cropped area, and population, districts that had a greater cereal output in 1970-1985 also have a greater number of markets today.

TABLE B.1:FACTORS INFLUENCING MARKET CONSTRUCTION

Dependent Variable:	Number of markets in 2005-2014	
cereal output	35.7487	
	(10.1235)	
cereal yield		0.3723
		(0.1018)
crop area	-0.0342	0.3315
	(0.1277)	(0.1198)
total area	0.4074	0.4083
	(0.0981)	(0.0979)
population	0.2400	0.2290
	(0.1062)	(0.1067)
N	251	251
R-Squared	0.67	0.67

Notes: All observations are at district level. All variables are in logs. All right hand side variables are district level averages over 1970-1985. All regressions have state fixed effects. OLS standard errors in parenthesis.

C Placebo Tests

In this section, I provide regression results to show that there are no discontinuities in other observable factors that might influence farmer prices. I run specifications like (4), but regress observable factors on competition instead of farmer prices.

In panel A, I regress log of potential yields of various crops on competition to show that there is no discontinuity in crop productivity at the border. In panel B, I regress log of land elevation and log on distance to the district headquarter on competition. Asher, Nagpal and Novosad (2017) shows that distance to district headquarter effects delivery of public goods and services to villages.

Differences in public goods like roads can influence farmer prices and can be a concern if correlated with spatial competition. Column 2 of Panel B allays this concern.

PANEL A

	$\Delta \log (\text{agro-climatically attainable yield})$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Rice	Wheat	Maize	Millet	Barley	Sorghum	Soybean
$\Delta \text{ comp}$	0.007 (0.020)	-0.016 (0.029)	-0.002 (0.006)	0.003 (0.005)	-0.003 (0.026)	0.002 (0.006)	0.009 (0.008)
Obs	173	173	173	173	173	173	173
State-Pair FE	✓	✓	✓	✓	✓	✓	✓
Kernel	Uniform	Uniform	Uniform	Uniform	Uniform	Uniform	Uniform

Notes: Brackets contain Conley (1999) standard errors that allow for spatial and serial correlation of the error terms. Arbitrary spatial correlation in error terms is allowed within 100 km from the geographic midpoint of the market pairs. Bandwidth for market pairs is 30 km. Data on Agro-Climatically attainable yield from FAO GAEZ and averaged for a 50 km zone within the state around each market. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

PANEL B

	(1)	(2)
	$\Delta \log(\text{elevation})$	$\Delta \log(\text{dist to district HQ})$
$\Delta \text{ comp}$	0.053 (0.076)	0.301 (0.820)
Obs	173	173
State-Pair FE	✓	✓
Kernel	Uniform	Uniform

Notes: Brackets contain Conley (1999) standard errors that allow for spatial and serial correlation of the error terms. Arbitrary spatial correlation in error terms is allowed within 100 km from the geographic midpoint of the market pairs. Bandwidth for market pairs equals 30 km. Data on land elevation from NASA SRTM and averaged around a 50 km zone within the state for each market. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

D Proof of Theorem 1

To show that there is unique fixed point of (11):

$$p^f(m) = (1 - \delta) \max_{k \in \mathcal{M} - \{m\}} \left\{ \frac{p^f(k)}{\tau_{mk}} \right\} + \delta p_m^r \quad (\text{D.1})$$

Define $p^f : \mathcal{M} \rightarrow [0, \max_m p_m^r]$. Note that $\mathcal{M} = \{1, 2, \dots, M\}$, $\tau : \mathcal{M} \times \mathcal{M} \rightarrow [1, \infty]$, $\delta \in (0, 1)$ and $p_m^r \in \mathbb{R}_+$.

D.1 Existence

Since the max function is continuous, the RHS of (D.1) is a continuous function. Therefore, Brouwer's Fixed Point Theorem guarantees existence of a $p^{f*}(\cdot)$ that solves the system of equations defined in (D.1). Further, since $\delta \in (0, 1)$ and $p_m^r \in \mathfrak{R}_+$, it must be that $p^{f*}(\cdot) > 0$.

D.2 Uniqueness

Define $T(f)(s) = (1 - \delta) \max_{s' \in \Gamma(s)} \left\{ \frac{f(s')}{\tau(s, s')} \right\} + \delta p_s^r$. I will verify Blackwell's sufficient conditions to show that $T(\cdot)$ is a contraction.

1. Monotonicity: Let $f(s) \geq g(s)$ for all s .

$$\begin{aligned} (1 - \delta) \max_{s' \in \Gamma(s)} \left\{ \frac{f(s')}{\tau(s, s')} \right\} + \delta p_s^r &\leq (1 - \delta) \max_{s' \in \Gamma(s)} \left\{ \frac{g(s')}{\tau(s, s')} \right\} + \delta p_s^r \\ \max_{s' \in \Gamma(s)} \left\{ \frac{f(s')}{\tau(s, s')} \right\} &\leq \max_{s' \in \Gamma(s)} \left\{ \frac{g(s')}{\tau(s, s')} \right\} \end{aligned}$$

Since $f(s) \geq g(s) \forall s$, let us say $s^* = \arg \max_{s'} \frac{g(s')}{\tau(s, s')}$, then $\frac{f(s^*)}{\tau(s, s^*)} \geq \frac{g(s^*)}{\tau(s, s^*)}$. It must be true that

$$\max_{s' \in \Gamma(s)} \left\{ \frac{f(s')}{\tau(s, s')} \right\} \geq \frac{f(s^*)}{\tau(s, s^*)} \geq \frac{g(s^*)}{\tau(s, s^*)} = \max_{s' \in \Gamma(s)} \left\{ \frac{g(s')}{\tau(s, s')} \right\}$$

Therefore,

$$T(f)(s) \geq T(g)(s)$$

2. Discounting: Let $a \in \mathbb{R}_+$.

$$\begin{aligned} T(f + a)(s) &= (1 - \delta) \max_{s' \in \Gamma(s)} \left\{ \frac{f(s') + a}{\tau(s, s')} \right\} + \delta p_s^r \\ &\leq (1 - \delta) \max_{s' \in \Gamma(s)} \left\{ \frac{f(s')}{\tau(s, s')} \right\} + \delta p_s^r + (1 - \delta) a \\ &= T(f)(s) + (1 - \delta) a \end{aligned}$$

The second line follows because $\tau(\cdot, \cdot) \geq 1$.

E Proof of Lemma 1

Let \mathcal{M} denote the set of markets in a region. When border restrictions are removed, let the set of markets available to farmers in this region increases to $\mathcal{N} \supset \mathcal{M}$. Denote equilibrium farmer prices before and after border restrictions are removed by $p^1(t)$ and $p^2(t)$ respectively, where $t \in \mathcal{M}, \mathcal{N}$.

Suppose $\exists m \in \mathcal{M}$, such that $p^2(m) < p^1(m)$. This implies

$$\begin{aligned} (1 - \delta) \max_{k \in \mathcal{N} - \{m\}} \left\{ \frac{p^2(k)}{\tau_{mk}} \right\} + \delta p_m^r &< (1 - \delta) \max_{j \in \mathcal{M} - \{m\}} \left\{ \frac{p^1(j)}{\tau_{mj}} \right\} + \delta p_m^r \\ \iff \max_{k \in \mathcal{N} - \{m\}} \left\{ \frac{p^2(k)}{\tau_{mk}} \right\} &< \max_{j \in \mathcal{M} - \{m\}} \left\{ \frac{p^1(j)}{\tau_{mj}} \right\} \\ \iff \max_{k \in \mathcal{M} - \{m\}} \left\{ \frac{p^2(k)}{\tau_{mk}} \right\} &< \max_{j \in \mathcal{M} - \{m\}} \left\{ \frac{p^1(j)}{\tau_{mj}} \right\}. \end{aligned}$$

The last line follows from $\mathcal{M} \subset \mathcal{N}$. Because of uniqueness it must be that

$$\max_{k \in \mathcal{M} - \{m\}} \left\{ \frac{p^1(k)}{\tau_{mk}} \right\} < \max_{j \in \mathcal{M} - \{m\}} \left\{ \frac{p^1(j)}{\tau_{mj}} \right\}.$$

This implies that

$$p^1(k^*) < p^1(k^*),$$

which is a contradiction. Therefore, it must be that $p^2(m) \geq p^1(m) \forall m \in \mathcal{M}$.

F Change in Production in Response to a Change in Prices

Let p denote prices that farmers get and w^k the price of the k th intermediate input. The production function is given in (5). The profit maximization problem for a farm can be written as follows where I omit the f subscript:

$$\max_{x^k} p \tilde{A} \left(h^\gamma l^\nu \prod_{k=1}^K (x^k)^{\alpha_k} \right) - \sum w^k x^k$$

The First Order Condition is:

$$p y \alpha_k = w^k x^k \tag{F.1}$$

Therefore:

$$\frac{\alpha_k}{\alpha_j} = \frac{w^k x^k}{w^j x^j}$$

$$x^j = \left(\frac{\alpha_j}{w^j} \right) \left(\frac{w^k x^k}{\alpha_k} \right)$$

From the first order condition (F.1), we can write:

$$p \left(\frac{\alpha_k}{x^k} \right) \left(\tilde{A} h^\gamma l^\nu \prod_{j=1}^K (x^j)^{\alpha_j} \right) = w^k$$

$$\tilde{A} h^\gamma l^\nu p \left(\frac{\alpha_k}{x^k} \right) \left(\prod_{j=1}^K (x^j)^{\alpha_j} \right) = w^k$$

$$\tilde{A} h^\gamma l^\nu p \left(\frac{\alpha_k}{x^k} \right) \left(\frac{w^k x^k}{\alpha_k} \right)^{1-\gamma-\nu} \left(\prod_{j=1}^K \left(\frac{\alpha_j}{w^j} \right)^{\alpha_j} \right) = w^k$$

$$\tilde{A} h^\gamma l^\nu p \left(\frac{\alpha_k}{x^k} \right)^{\gamma+\nu} \left(\prod_{j=1}^K \left(\frac{\alpha_j}{w^j} \right)^{\alpha_j} \right) = (w^k)^{\gamma+\nu}$$

$$\tilde{A} h^\gamma l^\nu p \left(\frac{\alpha_k}{w^k} \right)^{\gamma+\nu} \left(\prod_{j=1}^K \left(\frac{\alpha_j}{w^j} \right)^{\alpha_j} \right) = (x^k)^{\gamma+\nu}$$

$$\Upsilon p^{\frac{1}{\gamma+\nu}} = x^k$$

where, Υ is a constant.

Now, new output to old output ratio is:

$$\frac{y'}{y} = \frac{\prod_{k=1}^K (x^k)^{\alpha_k}}{\prod_{k=1}^K (x^k)^{\alpha_k}}$$

$$= \prod_k \left(\frac{(x^k)'}{x^k} \right)^{\alpha_k}$$

$$= \prod_k \left(\frac{p'}{p} \right)^{\frac{\alpha_k}{\gamma+\nu}}$$

$$= \left(\frac{p'}{p} \right)^{\frac{1-(\gamma+\nu)}{\gamma+\nu}}$$

F.1 A Note on Calibrating $\gamma + \nu$

From the first order condition (F.1),

$$\begin{aligned} py \sum_{k=1}^K \alpha_k &= \sum_{k=1}^K w^k x^k \\ \iff py(1 - \gamma - \nu) &= \sum_{k=1}^K w^k x^k \end{aligned}$$

Therefore, denoting wage rate by w^l and rental rate of land by w^h , expenditure on fixed factors must be:

$$w^l l + w^h h + \pi = (\gamma + \nu) py,$$

and therefore

$$\gamma + \nu = \frac{w^l l + w^h h + \pi}{w^l l + w^h h + \sum_{k=1}^K w^k x^k + \pi}.$$

G Algorithm to Compute Equilibrium Prices

1. Start with a distribution of retail prices $\{p_0^r\}$, associated farmer prices, and crop output $\{q_0^r\}$.
2. Remove border restrictions to trade and compute the new farmer prices and the anticipated output. Given the output and the demand elasticities, find the anticipated retail prices $\{p_A^r\}$.
3. Guess the new retail prices $\{p_{guess}^r\}$ as a fraction of the anticipated retail prices.
4. Find the associated farmer prices and quantities.
5. Find the actual change in retail prices $\{p_1^r\}$. If the actual retail prices are equal to the guess, then stop.
6. Otherwise update $p_{guess}^r = p_o^r + 0.2 \times (p_{guess}^r - p_1^r)$. Go to Step 4.

H Sensitivity to Allocation of Crop Choice to Grid Cells

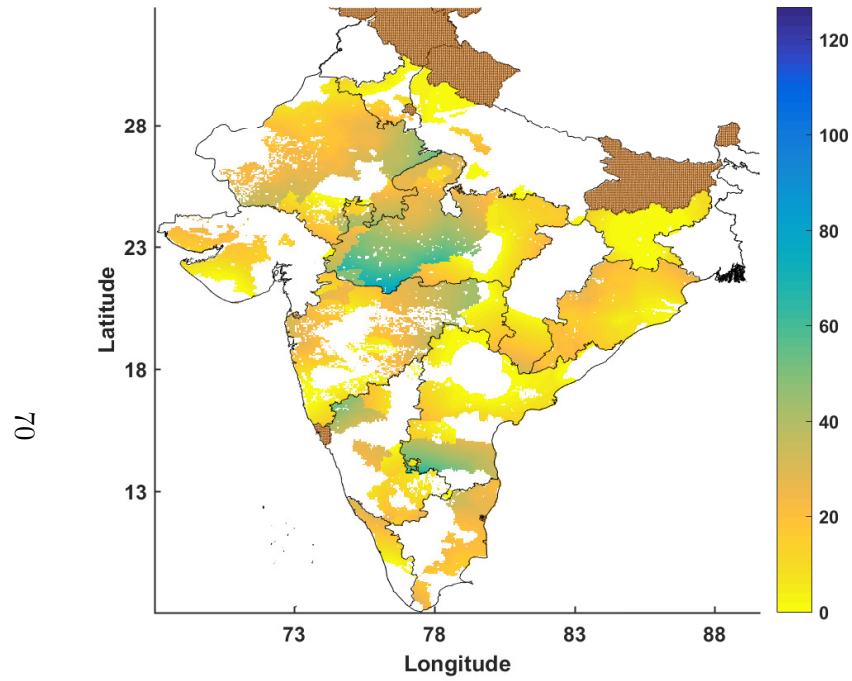


Figure H.1: % Increase in Prices - Fall Season
For the main crop grown in the district.

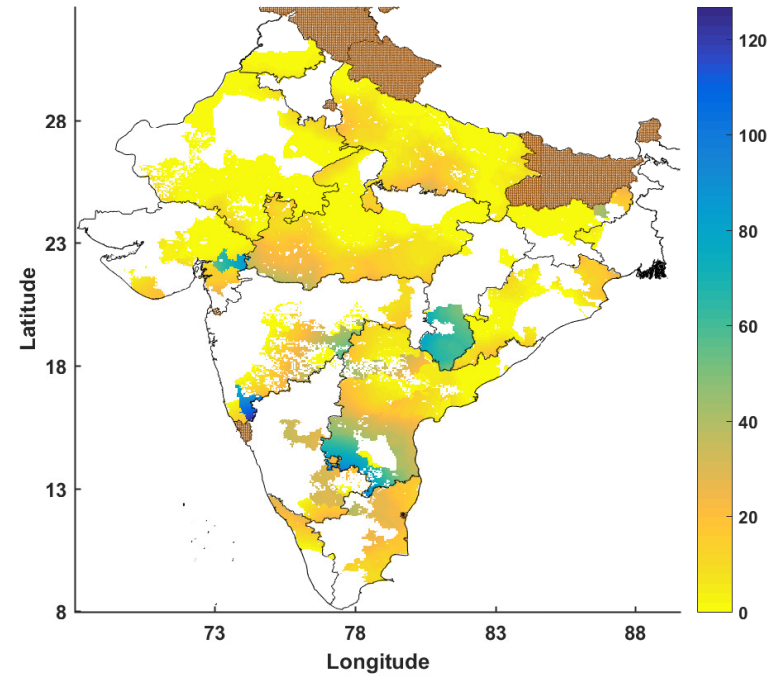


Figure H.2: %Increases in Prices - Spring Season
For the main crop grown in the district.

Notes: Brown color represents no data and white represents zero gains.

I Data Appendix

I.1 District Names

To maintain consistency across time periods, I appropriate all data to 2001 census boundaries of Indian districts. That is, districts that got bifurcated after 2001 are assigned to their parent districts.

I.2 Organizing the Data on Farmer Prices and Retail

Each observation in the farmer prices data provides the market level mode price for a given month for each commodity that was sold in any market. There is a variable that has information about the variety of the commodity. The village, district, and state of the market is also recorded.

I cleaned the data on farmer prices in the following steps. First, I winsorized the top and bottom 1% of the outlier observations. Unfortunately, most of the data (by the number of observations) is under the variety “other” and hence, this variable is not of much use. In the second step to maintain homogeneity of commodities, I dropped high-value varieties like “Super Fine Rice,” “Basmati Rice,” “Dehradun Rice,” “Sharbati Wheat,” “Fine Wheat” etc. I also dropped Maize and Millets meant for “Cattle Fodder.” Third, I corrected for Indian names of villages such that there is no duplication and then I geocoded the markets using the names of the state, district, and village in Google Maps API.

The unit of observation in the retail price data is a commodity-district-month-year. Data are recorded for three types of rice—fine, medium, and coarse. I use the data on medium and coarse varieties of rice and compute the median price across varieties. I fixed district names to match those in the farmer price data. Then I removed the outlier observations, by manually examining district-crop level data. I did not follow the 1% winsorizing rule since it would drop more observations than required. Then I computed district-crop-month level median prices, which I use as my final data on retail prices.

I.3 Crops-State Combinations in Structural Estimation

For structural estimation, I include data from the following states. This choice is meant to exclude states that do not produce a particular crop. To determine this I make use of Agriculture Statistics at a Glance published each year by the Ministry of Agriculture, Government of India.

Crop	States Excluded
Rice	Rajasthan
Wheat	Andhra Pradesh, Chhattisgarh, Kerala, Odisha, Tamil Nadu
Barley	All except Haryana, Madhya Pradesh, Punjab, Rajasthan, Uttar Pradesh
Maize	Chhattisgarh, Jharkhand, Kerala, Odisha, Punjab, West Bengal
Sorghum	Chhattisgarh, Kerala, Odisha, Tamil Nadu, West Bengal
Soybean	All except Madhya Pradesh, Maharashtra, Karnataka, Rajasthan
Finger Millet	Haryana, Kerala, Punjab, Rajasthan, Uttar Pradesh
Pearl Millet	Chhattisgarh, Jharkhand, Kerala, Odisha, Punjab, West Bengal

J Minimum Support Prices and Farmer Prices

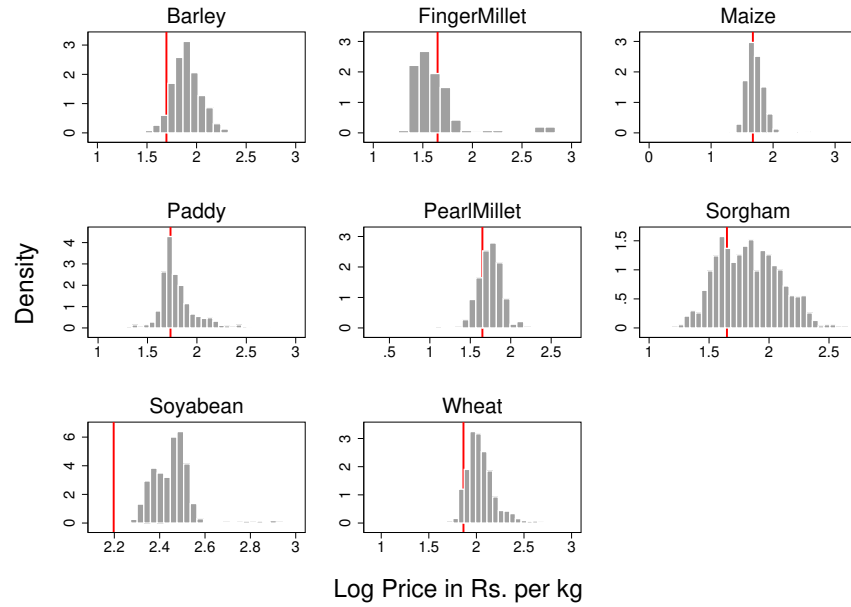


Figure J.1: Distribution of Prices and MSPs in Crop Year 2004–05

Notes: This figure plots the distribution of log prices across markets for the sample crops in the 2004–05 crop year. Vertical lines show minimum support prices (MSPs).

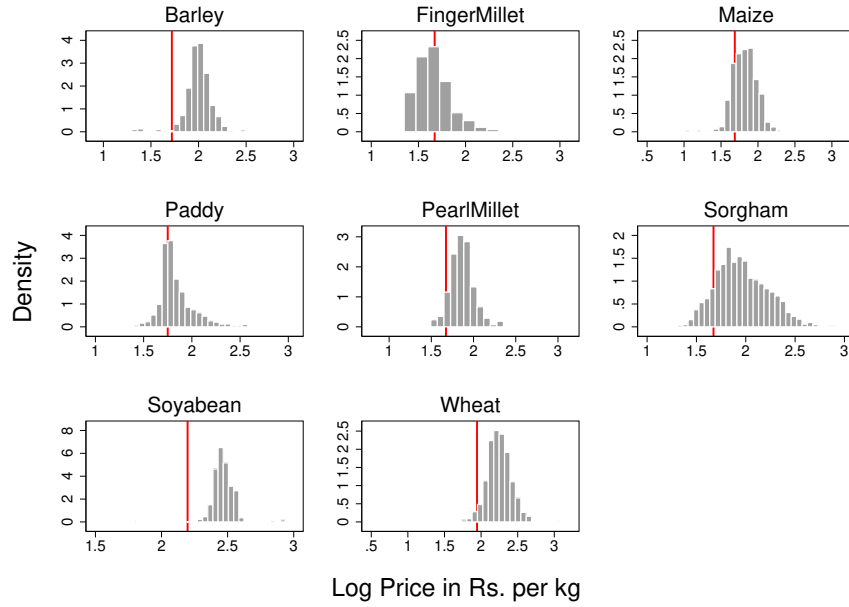


Figure J.2: Distribution of Prices and MSPs in Crop Year 2005–06

Notes: This figure plots the distribution of log prices across markets for the sample crops in the 2005–06 crop year. Vertical lines show minimum support prices (MSPs).

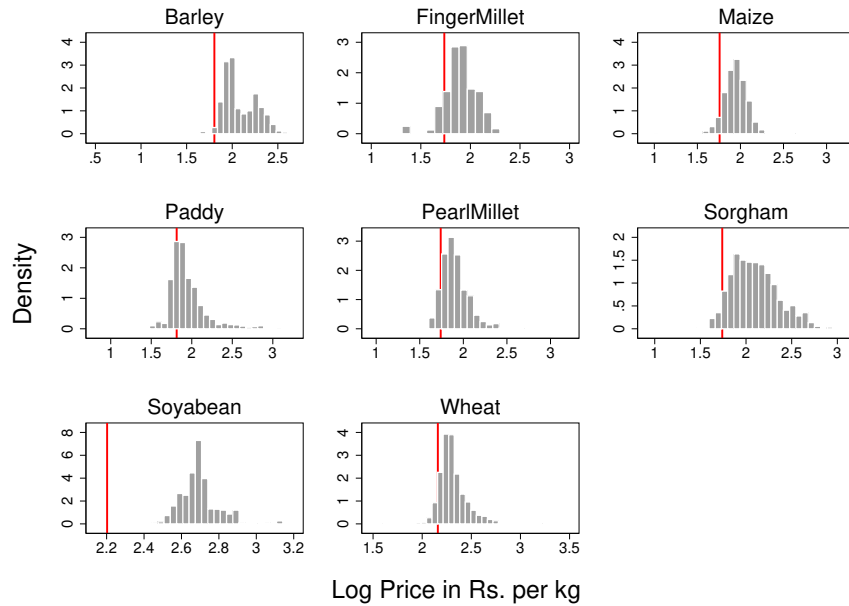


Figure J.3: Distribution of Prices and MSPs in Crop Year 2006–07

Notes: This figure plots the distribution of log prices across markets for the sample crops in the 2006–07 crop year. Vertical lines show minimum support prices (MSPs).

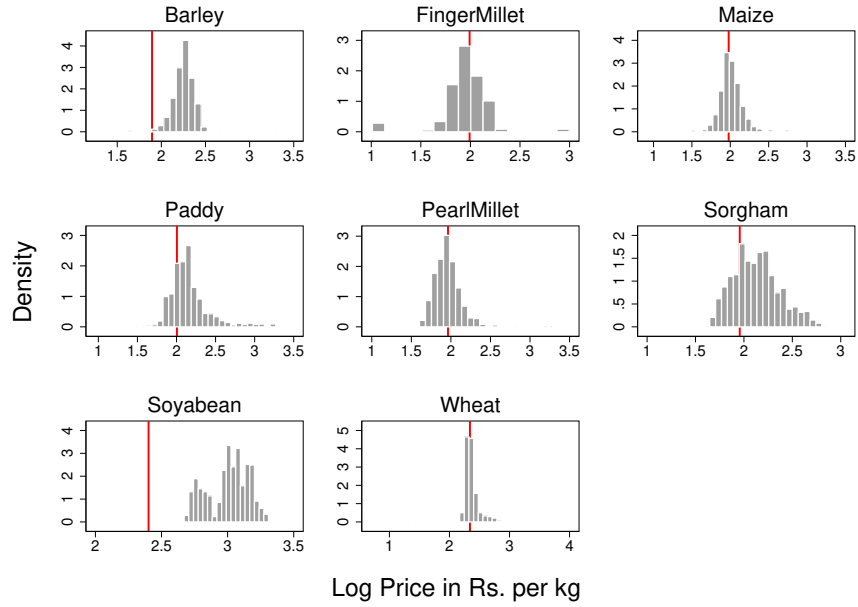


Figure J.4: Distribution of Prices and MSPs in Crop Year 2007–08

Notes: This figure plots the distribution of log prices across markets for the sample crops in the 2007–08 crop year. Vertical lines show minimum support prices (MSPs).

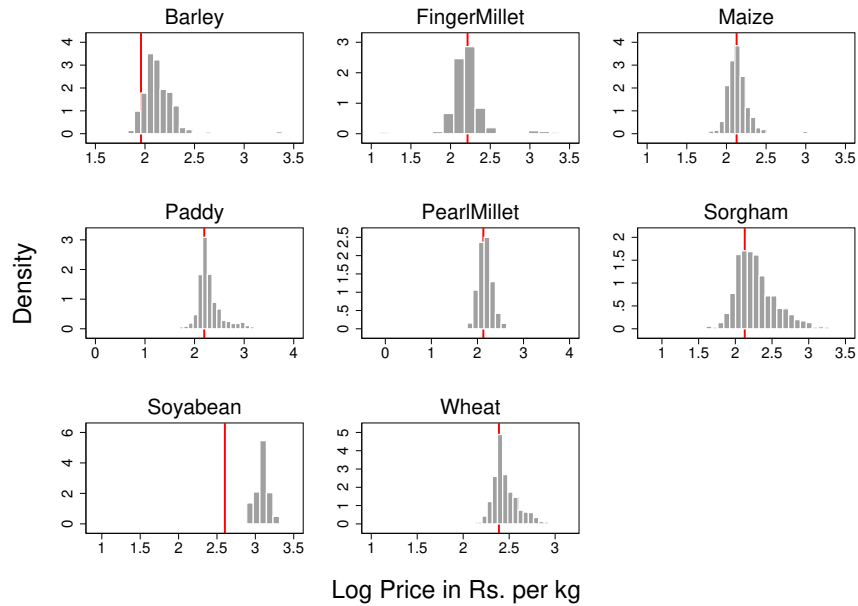


Figure J.5: Distribution of Prices and MSPs in Crop Year 2008–09

Notes: This figure plots the distribution of log prices across markets for the sample crops in the 2008–09 crop year. Vertical lines show minimum support prices (MSPs).

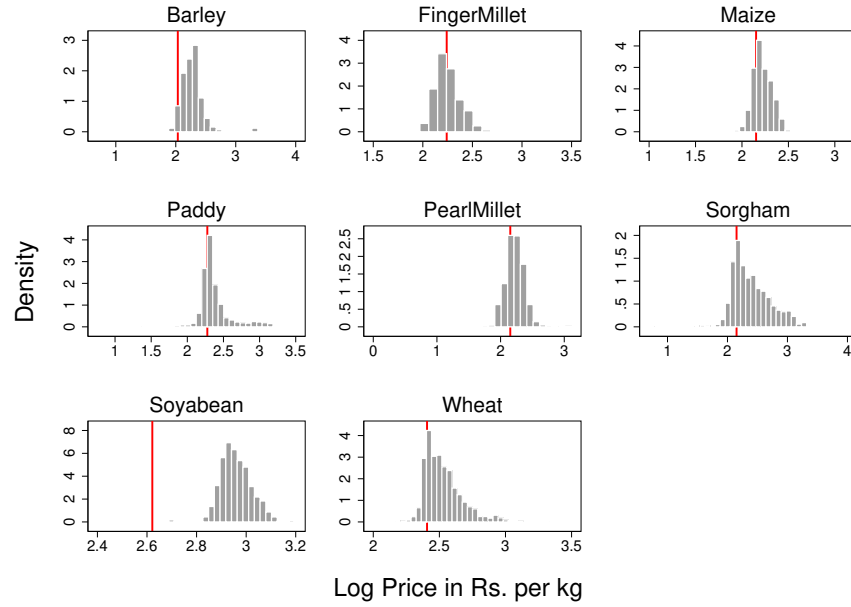


Figure J.6: Distribution of Prices and MSPs in Crop Year 2009–10

Notes: This figure plots the distribution of log prices across markets for the sample crops in the 2009–10 crop year. Vertical lines show minimum support prices (MSPs).

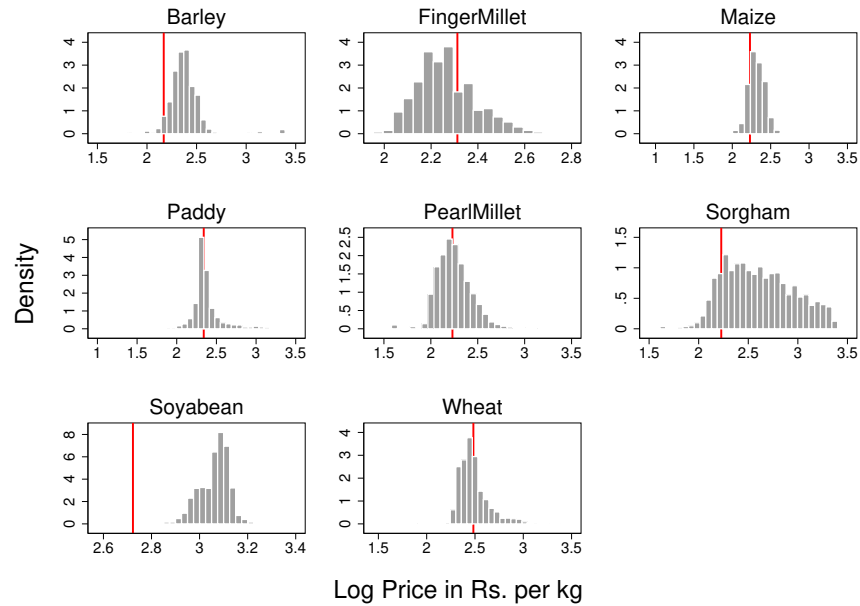


Figure J.7: Distribution of Prices and MSPs in Crop Year 2010–11

Notes: This figure plots the distribution of log prices across markets for the sample crops in the 2010–11 crop year. Vertical lines show minimum support prices (MSPs).

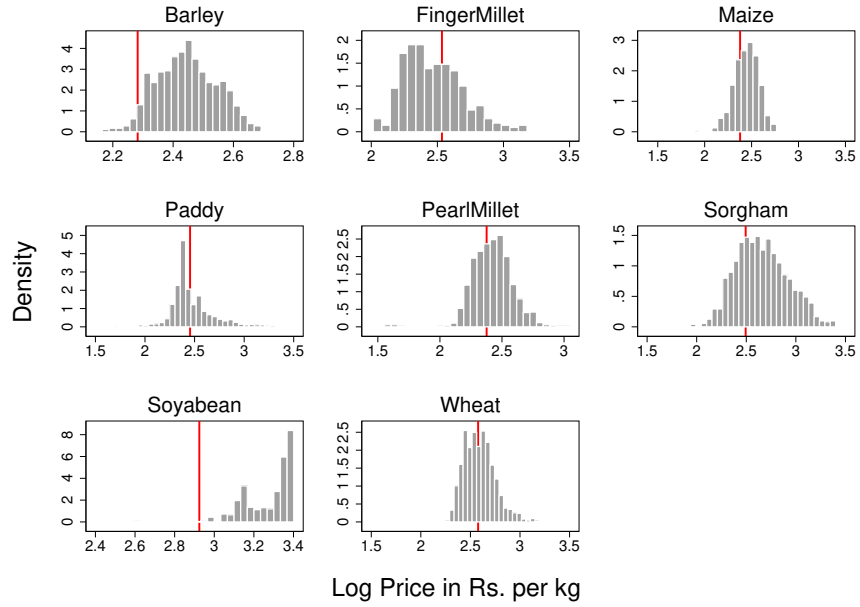


Figure J.8: Distribution of Prices and MSPs in Crop Year 2011–12

Notes: This figure plots the distribution of log prices across markets for the sample crops in the 2011–12 crop year. Vertical lines show minimum support prices (MSPs).

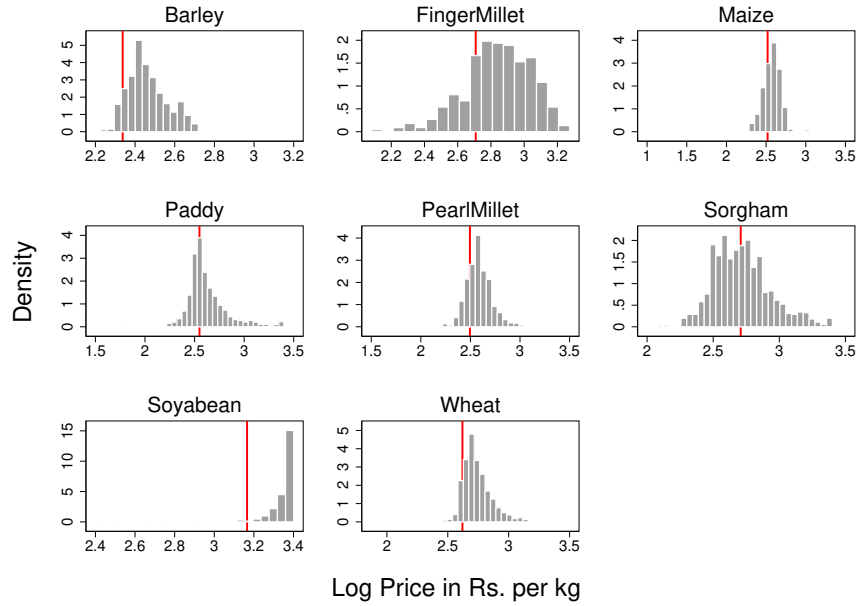


Figure J.9: Distribution of Prices and MSPs in Crop Year 2012–13

Notes: This figure plots the distribution of log prices across markets for the sample crops in the 2012–13 crop year. Vertical lines show minimum support prices (MSPs).

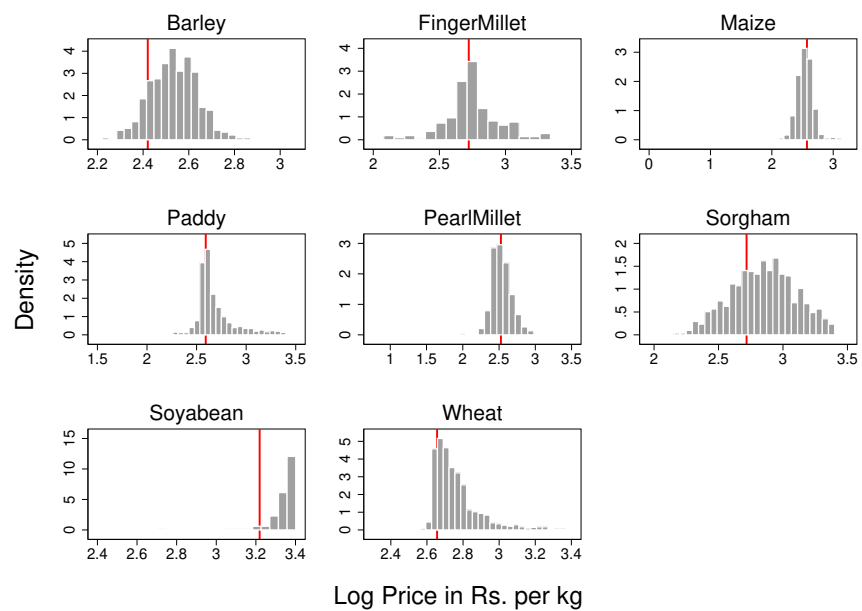


Figure J.10: Distribution of Prices and MSPs in Crop Year 2013–14

Notes: This figure plots the distribution of log prices across markets for the sample crops in the 2013–14 crop year. Vertical lines show minimum support prices (MSPs).