Price Regulation and the Adoption–Innovation Trade-off*

Felipe Berrutti† Matteo Ruzzante‡
(*Job Market Paper*)

October 14, 2025

[Click here for latest version]

Abstract

Regulating technology prices can raise adoption yet deter innovation. In India, price controls on genetically engineered (GE) cotton seeds induced this trade-off. Leveraging the policy's differential timing across states, we show that mandated price reductions accelerated adoption of GE seeds by farmers. Although seed supply kept pace, innovation stalled: fewer new varieties were introduced. Using newly assembled data from experimental field trials across India, we show that agronomic yields of new varieties fell in price-controlled states. To quantify the welfare implications of price and yield effects, we develop and estimate a structural model of demand and supply for seeds with endogenous product attributes. While the policy raised farmers' surplus, especially among the poor, ignoring innovation responses in equilibrium vastly overstates their welfare gains. We use the estimated model to assess alternative policies that better balance adoption and innovation incentives. For a given public budget, incentives for seed developers tied to the productivity of new varieties achieve the highest welfare for farmers.

* This draft benefited from comments provided by seminar participants at DIL, Northwestern, UIowa, Universidad de la República, ICABR Conference, LSE Environment Week, and Marco Fanno Alumni Workshop. Special thanks go to Lori Beaman, Vivek Bhattacharya, Douglas Gollin, Igal Hendel, Gastón Illanes, Benjamin Jones, Dean Karlan, John Loeser, Joel Mokyr, Francisco Pareschi, Dev Patel, Jacopo Ponticelli, Carl Pray, Robert Porter, Mar Reguant, and Christopher Udry for their valuable feedback and guidance. We are grateful for the time and generosity that numerous entrepreneurs, biotechnology researchers, plant breeders, and seed producers in the Indian cotton industry offered to us throughout the development of this project. We thank Matin Qaim for sharing his survey data on Indian cotton farmers. Financial support was provided by the Global Poverty Research Lab, the Weiss Fund for Research in Development Economics at the University of Chicago, the Graduate School at Northwestern University, the Northwestern Buffett Institute for Global Affairs, and the Susan Schmidt Bies Prize for Doctoral Student Research on Economics and Public Policy.

† *Analysis Group*, 650 California Street, Floor 23, San Francisco, California 94108. Email: Felipe.Berrutti@analysisgroup.com.

‡ *Northwestern University*, Economics Department, 2211 Campus Drive, Evanston, Illinois 60208. Email: matteo.ruzzante@u.northwestern.edu. Institutional job-market website: sites.northwestern.edu/mrb4682. Personal website: matteo-ruzzante.com.

1 Introduction

Under-adoption of modern technologies and under-provision of innovation drive productivity gaps in developing economies (Parente and Prescott, 1994; Basu and Weil, 1998; Acemoglu and Zilibotti, 2001). Efforts to close these gaps often focus on promoting the broad diffusion of productive technologies. However, policies aimed at increasing adoption may have unintended consequences on innovation. In this paper, we provide empirical evidence on this adoption–innovation trade-off in the context of price regulation and quantify its welfare implications. Given budget constraints and limited state capacity in low-income countries, regulating the price of *existing* technologies is an appealing policy to encourage adoption. Yet regulated prices may lower incentives to develop *superior* technologies, discouraging technology providers from investing in R&D and compromising product quality.

The tension between adoption and innovation incentives is salient in developing-country agriculture. On the demand side, smallholding farmers face multiple barriers to adopting modern technologies, such as improved seeds (Foster and Rosenzweig, 2010; Suri and Udry, 2022). Policies that lower user prices are a natural lever to raise adoption and they could yield large welfare gains. On the supply side, agricultural output growth has come from incremental, locally adapted technological advances (Evenson and Gollin, 2003; Gollin et al., 2021). While most frontier technologies originate in rich countries (e.g., genetic engineering in the US), they require adaptation to local geo-climatic conditions when deployed to non-original contexts (Pardey et al., 2010; Moscona and Sastry, 2025). In addition, seed varieties must be continually updated to cope with changes in the biological environment and maintain productivity gains over time (Olmstead and Rhode, 2002, 2008). Regulating prices may disrupt this flow of innovation and, ultimately, reduce farmers' welfare in equilibrium.

We study the equilibrium effects of price controls on genetically engineered (GE) cotton seeds in India, the largest cotton-producing nation globally and a hub for biotechnology research in South Asia. We leverage state-specific caps on the retail price of Bt cotton seeds, a group of varieties embodying a genetic technology that confers pest resistance and so reduces crop loss. The Bt technology alone is not directly useful to farmers: domestic seed firms must embed it in locally adapted varieties for India's ecologies. Price controls were enacted in 2006 in three Indian states (Andhra Pradesh, Gujarat, Maharashtra) and expanded nationwide in 2015. The unanticipated timing of the policy across states and the localized nature of innovation generate variation to identify how farmer adoption and firm innovation respond to lower prices.

Using a difference-in-differences (DiD) design and farmer-level longitudinal data, we evaluate the impact of the price caps on the demand side. To analyze supply-side responses, we bring together administrative data on seed companies and newly digitized records on GE products, including regulatory approvals and experimental field trials. Finally, we develop a structural model of demand and supply for the cotton seed market. Our model can separately recover farmers' willingness to pay for seed price and quality; profit-maximizing firms optimally set both price and quality. This allows us to: (i) decompose the overall effect of the policy on farmers' welfare, not only through price changes but also through endogenous quality adjustments; (ii) simulate counterfactual scenarios with alternative subsidy schemes.

Reduced-Form Results. The policy achieved its stated goal of making seeds more affordable, thereby accelerating the diffusion of Bt cotton. We find that farmers in states *with price controls* pay 40% less for cotton seeds than farmers in states *without controls*, closely matching the government-mandated caps. Although price controls were implemented at the state level, their impact was amplified nationally through the renegotiation of royalty fees between the Bt technology provider (i.e., Monsanto) and its licensees (i.e., downstream seed firms engaged in breeding and marketing cotton varieties that incorporate Bt). The renegotiated royalties applied uniformly to all licensees, regardless of where seeds were sold, and reduced Bt seed prices by 70% from pre-policy levels in *every* state. Because the policy bundles regulated retail prices with a nationwide cut in royalties, our cross-state reduced form likely understates its total impact on affordability and farmer welfare.

In price-controlled states, the drop in farm-gate prices increased the adoption of Bt cotton by 23 percentage points (pp), about 30% relative to observed trends in other states. Adoption delivered substantial benefits to farmers. Insecticide and labor expenditures fell by over 25% and 38%, respectively, a few years after the policy was enacted, likely a consequence of learning how to use the new technology. Total production costs decreased by 24%, generating a massive surge in cotton cultivation through the entry of new farms. The effects are persistent over time and highly robust to contemporaneous state-level shocks and cross-state spillovers.

Despite initial gains among technology end-users, price controls may have unintended consequences on technology providers, undermining private incentives to supply sufficient *quantity* or invest in product *quality*. In India's cotton seed market, pricing is the primary channel through which firms recoup R&D costs and appropriate innovation rents. Consequently, retail price caps compress the profit margins that fund the development of locally adapted varieties embodying the Bt technology.

To investigate effects on the supply side of the market, we use a proprietary database on private seed companies and find no evidence of seed shortages or market destruction attributable to price controls. The sales of cotton seeds do not decline relative to other agricultural inputs in the short term. Nevertheless, three years after the implementation of the policy, we observe an outright halt in the introduction of genetic crop technologies, coupled with a sharp fall in the number of new varieties developed by domestic cotton seed firms, despite limited market exit. As the pace of product innovation slows down, we document a decrease in varietal replacement rates at the farm level, resulting in a gradual increase in the age of seed varieties planted by cotton farmers.

To study innovation incentives, we take advantage of a key feature of our setting: seed innovation is incremental and highly localized. We compile a novel dataset on agronomic performance at the product level, based on experimental field trials from third-party regulatory testing. This allows us to observe over 600 cotton seed varieties tested in different states, before and after the policy, and so measure changes in innovation output over space and time. We obtain two main findings. First, we exploit the repeated evaluation of identical varieties in identical field stations *across time* and provide an empirical test of the "Red Queen hypothesis".[1] We find that cotton varieties experience rapid productivity decays, losing 6 to 7% of their yield per year. While unrelated to the policy, this naturally-occurring decay highlights the need for seed developers to regularly update existing varieties to maintain agronomic performance.

Second, we leverage the fact that varieties are tested *across multiple states* to identify the causal effect of price controls on local innovation output, comparing experimental field trials in price-controlled states versus other states in a DiD design. We find that agronomic yields of newly released Bt varieties fell by 30% in price-controlled states. The yield drop caused by the policy is sizable, especially when compared to agronomic estimates of yield gains from (i) the Bt technology alone in India (58% on average, up to 80% under high pest pressure, Qaim, 2003) and (ii) the introduction

[1] Seed varieties gradually develop increased vulnerability to pests and pathogens over time, which inevitably compromises their productivity: a stylized fact known in evolutionary biology as the "Red Queen hypothesis". This stands in stark contrast to other areas of technological innovation, such as manufacturing and industrial automation, where advancements do not experience a decline in efficiency but rather get supplanted by superior alternatives. Endogenous technological obsolescence also manifests in some spheres of health and pharmaceutical research, which deal with the prevention or mitigation of co-evolving diseases (e.g., due to antimicrobial resistance, Dubois and Gökkoca, 2025).

of a new cotton variety (16-21%, based on our data). The effects take three years to fully materialize, are not driven by differential trends preceding the policy, and remain robust to a wide array of confounding factors and empirical strategies. Qualitative interviews suggest that this yield drop stems from lower research budgets and a strategic reallocation of innovation effort in response to policy-induced profitability differentials across states. All in all, our reduced-form results reveal that price caps curtailed the innovation pipeline of seed firms, leading to a distortion in the development of new, location-specific varieties over the long run.

Structural Model and Estimation. We develop a structural model of demand and supply for cotton seeds in order to estimate counterfactual scenarios and quantify the welfare implications of price and quality responses altogether. Our model allows for rich consumer preference heterogeneity and, importantly for our empirical application, two endogenous product attributes. On the demand side, farmers optimally choose between brands of cotton seeds (and outside option crops) by considering their retail *price* and expected *yield*. On the supply side, oligopolistic firms sell differentiated products, endogenously setting *both* prices *and* yields to maximize profits. Production costs depend on yields, creating a trade-off between meeting farmer demand for high-yielding varieties and the costs required to develop and commercialize them.

We estimate demand by combining unique, nationally-representative individual choice data with aggregated statistics on market shares (à la Grieco et al., 2025) and using market structure instruments alongside an extensive set of fixed effects. Our demand estimates suggest that farmers are sensitive to both prices and yields, with average elasticities of 3.3. Smallholders are more sensitive to prices than to yields, in line with wedges between perceived and actual prices and heterogeneous returns to technology adoption, which are typical of agricultural markets in developing countries (Suri, 2011). On the supply side, we specify and estimate parametric cost functions that depend on yields, from which we recover marginal costs and implied markups. To address potential endogeneity arising from unobserved cost shocks, we instrument for yields using variation in demand shifters, i.e., exogenous shocks that affect farmers' demand but not firms' costs. Finally, we use the estimated model primitives to find equilibrium outcomes under a set of counterfactual scenarios.

Welfare and Counterfactuals. Compared to a scenario with no price regulation, the observed policy increases farmer surplus by making Bt products more affordable and thus broadening their adoption. Price controls disproportionately benefit the poor, who are more price-sensitive and less likely to adopt Bt in the absence of the policy. Nevertheless, consistent with our reduced-form evidence, the endogenous quality adjustments undertaken by firms are quantitatively important: the decline in product yields under price regulation offsets 31% of the welfare gains for farmers, relative to a naïve benchmark where firms do not respond and keep yields fixed.

The positive impact of the policy is a direct result of the large magnitude of the price drops. These were attainable since the policy effectively induced a renegotiation of Bt royalties (to compensate Monsanto for its genetic technology) and enabled local firms to keep operating in the market. In light of this, we further decompose the price change into two components: one attributable to lower royalties on Bt, and the other to the retail price ceiling. Our structural estimates imply that two-thirds of the observed price reductions are explained by lower royalties. Motivated by this finding, we allow for broader responses to the diminished returns on upstream innovation. Since genetic crop technologies take longer to become obsolete than hybrid varieties, we couple the drop in royalties with estimates on the future yield losses arising from the observed halt in genetic advances. Yield loss estimates, informed by India-specific agronomy and entomology literature, are large enough to reverse the policy's initial benefits, ultimately turning its net welfare impact on farmers negative.

We design and assess alternative policies that seek to mitigate the trade-off between product affordability and technological innovation. We compare price caps to more commonly used input subsidies to farmers and innovation subsidies to firms. The key difference is that, in these counterfactuals, the government bears the fiscal burden of sustaining adoption, rather than shifting it to the private sector. To replicate the equilibrium prices achieved under regulation, the government would have needed to implement a 55% linear subsidy. The high fiscal cost of such a policy may help explain why budget-constrained state governments in India opted for price caps instead. Allocating the same budget to seed firms as a performance-based grant, proportional to realized yields, results in higher prices for farmers but delivers dramatic yield gains and greater welfare than the farm subsidy. As the channels of welfare gains differ, so do their distributional incidence. Farm subsidies raise surplus through lower prices (while preventing yield losses); firm grants do so through quality upgrades (that come with higher prices). Our counterfactuals imply that, at equal levels of aggregate farmer welfare, the poorest are better off under subsidies than under innovation grants. Hence, even if fiscally costly policies can balance adoption and innovation incentives *in the aggregate*, the planner's choice entails an equity–efficiency trade-off *across the distribution* of farmers' wealth.

Related Literature. Our paper contributes to the long-standing literature on technological adoption and productivity gaps in developing-country agriculture (reviewed by, e.g., Jack, 2013, Suri and Udry, 2022). We study the introduction of a frontier technology in cotton, a leading cash crop suitable for cultivation in several developing countries due to its drought-tolerant nature. Unlike the Green Revolution for food staples, advances in cotton seed technology are driven by for-profit innovation by private firms. In the context we study, technology uptake is found to be highly responsive to a reduction in input prices, motivating large gains from government intervention.[2] However, farmers' welfare gains from increased adoption are ultimately shaped by firms' endogenous innovation responses. While the literature has focused on *demand-side* factor-price distortions and misallocation in agriculture,[3] our findings point to a complementary explanation: productivity gaps may persist not only because farmers fail to adopt available technologies, but also because *supply-side* market distortions curb firms' incentives to innovate and provide improved technologies.

We also contribute to work on transgenic crops in developing economies and, more specifically, to the ongoing debate on Bt cotton in India. We focus on reallocation *within agriculture* induced by input price changes, rather than on reallocation – of labor (Bustos et al., 2016, 2023) and capital (Bustos et al., 2020) – *across sectors* associated with input-embodied genetic technologies. We highlight how affordability can shape technological diffusion, the spatial distribution of transgenic crops, and their consequent impacts on farming and productivity.[4] Our results are consistent with heterogeneous impacts of Bt across regions and over time in India (e.g., Kathage and Qaim, 2012; Plewis, 2019; Kranthi and Stone, 2020).

Relatedly, there is growing interest in technological "appropriateness" and the ambiguous consequences of applying frontier innovations developed in (and for) rich

---

2 This is often not the case. Macours (2019) reviews experimental studies of input subsidy programs in developing countries and finds that the adoption of yield-enhancing technologies is low even when heavily subsidized, unveiling limited demand among a large share of smallholding farmers. Evidence on the effects of input price controls in agriculture, as an alternative to input subsidies, is scant.

3 See, e.g., Adamopoulos and Restuccia (2014), Shenoy (2017), Gollin and Udry (2021), Adamopoulos et al. (2022), Chen et al. (2023); reviewed by Ghatak and Mookherjee (2025). On intermediate inputs, specifically, see Restuccia et al. (2008) and Donovan (2021); on India, see Foster and Rosenzweig (2022), Chakraborty et al. (2025), and Bolhuis et al. (forthcoming).

4 Hansen and Wingender (2023) show that the cultivation of GE varieties raised agricultural yields globally, with the largest effects concentrated on cotton and warmer climates, while keeping harvested area constant. Qaim and Zilberman (2003) note that heightened pest pressure and low crop protection in developing countries and tropical regions, such as South Asia and sub-Saharan Africa, imply higher yield advantages from GE cotton than in high-income countries and temperate zones.

countries to lower-income settings (Akerman et al., 2025; Lerner et al., 2025; Moscona and Sastry, 2025). We zoom in on the microeconomic actors, who effectively adapt a global technology to suit local contexts. These actors encompass domestic seed firms, whose core task is crafting locally adapted varieties, often using genetic technologies licensed from multinational companies. Understanding how the market structure and profit incentives of these intermediary firms affect technological progress in developing countries is a crucial issue for the design of agricultural and innovation policies, which has yet to be systematically investigated.[5] At the same time, optimal policy hinges on public preferences for redistribution: prioritizing access for the poorest farmers versus quality upgrades that accelerate productivity growth.

Our results are relevant to a broader literature on the welfare implications of price and IP regulation. Prior empirical work has explored the consequences of price controls and deregulation policies in technological markets, such as pharmaceuticals, focusing on access, production costs, industry entry, R&D, introduction of new products, and advertising.[6,7] We contribute on two fronts: causal identification and direct measurement of innovation responses. Identification leverages unanticipated changes in price regulation across Indian states, in a setting where innovation is incremental and highly localized. Measurement is enabled by physical yields in experimental field stations. The findings from both our reduced-form and structural analysis illustrate the fundamental trade-off between technological adoption and innovation. Ignoring the re-optimization undertaken by technology providers in response to price regulation would lead to a significant overestimation of long-term welfare impacts on technology end-users. This is particularly relevant to the agricultural sector, where seed technologies are not universally transferable nor readily imported off the shelf.

To evaluate welfare impacts and policy counterfactuals, we adapt structural tools from empirical industrial organization. Our approach builds on Ciliberto et al. (2019),

[5] The nascent literature on adaptation responses to climate change in developing countries has mostly focused on farmers' decisions, such as adopting existing technologies or changing agricultural practices (Emerick et al., 2016; Glennester and Suri, 2018; Kala, 2019; Aker and Jack, 2023; Kondylis et al., 2024; Lane, 2024; Patel, 2025; reviewed by Carleton et al., 2024). By contrast, our results shed light on the decisive role of the private sector in developing new technologies for climate adaptation: we show that negative profit shocks can hamper innovation and erode product quality in a developing-country agricultural technology market, such as Indian cotton.

[6] A non-exhaustive list of papers on these outcomes includes Kyle (2007), Filson (2012), Cockburn et al. (2016), Dubois and Lasio (2018), Dubois et al. (2022), Maini and Pammolli (2023), Ji and Rogers (2024), Hristakeva et al. (2025); on India, see Chaudhuri et al. (2006), Dean (2023), Gupta and Cao (2024).

[7] Motivated by seminal models of endogenous growth, where profit incentives to innovate drive long-run technological progress (Romer, 1990; Grossman and Helpman, 1993; Howitt and Aghion, 1998), a concurrent strand of the literature has attempted to estimate the elasticity of innovation to market size, with a particular focus on the pharmaceutical industry (Acemoglu and Linn, 2004; Finkelstein, 2004; Blume-Kohout and Sood, 2013; Dubois et al., 2015; Myers and Pauly, 2019).

who propose a discrete-choice model of seed demand for GE crop varieties in the US. We depart in three ways: (i) we analyze a developing-country market; (ii) we allow for heterogeneity in farmers' willingness to pay for seed price and quality (measured by yields); (iii) we incorporate the supply side by explicitly modeling firms' joint decisions over price and quality in a price-controlled environment.[8]

Outline. Our paper is organized as follows. Section 2 describes the context and policy we study, while Section 3 details the data sources we use in the empirical analysis. Section 4 and 5 present reduced-form and structural evidence, respectively, on both sides of the market. Section 6 assesses the welfare implications of observed and alternative policies. Section 7 concludes.

2 Setting

2.1 Cotton Agriculture in India

We study the introduction of a technological innovation into cotton farming in India. India is the world's largest cotton producer and accounts for 20% of global output. Cotton is an important crop in Indian agriculture: it directly employs over 6 million farmers and indirectly involves 45 million people in related activities, such as fiber processing and textile manufacturing (USDA, 2024). Cotton is planted in the North, Central, and South zones of India. The northern zone cultivates cotton using irrigation and short-duration varieties, whereas the central and southern zones typically grow rain-fed cotton, under different pest environments, disease risks, and crop rotation systems. This wide spatial heterogeneity underscores the importance of localized innovation and fine-tuned adaptation to India's distinct agro-ecological requirements.

The productivity of cotton agriculture in India continues to lag far behind the rest of the world (Appendix Figure A1). The average cotton yield is around 1,270 kilograms per hectare, slightly over half of the global average and substantially less than the world's leader China (6,635) or the United States (2,840). The drivers of low yields are a subject of debate, with explanations ranging from the small size of farming operations (the average size of cotton holdings in India is 1.5 hectares), under-adoption

[8] Our empirical framework is inspired by an expanding literature on the importance of allowing for endogenous product attributes and quality choice in equilibrium models (Mazzeo, 2002; Draganska et al., 2009; Wollmann, 2018; Crawford et al., 2019; Fan and Yang, 2020; Barahona et al., 2023; Atal et al., 2025). In particular, our parametrization of the marginal and fixed cost functions follows the approach of Fan (2013) and Barwick et al. (2024).

of modern inputs, irrigation, and machinery, high row space between cotton plants for manual pest control, and lack of innovation by seed-producing companies (Blaise and Kranthi, 2019). In February 2024, a *Parliamentary Panel Report on the Cotton Sector* pointed to "outdated Bt seed technology, whitefly and pink bollworm infestation" as main culprits and noted that "the country is in dire need of varieties of cotton seeds/plants that are adaptive/suitable for our soil and climatic conditions".

Seed Industry: Market and Cost Structures. Crop varietal development as well as seed production and distribution are done by both the public and private sector. The public sector operates mainly through agricultural universities that develop new varieties and government agencies that multiply and distribute them to farmers. Over time, however, the private sector has come to dominate the market: according to Spielman et al. (2014), it accounts for 76% of seed sales in India, with an even larger share among cash crops and hybrid varieties. Cotton, in particular, is almost exclusively supplied by private companies.[9] Unlike open-pollinated varieties, hybrids have biological properties that are akin to "built-in" IP. This is because seeds saved from hybrid crops do not exhibit the same vigor as the parent generation, preventing farmers from recycling them. Such a feature enables firms to recoup their investments in varietal development and likely explains the growth in private-sector involvement since the 1980s (Murugkar et al., 2006).

Cotton seed firms in India are mostly local, family-owned, and medium-sized companies. The market is oligopolistic: in 2014, the largest firm held less than 15% of the market, while the top 3-firm and 5-firm concentration ratios were 42% and 62%, respectively (Appendix Table E1). Since importing seeds for major crops like cotton is not permitted in India, these firms manage vertically integrated operations: in-house R&D for hybrid development, seed production via contracted growers, post-harvest processing and quality controls at owned plants. Multinational biotech firms participate in the Indian seed industry as well, mostly by licensing biotechnologies, such as GE traits, to domestic seed firms, rather than engaging directly in seed production and marketing. These traits are a key input for producing seeds, acting as a productivity-enhancing technology that is embodied in the final product.

The seed industry is the most research-intensive segment of Indian agriculture and leads private R&D investment in the sector. Although comprehensive data on cotton-specific R&D do not exist, research intensity – proxied by R&D expenditure as a

[9] For instance, in the sample of cotton farmers described below and employed in the reduced-form analysis, 98% of the respondents purchase officially labeled F1 hybrid seeds from an agro-input dealer.

percentage of sales – was estimated at 7% for the overall Indian seed industry in 2009, far exceeding that of fertilizers (0.1%), pesticides (1.1%), and agricultural machinery (1.2%) (Pray and Nagarajan, 2012).[10] Once hybrids are developed, seed production is outsourced to contract farmers, who are provided with parent seeds and tasked with performing hybridization. According to internal cost data we obtained from two market leaders, contract farmers are paid a fixed procurement rate of 35-40% of the final price of a seed packet. Processed and packaged seeds are then sold through a network of agrodealers, who capture a retail margin of 10-15%. Appendix Table A1 provides an itemized breakdown of the cost structure of these two firms.

Downstream and Upstream Innovation: Hybrid Breeding and Genetic Engineering. Domestic seed companies innovate through breeding, i.e., crossing different hybrid varieties to enhance the characteristics of the offspring. As we further detail for the case of cotton in Appendix B, seed development is highly geography-specific, a natural consequence of the diverse growing conditions and ecological environments across India. Most firms run a portfolio of multiple breeding programs, each targeted to distinct agro-ecological segments. This innovation process is time-consuming and requires substantial investment: developing a new variety can take over five years, including the process of selecting parent lines, cross-pollinating plants, testing in field trials, and undergoing regulatory approvals.

Private investment in cotton seed development greatly expanded in the early 2000s, mostly driven by varietal improvement to incorporate the Bt (*Bacillus thuringiensis*) trait into local varieties of cotton. Bt cotton refers to a set of insect-resistant cotton varieties. These varieties have been genetically engineered to produce a toxin that damages the midgut lining of bollworms without requiring the targeted application of pesticides. Since 2000, the Indian authorities have approved 6 genetic events and over 1,000 hybrid Bt varieties for cultivation.[11]

[10] This compares to a global average of 10.5% R&D intensity on seeds during the same year (Fuglie et al., 2011). High rates of investment may reflect both the need for local adaptation and the impact of major scientific breakthroughs in seed biotechnology (Pray and Nagarajan, 2013, 2014). During the study period, over 50 companies were actively breeding cotton in India, 37 of which incorporating Bt traits into their hybrid lines (Pray and Nagarajan, 2010). Refer to the next footnote for terminology.

[11] More descriptive statistics are provided in Section 4.4. "Genetic events", or "traits", refer to artificially induced changes in the plant's genome that result from the introduction of genetic material, e.g., conferring resistance to target pests or herbicides. "Hybrid varieties" indicate plants that are bred by crossing two different parent lines or varieties to create offspring that possess desirable traits from both parents. Many different Bt cotton varieties incorporate the same genetic event, which is transferred from the soil-dwelling bacterium Bt and allows plant cells to produce crystal insecticidal proteins. To ensure good yields, however, Bt traits must be paired with genotypes that are well-adapted to local agro-ecological environments through breeding.

The Bt technology was developed by Monsanto in the early 1990s in the US under the trade name Bollgard®. Following illegal cultivation and the spread of unapproved cultivars, Bt was formally introduced to India in 2002 through a joint venture between Monsanto and the domestic company Mahyco Seeds. The technology was granted a temporary monopoly and licensed to other seed-producing companies (33 at its peak) that, in turn, backcrossed it into their proprietary hybrid breeding lines. Licensing contracts required firms to make an initial lump-sum payment and to then pay a per-packet trait fee, which were nationally determined; they did not fix the final seed price. In addition, licensees were expected to invest in equipment, testing facilities, and greenhouses for field trials. The technology was not widely adopted until 2006, when it started to rapidly gain favor with farmers. By the early 2010s, over 90% of the cotton acreage in India was planted with Bt hybrid varieties (Choudhary and Gaur, 2015). In 2025, Bt cotton remains the only GE crop available for commercial cultivation in India.

2.2 Bt Cotton Seed Price Controls

> *«[T]he traders in cotton seed including transgenic seed are exploiting poor farmers by collecting exorbitant prices; [...] it has become imperative on the part of the State to regulate the supply, distribution and sale of cotton seeds by fixing the sale price in the interests of the farmers in the State.»*
>
> — Andhra Pradesh Act No. 29 of 2007

Concerns over the high price of Bt cotton seeds during the initial stages of technology release led some states to cap the retail price of a packet of cotton seeds. In the early 2000s, a seed packet of 450 grams of conventional cotton was sold at 450-650 Indian rupees (₹) while Bt hybrids were introduced to the Indian market at an initial price of ₹1,600-1,800, around 33-37 US dollars, or $169-190 in purchasing power parity (PPP) terms. A large fraction of this cost, namely ₹1,100, consisted of royalties paid to Monsanto for the Bt technology. There is extensive journalistic coverage of both farmers and local seed companies expressing grievances over these licensing fees, which were conceived as an unfair trade practice.[12]

[12] Other historical accounts on this period can be found in Sadashivappa and Qaim (2009), Pray and Nagarajan (2010), and Menon and Uzramma (2017). Moreover, Newell (2007) noted that Monsanto faced heightened public and political scrutiny due to its active commitment to publicize the benefits of biotechnology in India: social activist groups and legal campaigns accused the company of engaging in biopiracy, using the terminator technology, and undertaking unauthorized trials.

In 2006, as a result of a complaint filed by the state government of Andhra Pradesh, the Monopolies and Restrictive Trade Practices Commission ruled that the state government could set the price of Bt cotton seeds under the Essential Commodities Act.[13] Although Mahyco Monsanto Biotech appealed the decision, the price cap was put in place starting in the 2006 growing season. The retail price of Bt cotton seeds containing the *Bollgard 1* (BG-I) trait was capped at ₹750 per packet. The bordering state governments of Maharashtra and Gujarat followed suit, imposing identical price ceilings during the same growing season. In the state of Madhya Pradesh, price controls were announced but later withdrawn due to a legal defeat in the High Court.

Price control acts led to a national renegotiation of the Bt technology fees, which were slashed to ₹150 per packet for all licensees and all states. The price caps were updated in later years to accommodate the introduction of an updated Bt, known as *Bollgard 2* (BG-II), a technology developed by Monsanto using a double-gene construct to provide improved insect resistance. At last, in December 2015, the Cotton Seed Price (Control) Order was issued by the federal government, stipulating a national maximum retail price of cotton seeds: the only crop with such regulation in India.[14] As of 2025, this nationwide cap is still in place and is regularly updated by a ministerial committee in advance of the cotton planting season.

3 Data

3.1 Technological Demand and Agricultural Outcomes

Our main data source for studying how price regulation affects farmers' technology adoption is a longitudinal survey from Kathage and Qaim (2012).[15] Four waves were collected on a biennial basis between 2002 and 2008, tracking 533 households in 4 states, 10 districts, and 63 villages. The sampling is representative of cotton farmers in central and southern India at baseline (2002), and the survey instrument encompasses

[13] Anecdotally, Andhra Pradesh was the first state to propose the price-control policy due to the significant role played by its civil society and farmers' groups (Peschard and Randeria, 2020). Environmental activists opposed the introduction of GE crops based on biosafety and seed sovereignty concerns (e.g., Shiva et al., 1999; Sahai, 2002). In addition, Bt was accused of lying behind the phenomenon of farmer suicides by contributing to indebtedness via crop failure (e.g., Gruère et al., 2008).

[14] The rationale behind extending the price controls to the entire nation lies in the large spatial variation that state-specific caps had resulted in. Verbatim from the federal law: "fixation of sale price by multiple authorities resulted in fixation of different prices in different States and necessitated fixing of uniform prices for Bt cotton seeds across the country."

[15] This dataset has been validated and used to quantify the effect of Bt cotton on various household outcomes, including pesticide use and poisoning (Krishna and Qaim, 2012 and Kouser and Qaim, 2011, respectively), as well as on the agro-ecosystem (Veettil et al., 2017). Nevertheless, the empirical exploration of seed pricing and its economic implications remain unexplored in the existing literature.

a wide array of questions on agricultural inputs and output, with a strong emphasis on Bt-related outcomes.[16] Importantly for our empirical exercise, the sample covers four states, two of which had price controls (Andhra Pradesh and Maharashtra) and two of which did not (Karnataka and Tamil Nadu), over two pre- and two post-event periods. Thus, we use these data for *causal identification of partial equilibrium effects* among incumbent cotton farmers, excluding selection on entry into cotton farming.

Despite the detailed nature of this panel dataset, its sample size is relatively small to capture sufficient variation in seed choice and identify substitution patterns across products. To supplement this, we draw on a unique, proprietary dataset from the Chennai-based Francis Kanoi Marketing Research, known as the Cotton Crop Track. Spanning four waves – from 2002, the year of Bt introduction in India, to 2014, by which time adoption rates had surpassed 95% – these data contain a large, repeated cross-section of around 20,000 farms and 900 villages per wave. The sample is stratified by landholding and designed to be nationally representative of cotton farming in each survey year. Besides plot-level seed choice, the survey measures prices, quantities, acreage, and yields. Seed companies were reported to regularly subscribe to this dataset so as to obtain market share estimates for both their products and those of their competitors. The richness of the data and its widespread use by the industry make it particularly well-suited for the *structural estimation of a discrete choice model of seed demand*.

In order to expand our analysis to additional outcomes and to observe switching across crops – including selection into cotton – we rely on the publicly available Cost of Cultivation/Production Survey (CCS). This survey scheme reaches roughly 8,000 farmers every year and is implemented by state agricultural universities across India, under the coordination of the Directorate of Economics and Statistics in the Ministry of Agriculture and Farmers Welfare (DESMOA). Launched in 2000, the CCS is a rotating panel with full replacement every three years, allowing researchers to follow the same set of farm households over three time periods and a repeated cross-section beyond such interval.[17] Plot-level information on input prices and quantities

[16] The survey records the specific type of cotton variety planted – a level of granularity rarely found in agricultural surveys in developing countries – and its corresponding farm-gate price. If a household has two plots where it plants two different varieties, say a conventional seed of cotton and a hybrid Bt, input/output data are collected for both varieties at the household-plot level.

[17] Farmers in CCS are selected through a three-stage stratified random sampling. First, *tehsils* or sub-districts (75 every time the sample is refreshed) are allocated to the different agro-economic zones of a state in proportion to the area under cultivation; the survey only covers the principal crops in the zone concerned. Second, a single village or a cluster of neighboring hamlets around a nucleus village in the *tehsil* is sampled following the same proportionality criterion. Operational holdings provide the third and ultimate sampling unit: farms are stratified into five groups according to their total area under

as well as any other operational expenses are measured in each farming season.

Finally, we have Crop Production Statistics (CPS) on total area and output at the district level. These two aggregates are reported by DESMOA for each farming season between 1997 and 2022, allowing us to track the spatial distribution of cotton farming, including entry and exit, over our study period.

3.2 Product Sales and Technological Supply

To assess the effect of price regulation on the supply side of the Indian seed market, we turn to administrative data from multiple sources. We observe the performance of seed companies through Prowess, a proprietary database from the Centre for Monitoring Indian Economy (CMIE). The data are built from periodical reports and filings on the universe of listed companies and a larger set of unlisted companies since 1990. The main variable used in our analysis is product-by-firm-level sales.[18]

We measure innovation by seed companies through market release of new varieties and realized product quality. First, we scraped data on GE crop approvals worldwide from the International Service for the Acquisition of Agri-biotech Applications (ISAAA). Based on publicly available decision documents of each approving country, the Biosafety Clearing House of the Convention on Biological Diversity, and scholarly articles, this database specifies the genetic trait, agronomic function, technology developer, country, and year. Second, we digitized the official list of all commercially released varieties of Bt cotton in India, with information on the producer, the GE technology, and the target geographic zone (i.e., southern, central, or northern).

We hand-coded information on: (i) the permissions to conduct field trials of GE varieties, the location of such trials, and the number of varieties tested for the Indian market from the Genetic Engineering Appraisal Committee (GEAC), i.e., the authority of the Ministry of Environment, Forest and Climate Change in charge of approving large-scale trials and commercial release in India;[19] (ii) yields of hybrid cotton

cultivation and then two farms are randomly selected from each stratum, for a total of 10 farms per village, or 750 per zone (DESMOA, 2008).

[18] Out of a total of 46,976 companies, Prowess (February 2024 database vintage) contains 553 firms selling seeds, of which 215 selling cotton seeds, and 57 firms in the agricultural research and/or seed sector, of which 31 producing and marketing cotton seeds. The data do not allow to study firm entry decisions because joining Prowess may be the result of either entry, listing, or a voluntary decision of the CMIE.

[19] In practice, we have official minutes on the universe of meetings held by GEAC (downloaded from `http://geacindia.gov.in/decisions-of-GEAC-meetings.aspx`). These are organized in order to consider applications and other policy issues related to GE. The proceedings record allows one to identify the specific seed company, hybrid variety, and underlying genetic event under evaluation as

varieties developed by private companies for the Indian market from a large set of agronomic field trials, run by the Indian Council of Agricultural Research (ICAR).[20] To the best of our knowledge, this is the first comprehensive, micro-level dataset on research effort and output at the product level for any developing country.

4 Reduced-Form Evidence

4.1 Demand

Identification and Estimation. The regulatory variation, which is generated by the differential timing of the price-control policy across states, allows us to identify the effects of the policy on farmers. We do so by considering a simple DiD design with binary treatment and comparing the change in outcomes in states with formal price controls (the treatment group) versus states without price controls (the status-quo group) over the same time period. This approach recovers the causal impact of the policy, namely the average treatment effect on the treated (ATT), provided that the following three identifying assumptions hold: (i) there is no spillover from treated to status-quo states (*Stable Unit Treatment Value Assumption* or SUTVA); (ii) potential outcomes would have evolved along a similar trajectory across states in the absence of the treatment (*parallel trends*); (iii) the treatment has no effect before its actual implementation (*no anticipation*).

Our DiD captures *relative* effects by comparing states that are directly affected by price controls to those that are not. Yet the latter group may be indirectly affected through either treatment contamination (say, due to smuggling of cheaper seeds) or general equilibrium effects at the national level (say, due to changes in royalty fees or supply shortages). While we test for spillovers later in this section, *aggregate* effects are explored through the structural model in Section 5.

As an estimating equation, we use the following event-study model

$$Y_{i,t} = \alpha_{s(i)} + \alpha_t + \sum_{\tau \neq -1} \beta_\tau \cdot PriceCap_{s(i)} \cdot \mathbb{1}\{t = \tau\} + \varepsilon_{i,t} \tag{1}$$

well as the decision made by the regulatory authority.

[20] These coordinated trials were conducted by public-sector agronomists as part of a national program, known as All India Coordinated Research Project on Cotton (AICRP). The project comprises a network of 17 state agricultural universities and 22 local research centers. Trials were aimed at testing the agronomic performance of cotton hybrids and informing regulatory authorities. Therefore, we see these data as an independent, third-party evaluation, which is consistent across testing locations and not directly affected by endogenous choice by farmers. Moreover, firms had no control over the assignment of testing locations within a regulatory geographic zone, making endogenous selection into the trials very unlikely (Tandon et al., 2015).

where Y measures the outcome of interest for a household i, living in state s, in year t. $PriceCap$ is our treatment indicator, which is equal to one for states that regulated cotton seed prices in 2006, i.e., Andhra Pradesh, Gujarat, and Maharashtra, and to zero otherwise.[21] As noted in Section 2.2, the introduction of price controls and the specific level at which prices were capped were widely unanticipated. Our basic specification includes two-way fixed effects (TWFE), capturing time-invariant characteristics of each state and time-specific shocks that are common across states; whenever the data structure allows it, we augment Equation 1 with district, village, and household fixed effects.τ indicates the number of periods before or after states regulated prices. In order to estimate the model, we normalize the β coefficients with respect to a reference pre-event time period, i.e., 2005 (or $\tau = -1$). The error term, ε, allows for arbitrary intra-cluster correlation within either village or state.[22]

Besides shedding light on treatment effect dynamics over time, the event-study specification enables us to examine the existence of parallel trends in the pre-periods, i.e., before the policy comes into effect. In particular, we test whether the coefficients β_τ are statistically different from zero for each $\tau < -1$. A second potential concern for identification in our setting is that farmers from status-quo states may cross the border into treated states in order to purchase seeds at lower prices. This, in turn, is likely to bias our estimates toward zero. Therefore, using GPS coordinates, we identify villages that are located on the border with treated states (henceforth, 'spillover villages'). Such villages are then either included in the treatment group or dropped from the estimation sample, allowing us to mitigate the potential bias introduced by spillover effects.

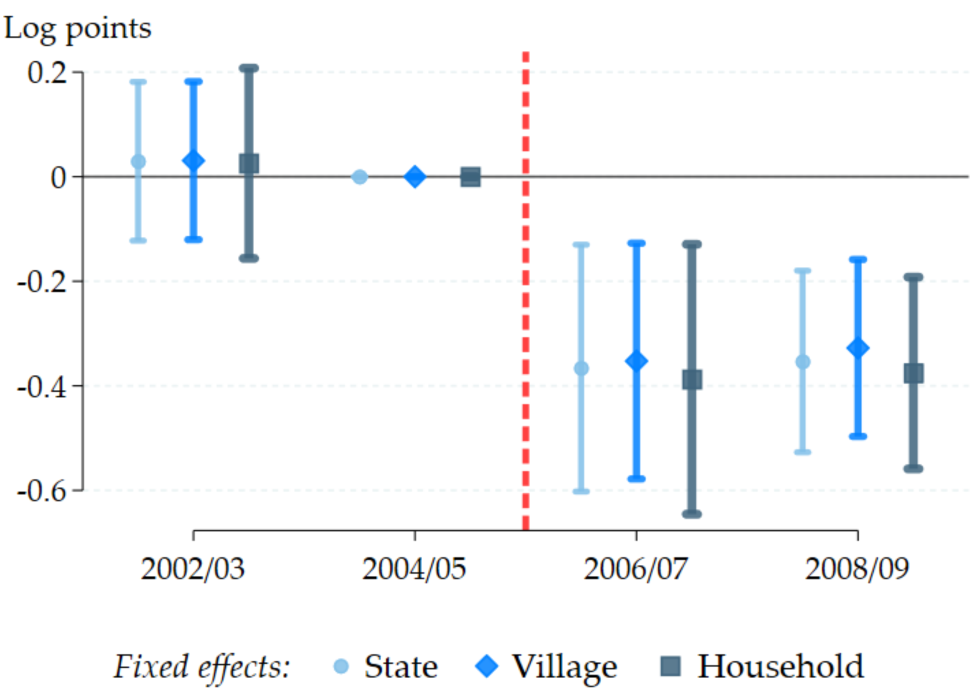
First Stage. We begin our empirical analysis by evaluating whether the policy achieved its stated goal of reducing cotton seed prices at the farm gate. We do

[21] The treatment indicator includes the area under the newly formed state of Telangana, which originated from the bifurcation of Andhra Pradesh in 2014. The *Telangana Cotton Seeds Act 2007* was adapted from the one of Andhra Pradesh, after the administrative reorganization.

[22] Within-cluster dependence in our setting arises from both sampling variability and treatment assignment. In the different datasets described in Section 3.1, households are randomly sampled from a set of *villages* (at minimum 63 in the panel survey from Kathage and Qaim, 2012), providing a natural source of uncertainty about population parameters and, therefore, grouping of observations. On the other hand, given that the treatment status is defined by state governments, a design-based approach would suggest clustering standard errors at the *state* level. However, the latter approach only accounts for between-cluster variation in treatment assignments and restricts the number of clusters to as few as four in some estimations, making clustered inference subject to important shortcomings. In Appendix C.1.1, we empirically address this trade-off by testing for the level of clustering (Cai, 2023) and find the *village* level to be the most appropriate one. For robustness, all tables in the paper present confidence intervals adjusting for clustering at either level. In Appendix C.2, we consider alternative procedures for inference with few clusters, including small-sample ad-hoc adjustments, wild bootstrap, bias-corrected variance and aggregation methods in the case of *state* clustering.

so by using self-reported data from our four-state panel survey, and we present event-study estimates in Figure 1. The pre-treatment coefficients provide reassuring evidence that prices were not growing at a differential rate across states prior to the implementation of the policy. After the 2006 event, prices dropped by 40% in states with price controls, compared to states without. The estimated coefficients are stable across fixed-effects specifications and statistically significant at the 1 percent level, regardless of the level of clustering as well as after adjusting standard errors for the small number of clusters (Appendix Table C2). Moreover, the results are robust to including 'spillover villages' in the treatment group or dropping them from the estimation sample (Appendix Table C3).

Figure 1. First-Stage Effects on Cotton Seed Prices



Notes: Event-study estimates with 95 percent confidence intervals based on least-squares regressions as in Equation 1: year and Bt seed fixed effects are included in all the models; additional fixed effects are indicated in the legend below the graph. Standard errors clustered at the village level. Unit of observation: household × plot × survey wave. Data from Kathage and Qaim (2012)'s panel survey. The outcome is expressed in natural logarithm, so that coefficients approximate percentage changes. The vertical red line signals the treatment timing. Full set of estimates in Appendix Table C3.

The short- and long-term DiD estimates, i.e., for $\tau = 0$ and $\tau = 1$, suggest that the price reduction was achieved in the first year of policy implementation and then sustained over time. This divergence becomes even more apparent when using nationally representative data over a longer time frame. Appendix Figure A2a plots the evolution of average costs paid by farmers for cotton seeds from the CCS: while prices in treated states were higher before the policy, this gap is completely eliminated in 2006. After a few years of adjustment, the trends diverge sharply between states with and without price controls, reaching a 20% gap by 2013.

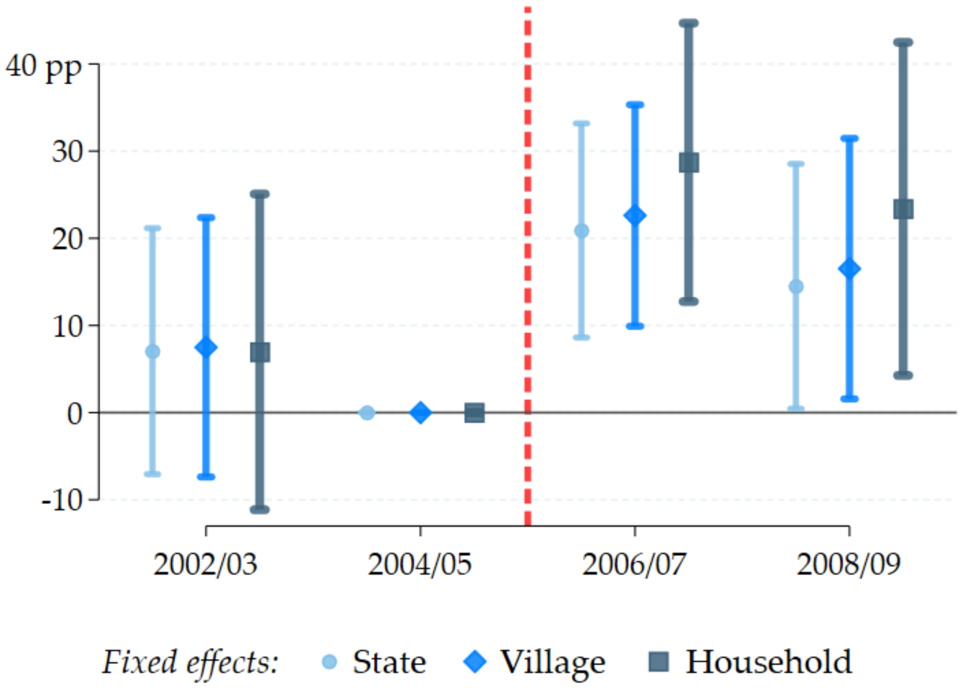
The ATT effects in price-controlled states are measured relative to changes in other states, overlooking any policy-induced price change that occurred nationwide. Appendix Figure A3 compares the distribution of Bt-cotton and conventional-cotton seed prices in the four-state panel before and after the policy. Prices declined across the board due to the national renegotiation of Bt royalty fees described in Section 2.2. However, states with price control acts experienced a much larger price reduction: almost all farmers paid below the price cap, whereas in the other states, a significant mass faced prices above it. This first-stage variation allows us to isolate the causal effect of the state-specific caps on farmers' and seed-producing firms' outcomes, which we examine in the remainder of this section. Nevertheless, the general drop in prices due to lower royalty fees remains a central aspect of the policy's overall effect. The structural model presented in the next section enables us to also account for the national decline in prices and decompose the role of each component, which is crucial for assessing welfare impacts of counterfactual policies.

Technological Adoption. The substantial reduction in cotton seed prices, resulting from the policy being studied, boosted farmers' adoption of the Bt technology. Households in price-controlled states were more likely to plant Bt seeds on their cotton plots in 2006 by 29 pp (Figure 2), i.e., almost a 50% increase compared to the sample mean in the pre-policy period and 38% over the counterfactual mean. Treated states maintained a significantly higher level of technology adoption in 2008, where the probability of adopting Bt was 23-pp higher. Again, the estimates are slightly higher when considering the potential spillover effects and survive alternative inference procedures to deal with few clusters (Appendix Tables C4 and C2).[23]

These effects are entirely driven by first-time adopters, i.e., farmers who had never used such technology in the past. Also, we observe that, following first adoption, farmers do not revert to conventional seeds, suggesting that Bt seeds were perceived as profitable by users and that the process of technology diffusion triggered by the

[23] We cannot test whether farmers reacted to the price drop in the nationally representative data because the CCS survey instrument only introduced a question on Bt in 2007, i.e., one year after the policy event. However, the post-treatment trends in Bt adoption across states in Appendix Figure A2b confirm the pattern in our main DiD estimates and align with the well-known S-shaped curve of diffusion (e.g., Griliches, 1957). Starting from similar levels in 2007, price-controlled states experienced a stark acceleration in the rate of adoption, reaching almost 100% in 2010, whereas the other states achieved an average adoption of 50% during the same period.

Figure 2. Effects on Bt Cotton Adoption



Notes: Event-study estimates with 95 percent confidence intervals based on least-squares regressions as in Equation 1: year fixed effects are included in all the models; additional fixed effects are indicated in the legend below the graph. Standard errors clustered at the village level. Unit of observation: household × plot × survey wave. Data from Kathage and Qaim (2012)'s panel survey. The outcome is an indicator variable equal to one if the household planted a Bt cotton variety in the plot and equal to zero otherwise. The vertical red line signals the treatment timing. Full set of estimates in Appendix Table C4.

policy was long-lasting. At the intensive margin, we find that price controls raise the rate of input usage: the physical quantity of seeds planted per acre increased by 24.4% in 2006 and 39.4% in 2008 among treated states (Appendix Table C5).

Substitute Inputs. The main benefit from adopting Bt seeds is that they provide enhanced protection against certain insect pests, particularly the cotton bollworm. Therefore, we expect farmers to spend less on insecticides, especially on those aimed at preventing bollworm infestations. Estimates in Appendix Table C6 reflect such adjustment, though the negative effects only appear after a learning phase: while we find no statistically significant difference in 2006, insecticide expenditures in price-controlled states decrease by 33% in 2008.[24] This late decline is driven by reduced use at the extensive margin of insecticides against the American and spotted bollworm

[24] The insufficient adjustment in pesticide use is in line with the model of selective learning proposed by Ghosh (2019) for the case of Bt cotton, where farmers tend to acquire information by observing average productivity of each input, rather than of specific input combinations. This is a more general phenomenon in agriculture, where adoption decisions are connoted by multidimensionality and interdependence due to the costs of re-optimizing inputs and practices often associated with new technologies (Laajaj and Macours, 2024).

complex, which are the explicit target of the Bt toxin (Appendix Table C7).

A complementary advantage of planting Bt cotton seeds, coupled with diminished insecticide application during crop growth, is the potential decrease in labor demand for pest control activities and cotton boll picking on the farm.[25,26] In line with this, Appendix Figure C1 shows that the policy treatment reduced the number of working hours from hired labor, which is largely carried out by casual workers, by up to 40%.[27] Again, these effects are not immediate and only manifest after a few years of technological learning. On the contrary, we do not find any effects on household labor (Appendix Table C8). The reduction in hired labor hours generates a similarly sized decrease in labor expenses following the policy (Appendix Table C9).

Production Costs. Lower seed prices, combined with fewer expenditures on insecticides and labor, result in large cost reductions for cotton farmers. Figure 3 plots event-study estimates on the cost of cotton seeds and on the total cost of cultivating cotton as measured by the national CCS.[28] The dynamic treatment effects demonstrate the persistent impact of the policy. Despite cotton seed costs making up a relatively small fraction of the overall cost of cultivation (on average, 20%), the higher use of the Bt technology is associated with an estimated decrease in the overall cost of 24%, pooling post-treatment periods (Appendix Table C10). These results are robust to using a triple DiD strategy, which considers crops other than cotton (in the same state) as additional counterfactual. This assuages concerns that the estimated effects are driven by other state-level shocks, such as contemporaneous changes in agricultural policies. We detail the alternative empirical strategy and compare the respective estimates in Appendix D.

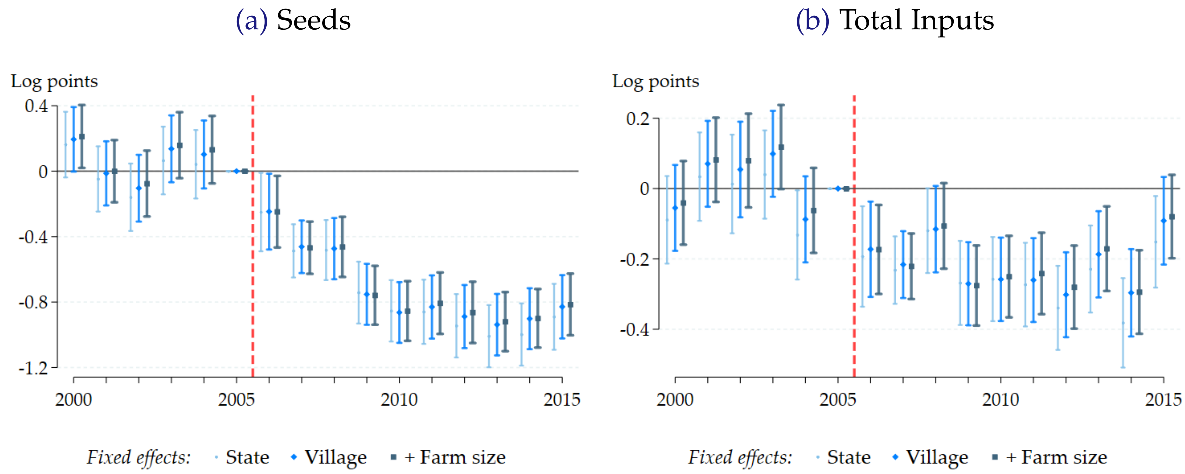
[25] Indian cotton varieties are characterized by staggered blooming patterns. While conventional cotton required as many as five to six picking operations, which are almost entirely performed by hand in India, Bt hybrid varieties were found to have more synchronous fruiting and maturity, reducing the number of pickings needed to as few as three in some cases (Mayee et al., 2004).

[26] For this part of the analysis, we are not able to continue using the four-state panel survey because it lacks comprehensive reporting of labor inputs. Therefore, we momentarily shift to the CCS data, where trends in cotton seed costs and Bt adoption are consistent with our four-state panel.

[27] This is a context with a relatively low level of mechanization: in our estimation sample, 55.7% of cotton producers own animal labor and 28.1% own some form of agricultural equipment. Cotton farming in India, especially at weeding and picking harvest time, is largely performed by female manual labor: 96.7% of cotton farms employ casual labor and 17.7% have permanent employees.

[28] We continue to use CCS for two reasons: (i) it is explicitly designed to capture the overall cost of cultivating a crop, including cotton, across Indian states – the survey instrument encompasses any source of expenditure faced by farmers, from purchasing agricultural inputs to paying for animal and human labor, from covering irrigation and machinery charges to any other rent for land; (ii) our four-state panel survey does not include exhaustive measures of labor costs, which are the main components of total costs in this context.

Figure 3. Effects on Cotton Costs per Acre



Notes: Event-study estimates with 95 percent confidence intervals based on least-squares regressions as in Equation 1: year and season fixed effects are included in all the models; additional fixed effects are indicated in the legend below the graph. Standard errors clustered at the village level. Unit of observation: household × parcel × plot × season × year. Data from the Cost of Cultivation/Production Survey. The outcome is expressed in natural logarithm, so that coefficients approximate percentage changes from the pre-treatment comparison period (i.e., 2005). The vertical red line signals the treatment timing. Estimates pooling pre- and post-treatment periods are in Appendix Table C10.

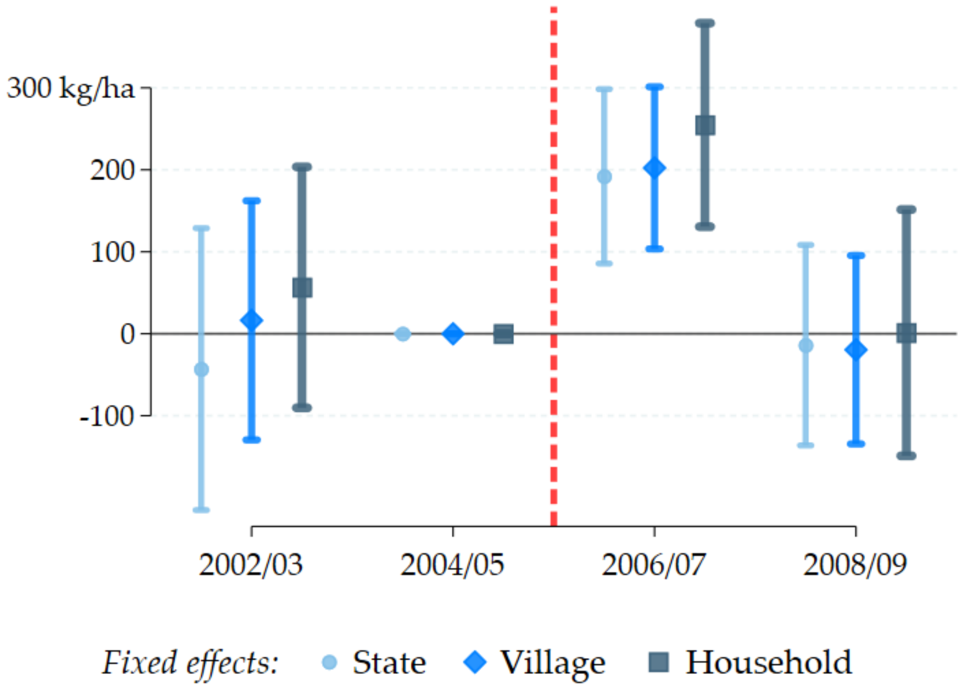
Agricultural Output. The adoption pattern in response to the government-mandated price reduction suggests that cotton farmers view the Bt technology as a profitable investment. We leverage our production data in order to test whether productivity gains materialize in the aftermath of the policy. Figure 4 shows that the policy treatment had a significant and sizable effect on cotton yields, which went up by 255 kilograms per hectare (p<0.001) in the first period, i.e., around 35% over the counterfactual mean.[29] However, this effect completely disappears in the second survey wave with no detectable difference between treated and status-quo states.

The lack of a lasting impact may be explained by a combination of factors: first, as pointed out by Herring (2013), the yield effects induced by the Bt gene are tightly connected to the extent of pest pressure observed during the agricultural season, which, in turn, depends on many environmental factors, first and foremost the weather.[30] Analogously to how irrigation water insures against the lack of rain, Bt provides a form of insurance policy for cotton farmers, where the returns from investing in Bt increase with stochastic bollworm pressure. Although not representative of our sample of villages, the crop-pest-weather database from ICAR's Central Research Institute for Dryland Agriculture indicates that bollworm pressure was lower in the 2008/2009

[29] The size of the implied treatment-on-the-treated effect of Bt adoption on yields is statistically indistinguishable from the 668 kg/ha increase estimated by Qaim and Zilberman (2003) using on-farm field trials conducted before the commercial approval of Bt.

[30] For a theoretical generalization of this argument, see Rosenzweig and Udry (2020).

Figure 4. Effects on Farm-Level Cotton Yields



Notes: Event-study estimates with 95 percent confidence intervals based on least-squares regressions as in Equation 1: year fixed effects are included in all the models; additional fixed effects are indicated in the legend below the graph. Standard errors clustered at the village level. Unit of observation: household × plot × survey wave. Data from Kathage and Qaim (2012)'s panel survey. Farm-level yields are calculated as self-reported production (in kilograms) over self-reported cultivated area (in hectares). The vertical red line signals the treatment timing. Full set of estimates in Appendix Table C11.

season than in 2006/2007 in Maharashtra, one of the price-controlled states (no data is available for other states in those years). Second, it is possible that the low cost of cultivating cotton induces farmers to expand their production to marginal – and potentially less productive – land: we provide some evidence on the latter channel in the following Section 4.2. Third, and perhaps most importantly, price controls may have distorted the incentives for private seed firms to supply adequate quantity or maintain the quality of Bt cotton varieties: these potential margins of adjustment are at the core of Section 4.3 and 4.4, respectively.

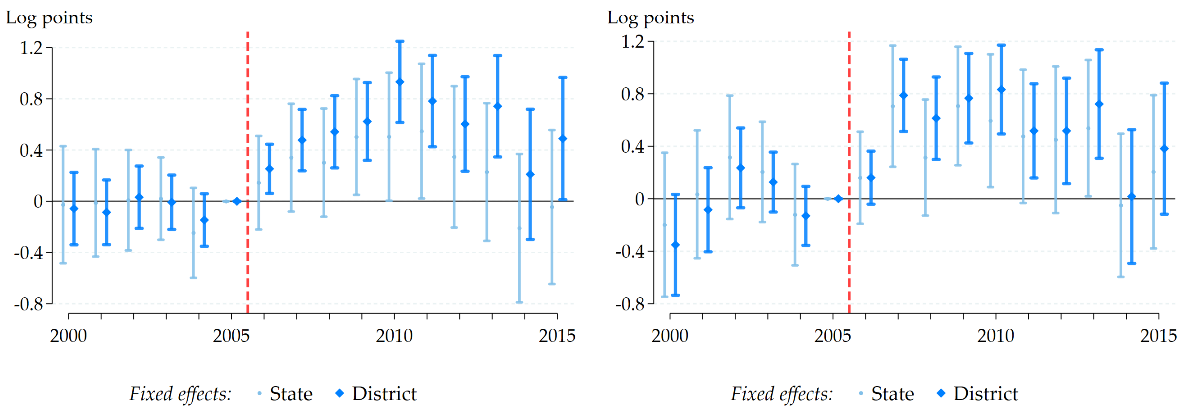
4.2 Farm Entry

The spatial differentiation in farm-gate prices of Bt seeds, generated by the statewide policies enacted in 2006, alters the costs of cultivation among incumbent cotton farmers, as shown so far. These lower costs, at the same time, are likely to affect cotton production across states (i) by attracting new farmers, who had not planted cotton in the past, and/or (ii) by changing the amount of cultivated land devoted to cotton among incumbent farmers, who were already cultivating it before the policy. In this subsection, we unpack these two mechanisms empirically.

Getting on the Cotton Bandwagon. We consider the universe of districts in India's cotton-growing states and estimate event-study models on *total* cotton acreage and production as in Equation 1. The area devoted to cotton cultivation steadily increases in price-controlled states (Figure 5a): the ATT effects grow over time, reaching a peak of 90% in 2010, i.e., five years after the onset of the policy. Cotton production also increases in treated districts, but the effect dynamics are notably different: production spikes in the immediate aftermath of price controls but then plateaus rather than tracking the continuous rise in acreage (Figure 5b).[31] While production *per acre* initially rises, this divergence leads to a long-run reversal in productivity (Appendix Table C12).

Figure 5. Effects on District-Level Outcomes

(a) Cotton Acreage (b) Cotton Production



Notes: Event-study estimates with 95 percent confidence intervals based on least-squares regressions as in Equation 1: year and season fixed effects are included in all the models; additional fixed effects are indicated in the legend below the graph. Standard errors clustered at the district level. Unit of observation: district × season × year. Data from the Crop Production Statistics. Sample: cotton-growing states. The outcome is expressed in natural logarithm, so that coefficients approximate percentage changes from the pre-treatment comparison period (i.e., 2005). The vertical red line signals the treatment timing. Full set of estimates in Appendix Table C12.

The expansion in cotton acreage induced by cheaper seeds arose disproportionately on marginal land, i.e., districts with lower suitability for Bt cotton. We provide evidence on this by interacting the district-level average of agro-climatic potential yields for cotton with our treatment indicator in Equation 1.[32] The estimated heteroge-

[31] See Appendix Figure A4a and A4b for similarly-shaped trends in cotton seed and cotton lint production, respectively, nationwide. During this period, India transitioned from being a net importer of cotton to becoming a major exporter (Appendix Figure A4c).

[32] 'Potential yields' are based on the 1981-2010 values under irrigation conditions and high input levels from the Global Agro-Ecological Zoning (GAEZ) database, version 4, compiled by the Food and

neous treatment effects in Appendix Table C13 (Columns 1-3) reveal that districts with inherently low suitability for cultivating cotton with improved seeds drive the overall expansion in acreage. On the other hand, the effects on cotton output in low-suitability districts are smaller and statistically indistinguishable from zero (Columns 4-6), implying that increased acreage did not yield proportional gains in output in low-suitability districts (Columns 7-9). We take these results as suggestive that price controls can introduce spatial misallocation in cropping patterns, reducing the productive efficiency gains from the diffusion of new technologies.[33]

The higher cotton acreage is explained by households' decision to cultivate cotton at the extensive margin, rather than by increased area under cotton cultivation at the intensive margin. First, using the CCS nationally representative sample, we find that farmers in price-controlled states are more likely to cultivate cotton in the aftermath of the policy (Appendix Figure C2). The ATT effects, which peak at 6.2pp in 2009, are large when compared to the sample mean in the pre-event period, i.e., 13.7% (namely, 21.8% in treated states and 8.8% in the comparison group), and their gradual increase over time mirrors the aggregate expansion in cotton acreage shown above. On the other hand, the positive effect on cotton acreage among this set of farmers is small and not statistically significant (Appendix Tables C14). We replicate the same extensive-margin finding among our panel baseline survey sample, where households interviewed in Wave 1 are 22.6 and 11.8 pp more likely to keep cultivating cotton in Wave 3 and 4, respectively (Appendix Table C15).[34] Again, the intensive-margin estimates on cultivated area among cotton farmers are statistically insignificant (Appendix Table C16).

4.3 Product Sales

The boost in demand, triggered by the observed reduction in Bt cotton prices, can have two possible effects on the quantity of seeds supplied in equilibrium. With price

Agriculture Organization of the United Nations (FAO) using historical climate data. We split the sample by either terciles or quartiles in order to have a discrete measure of cotton suitability.

[33] Anecdotal evidence from this time period supports this interpretation: especially in southern Andhra Pradesh and central Maharashtra, farmers converted crop acreage, originally devoted to maize, millet, sorghum, and soybeans to cotton. Rather than exploiting new black soils, which are naturally favorable for the cultivation of cotton, this cropping shift relied on "lighter" soils with much lower clay content. In line with this, Blaise and Kranthi (2019) assert that up to 30% of Indian cotton is cultivated on marginal environments.

[34] Note that, due to the sampling strategy, every household in the first wave of data collection reported a positive cotton area. In the following three waves, 4.6%, 17.6%, and 5.7% of the interviewed households, respectively, did not cultivate cotton in any of their plots. In other words, the variation captured by our ATT estimates is partially explained by the differentially higher exit of cotton farmers in states without price controls rather than by re-entry in treated states.

caps in place, firms face lower per-unit margins on Bt cotton seeds. If these margins are too low to cover production and distribution costs, seed firms may respond by cutting supply in treated states, leading to shortages. If instead margins remain positive, the surge in demand may lead firms to increase the quantity of seeds they put on the market (while adjusting on other dimensions, such as quality). Which of these forces prevails is an open empirical question, which we address in this subsection.

We start by noticing that, as described in Section 2.1, the largest firms involved in the Indian cotton seed business have *national* market coverage. This invalidates the methodology used for estimating treatment effects on demand by exploiting differential policy exposure.[35] By contrast, to establish a relevant counterfactual for the supply side, we compare the evolution of cotton seed sales (treated products) to either all other agricultural inputs or just seeds for other crops (status-quo products) and so estimate the following event-study models

$$Y_{j,p,t} = \alpha_j + \alpha_p + \alpha_t + \sum_{\tau \neq -1} \beta_\tau \cdot CottonSeed_p \cdot \mathbb{1}\{t = \tau\} + \varepsilon_{j,p,t} \tag{2}$$

where j, p, and t index a company, product, and year, respectively. Y is the monetary value of sales, while $CottonSeed$ is an indicator variable that is equal to one if the item sold p is cotton seeds and zero otherwise.[36] Our estimation sample includes any agricultural input or seed seller in the Prowess database, and we allow for intra-cluster correlation within company. The identifying assumptions are the same as in Section 4.1: no anticipation, no spillover, and parallel trends.

Appendix Figure C3 plots dynamic DiD estimates from Equation 2. The sales of cotton seeds do not grow less relative to other inputs or seeds. If anything, in the short aftermath of the policy event, they significantly increase as compared to status-quo products. Treatment effects remain positive but converge to zero from $\tau = 4$ onward.[37] Instead of fueling black markets, it seems that reduced prices on Bt cotton seeds allowed more farmers to purchase authorized – rather than illegally-bred –

[35] As an additional constraint on such analysis, we do not have firm data on sales *at the state level*. However, we should note that treated states accounted for 62% of total cotton production in the pre-treatment period. This share went up as a result of the policy.

[36] Prowess reports information both at the company and at the company-product level. This means that we can distinguish the sales of cotton seeds from the sales of other products, such as agricultural chemicals or non-cotton seeds, within the same company. We take advantage of this feature of the data in order to construct our treatment exposure variable and comparison groups.

[37] We find similar effects on physical quantities (Appendix Table C17). However, the sample employed to estimate effects on quantities is much smaller than and partially not overlapping with the one of monetary values, due to missing data. In Columns (5-6), we re-estimate the regressions on sale values on the sample with non-missing data on quantities and find short-term effects of comparable magnitude.

seeds, leading to a stark decline in the spread of unapproved cultivars after 2006 (Pray and Nagarajan, 2010).[38]

These results suggest that the maximum sales price imposed by some state governments allowed firms to experience a short-term growth in the volume of cotton seeds sold to farmers at capped, but still profitable, prices. This is consistent with our demand-side evidence on increased Bt adoption and farm entry into cotton. Although the policy guarantees positive profits for cotton seed firms, thereby avoiding market destruction, the prospective margins may still prove insufficient for firms to make significant long-term investments in new technologies. We delve into this question in the final part of our reduced-form analysis.

4.4 Technological Innovation

Investing in innovation is a firm's intertemporal decision, which depends not only on its short-term profitability, but also on the expected stream of future returns. By decreasing the net present value of new products, price regulation can distort the incentives for conducting research and developing seed varieties with superior performance. This is particularly true in contexts with low and uncertain IP protection, such as the one under study, where pricing stands as the only device for firms to recover R&D costs and earn innovation rents. However, given that product development is a lengthy process, requiring years of experimentation, plant crossing, and regulatory trials for approval, innovation responses might take some time to materialize.

We start by looking at these supply responses descriptively through an analysis of all the GE crop events and cotton varieties released to the Indian market.[39] Then, we leverage temporal and spatial variation in agronomic performance of such varieties,

[38] The initial cap of ₹750 per packet was set below the prevailing prices of illegal Bt cotton hybrids (e.g., ₹920 in Gujarat in 2003, Ramaswami et al., 2008; ₹1,190, on average, across Indian states in 2004, according to the nationally-representative data we use in this paper). The market share of Navabharat, the main seller of unapproved seeds containing the Bt gene, went from 2.2% in 2004 to 1.4% in 2008 and 0% in 2013 (Appendix Table E1). For Gujarat, the state where illegal Bt was most common, the market share plunged from 24.1% in 2004 to 3.5% in 2008.

[39] In India, GE organisms like Bt cotton must be approved by the government before being legal for commercialization (Ahuja, 2018). Until January 2009, firms that wanted to put a new variety on the market had to submit an application with data from greenhouse tests. The Review Committee on Genetic Manipulation (RCGM) would decide whether firms could carry out confined trials on fields (known as biosafety research level 1, or BRL-1). After these were approved, the application would be forwarded to the GEAC, which approved or denied requests for further large-scale trials (biosafety research level 2, or BRL-2) and, ultimately, commercialization. In 2009, the authorities streamlined the approval procedure and created the current "event-based approval mechanism". This eliminated the need for BRL-2 trials for new varieties containing one of the four genetic modifications that had been cleared by the authorities. Varieties containing new genetic events still have to undergo BRL-2 tests.

as tested in experimental field trials, in order to quantify their rate of productivity decay and identify policy impacts on product quality, as proxied by cotton lint yields.[40]

Descriptive Evidence on Product Innovation and Varietal Aging. We use administrative data from the ISAAA and GEAC containing the universe of all genetic events and hybrid varieties of Bt cotton, respectively, that were approved for commercialization in India. Given that approvals take place at the supra-state level, we are not able to continue using a DiD design to make any causal claim.

The first event of Bt cotton, *Bollgard 1*, was approved for public use in India in 2002 and was followed by the introduction of a few competing genetic traits. The approval of new events completely stopped in 2009: this is an unusual pattern for GE crops, as is evident in Appendix Figure A5. Across the world, countries that introduce GE cultivation typically keep on updating their technology over time. This mostly happens through the entry of new events developed and licensed by foreign companies. On the contrary, Monsanto India did not introduce successive improvements to its Bt technology, which were instead launched in other countries during the same period (e.g., *Bollgard 2 with Roundup Ready Flex* and *Bollgard 3*).[41]

Appendix Figure A7 documents that, following the introduction of Bt in India, there was an uptick in cotton varietal development as many local companies tried to introgress Bt into their own varieties. However, innovation in hybrid cotton came to a halt in 2011, around five years after the onset of state-wide price controls: both the number of seed varieties approved in a year and the number of companies releasing varieties decreased substantially. This is consistent with an earlier reduction in the number of applications submitted by firms (Appendix Figure A8), which is driven by fewer trials being conducted in central and southern India and a three-year pipeline for regulatory approval.[42]

As innovation slows down, we use detailed information from ICRISAT's Village Dynamics in South Asia (VDSA) panel to document that farmers use older seed

[40] "Lint yield" is defined as the quantity of cotton fiber (per hectare), which is obtained from harvested production after separating it from cotton seeds.

[41] In Appendix Figure A6, we provide additional descriptive evidence of a negative impact on the returns to genetic innovation, by plotting the evolution of royalties paid to Monsanto. After the 2006 state-wise price controls, royalties continued to grow in aggregate, yet at a much lower rate compared to the pre-policy period. Only after price regulation was expanded to the entire nation and trait fees entirely eliminated, we see royalties plummeting and reaching zero in 2021.

[42] Due to the simplification in the approval procedure, explained in Footnote 39, there exists no public information on the number of large-scale trials from 2009 onward.

varieties.[43] Appendix Figure A9 shows that, before the price control, the average number of years elapsed from market release to planting was less than one. On the other hand, starting in 2006, seed varietal age steadily rises, reaching 6 years in 2014.

The descriptive evidence provided so far does not necessarily imply that the reduction in innovation and varietal replacement is causally linked to the price-control policy. First, after a few years of intense experimentation and product proliferation, the market may have been saturated. Second, although firms introduce fewer new products, such products may come with higher quality, e.g., because of learning by doing. In the remainder of this section, we rule out both of these hypotheses by exploring productivity dynamics in experimental field trials run by the regulatory authority and unaffected by farmers' endogenous decisions.

Quantifying Productivity Decay Over Time. We leverage the fact that seed varieties are repeatedly evaluated in agronomic trials over time and construct a measure of varietal age, Age, which is equal to the number of years elapsed since a variety was tested for the first time. Therefore, we estimate the following fixed-effects regression

$$\log\left(Yields_{v,l,t}\right) = \alpha_v + \alpha_l + \rho \cdot Age_{v,t} + \varepsilon_{v,l,t} \tag{3}$$

where v indexes a seed variety, tested in location l and year t. Importantly, the inclusion of variety and location fixed effects, α_v and α_l, allows us to compare the same variety in the same field station over time. Standard errors are clustered at the variety level.

Table 1 shows that cotton seed varieties lose an average of 6-7% of lint yields every year: a phenomenon that is known in evolutionary biology as the "Red Queen hypothesis" and is mostly due to the ever-changing pest environment (Footnote 1). The regression coefficients are stable across specifications, suggesting that the estimated yield loss is not driven by endogenous differences in testing conditions.[44]

Causal Evidence on Reduced Productivity in Price-Controlled States. We exploit a second feature of our data and empirical setting: agronomic trials of seed varieties are carried out in multiple experimental stations across India, which include states

[43] We do not use this data in the reduced-form analysis because, up to 2008, it is only available for six villages in the states of Andhra Pradesh and Maharashtra (both price-controlled). However, unlike the other datasets, the survey instrument provides finer detail on agricultural inputs, allowing for a more precise measurement of specific seed varieties and, therefore, their varietal age.

[44] We further validate these results by looking at self-reported yields at the farm level from the ICRISAT-VDSA data. Using data from ICRISAT, we estimate a productivity decay of 8% per year (results not in the paper yet). However, despite our final specification controls for household and year fixed effects, these results should be taken with caution as cotton variety replacement rates may be an endogenous choice by farmers, introducing selection bias in our estimates.

Table 1. Seed Varietal Aging and Productivity Decay

	(1)	(2)	(3)	(4)	(5)	(6)
Years elapsed since first trial	-0.049**	-0.031*	-0.066**	-0.060*	-0.064**	-0.068**
	(0.022)	(0.016)	(0.031)	(0.031)	(0.031)	(0.028)
Number of observations	6,760	6,760	6,760	6,760	6,760	6,760
Number of clusters	619	619	619	619	619	619
Adjusted R-squared	0.007	0.088	0.322	0.329	0.359	0.404
Year fixed effects		✓				
Variety fixed effects			✓	✓	✓	✓
Variety zone fixed effects				✓	✓	✓
Trial state fixed effects					✓	✓
Trial location fixed effects						✓

Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: seed variety × year × trial. Data digitized from the Bt reports of the All India Coordinated Research Project on Cotton of the Indian Council of Agricultural Research. All regressions are least squares, as in Equation 3 with fixed effects (indicated in the last five rows of the table) and standard errors clustered at the seed variety level (in parentheses). The outcome, i.e., lint yields, is expressed in natural logarithm, so that coefficients approximate percentage changes.

under price regulation and states with no such policy. Since innovation is location-specific in this context (Appendix B), we consider an identification strategy analogous to the demand-side analysis in Section 4.1.[45] Returning to an event-study design, we compare the agronomic performance of new seed varieties across states using the following event study

$$\overline{Y}_{v,l} = \alpha_l + \alpha_{t_v} + \sum_{\tau \neq -1} \beta_\tau \cdot PriceCap_{s(l)} \cdot \mathbb{1}\{t_v = \tau\} + \varepsilon_{v,l} \tag{4}$$

where $\overline{Y}$ averages the lint yield of a seed variety v (produced by company j), tested in location l (in state s), across all field trials up to its official year of market release, t_v; the other variables are the same as in Equation 1.[46] We further augment the TWFEs, i.e., the location of the field trial and the year of seed variety release, with company and company-by-year fixed effects, $\alpha_{j(v)}$ and $\alpha_{j(v)} \times \alpha_{t_v}$, as well as zone-by-year fixed effects, $\alpha_{z(l)} \times \alpha_{t_v}$. These allow us to capture time-invariant characteristics of an innovator, innovator-specific trends in quality (e.g., due to learning by doing), and regional shifts in growing conditions (e.g., due to climate change), respectively.

45 Technology spillovers are inhibited by both the sharp heterogeneity in geo-climatic conditions over space, which makes plant breeding programs highly localized, and the productivity decay of seed varieties over time, which rapidly reduces the knowledge benefits between rival companies that operate in similar areas. Moreover, reproducing hybrid seeds requires access to proprietary and closely guarded parental lines, which are difficult for competing firms to reverse-engineer.

46 For robustness, we consider other moments of the yield's distribution, such as the sample minimum, median, and maximum. Also, instead of averaging different field trials over time, we take either the first trial ever conducted on a certain variety or the last one before release.

Standard errors are clustered at the company or state level.

Figure 6 plots event-study coefficients, providing evidence of parallel trends prior to the policy event and large reductions in lint yields in its aftermath.[47] In line with the length of the regulatory process for release, the effects fully arise only three years after prices are regulated, reaching an average of -218 kilograms per hectare, i.e., -30% compared to the counterfactual mean. Our results are robust to considering alternative procedures for clustered inference (Appendix Table C18), to controlling linearly for varietal age at testing (Column 7 of Appendix Table C18), to logging the outcome variable (Appendix Figure C4), to using other moments of the yield distribution (Appendix Figure C5), and to the stage of field trials (Appendix Figure C6). In fact, both varietal age at release and the probability of approval are balanced by policy exposure (Appendix Figure C7), suggesting that the effects are not driven by changes in either the timing of product release or the regulatory process. The inclusion of a seed-variety fixed effect in Column (8) of Appendix Table C18 restricts the identifying variation to varieties that are tested in at least one treated and one status-quo state: reassuringly, our results hold.

We hypothesize that firms located in (and, therefore, more likely to sell to) price-controlled states and firms with smaller market shares (whose profit margins rely on cotton to a higher extent and whose investments take more years to be recouped) have sharper reactions to the policy. The heterogeneous treatment effects estimated in Appendix Table C19 support our hypothesis: the reductions in agronomic yields among ex ante more exposed firms are larger, averaging 40% in several specifications.

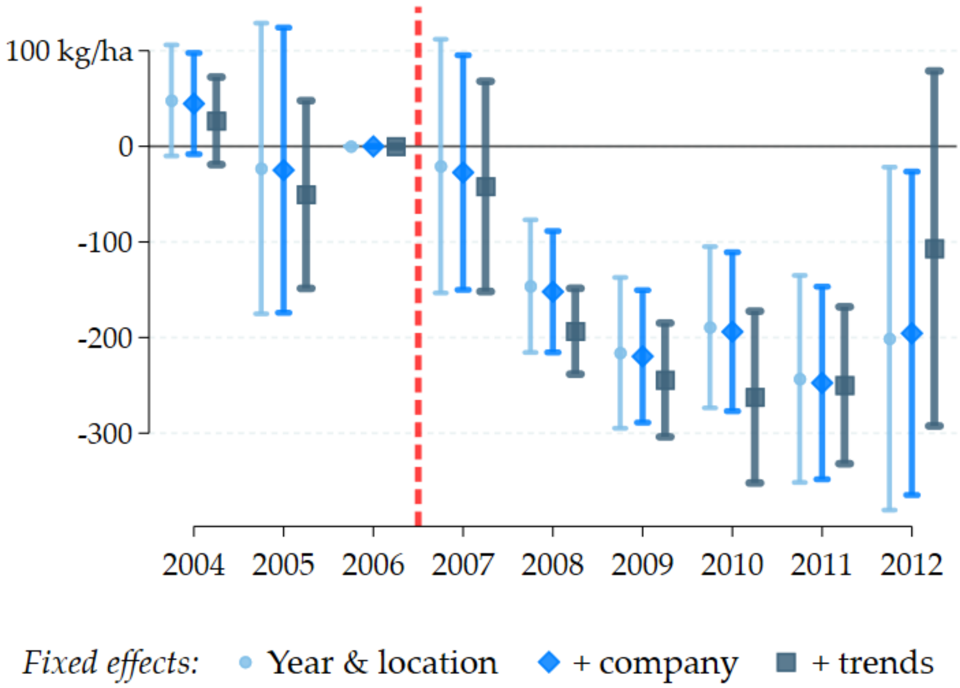
The observed effect on agronomic yields likely stems from within-firm targeting of research effort. Qualitative evidence from unstructured interviews we conducted with company representatives and plant breeders suggests that this reduction was the result of managerial decisions to diversify research portfolios toward non-regulated crops, cut down budgets and staff for cotton, disinvest from improvements in genetic inputs (e.g., through selection of new germplasm and identification of DNA molecular markers) and applied breeding techniques.[48]

To understand the magnitude of our results, we compare them to the productivity

[47] Consistent with a negative reaction to the policy, these effects are due to a decrease in lint yields in price-controlled states, rather than an increase in the comparison group (Appendix Figure A10).

[48] For instance, in a recent interview, the chief technology officer of Mahyco stated that "we have drastically reduced staff, investment and activities. [...] We reduced 50-plus staff who were working on GM crop-related matters. We have cut our funding by 70%. Whatever we had at BRL (Biosafety Research Level)-I or BRL-II, stages we have put on the shelf." ('Govt policies have shrunk Mahyco's ag-biotech research spending by 70%', *Financial Express*, March 6, 2019).

Figure 6. Effects on Agronomic-Trial Yields of Cotton Varieties



Notes: Event-study estimates with 95 percent confidence intervals based on least-squares regressions as in Equation 4: fixed effects are indicated in the legend below the graph. Standard errors clustered at the company level. Unit of observation: seed variety × company × trial location. Data digitized from the Bt reports of the All India Coordinated Research Project on Cotton of the Indian Council of Agricultural Research. The outcome is expressed in kilograms per hectare. The vertical red line signals the treatment timing. Estimates pooling pre- and post-treatment periods and using alternative inference procedures are in Appendix Table C18.

gains from two types of innovation: (i) downstream, incremental improvements from the introduction of new hybrid varieties; (ii) upstream, radical breakthroughs such as the Bt genetic modification. First, we estimate the yield advantage of a new variety (i.e., one being tested for the first time) relative to the pool of existing varieties, conditional on location and year fixed effects:

$$\log\left(Yields_{v,l,t}\right) = \alpha_l + \alpha_t + \eta \cdot \mathbb{1}\left\{Age_{v,t} = 0\right\} + \varepsilon_{v,l,t} \tag{5}$$

We find that a new hybrid delivers an average increase in yields of 16 to 21% (Appendix Table C20). Given that Indian seed companies release about one hybrid per state each year, this implies that our estimated treatment effects in Figure 6 are equivalent to two years of typical hybrid breeding effort for a state. Second, because our data only contain Bt varieties, we draw on previous agronomic assessments: Qaim (2003) reports a 58% yield advantage of Bt (80% under high pest pressure), using four years of field trials of three hybrids *with and without* the Bt gene.[49] This implies

[49] Developed by Mahyco, the three Bt hybrids (MECH-12, MECH-162 and MECH-184) were the first

that the innovation response to price regulation offsets over half of the productivity gains attributable to the Bt technological breakthrough itself. In this sense, maximizing adoption of an existing technology through capped prices can mislead policy if productivity losses from reduced downstream innovation are ignored.

5 Structural Model

The reduced-form evidence unequivocally indicates that the price-capping policy lowered retail prices but also lowered the agronomic yields of GE cotton seeds. A key empirical question, however, remains: what would prices and yields have been without the policy? Given that, as reviewed in Section 2.2, price controls triggered a national renegotiation of Bt technology fees, we cannot identify counterfactual seed prices and thus farmer choice from the data alone: our DiD estimates only recover the relative difference in prices – and yields – between states with and without price control acts. In order to go beyond this evidence and quantify the overall welfare impact of the policy on farmers, we need to specify a model of demand and supply for cotton seeds. In this section, we develop and estimate such a model and present our empirical estimates. Appendix E discusses the data construction and estimation strategy in more depth.

5.1 Demand

We model demand for seeds using a random coefficients discrete choice model of differentiated products, where farmers choose the seed that maximizes their indirect utility. We deliberately opt to model seed demand through a utility model instead of through profit maximization for three reasons.

First of all, there is widespread evidence that smallholding farmers in developing countries face substantial frictions to profit maximization due to incomplete markets for credit, insurance, labor, and land (Udry, 1999; LaFave and Thomas, 2016; Dillon and Barrett, 2017; Jones et al., 2022). Therefore, seed choice probabilities arising from assuming profit maximization may not be appropriate to think about counterfactual choices made by farmers, who face binding constraints.[50] Our utility model, instead,

to be approved for commercialization to the Indian market. Before Mahyco started sub-licensing the genetic trait to other seed companies, these hybrids were considered a key benchmark for gauging the expected effects of Bt.

50 Existing econometric tools to estimate production functions are built on the assumption that firms optimally choose intermediate inputs, based on their unobserved productivity and on the choice of other inputs. Input market frictions, which are pervasive in developing-country agriculture, are likely to distort input choice and invalidate these methods (Shenoy, 2021; de la Parra and Shenoy, 2025).

is agnostic about farm production primitives and recovers underlying preferences for product characteristics that are consistent with the data.

The second reason is that, if farmers are heterogeneously productive, correctly specifying a profit maximization model would require identifying individual productivity to predict profits under alternative seed choices. This is particularly relevant in the case of GE cotton – a technology that is associated with major adjustments in other productive inputs, such as labor and pesticides – and consistent with the results in Section 4.1. Substantial data limitations, namely due to labor inputs not being measured in our main datasets, constrain our ability to compute counterfactual profits.

Finally, our discrete choice framework, where price enters linearly into indirect utility, can be microfounded through basic assumptions on the production function of farmers, as in Ciliberto et al. (2019). These authors show that, under a seed-specific production technology with constant returns to scale and a fixed proportion of land and seeds, the individual objective of profit maximization is consistent with utility maximization.

Economic Primitives. We assume a farmer f in market m (i.e., a district-year pair) chooses the brand b that maximizes their indirect utility u, given by

$$u_{fbm} = \underbrace{\alpha \cdot p_{bm} + \gamma \cdot y_{bm} + \xi_m + \xi_{bm}}_{\text{mean utility } \equiv \delta_{bm}} + \underbrace{\mu^z_{fbm}(p_{bm}, y_{bm}; \theta^z)}_{\substack{\text{farmer-specific} \\ \text{deviation}}} + \varepsilon_{fbm} \tag{6}$$

where p is price and y is yield. ξ_m is a market fixed effect, which we include to control for heterogeneity in output and complementary input prices across time and space. ξ_{bm} is a brand-market specific unobserved utility, containing product attributes that are observable to the farmer but not to the econometrician. We allow for *observed* heterogeneity in individual preferences over product characteristics through the term $\mu^z_{fbm} \equiv z_{fm} \cdot (\theta^z_1 \cdot p_{bm} + \theta^z_2 \cdot y_{bm})$, i.e., a linear combination of price and yield with a farmer-specific variable z (measuring plot size), whose effect on utility is parameterized by the vector θ^z. We use plot size because it provides a parsimonious proxy for a farmer's scale of production, complementary technology use and cropping practices, wealth and related demographics (Appendix Table E4). ε_{fbm} is the standard *unobserved* idiosyncratic preference shock.

We specify the outside option of farmers as growing other crops that are not cotton (see Appendix E.1). The utility of not purchasing any cotton seed is assumed to be

$$u_{f0m} = \beta_1 \cdot \Pi_{0m} + \beta_2 \cdot t_m + \mu^z_{f0m}(t_m; \theta^z) + \mu^\zeta_{f0m}(\theta^\zeta) + \varepsilon_{f0m} \tag{7}$$

where we allow the utility of the outside option, $b = 0$, to vary across markets through its dependence on Π_0, the per-hectare profits of growing crops different from cotton. We include a linear time trend, t, to reflect the impact of cotton biotechnological diffusion on the profitability of rival crops throughout our sample period. We further allow this process to be heterogeneous across farmers through $\mu^z_{f0m} \equiv z_{fm} \cdot \theta^z_3 \cdot t_m$, a linear combination of the time trend with plot size. Finally, we include a random coefficient on the constant term, $\mu^\zeta_{f0m} \equiv \theta^\zeta \cdot \zeta_{fm}$, to account for *unobserved* heterogeneity in the value of outside goods across farmers, e.g., due to underlying differences in productivity of cotton versus other crops.

Identification. A key identification challenge arises when estimating the preferences for price, as seed firms may factor unobserved product attributes into their pricing strategies, generating correlation between p_{bm} and ξ_{bm}. To deal with this endogeneity, we instrument price using two variables: the number of rival varieties offered in the market (i.e., a district-year pair) and the number of own varieties offered by a brand in the market.[51] We are compelled to use market structure instrumental variables (IVs) instead of more traditional cost shifters due to the specific nature of the production process of seeds. Given that fiber envelops the seeds in cotton, producing cotton seeds is equivalent to producing cotton lint. Therefore, any cost shocks to firms are also demand shocks for farmers, rendering them invalid for identification of demand. Furthermore, the underlying variation in relative prices and products offered arising from the policy (both in markets with and without the policy, given the national drop in technology fees) contributes to identifying preferences for price. The estimates of the IV first stage are reported in Appendix Table E5.

Preferences for yield are identified under the assumption that observed yields are uncorrelated with unobserved demand shocks, conditional on market fixed effects. In other words, our identification strategy leverages the residual variation in yields within a district-year pair, i.e., our definition of a market in this setting. We argue that this is not endogenous to other product attributes given the uncertain output of cross-breeding, due to environmental factors beyond the control of innovators, and the fact that firms likely target their innovation adjustments at a higher geographic

[51] The number of rival varieties proxies for competitive pressure *across brands*. The number of own varieties aims to capture cannibalization *within brand*: when a firm introduces a high-yielding variety, it may price down parts of its portfolio to expand market share. These shifters affect price through market structure rather than through unobserved time-varying demand shocks ξ_{bm}. Two features support the exclusion restriction: (i) a rich set of fixed effects (ξ_m) absorb systematic differences in demand shifters at the market level (farmers' characteristics, labor availability, land quality, transportation infrastructure, weather, etc.) (ii) regulatory approval takes place at the zone level, so cross-district variation in variety counts is unlikely to be strategic with respect to local demand shocks.

level: our supply model assumes that firms set yields at the state level in the next subsection.

Estimation. We estimate the model using the conformant likelihood estimator with exogeneity restrictions (CLEER) proposed by Grieco et al. (2025). Individual choice data and product characteristics are drawn from Francis Kanoi Marketing Research's Cotton Crop Track. The data are nationally representative and therefore allow us to compute seed market shares, albeit with some sampling error. CLEER combines the likelihood of two mixed logit estimators for these farmer-level and product-level data with the exogeneity restrictions proposed above on product characteristics. It has been shown to be efficient and to converge at the fastest rate given the identifying variation in the data. Importantly, unlike standard BLP approaches (Berry et al., 2004), CLEER accounts for sampling variability in market shares. We review the details of the estimation in Appendix E.2.

Demand estimates are shown in Table 2. Columns (1) and (2) present the results without and with market fixed effects, while Column (3) reports our preferred specification using IVs. We obtain coefficients for price and yield that have the expected sign and are statistically significant at the 1 percent level. Our non-linear estimates on demographic interactions suggest an interesting dimension of preference heterogeneity: farmers with larger plots, who are wealthier, more likely to use modern complementary inputs and so to disproportionately benefit from high-yielding varieties, are less sensitive to prices and more sensitive to yields. Namely, a one standard deviation change in plot size implies a 10 percent change in the price coefficient. Also, as plot size increases, farmers reveal a stronger preference for inside goods (i.e., for cotton); this latter heterogeneity is likely explained by their investment and specialization in cotton production, making them less likely to switch out due to other shocks during this time period. On the other hand, the random coefficient on the constant is statistically insignificant.

The distribution of implied demand elasticities is plotted in Appendix Figure E2. On average, farmers are equally sensitive to price and to yields: elasticities are 3.26 and 3.29, respectively.

5.2 Supply

Estimating counterfactual prices and yields under alternative regulatory regimes requires fully specifying a model of conduct for seed-producing firms that optimally

Table 2. Structural Demand Estimates

	(1)	(2)	(3)
Seed price ('00 ₹)	-0.043***	-0.049***	-0.420***
	(0.011)	(0.009)	(0.031)
Physical yield ('00 kg/ha)	0.231***	0.309***	0.314***
	(0.008)	(0.007)	(0.010)
Outside option log-profits-per-ha	-0.732***	-0.495***	-0.247***
	(0.046)	(0.073)	(0.085)
Time trend	0.059***	-0.003***	-0.002***
	(0.011)	(0.000)	(0.000)
Plot size ('0 ha) × Price	0.029***	0.029***	0.026***
	(0.005)	(0.005)	(0.005)
Plot size ('0 ha) × Yield	-0.000	-0.000	0.028***
	(0.005)	(0.005)	(0.005)
Plot size ('0 ha) × Time trend	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)
Random coefficient on constant	0.033	0.033	0.006
	(0.075)	(0.075)	(0.028)
Number of micro-consumers	628,143		
Number of markets	240		
Market fixed effects		✓	✓
Instrument variables for price			✓

Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: household × plot. Data on choice, product characteristics, and plot size (for cotton farmers) from Francis Kanoi Marketing Research's Cotton Crop Track. Data on profits and plot size (for non-cotton farmers) sampled with replacement from the Cost of Cultivation/Production Survey to match shares from the Crop Production Statistics. Estimates are based on the conformant likelihood estimator with exogeneity restrictions by Grieco et al. (2025). Robust standard errors in parentheses. Markets are defined as district-year pairs. The instrumental variables for prices included in Column (3) are the number of rival hybrid varieties offered in the market and the number of own hybrid varieties offered by a brand in the market. The estimates of their first stage are in Appendix Table E5. The sign of the coefficients on the outside-option parameters (i.e., β_1 and β_2 in Equation 7) is flipped to reflect their effect on the utility of inside goods.

set their prices and qualities.

Firms' Profit Maximization. In line with our data construction assumptions, we assume that brands are single-product firms at the market level and maximize total profits, Π.[52] In an oligopoly, the problem of firm b in market m can be written as

$$\max_{\substack{p_{bm}\in\mathcal{P}_m \\ y_{bm}\in\mathbb{R}^{++}}} \Pi_{bm} := \underbrace{\Big(p_{bm} - \mathbb{1}\{b \in \mathcal{B}t\}\, r_m - mc_{bm}(y_{bm})\Big) \cdot Q_{bm}(\vec{p}_m, \vec{y}_m)}_{\text{operating profits}} - \underbrace{FC_{bm}(y_{bm})}_{\text{fixed costs}} \tag{8}$$

where r is the value of the per-packet royalty or trait fees for Bt (paid by the firm to to Monsanto), $mc(\cdot)$ is the marginal cost of producing a packet of seeds, and Q is the number of packets sold.[53] Fixed costs, $FC(\cdot)$, capture the expenses that a firm has to sustain in order to achieve a certain level of yields through the development and release of new seed varieties.[54] The key difference between markets with and without price controls is the set $\mathcal{P}_m$. In markets without price controls, $\mathcal{P}_m = \mathbb{R}^{++}$. In markets with price controls, $\mathcal{P}_m = (0, \check{P}_m]$, $\check{P}$ being the cap set by state governments.

Building on Fan (2013) and Barwick et al. (2024), we assume a linear specification of the marginal cost function, such that

$$mc_{bm}(y_{bm}) = \omega \cdot y_{bm} + \xi_b + \kappa \cdot t_{\mathbb{1}\{b\in BT\}} + \nu_{bm}^{mc} \tag{9}$$

where ω is the slope of the marginal costs with regard to yields. We include brand fixed effects ξ_b to control for time-invariant differences in productive efficiency across companies and a linear time trend specific to Bt brands to capture potential cost reductions due to learning and technological advances in breeding GE cotton. ν^{mc} is an unobserved cost shock. Fixed costs are assumed to be non-linear in yields, so that

$$\frac{\partial FC_{bm}}{\partial y_{bm}} = \phi' + \phi'' \cdot y_{bm} + \nu_{bm}^{FC} \tag{10}$$

where ϕ' is the intercept of the slope of the fixed cost function, ϕ'' is the second derivative or convexity with regard to yields, and ν^{FC} is a firm-market specific shock. The latter parameterization reflects the notion that, as yields increase, the effort required to raise them further may grow at an increasing rate.

The first order conditions (FOCs) of the unconstrained profit optimization problem

[52] As noted by Ciliberto et al. (2019) for the case of US corn and soybeans, a "product" in the seed industry is best understood as a "product line" that evolves over time, depending on the underlying germplasm and the incorporation of GE traits: key dimensions of differentiation that are valued by buyers. Appendix E.1 provides further details on how we define and aggregate products using our survey data.

[53] While licensing contracts for GE traits are typically confidential in other settings, requiring additional assumptions to estimate trait fees (Moschini and Perry, 2024), we were able to obtain them as primary data from company records. In India, trait fees for Bt cotton are set nationally, so they only vary by market m, and not by Bt brand $b \in \mathcal{B}t$. They are equal to 0 for non-Bt brands.

[54] Our model is static. In theory, firms are forward-looking and maximize the discounted sum of future profits. We abstract from dynamic considerations and take the market environment as exogenous and constant over our time horizon. Under this assumption, the profit objective in Equation 8 provides a reduced-form approximation to the discounted objective. Accordingly, $FC(\cdot)$ should be treated as the per-period cost of attaining yield y, rather than the outcome of an intertemporal R&D decision.

each firm b solves in market m are:

$$[p_{bm}] : \quad \left(p_{bm} - \mathbb{1}\{b \in \mathcal{B}t\}\, r_m - mc_{bm}(y_{bm})\right)\left(\frac{\partial Q_{bm}(\vec{p}_m)}{\partial p_{bm}}\right) + Q_{bm} = 0 \tag{11}$$

$$[y_{bm}] : \quad \underbrace{\left(p_{bm} - \mathbb{1}\{b \in \mathcal{B}t\}\, r_m - mc_{bm}(y_{bm})\right)\left(\frac{\partial Q_{bm}(\vec{y}_m)}{\partial y_{bm}}\right)}_{\text{marginal revenue gain from changing yields}} = \underbrace{\frac{\partial mc_{bm}}{\partial y_{bm}} Q_{bm} + \frac{\partial FC_{bm}}{\partial y_{bm}}}_{\text{cost-side effect}} \tag{12}$$

$$= \omega \cdot Q_{bm} + \phi' + \phi'' \cdot y_{bm} + \nu_{bm}^{FC}$$

An important trade-off arises from a firm's optimal choice of yields: farmers value yields, but yields may be costly to provide. How costly? The ultimate goal of our supply model is to quantitatively address this question.

Estimation Strategy and Results. The empirical objects of interest are ω, ϕ', and ϕ''. In markets where firms are price-controlled, we cannot distinguish between constrained and unconstrained firms when they price at the cap.[55] Therefore, we estimate our supply-side parameters in unconstrained markets. We do so through the following two-step procedure. First, we invert the Nash-Bertrand single-product firm pricing FOC given demand estimates to recover marginal costs for each brand-market pair. Second, we estimate the three supply-side parameters jointly by generalized method of moments (GMM), using our model of marginal costs (Equation 9) and the yield FOC (Equation 12) as moments.

Given that cost shocks are observable to the firm but not to the econometrician, we instrument yields with two demand shifters: plot size, z, and outside option profits per hectare, Π_0. Firms have more incentives to provide yields in markets with larger average plot sizes, since these farmers value yield relatively more. If plot sizes are uncorrelated with supply shocks, this provides a valid instrument to estimate supply. Similarly, but with an opposite sign, the profitability of outside option crops predicts farm exit from cotton and thus lower overall demand for cotton seeds (Table 2).[56] Intuitively, our estimation strategy leverages the covariation of prices and yields to

[55] This empirical challenge is reflected in the data, where a small share of seeds are priced below the ₹750 and ₹930 caps – for BG-I and BG-II cotton, respectively – in price-controlled states (Appendix Figure A3). The fact that not all prices strictly adhere to the cap implies the potential existence of unconstrained firms.

[56] In line with this argument, we find a strong first stage (Appendix Table E6). The identifying assumption for the second stage to be valid is that cost shocks are mean independent of the instruments, $W^{\text{supply}} := (z, \Pi_0)$; formally, $\mathbb{E}\left[\nu^{mc} \mid W^{\text{supply}}\right] = 0$. The timing of seed manufacturing is key to ensuring the validity of our supply instruments: seed firms set yields and incur production costs before the actual realization of farmers' input costs and output prices. Once these conditions are known, it is no more expensive to sell seeds to larger farms or when other crops are more profitable. Moreover, the set of farmers that act as "seed multipliers" is likely a selected group with larger plots: these farmers, according to our demand estimates, are less sensitive to changes in the outside option.

recover the variable cost of production. On the other hand, the observed choices of yields across markets are informative of the curvature of fixed costs of production with regard to yield.

Table 3 reports the supply estimates with and without instruments. The GMM estimates using demand shifters in Column (3) are positive and statistically significant, confirming that providing quality is costly to firms, both at the margin and as a fixed cost investment. In contrast, the OLS estimates in Columns (1-2) potentially suffer of endogeneity and appear biased toward zero, suggesting that unobserved cost shocks and yields are indeed correlated. This could be due, e.g., to negative marginal cost shocks, such as adverse weather conditions or input shortages, prompting firms to reduce expenses on breeding for cost savings. On average, increasing yields by 100 kilograms per hectare costs firms an additional ₹43 per packet of seeds.

Using the GMM estimates, Appendix Figure E4 traces the slope of the fixed cost function: while it is always costly to provide higher quality, such cost grows with the level of yield, implying convex fixed costs in yields. As firms move toward the Indian productivity frontier, it becomes more and more expensive to innovate and increase productivity further. At the average level of yields, i.e., 1,100 kilograms per hectare, increasing yield by 100 kilograms raises fixed costs by about ₹2.5 million.

As a final step, we project implied marginal costs onto yields, brand fixed effects, and the Bt time trend to recover brand-year-specific marginal costs in constrained markets. The full details of this final step can be found in Appendix E.3. Appendix Figure E3 reports our estimates of marginal costs by type of brand. We recover an average markup of 30%, which is in line with gross margins we obtained from companies' internal cost data. In addition, prices implied by our structural estimates of marginal costs are strongly correlated with observed prices in the data (coefficient of correlation of 0.90).

6 Welfare and Counterfactual Policy Analysis

The main goal of the structural analysis was to understand how the *price-cap* policy, which was enacted by some Indian state governments in 2006, changed farmer welfare compared to a benchmark scenario with *no* policy. With the model parameters of demand and supply estimated, we can now answer this question. First, we can solve for optimal prices and yields set by firms, recover counterfactual farmer choices and obtain a corresponding measure of utility. Then, we can use the model to evaluate

Table 3. Structural Supply Estimates

	(1)	(2)	(3)
Estimation method:	*OLS*	*Brand fixed effects*	*GMM*
Marginal cost slope, ω	-0.038	-0.015	0.434***
	(0.030)	(0.027)	(0.097)
Fixed cost slope intercept , ϕ'	-55,129.6	-53,535.5	-12,442.6
	(34,716.9)	(33,628.0)	(14,690.5)
Fixed cost convexity, ϕ''	11,269.5***	10,962.1***	3,359.5**
	(3,805.9)	(3,688.6)	(1,326.6)
Number of observations	226		

Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: brand × state × year. All regression specifications (in the column header) include a linear Bt-specific time trend. Robust standard errors in parentheses. In Column (3), we use a two-step GMM estimator with moment conditions from Equations 9 and 12 and a heteroskedasticity-robust weight matrix. The instrumental variables – assumed to be orthogonal to the moment conditions – are plot size, outside option log-profits-per-hectare, and their interaction. The estimates of their first stage are in Appendix Table E6. Costs are in 100 Indian rupees (₹), while yields are in 100 kilograms per hectare. Data on the inside good from Francis Kanoi Marketing Research's Cotton Crop Track. Data on profits and plot size (for non-cotton farmers) sampled with replacement from the Cost of Cultivation/Production Survey to match shares from the Crop Production Statistics.

alternative policies, such as subsidies, in terms of welfare impact and fiscal cost.

Set-up and Methodology. We start by considering a counterfactual scenario where firms are able to freely price and compare it to the observed price caps. Counterfactual prices and yields are estimated under two assumptions. The first assumption is that, without the caps, equilibrium prices and yields are those arising from the profit-maximization FOCs (Equations 11 and 12). Second, absent price controls, technology fees would not have been renegotiated between the technology provider, i.e., Monsanto, and the domestic seed firms. Therefore, we set trait fees back to ₹1,100 per packet for all markets after 2006. To solve for optimal prices and yields, we use an iterative quasi-Newton algorithm that jointly updates product attributes using FOC violations and heavily penalizes pricing above the cap in regulated markets (detailed in Appendix E.4).

The model provides a strong fit of the data under price caps. Appendix Figure E5 contrasts equilibrium outcomes from the structural model with the observed values in the data among regulated markets. For both prices and yields, the correlation is around 85%. Notably, our algorithm captures the mass of prices at the regulatory caps, ₹750 in 2009 and ₹930 in 2013. We further assess our model's performance by estimating ATT effects using structural outcomes rather than the actual data. Even if the policy shock was not used for estimating the structural model and the sampling of farmers and products differs from the one in Section 4, the model-based DiD estimates in Appendix Table E7 align well with the reduced-form ones. As the price cap is relaxed – holding royalty fees constant – we find that optimal prices increase, suggesting that firms were indeed constrained in their pricing decisions under the current policy; consequently, optimal yields also increase (Appendix Figure E6).

As a key welfare metric to assess the overall policy impact from these concomitant changes, we compute farmer surplus, FS, in markets with price controls using the logit formula,

$$FS(\theta) = \frac{1}{\alpha + \theta_p(\theta)} \log \left[1 + \sum_b \exp\left(u_b(p, y; \theta)\right)\right] \tag{13}$$

Prices and yields both affect utility and, hence, farmer surplus. The extent to which they do is determined by the heterogeneous farmers' preferences for price and yield that we have estimated in Section 5.1. The extent to which firms react to the policy and adjust product attributes in equilibrium is mediated by the structure of the market and the cost that firms face in order to provide yields from Section 5.2. Importantly, our welfare metric is measured relative to the zero-utility outside option.[57] We then integrate out over the empirical distribution of farmers' preferences to obtain aggregate farmer surplus,

$$FS = \int FS(\theta)\, dF(\theta)$$

Welfare Impact and Decomposition. Table 4 presents our welfare estimates, along with the main equilibrium outcomes, under several counterfactual scenarios. Compared to a benchmark scenario with no policy (Row 1), the observed price controls (Row 2) increase farmer surplus by about ₹17.4 billion, or ₹584 per household. This represents up to 30% of the average cost of cotton seeds and is more than double the no-policy surplus. Consistent with heterogeneity in preferences and baseline adoption, welfare gains are larger for poorer farmers (who are more price- than yield-

[57] This is an imperfect measure of welfare. As discussed above, many farmers are likely not making choices that are consistent with profit maximization, either due to behavioral biases or to effectively binding constraints. Our consumer-based metric reflects a farmer's "perceived surplus", as inferred from their revealed preferences, and is only determined by equilibrium changes within the seed market. It excludes general equilibrium effects in related input markets (e.g., Bt-induced labor and pesticide savings), which could amplify the welfare gains from cheaper seeds, nor in output markets, which could reduce them (e.g., increased cotton supply depressing output prices and altering the profitability of outside-option crops). Furthermore, we abstract from the potential health and environmental benefits from reduced pesticide use as well as concerns about losses in crop biodiversity.

sensitive and less likely to plant Bt cotton in the absence of the policy): the relative change in farmer surplus under observed price controls is always positive but declines monotonically with plot size (Appendix Figure E7a).[58]

Table 4. Welfare Estimates Under Counterfactual Policies

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Equilibrium outcomes				*Farmers surplus*		*Fiscal cost*
Scenario ↓	Average prices	Average yields	Share of cotton	Share of Bt adoption	Total	Per household	Total
1) Benchmark: no policy $\rightsquigarrow p^{\star}, r^{\text{high}}, y^{\star}$	1,532	1,129	1.6%	39%	10B	362	
2) Price controls: observed policy $\rightsquigarrow p\downarrow, r^{\text{low}}, y\downarrow$	712	1,068	14.5%	95.3%	27B	945	
Welfare decomposition							
3) Ignoring endogenous quality adjustments $\rightsquigarrow p\downarrow, r^{\text{low}}, y^{\star}$	712	1,129	20%	94.7%	35B	1,237	
4) Only considering renegotiation of royalty fees $\rightsquigarrow p\downharpoonright, r^{\text{low}}, y^{\star}$	960	1,129	8.6%	99.2%	20B	706	
Benchmarks with reaction in upstream innovation							
5) Release of herbicide-tolerant technology $\rightsquigarrow p^{\star}, r^{\text{high}}, y^{\star}\cdot 1.21$	1,532	1,332	3.7%	41.5%	14B	513	
6) Development of pink-bollworm resistance $\rightsquigarrow p^{\star}, r^{\text{high}}, y^{\star}\cdot 1.75$	1,532	1,615	20.4%	54.9%	38B	1,294	
Alternative policies (government-financed)							
7) Linear farm subsidy $\rightsquigarrow (1-\tau)\cdot p^{\text{sub}\rightarrowtail f}, r^{\text{high}}, y^{\text{sub}\rightarrowtail f}$	712	1,144	20.4%	91.2%	35B	1,217	63B
8) Partial innovation firm subsidy $\rightsquigarrow p^{\text{sub}\rightarrowtail j}, r^{\text{high}}, (1+\psi)\cdot y^{\text{sub}\rightarrowtail j}$	1,760	2,729	19.2%	88.5%	35B	1,115	36B
9) Full innovation firm subsidy	1,918	3,199	32.5%	93.5%	54B	1,847	63B

Notes: Equilibrium outcomes, p for *prices* and y for *yields*, are consistent with our structural model in Section 5 and estimated through the quasi-Newton algorithm described in Appendix E.4. Share of Bt adoption is defined as the total market share of Bt brands with respect to the total market share of the inside good (i.e., cotton). * refers to the equilibrium under the benchmark of no policy, ↓ under the observed policy of price controls, ⇂ under royalty cuts alone (i.e., without specific retail price caps), $^{\text{sub}\rightarrowtail f}$ under linear farm subsidy, and $^{\text{sub}\rightarrowtail j}$ under innovation firm subsidy. $r^{\text{high}} = 1,100$ are *royalty fees* with no price regulation, while r^{low} are those observed under regulation. Farmer surplus is based on the logit formula in Equation 13; 'Per household' is equal to 'per farmer' surplus multiplied by the average number of seed packets purchased in price-controlled markets. The government cost for the linear farm subsidy is calculated by: (i) multiplying a 54.9% subsidy rate by the market-level equilibrium price, the market-level equilibrium share of cotton, and the total number of seed packets sold; (ii) summing across price-controlled markets. The 'partial' and 'full' innovation firm subsidy are welfare-neutral and budget-neutral, respectively, compared to the linear farm subsidy. All monetary values are in nominal Indian rupees (₹).

Ignoring endogenous quality adjustments would substantially overstate farmers' welfare under the policy. The naïve impact (i.e., assuming no change in yields, Row 3) is about ₹25.3 billion, or ₹875 per household. The local innovation lost – due to the observed re-optimization undertaken by domestic private firms (Row 2) – reduces this naïve impact by 31%. Conversely, we notice that, had the quality drop been roughly proportional to the price drop (i.e., had the elasticity of quality to price been around one), the policy would have been welfare-neutral for farmers. What prevented a proportional drop?

As discussed earlier, the policy under study resulted in a bundled treatment: (i) state-level retail price caps and (ii) a nationwide renegotiation of Bt royalty fees, precipitated by the caps themselves. The model allows us to unbundle the contributions of these two components. Row 4 shows that roughly two-thirds of the total price drop – driving 41.1% of the associated farmer surplus increase – is explained by the royalty cut, with the remaining third attributable to the retail ceiling. This distinction is critical to understanding both the policy's feasibility and its long-term implications. The initial cap was set at ₹750, well below the pre-policy royalty fees of ₹1,100, making it

[58] This change includes farms that enter into cotton as a result of the policy, mirroring our reduced-form evidence in Section 4.2. If we subset the sample to cotton growers in the benchmark scenario of no policy, the result holds (Appendix Figure E7b).

clear that downstream suppliers could not have kept operating under the mandated price ceiling without an upstream concession. In other words, the viability of the policy hinged on a de facto IP appropriation: without it, Bt products would have been pulled from the market. This matters for welfare: part of the observed gains for farmers reflects a transfer of resources from the GE technology provider, rather than improved market efficiency. While the price-cap component compressed the margins of downstream local seed firms, the royalty-cut component did not. Preserved margins and a surge in demand under the latter may have buffered local firms' incentives to reduce innovation, helping explain why the drop in quality was smaller than the drop in prices.

6.1 Accounting for Reaction of Upstream Innovation

The reduction in royalty payments to Monsanto was central to the welfare gains experienced by Indian farmers. However, this may have come at a cost not captured by our model: diminished incentives for upstream innovation and delayed release of new technologies to the Indian market. In practice, we observe that global advancements that became available elsewhere – most notably, Bt varieties with herbicide tolerance – were not deployed to India during this period (see descriptive evidence in Section 4.4). At the same time, domestic firms did not develop GE traits with improved resistance to emerging pests, such as the PBW, against which existing Bt traits had become increasingly ineffective (Tabashnik et al., 2013; Tabashnik and Carrière, 2019). Concerns over royalty reductions were prominent among technology providers, as evident in Monsanto's 10-K filings and in statements from managers of competing biotechnology firms (Appendix Table A2 and A3, respectively).

How do our welfare estimates change if we assume that the stagnation in upstream innovation was caused by price regulation? While we cannot credibly identify this causal link, we draw on the agronomy and entomology literature to obtain reasonable estimates of yield gains from "missing products" that never reached the Indian market. We focus on two technological innovations: (i) stacking a herbicide-tolerant (HT) trait onto Bt and (ii) incorporating genetic resistance to the PBW.[59] We augment optimal yields in the benchmark scenario of no policy to account for these innovations

[59] Given that no new genetic technologies have been released to India since 2009, no field trials of such technologies are publicly available. For herbicide tolerance, we use the most recent assessment by Gharde and Singh (2018), who quantifies an 18% average cotton yield loss from weeds in India, using data from more than a thousand farmer field trials between 2003 and 2014. For PBW resistance, we consider additional losses of 25%, based on roving field surveys of 83 villages and boll damage data by Fand et al. (2019).

(Rows 5-6) and compare the resulting welfare to that from the price-control policy (Row 2). Accounting for the release of an existing technology, such as HT cotton, partially offsets increases in farmer surplus. Further accounting for the development and introduction of a technological breakthrough, such as PBW-resistant cotton, turns welfare impacts on farmers negative. While speculative, these counterfactuals illustrate how reduced upstream innovation could reverse the short-run welfare gains from price regulation. On the other hand, in the lack of price intervention, the new technologies would have diffused at a much lower rate.

6.2 Designing Alternative Policies: Subsidizing Adoption or Innovation?

We conclude by comparing price controls to a set of policies aimed at mitigating the trade-off between product affordability and technological innovation. To direct technological progress without resorting to price regulation, a government can subsidize either technology users (i.e., farmers) or providers (seed-producing firms). We consider two main alternative policies: (i) a linear subsidy to cotton farmers, perhaps the most commonly used intervention to raise technology adoption in developing countries, and (ii) targeted R&D grants to seed firms, a supply-side industrial policy designed to strengthen private innovation incentives.[60]

Farm Input Subsidies. We insert a wedge between the price paid by farmers and the price received by firms so that operating profits in Equation 8 become $\left(p_{bm} - r_m - mc_{bm}(y_{bm})\right) \cdot Q_{bm}\left((1-\tau)\vec{p}_m, \vec{y}_m\right)$. We search over a grid of subsidy rates to find the linear schedule that generates the average equilibrium price observed under price controls. Using this subsidy rate, we calculate the corresponding budget cost for the government under the assumption of no administrative or implementation costs.

By increasing firms' profits and boosting demand for cotton seeds, the farm subsidy fully offsets the reduction in yields observed under price controls (Table 4, Row 7). The estimated increase in farmer surplus is large, reflecting the combined benefit from cheaper seeds and stable yields. However, this policy imposes a substantial fiscal burden: achieving the same average price as under price regulation would require a 54.9% linear subsidy. The implied budgetary cost for the government amounts to

[60] We acknowledge that, when designing agricultural and innovation policies, governments may respond to political economy – rather than budgetary and capacity – constraints (Acemoglu and Robinson, 2013; Dercon, 2024). Our counterfactual policy analysis assumes that governments have access to an unrestricted set of policy measures and aim to maximize farmers' surplus, instead of responding to political incentives or other competing objectives.

4.4% of India's nationwide provision for fertilizer subsidy (RBI, 2009, 2014), equivalent to 15.2% of the portion allocated to price-controlled states or 175.2% of the portion effectively accruing to cotton farmers (Praveen et al., 2017). The prohibitive cost of farm input subsidies may help explain why price caps were the preferred policy.

Firm Innovation Subsidies. While effective at restoring yields, and so increasing farmer surplus, input subsidies were found to entail exceptionally large fiscal costs given the sheer size of the farming population in India. Including administrative costs for enforcement and verification, and potential risks of leakage or elite capture, would make the fiscal bill even larger. This motivates a final alternative: subsidizing the much smaller set of domestic seed-producing firms. Specifically, we consider performance-based grants, designed to incentivize these firms to upgrade product quality.

Pull incentives, such as prizes and advance market commitments, are especially appealing in settings like agriculture (Kremer and Zwane, 2005, 2006), where productivity is directly observable.[61] Agronomic yield data from regulatory trials could be readily used to inform targeting to high-performing firms, avoiding the information and implementation challenges associated with targeting farmers. While farmer targeting may respond to equity concerns, the planner's objective in firm targeting is clear: direct resources to the most productive innovators. That said, *potential* yields from experimental stations should be treated with some caution, as they may miss other dimensions of quality that are valued by smallholder farmers in *real-life* conditions (Macours, 2019).

To search for the counterfactual grant, we compute the welfare effect of a set of grants that differ in their budget cost (Appendix E.4). Within the structure of our model, the grants are allocated in proportion to a firm's yields (relative to competitors in a market) and used to lower the slope of its fixed costs in Equation 10. With the innovation budget fixed, the grant operates as a zero-sum tournament: in each market, firms compete on quality to secure a larger allocation. We then simulate prices and yields given this grant. In Row 8, we assume that the government chooses the grant that, given this allocation rule, replicates the welfare effect of the farm subsidy estimated above. This grant is equal to 56% of the farm-subsidy budget.

[61] Pull programs are also appealing for other technological products with under-provision of innovation, notably pharmaceuticals and vaccines in developing countries (Kremer and Glennerster, 2004; Kremer and Williams, 2010).

Consumer preference heterogeneity shapes the distributional incidence of the two policy alternatives: the poorest smallholders are more likely to prefer price-reducing input subsidies, whereas larger farmers favor yield-enhancing innovation grants (Appendix Figures E7c and E7d). Using the full farm-subsidy budget to fund R&D grants delivers even larger welfare gains for farmers (Row 9). Rather than through cheaper seeds, these welfare gains are achieved by increasing yields at a higher rate than prices. This innovation policy drives technological progress in Indian cotton, while delivering greater surplus per unit of public spending across the distribution of farms.

7 Conclusion

We study the technological and welfare trade-offs induced by price regulation in the context of Indian agriculture. In 2006, three states imposed a maximum retail price on cotton seeds embodying Bt, a genetic technology designed to enhance pest resistance and so protect harvestable yield from biotic stress, while reducing costs of cultivation. We leverage this unique policy natural experiment to identify causal impacts on technological adoption and innovation. We combine several sources of farmer-level data and empirically show that the resulting reduction in farm-gate prices generated a rapid and lasting penetration of Bt seeds among cotton farmers. We further provide evidence across agricultural inputs and output that the Bt technology had a transformative impact on cotton farming in India.

We conjecture that, despite initially ensuring the affordability of the existing technology and driving its virtually universal adoption among end-users, price regulation can reduce the incentives for doing research and developing novel seed varieties among technology providers over the long term. Maintaining a steady flow of innovation holds distinct significance within the context of agricultural production, since technological advances in crop variety development are seldom global, require location-specific adaptation, and become obsolete in the face of evolving environmental forces, such as climate change and the emergence of new diseases, pests, and weeds. In India, local adaptation of cotton varieties is performed by the private sector; in much of Africa, the need may be even more acute given pervasive heterogeneity and near-zero breeding capacity (Suri and Udry, 2022).

We examine the supply-side response by drawing from administrative data on seed companies and newly assembled data on GE varieties. Our results provide suggestive evidence that the increased volume of sales allowed firms to remain in the market in the short term. Nonetheless, we provide both descriptive and causal evidence indicating reduced product introduction and distorted innovation, respectively, in the long term: a couple of years after the policy, firms started releasing fewer seed varieties, which exhibited lower agronomic yields in price-controlled states. The policy-induced yield loss erases more than half of the productivity gains expected from adopting the GE technology in those states.

A structural model of the cotton seed market allows us to assess the welfare implications of price and quality responses in equilibrium, and compare them to alternative policies. From a broader perspective, our findings highlight the importance of endogenous supply responses and quality adjustment to evaluate the equilibrium effects of public policies in technological markets. Neglecting this margin of adjustment can lead researchers and policymakers to overestimate the long-term gains from demand-driven market interventions, such as price regulation.

Our welfare analysis abstracts from several dimensions of agricultural technology adoption and innovation. First, we do not consider equilibrium effects in complementary input markets, such as labor and pesticides. Second, we do not take into account the role of intermediary suppliers of seeds, such as agro-dealers (e.g., Dillon et al., 2025), nor do we allow for firm entry and exit decisions. Ours is a retrospective analysis on farmers: we do not model firms' R&D and dynamic investments, which would require either more comprehensive data on innovation inputs or additional structure on firm conduct. Lastly, we abstract away from imitative and intertemporal innovation spillovers across firms and markets (e.g., Jones and Summers, 2022). Despite these limitations, our empirical framework offers a stylized model of product targeting to local ecologies, which is designed to capture the core trade-off between affordability and innovation in the seed market. This trade-off is likely relevant to a range of other industries, such as green energy, fintech, and pharmaceuticals, where the development and widespread adoption of affordable innovations are essential for economic development and well-being. Quantifying the importance of these complementary forces and understanding the effects of price interventions in non-agricultural markets offer promising directions for future research. ◆

References

Acemoglu, Daron and Joshua Linn (2004) "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry," *The Quarterly Journal of Economics*, 119 (3), 1049–1090, 10.1162/0033553041502144.

Acemoglu, Daron and James A Robinson (2013) "Economics versus Politics: Pitfalls of Policy Advice," *Journal of Economic Perspectives*, 27 (2), 173–192, 10.1257/jep.27.2.173.

Acemoglu, Daron and Fabrizio Zilibotti (2001) "Productivity Differences," *The Quarterly Journal of Economics*, 116 (2), 563–606, 10.1162/00335530151144104.

Adamopoulos, Tasso, Loren Brandt, Jessica Leight, and Diego Restuccia (2022) "Misallocation, Selection, and Productivity: A Quantitative Analysis with Panel Data from China," *Econometrica*, 90 (3), 1261–1282, 10.3982/ECTA16598.

Adamopoulos, Tasso and Diego Restuccia (2014) "The Size distribution of Farms and International Productivity Differences," *American Economic Review*, 104 (6), 1667–1697, 10.1257/aer.104.6.1667.

Ahuja, Vibha (2018) "Regulation of Emerging Gene Technologies in India," *BMC Proceedings*, 12 (8), 14, 10.1186/s12919-018-0106-0.

Aker, Jenny C and B Kelsey Jack (2023) "Harvesting the Rain: The Adoption of Environmental Technologies in the Sahel," *Review of Economics and Statistics*, 1–52, 10.1162/rest_a_01404.

Akerman, Ariel, Jacob Moscona, Heitor S Pellegrina, and Karthik Sastry (2025) "Public R&D Meets Economic Development: Embrapa and Brazil's Agricultural Revolution," Unpublished manuscript.

Atal, Juan Pablo, José Ignacio Cuesta, and Morten Sæthre (2025) "Quality Regulation and Competition: Evidence from Pharmaceutical Markets," Unpublished manuscript.

Barahona, Nano, Cristóbal Otero, and Sebastián Otero (2023) "Equilibrium Effects of Food Labeling Policies," *Econometrica*, 91 (3), 839–868, 10.3982/ECTA19603.

Barwick, Panle Jia, Hyuk-Soo Kwon, and Shanjun Li (2024) "Attribute-based Subsidies and Market Power: an Application to Electric Vehicles," Unpublished manuscript.

Basu, Susanto and David N Weil (1998) "Appropriate Technology and Growth," *The Quarterly Journal of Economics*, 113 (4), 1025–1054, 10.1162/003355398555829.

Berry, Steven, James Levinsohn, and Ariel Pakes (2004) "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *Journal of Political Economy*, 112 (1), 68–105, 10.1086/379939.

Blaise, Desouza and Keshav Raj Kranthi (2019) "Cotton Production in India," in Jabran, Khawar and Bhagirath Singh Chauhan eds. *Cotton Production*, Chap. 10, 193–215, New York City, New York: John Wiley & Sons, 10.1002/9781119385523.ch10.

Blume-Kohout, Margaret E and Neeraj Sood (2013) "Market Size and Innovation: Effects of Medicare Part D on Pharmaceutical Research and Development," *Journal of Public Economics*, 97, 327–336, 10.1016/j.jpubeco.2012.10.003.

Bolhuis, Marijn A, Swapnika R Rachapalli, and Diego Restuccia (forthcoming) "Misallocation in Indian Agriculture," *American Economic Journal: Macroeconomics*, 10.1257/mac.20210358.

Bustos, Paula, Bruno Caprettini, and Jacopo Ponticelli (2016) "Agricultural Productivity and Structural Transformation: Evidence from Brazil," *American Economic Review*, 106 (6), 1320–1365, 10.1257/aer.20131061.

Bustos, Paula, Gabriel Garber, and Jacopo Ponticelli (2020) "Capital Accumulation and Structural Transformation," *The Quarterly Journal of Economics*, 135 (2), 1037–1094, 10.1093/qje/qjz044.

Bustos, Paula, Juanma Vincenzi, Joan Monràs, and Jacopo Ponticelli (2023) "Industrialization without Innovation," Unpublished manuscript.

Cai, Yong (2023) "A Modified Randomization Test for the Level of Clustering," *Journal of Business & Economic Statistics*, 1–13, 10.1080/07350015.2023.2261567.

Carleton, Tamma, Esther Duflo, B Kelsey Jack, and Guglielmo Zappalà (2024) "Adaptation to Climate Change," in *Handbook of the Economics of Climate Change*, 1, 143–248: Elsevier, 10.1016/bs.hesecc.2024.10.001.

Chakraborty, Pubali, Anand Chopra, and Lalit Contractor (2025) "The Equilibrium Impact of Agricultural Support Prices and Input Subsidies," Unpublished manuscript.

Chaudhuri, Shubham, Pinelopi K Goldberg, and Panle Jia (2006) "Estimating the Effects of Global Patent Protection in Pharmaceuticals: A Case Study of Quinolones in India," *American Economic Review*, 96 (5), 1477–1514, 10.1257/aer.96.5.1477.

Chen, Chaoran, Diego Restuccia, and Raül Santaeulàlia-Llopis (2023) "Land Misallocation and Productivity," *American Economic Journal: Macroeconomics*, 15 (2), 441–465, 10.1257/mac.20170229.

Choudhary, Bhagirath and Kadambini Gaur (2015) "Biotech Cotton in India, 2002 to 2014," ISAAA Series of Biotech Crop Profiles. ISAAA: Ithaca, NY.

Ciliberto, Federico, GianCarlo Moschini, and Edward D Perry (2019) "Valuing Product Innovation: Genetically Engineered Varieties in US Corn and Soybeans," *The RAND Journal of Economics*, 50 (3), 615–644, 10.1111/1756-2171.12290.

Cockburn, Iain M, Jean O Lanjouw, and Mark Schankerman (2016) "Patents and the Global Diffusion of New Drugs," *American Economic Review*, 106 (01), 136–164, 10.1257/aer.20141482.

Crawford, Gregory S, Oleksandr Shcherbakov, and Matthew Shum (2019) "Quality Overprovision in Cable Television Markets," *American Economic Review*, 109 (3), 956–995, 10.1257/aer.20151182.

Dean, Emma Boswell (2023) "Who Benefits from Pharmaceutical Price Controls? Evidence from India," Unpublished manuscript.

Dercon, Stefan (2024) "The Political Economy of Economic Policy Advice," *Journal of African Economies*, 33 (Supplement_2), ii26–ii38, 10.1093/jae/ejae027.

Dillon, Andrew, Travis J Lybbert, Hope Michelson, and Jessica Rudder (2025) "Agricultural Input Markets in Sub-Saharan Africa: Theory and Evidence from the (Underappreciated) Supply Side," *Annual Review of Resource Economics*, 17, 10.1146/annurev-resource-102324-122108.

Dillon, Brian and Christopher B Barrett (2017) "Agricultural Factor Markets in Sub-Saharan Africa: An Updated View with Formal Tests for Market Failure," *Food Policy*, 67, 64–77, 10.1016/j.foodpol.2016.09.015.

Directorate of Economics and Statistics of the Department of Agriculture, Cooperation and Farmers Welfare, Ministry of Agriculture and Farmers Welfare, Government of India (2008) "Manual on Cost of Cultivation Surveys," CSO-M-AG-02, `https://desagri.gov.in/wp-content/uploads/2021/06/manual_cost_cultivation_surveys_23july08_0.pdf`.

Donovan, Kevin (2021) "The Equilibrium Impact of Agricultural Risk on Intermediate Inputs and Aggregate Productivity," *The Review of Economic Studies*, 88 (5), 2275–2307, 10.1093/restud/rdaa084.

Draganska, Michaela, Michael Mazzeo, and Katja Seim (2009) "Beyond Plain Vanilla: Modeling Joint Product Assortment and Pricing Decisions," *Quantitative Marketing and Economics*, 7, 105–146, 10.1007/s11129-008-9047-7.

Dubois, Pierre, Olivier De Mouzon, Fiona Scott-Morton, and Paul Seabright (2015) "Market Size and Pharmaceutical Innovation," *The RAND Journal of Economics*, 46 (4), 844–871, 10.1111/1756-2171.12113.

Dubois, Pierre, Ashvin Gandhi, and Shoshana Vasserman (2022) "Bargaining and International Reference Pricing in the Pharmaceutical Industry," Unpublished manuscript.

Dubois, Pierre and Gökçe Gökkoca (2025) "Antibiotic Demand in the Presence of Antimicrobial Resistance," *Review of Economics and Statistics*, 1–46, 10.1162/rest.a.272.

Dubois, Pierre and Laura Lasio (2018) "Identifying Industry Margins with Price Constraints: Structural Estimation on Pharmaceuticals," *American Economic Review*, 108 (12), 3685–3724, 10.1257/aer.20140202.

Emerick, Kyle, Alain De Janvry, Elisabeth Sadoulet, and Manzoor H Dar (2016) "Technological Innovations, Downside Risk, and the Modernization of Agriculture," *American Economic Review*, 106 (6), 1537–1561, 10.1257/aer.20150474.

Evenson, Robert E and Douglas Gollin (2003) "Assessing the Impact of the Green Revolution, 1960 to 2000," *science*, 300 (5620), 758–762, 10.1126/science.1078710.

Fan, Ying (2013) "Ownership Consolidation and Product Characteristics: A Study of the US Daily Newspaper Market," *American Economic Review*, 103 (5), 1598–1628, 10.1257/aer.103.5.1598.

Fan, Ying and Chenyu Yang (2020) "Competition, Product Proliferation, and Welfare: A Study of the US Smartphone Market," *American Economic Journal: Microeconomics*, 12 (2), 99–134, 10.1257/mic.20180182.

Fand, Babasaheb B, VS Nagrare, SP Gawande, DT Nagrale, BV Naikwadi, Vrushali Deshmukh, Nandini Gokte-Narkhedkar, and VN Waghmare (2019) "Widespread Infestation of Pink Bollworm, *Pectinophora Gossypiella* (Saunders) (Lepidoptera: Gelechidae) on Bt Cotton in Central India: A New Threat and Concerns for Cotton Production," *Phytoparasitica*, 47, 313–325, 10.1007/s12600-019-00738-x.

Filson, Darren (2012) "A Markov-Perfect Equilibrium Model of the Impacts of Price Controls on the Performance of the Pharmaceutical Industry," *The RAND Journal of Economics*, 43 (1), 110–138, 10.1111/j.1756-2171.2012.00159.x.

Finkelstein, Amy (2004) "Static and Dynamic Effects of Health Policy: Evidence from the Vaccine Industry," *The Quarterly Journal of Economics*, 119 (2), 527–564, 10.1162/0033553041382166.

Foster, Andrew D and Mark R Rosenzweig (2010) "Microeconomics of Technology Adoption," *Annual Review of Economics*, 2 (1), 395–424, 10.1146/annurev.economics.102308.124433.

——— (2022) "Are There Too Many Farms in the World? Labor Market Transaction Costs, Machine Capacities, and Optimal Farm Size," *Journal of Political Economy*, 130 (3), 636–680, 10.1086/717890.

Fuglie, Keith, Paul Heisey, John L King, Kelly Day-Rubenstein, David Schimmelpfennig, Sun Ling Wang, Carl E Pray, and Rupa Karmarkar-Deshmukh (2011) "Research Investments and Market Structure in the Food Processing, Agricultural Input, and Biofuel Industries Worldwide," USDA-ERS Economic Research Report 130, United States Department of Agriculture.

Gharde, Yogita and PK Singh (2018) "Yield and Economic Losses Due to Weeds in India," Technical Bulletin No. 17. ICAR – Directorate of Weed Research, Jabalpur.

Ghatak, Maitreesh and Dilip Mookherjee (2025) "Misallocating Misallocation?" *Annual Review of Economics*, 17, 10.1146/annurev-economics-091624-051237.

Ghosh, Srijita (2019) "Multidimensional and Selective Learning: A Case Study of Bt Cotton Farmers in India," Unpublished manuscript.

Glennester, R and T Suri (2018) "Agricultural Technology and Nutrition: The Impacts of NERICA Rice in Sierra Leone," Unpublished manuscript.

Gollin, Douglas, Casper Worm Hansen, and Asger Mose Wingender (2021) "Two Blades of Grass: The Impact of the Green Revolution," *Journal of Political Economy*, 129 (8), 2344–2384, 10.1086/714444.

Gollin, Douglas and Christopher Udry (2021) "Heterogeneity, Measurement Error, and Misallocation: Evidence from African Agriculture," *Journal of Political Economy*, 129 (1), 1–80, 10.1086/711369.

Grieco, Paul LE, Charles Murry, Joris Pinkse, and Stephan Sagl (2025) "Optimal Estimation of Discrete Choice Demand Models with Consumer and Product Data," Unpublished manuscript.

Griliches, Zvi (1957) "Hybrid Corn: An Exploration in the Economics of Technological Change," *Econometrica*, 501–522, 10.2307/1905380.

Grossman, Gene M and Elhanan Helpman (1993) *Innovation and Growth in the Global Economy*, Cambridge, Massachusetts: MIT Press.

Gruère, Guillaume P, Purvi Mehta-Bhatt, and Debdatta Sengupta (2008) "Bt Cotton and Farmer Suicides in India. Reviewing the evidence," IFPRI Discussion Paper 00808, Washington, DC: International Food Policy Research Institute.

Gupta, Harsh and Shengmao Cao (2024) "Price Controls with Imperfect Competition and Choice Frictions: Evidence from Indian Pharmaceuticals," Unpublished manuscript.

Hansen, Casper Worm and Asger Mose Wingender (2023) "National and Global Impacts of Genetically Modified Crops," *American Economic Review: Insights*, 5 (2), 224–240, 10.1257/aeri.20220144.

Herring, Ronald J (2013) "Reconstructing Facts in Bt Cotton: Why Scepticism Fails," *Economic and Political Weekly*, 63–66.

Howitt, Peter and Philippe Aghion (1998) "Capital Accumulation and Innovation as Complementary Factors in Long-Run Growth," *Journal of Economic Growth*, 3, 111–130, 10.1023/A:1009769717601.

Hristakeva, Sylvia, Julie Holland Mortimer, and Eric Yde (2025) "The Effect of Price Caps on Pharmaceutical Advertising: Evidence from the 340b Drug Pricing Program," Unpublished manuscript.

Jack, B Kelsey (2013) "Market Inefficiencies and the Adoption of Agricultural Technologies in Developing Countries," White paper, Agricultural Technology Adoption Initiative, J-PAL (MIT) and CEGA (UC Berkeley).

Ji, Yunan and Parker Rogers (2024) "The Long-Run Impacts of Regulated Price Cuts: Evidence from Medicare," Unpublished manuscript.

Jones, Benjamin F and Lawrence H Summers (2022) "A Calculation of the Social Returns to Innovation," in Austan, Goolsbee and Benjamin F Jones eds. *Innovation and Public Policy*, Chap. 1, 13–60: University of Chicago Press, 10.7208/chicago/9780226805597.003.0002.

Jones, Maria, Florence Kondylis, John Loeser, and Jeremy Magruder (2022) "Factor Market Failures and the Adoption of Irrigation in Rwanda," *American Economic Review*, 112 (7), 2316–2352, 10.1257/aer.20210059.

Kala, Namrata (2019) "Learning, Adaptation, and Climate Uncertainty: Evidence from Indian Agriculture," Unpublished manuscript.

Kathage, Jonas and Matin Qaim (2012) "Economic Impacts and Impact Dynamics of Bt (*Bacillus Thuringiensis*) Cotton in India," *Proceedings of the National Academy of Sciences*, 109 (29), 11652–11656, 10.1073/pnas.1203647109.

Kondylis, Florence, John Loeser, Mushfiq Mobarak, Maria Jones, and Daniel Stein (2024) "Decentralizing Agricultural Demonstration to Accelerate Social Learning," Unpublished manuscript.

Kouser, Shahzad and Matin Qaim (2011) "Impact of Bt Cotton on Pesticide Poisoning in Smallholder Agriculture: A Panel Data Analysis," *Ecological Economics*, 70 (11), 2105–2113, 10.1016/j.ecolecon.2011.06.008.

Kranthi, Keshav Raj and Glenn Davis Stone (2020) "Long-Term Impacts of Bt Cotton in India," *Nature Plants*, 6 (3), 188–196, 10.1038/s41477-020-0615-5.

Kremer, Michael and Rachel Glennerster (2004) *Strong Medicine: Creating Incentives for Pharmaceutical Research on Neglected Diseases*: Princeton University Press.

Kremer, Michael and Heidi Williams (2010) "Incentivizing Innovation: Adding to the Tool Kit," in Lerner, Josh Lerner and Scott Stern eds. *Innovation Policy and the Economy*, 10, 1–17: The University of Chicago Press.

Kremer, Michael and Alix Peterson Zwane (2005) "Encouraging Private Sector Research for Tropical Agriculture," *World Development*, 33 (1), 87–105, 10.1016/j.worlddev.2004.07.006.

——— (2006) "Creating Incentives for Private Sector Involvement in Poverty Reduction: Purchase Commitments for Agricultural Innovation," in Kaul, Inge and Pedro Conceição eds. *The New Public Finance Responding To Global Challenges*, New York: Oxford University Press.

Krishna, Vijesh V and Matin Qaim (2012) "Bt Cotton and Sustainability of Pesticide Reductions in India," *Agricultural Systems*, 107, 47–55, 10.1016/j.agsy.2011.11.005.

Kyle, Margaret K (2007) "Pharmaceutical Price Controls and Entry Strategies," *Review of Economics and Statistics*, 89 (1), 88–99, 10.1162/rest.89.1.88.

Laajaj, Rachid and Karen Macours (2024) "The Complexity of Multidimensional Learning in Agriculture," Unpublished manuscript.

LaFave, Daniel and Duncan Thomas (2016) "Farms, Families, and Markets: New Evidence on Completeness of Markets in Agricultural Settings," *Econometrica*, 84 (5), 1917–1960, 10.3982/ECTA12987.

Lane, Gregory (2024) "Adapting to Climate Risk with Guaranteed Credit: Evidence from Bangladesh," *Econometrica*, 92 (2), 355–386, 10.3982/ECTA19127.

Lerner, Josh, Junxi Liu, Jacob Moscona, and David Y Yang (2025) "Appropriate Entrepreneurship? The Rise of China and the Developing World," Unpublished manuscript.

Macours, Karen (2019) "Farmers' Demand and the Traits and Diffusion of Agricultural Innovations in Developing Countries," *Annual Review of Resource Economics*, 11, 483–499, 10.1146/annurev-resource-100518-094045.

Maini, Luca and Fabio Pammolli (2023) "Reference Pricing as a Deterrent to Entry: Evidence from the European Pharmaceutical Market," *American Economic Journal: Microeconomics*, 15 (2), 345–383, 10.1257/mic.20210053.

Mayee, CD, Phundan Siivgh, Punit Mohan, and DK Agarwal (2004) "Evaluation of Bt Transgenic Intra-Hirsutum Hybrids for Yield and Fibre Properties," *Indian Journal of Agricultural Science*, 74 (1), 46–47.

Mazzeo, Michael J (2002) "Product Choice and Oligopoly Market Structure," *RAND Journal of Economics*, 221–242, 10.2307/3087431.

Menon, Meena and Uzramma (2017) *A Frayed History: The Journey of Cotton in India*: Oxford University Press.

Moschini, GianCarlo and Edward Perry (2024) "Innovation, Licensing, and Competition: Evidence from Genetically Engineered Crops," Unpublished manuscript.

Moscona, Jacob and Karthik Sastry (2025) "Inappropriate Technology: Evidence from Global Agriculture," Unpublished manuscript.

Murugkar, M., B. Ramaswami, and M. Shelar (2006) "Liberalization, Biotechnology and the Private Seed Sector: The Case of India's Cotton Seed Market," Discussion Paper 06-05. Indian Statistical Institute, New Delhi.

Myers, Kyle and Mark Pauly (2019) "Endogenous Productivity of Demand-Induced R&D: Evidence from Pharmaceuticals," *The RAND Journal of Economics*, 50 (3), 591–614, 10.1111/1756-2171.12289.

Newell, Peter (2007) "Biotech Firms, Biotech Politics: Negotiating GMOs in India," *The Journal of Environment & Development*, 16 (2), 183–206, 10.1177/1070496507300920.

Olmstead, Alan L and Paul W Rhode (2002) "The Red Queen and the Hard Reds: Productivity Growth in American Wheat, 1800–1940," *The Journal of Economic History*, 62 (4), 929–966, 10.1017/S0022050702001602.

——— (2008) *Creating Abundance: Biological Innovation and American Agricultural Development*, Cambridge, England: Cambridge University Press.

Pardey, Philip G, Julian M Alston, and Vernon W Ruttan (2010) "The Economics of Innovation and Technical Change in Agriculture," in *Handbook of the Economics of Innovation*, 2, 939–984: Elsevier, 10.1016/S0169-7218(10)02006-X.

Parente, Stephen L and Edward C Prescott (1994) "Barriers to Technology Adoption and Development," *Journal of Political Economy*, 102 (2), 298–321, 10.1086/261933.

de la Parra, Brenda Samaniego and Ajay Shenoy (2025) "Unpacking Entrepreneurial Ability," Unpublished manuscript.

Patel, Dev (2025) "Environmental Beliefs and Adaptation to Climate Change," Unpublished manuscript.

Peschard, Karine and Shalini Randeria (2020) "Taking Monsanto to Court: Legal Activism around Intellectual Property in Brazil and India," *The Journal of Peasant Studies*, 47 (4), 792–819, 10.1080/03066150.2020.1753184.

Plewis, Ian (2019) "Adopting Hybrid Bt Cotton: Using Interrupted Time-Series Analysis to Assess its Effects on Farmers in Northern India," *Review of Agrarian Studies*, 9 (2), 4–23, 10.22004/ag.econ.308317.

Praveen, KV, KS Aditya, ML Nithyashree, and A Sharma (2017) "Fertilizer Subsidies in India: An Insight to Distribution and Equity Issues," *Journal of Crop and Weed*, 13 (3), 24–31.

Pray, Carl E and Latha Nagarajan (2010) "Price Controls and Biotechnology Innovation: Are State Government Policies Reducing Research and Innovation by the Ag Biotech Industry in India?" *AgBioForum*, 13 (4), 297–307.

——— (2012) "Innovation and Research by Private Agribusiness in India," IFPRI Discussion Paper 01181, Washington, DC: International Food Policy Research Institute.

——— (2013) "Role of Biotechnology in Stimulating Agribusiness R&D Investment in India," *AgBioForum*, 16 (2), 104–111.

——— (2014) "The Transformation of the Indian Agricultural Input Industry: Has It Increased Agricultural R&D?" *Agricultural Economics*, 45 (S1), 145–156, 10.1111/agec.12138.

Qaim, Matin (2003) "Bt Cotton in India: Field Trial Results and Economic Projections," *World Development*, 31 (12), 2115–2127, 10.1016/j.worlddev.2003.04.005.

Qaim, Matin and David Zilberman (2003) "Yield Effects of Genetically Modified Crops in Developing Countries," *Science*, 299 (5608), 900–902, 10.1126/science.1080609.

Ramaswami, Bharat, Carl E Pray, and N Lalitha (2008) "The Limits of Intellectual Property Rights: Lessons from the Spread of Illegal Transgenic Seeds in India," Working Paper No. 182, Gujarat Institute of Development Research, Ahmedabad.

Reserve Bank of India (RBI) (2010) "Union Budget 2009-10: Review and Assessment," Monthly Bulletin, November 2009.

——— (2014) "Union Budget 2014-15: An Assessment," Monthly Bulletin, September 2014.

Restuccia, Diego, Dennis Tao Yang, and Xiaodong Zhu (2008) "Agriculture and Aggregate Productivity: A Quantitative Cross-Country Analysis," *Journal of Monetary Economics*, 55 (2), 234–250, j.jmoneco.2007.11.006.

Romer, Paul M (1990) "Endogenous Technological Change," *Journal of Political Economy*, 98 (5, Part 2), S71–S102, 10.1086/261725.

Rosenzweig, Mark R and Christopher Udry (2020) "External Validity in a Stochastic World: Evidence from Low-Income Countries," *Review of Economic Studies*, 87 (1), 343–381, 10.1093/restud/rdz021.

Sadashivappa, Prakash and Matin Qaim (2009) "Bt Cotton in India: Development of Benefits and the

Role of Government Seed Price Interventions," *AgBioForum*, 12 (2), 172–183.

Sahai, Suman (2002) "Bt Cotton: Confusion Prevails," *Economic and Political Weekly*, 1973–1974.

Shenoy, Ajay (2017) "Market Failures and Misallocation," *Journal of Development Economics*, 128, 65–80, 10.1016/j.jdeveco.2017.05.004.

——— (2021) "Estimating the Production Function under Input Market Frictions," *Review of Economics and Statistics*, 103 (4), 666–679, 10.1162/rest_a_00927.

Shiva, Vandana, Ashok Emani, and Afsar H Jafri (1999) "Globalisation and Threat to Seed Security: Case of Transgenic Cotton Trials in India," *Economic and Political Weekly*, 601–613.

Spielman, David J, Deepthi E Kolady, Anthony Cavalieri, and N Chandrasekhara Rao (2014) "The Seed and Agricultural Biotechnology Industries in India: An Analysis of Industry Structure, Competition, and Policy Options," *Food Policy*, 45, 88–100, 10.1016/j.foodpol.2014.01.001.

Suri, Tavneet (2011) "Selection and Comparative Advantage in Technology Adoption," *Econometrica*, 79 (1), 159–209, 10.3982/ECTA7749.

Suri, Tavneet and Christopher Udry (2022) "Agricultural Technology in Africa," *Journal of Economic Perspectives*, 36 (1), 33–56, 10.1257/jep.36.1.33.

Tabashnik, Bruce E, Thierry Brévault, and Yves Carrière (2013) "Insect Resistance to Bt Crops: Lessons from the First Billion Acres," *Nature Biotechnology*, 31 (6), 510–521, 10.1038/nbt.2597.

Tabashnik, Bruce E and Yves Carrière (2019) "Global Patterns of Resistance to Bt Crops Highlighting Pink Bollworm in the United States, China, and India," *Journal of Economic Entomology*, 112 (6), 2513–2523, 10.1093/jee/toz173.

Tandon, JP, SP Sharma, JS Sandhu, DK Yadava, KV Prabhu, and OP Yadav (2015) "Guidelines for Testing Crop Varieties under the All-India Coordinated Crop Improvement Projects," Indian Council of Agricultural Research, New Delhi, India, `https://icar.org.in/en/node/10186`.

Udry, Christopher (1999) "Efficiency and Market Structure: Testing for Profit Maximization in African Agriculture," in Ranis, Gustav and Lakshmi K Raut eds. *Trade, Growth and Development: Essays in Honor of T.N. Srinivasan*, Amsterdam, Netherlands: Elsevier Science.

United States Department of Agriculture (USDA), Foreign Agricultural Service (2024) "Cotton and Products Annual: India," Attaché Report IN2024-0017.

Veettil, Prakashan Chellattan, Vijesh V Krishna, and Matin Qaim (2017) "Ecosystem Impacts of Pesticide Reductions through Bt Cotton Adoption," *Australian Journal of Agricultural and Resource Economics*, 61 (1), 115–134, 10.1111/1467-8489.12171.

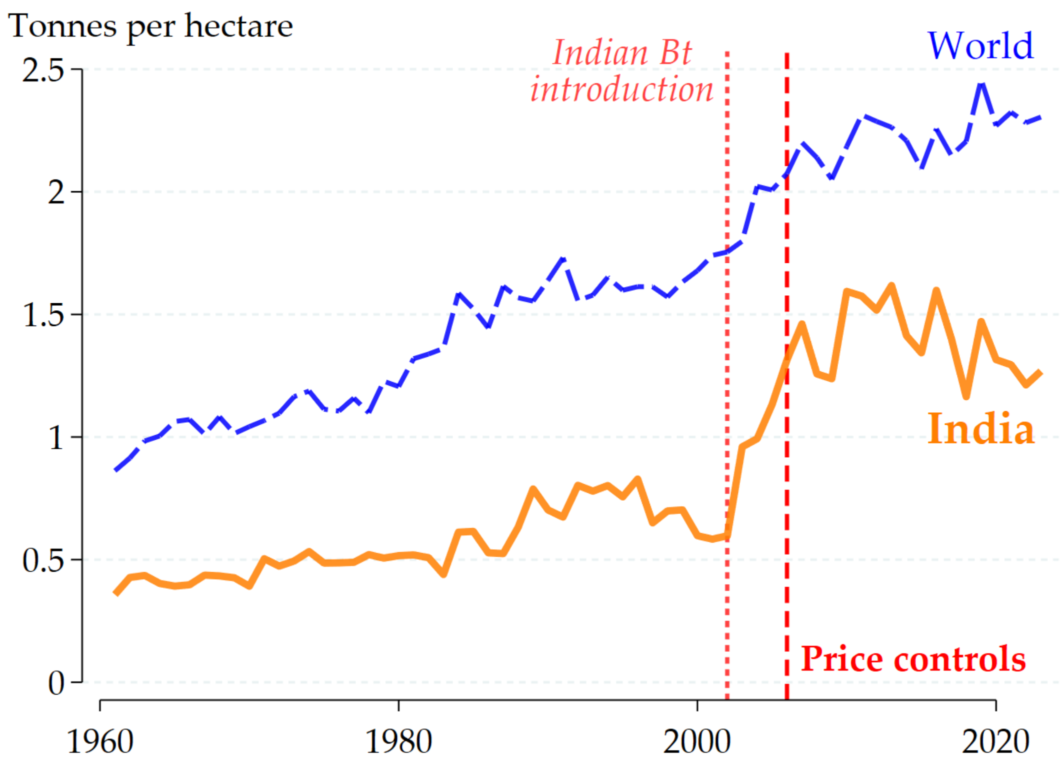
Wollmann, Thomas G (2018) "Trucks Without Bailouts: Equilibrium Product Characteristics for Commercial Vehicles," *American Economic Review*, 108 (6), 1364–1406, 10.1257/aer.20160863.

APPENDIX

for "Price Regulation of Agricultural Technology" by Felipe Berrutti and Matteo Ruzzante

A Descriptive Statistics

Figure A1. Average Seed Cotton Yields in India and in the World



Notes: Yields are defined as tonnes of harvested production (unginned) per hectare. Cotton production refers to cotton lint (ginned), i.e., fiber that has not been carded or combed. Data from FAOSTAT (Food and Agriculture Organization of the United Nations).

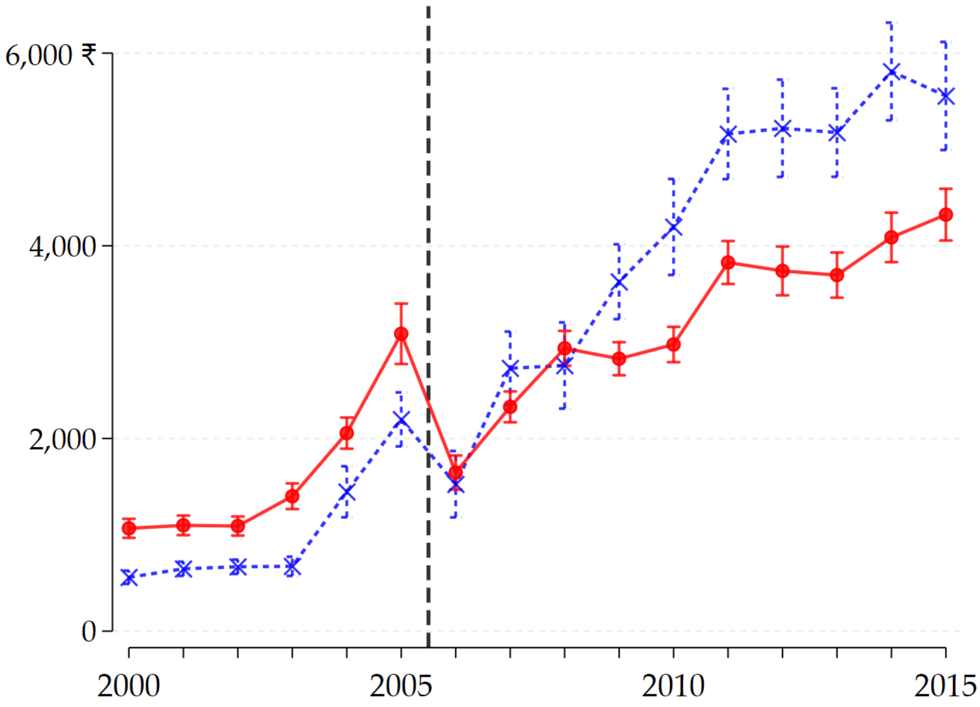
Table A1. Price Structure of a Bt Cotton Seed Packet

Item ↓	Company A (for 2016)	Company B (for 2018)
Procurement costs		
Procurement rate for contract farmers	330	285
Parent seeds	15	15
Sub-total	*345*	*300*
Overhead expenses		
R&D	35	60
Delinting, grading, ginning outturn testing, processing, and seed treatment	42	60
Insurance at ginning, storage, and processing	3	5
Refugia (non-Bt) seeds	20	20
Loading and transport	5	5
Storage	2	2
Loss on procurement due to low germination	5	5
Financial charges	15	15
Administrative expenses	57	57
Advertisement and promotion	10	65
Sales return	30	30
Seed discard		20
Sub-total	*224*	*344*
Trait fees		
	49	*39*
Margins		
Agro-dealer	75	115
Company	107	44
Sub-total	*182*	*159*
Total = price	**800**	**842**

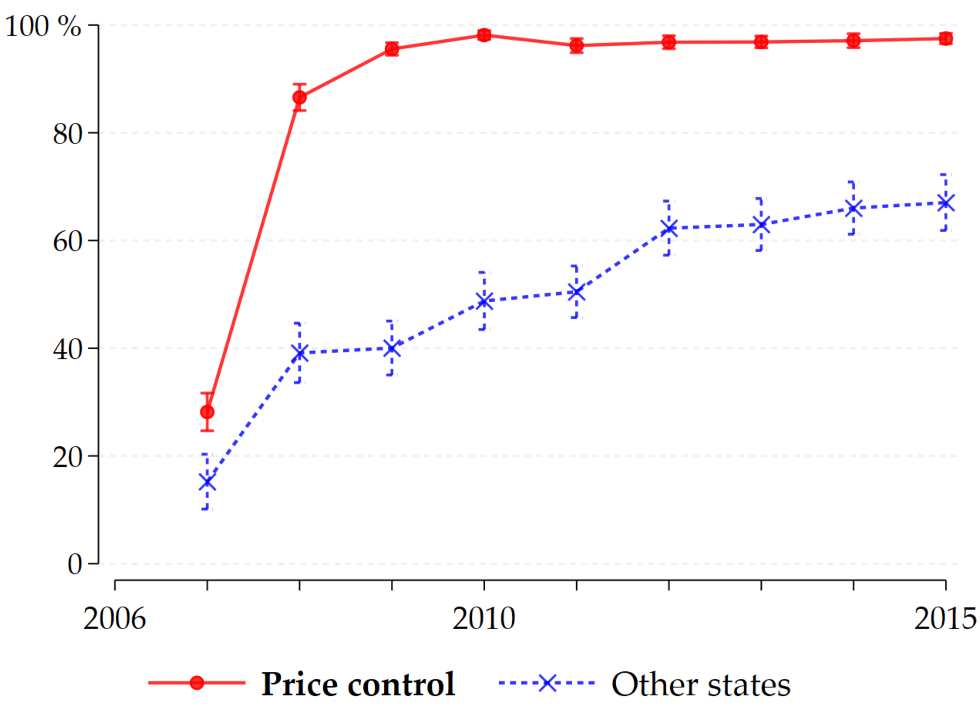
Notes: Internal cost data were obtained in May 2024 through personal communications with market leaders in the Indian cotton seed industry.

Figure A2. Trends in the National Sample

(a) Cotton Seed Costs



(b) Bt Cotton Adoption

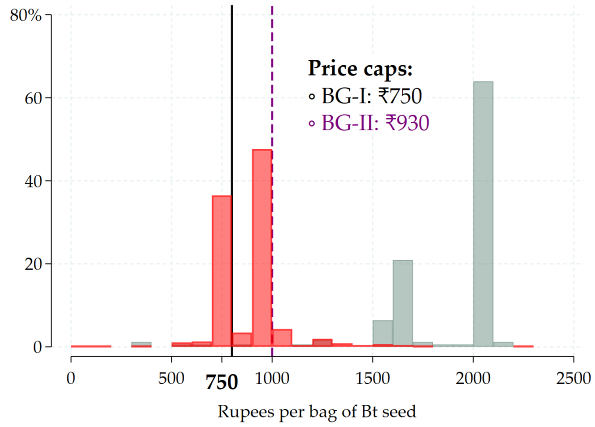


Notes: Means estimates with 95 percent confidence intervals based on least-squares regressions with standard errors clustered at the farmer level. Unit of observation: household × parcel × plot × season × year. Data from the Cost of Cultivation/Production Survey. The outcome in panel (a) is measured in Indian rupees (₹) per acre. The vertical dashed line signals the treatment timing. "**Price control**" includes all interviewed households in states with price control acts, i.e., Andhra Pradesh, Gujarat, Maharashtra, and Telangana (formed in 2014).

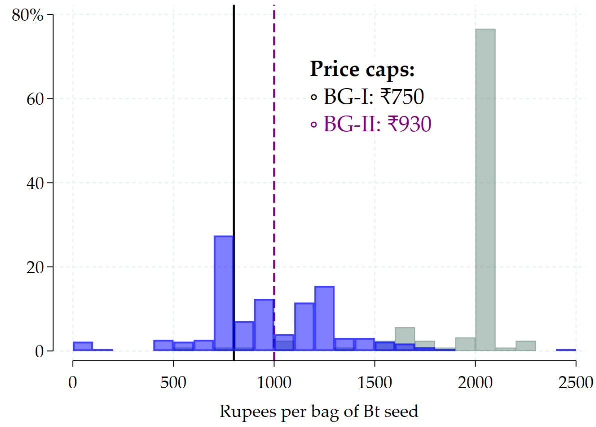
Figure A3. Distribution of Cotton Seed Prices by Treatment Exposure

Bt Cotton

(a) Price-Controlled States

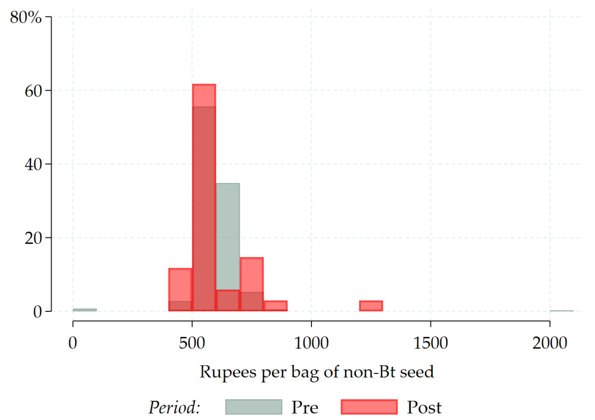


(b) Other States

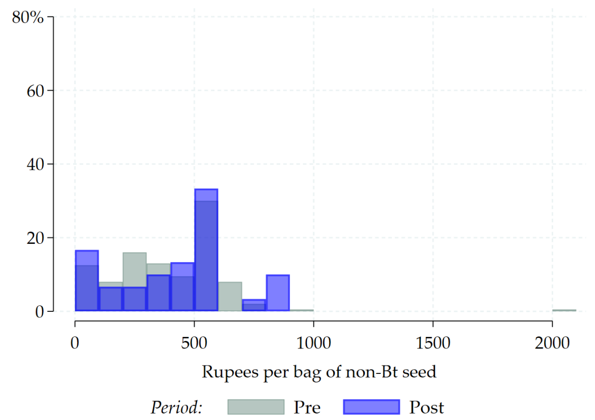


Conventional Cotton

(c) Price-Controlled States



(d) Other States



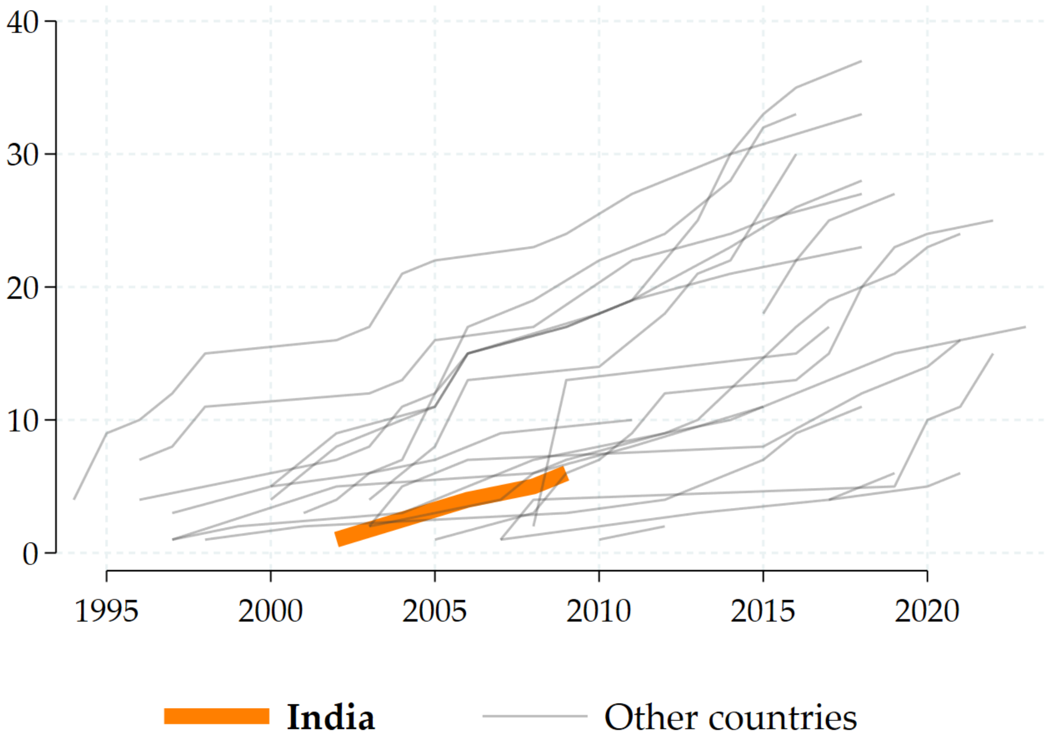
Notes: "Price-controlled states" in the left panels include all interviewed households in states with price controls (i.e., Andhra Pradesh and Maharashtra). "Other states" in the right panels are Karnataka and Tamil Nadu. "Pre" periods in gray are before the price control policy (i.e., 2002/2003 and 2004/2005). "Post" periods in red or blue are after (i.e., 2006/2007 and 2008/2009). Unit of observation: household × plot. Data from Kathage and Qaim (2012)'s panel survey. The outcome, in Indian rupees (₹), is rescaled to reflect the price for a 450-gram bag, the typical size of a hybrid cotton seed package (enough to plant an acre of land) in India. Note that Bt cotton seed packets typically include an additional 120 grams of non-Bt cotton seeds, intended to serve as refugia: this weight is excluded from the reported price calculation. "BG-I" and "BG-II" in panel (a) refer to *Bollgard 1* (Bt single-gene Cry1Ac) and *Bollgard 2* (double-gene Cry1Ac and Cry2Ab) cotton varieties, respectively.

Figure A4. Nationwide Trends in Indian Cotton

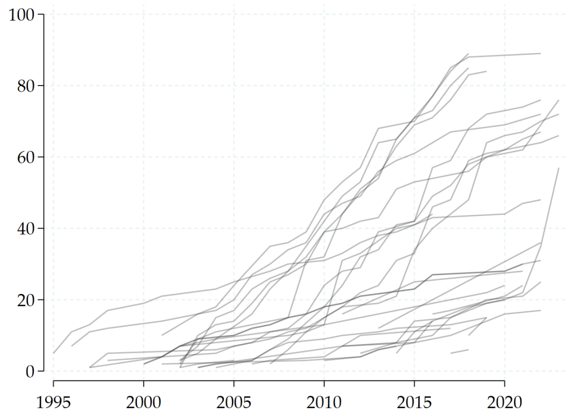
Notes: Lint production and import/export refer to cotton lint (ginned), i.e., fiber that has not been carded or combed, while seed production is considered unginned. Data from FAOSTAT (Food and Agriculture Organization of the United Nations).

Figure A5. Evolution of Genetically Modified Crop Event Approvals by Country

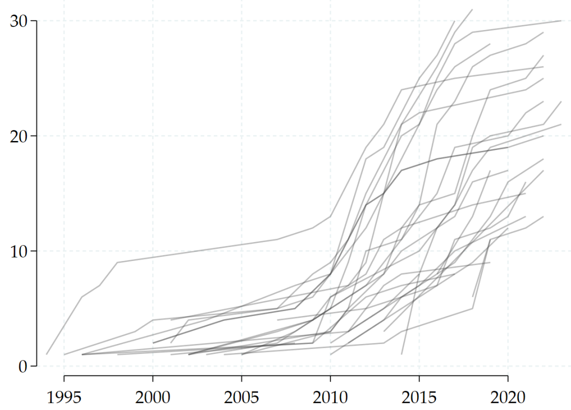
(a) Cotton



(b) Maize

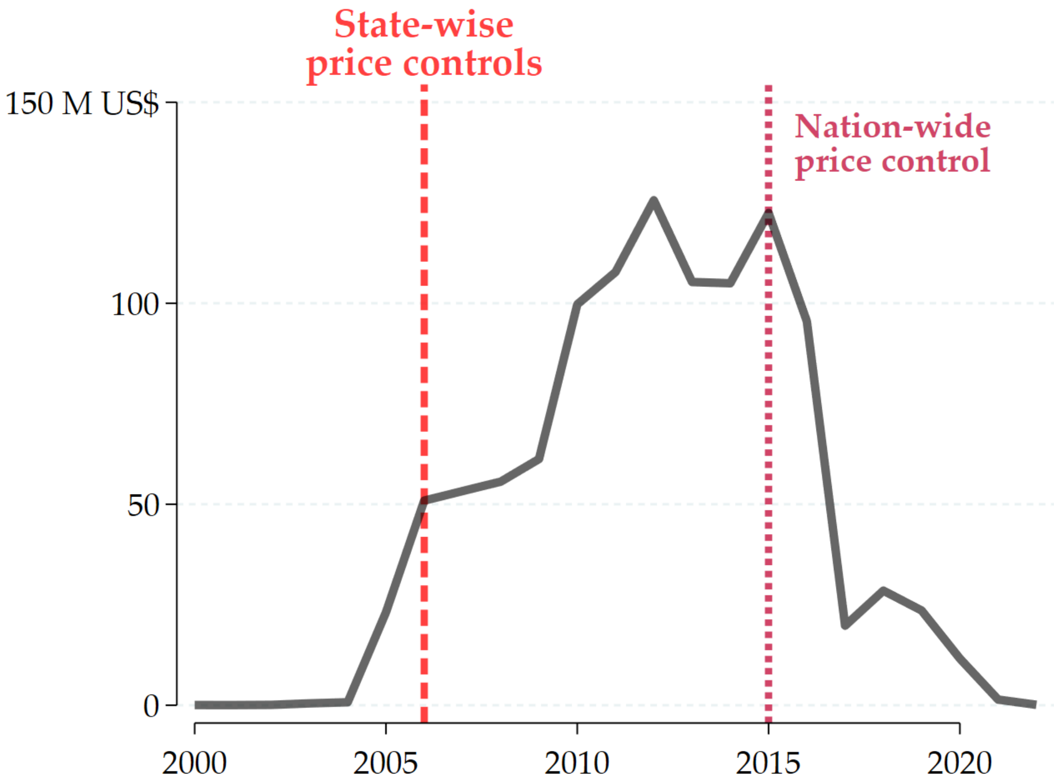


(c) Soybean



Notes: Data from the GM Approval Database (scraped from https://www.isaaa.org/gmapprovaldatabase/default.asp) of the International Service for the Acquisition of Agri-biotech Applications. Each gray line represents an approving country: it starts in the year with the first GM crop approval (e.g., 2002 for cotton in India) and ends with the latest year of GM crop approval (e.g., 2009 for cotton in India).

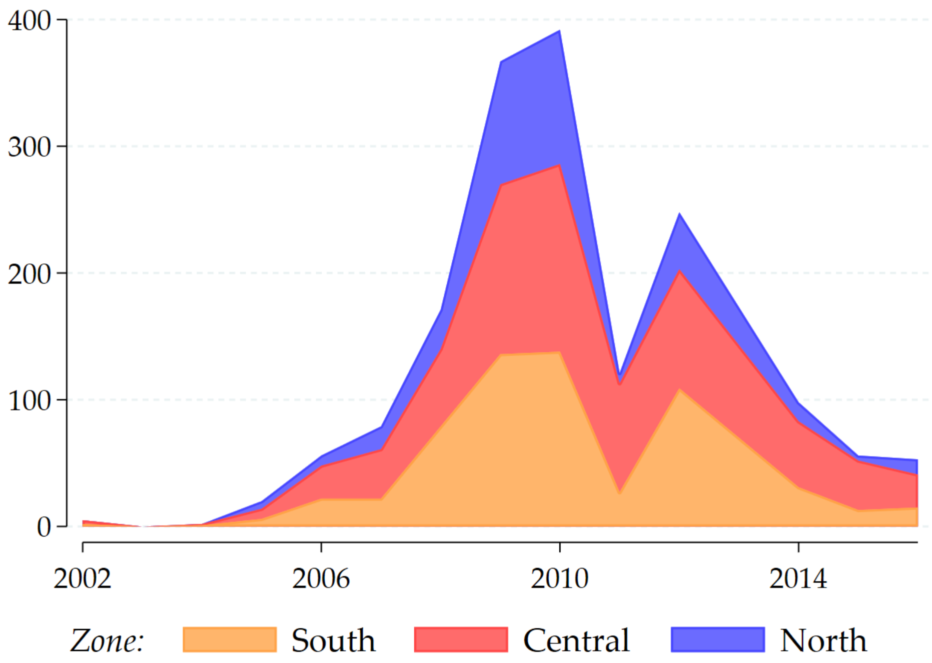
Figure A6. Evolution of Technology Royalties



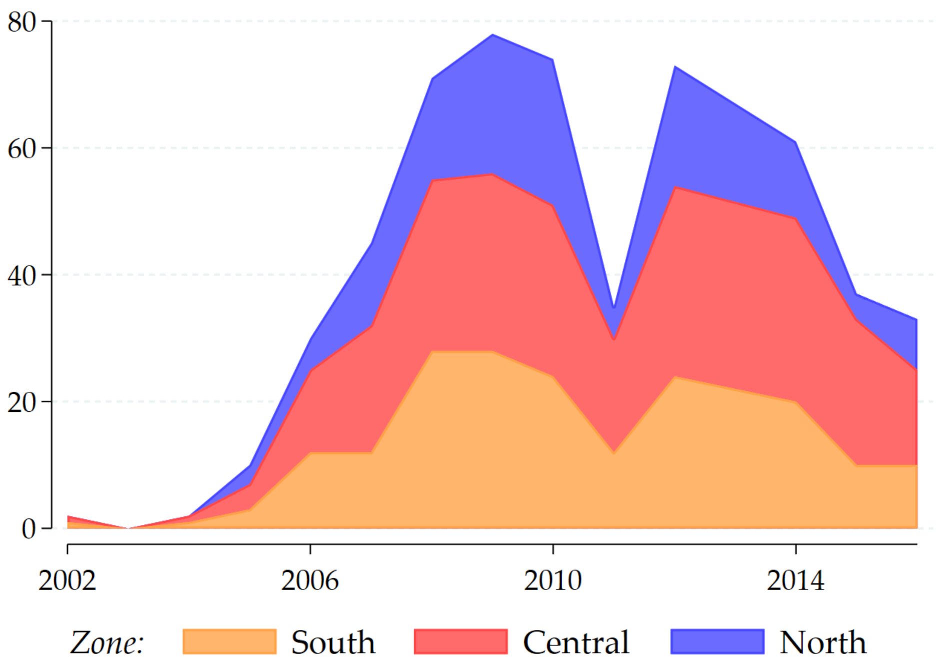
Notes: Data from the Prowess database of the Centre for Monitoring Indian Economy. The black line plots the yearly royalty fees for sub-licensing Bt genetic trait(s), as paid by domestic seed firms to Mahyco Monsanto Biotech.

Figure A7. Evolution of Seed Variety Approvals by Zone

(a) Number of Varieties Approved

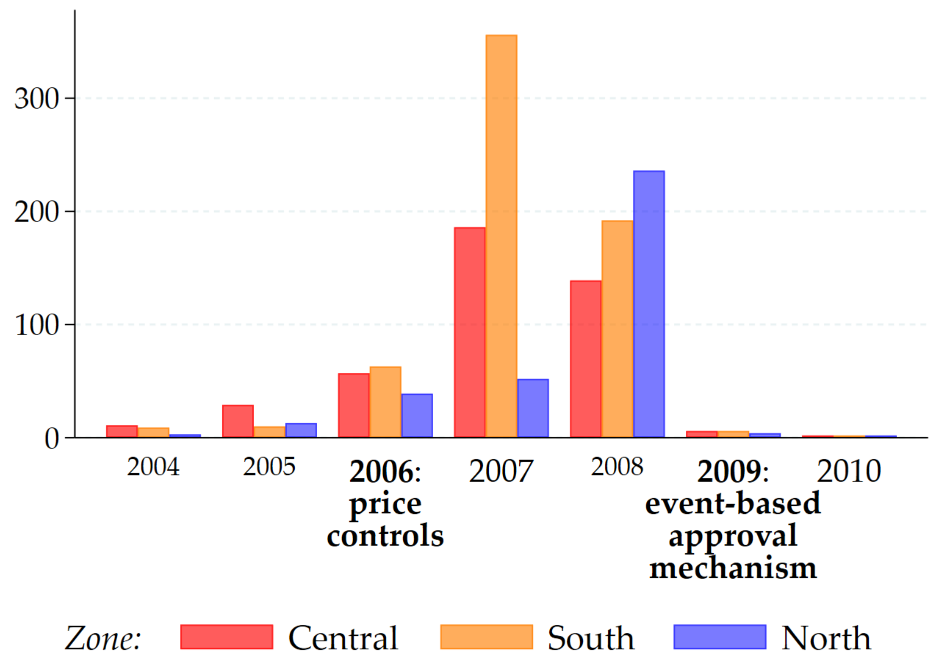


(b) Number of Companies Approving at Least One Variety



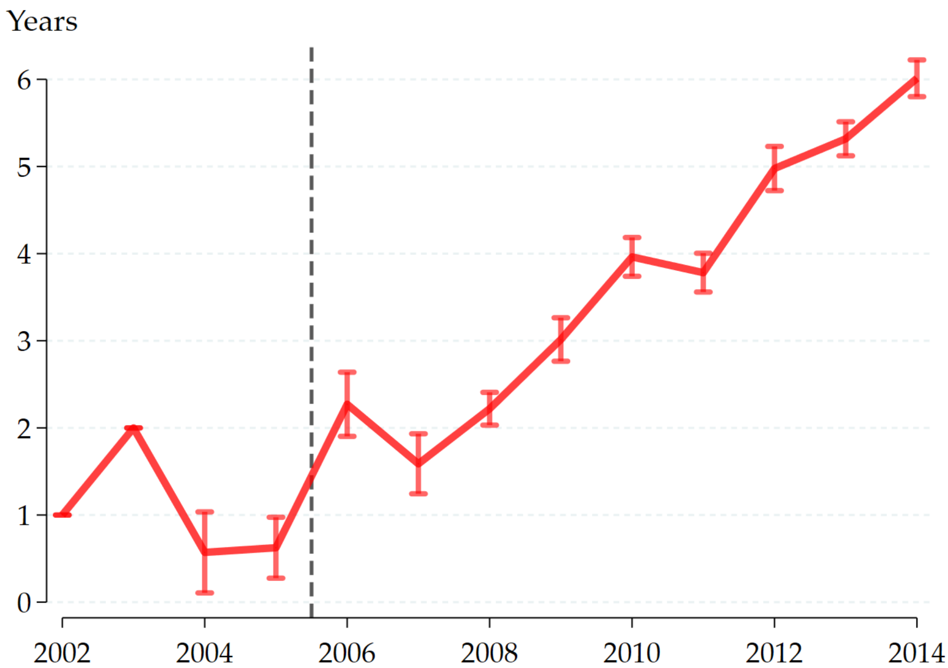
Notes: Data from the Genetic Engineering Appraisal Committee of the Ministry of Environment, Forest and Climate Change. The North zone includes the states of Haryana, Punjab, and Rajasthan; the Central zone includes the states of Gujarat, Madhya Pradesh, Maharashtra, and Odisha; the South zone includes the states of Andhra Pradesh, Karnataka, Tamil Nadu, and Telangana (formed in 2014).

Figure A8. Evolution of Applications to Conduct Field Trials by Zone



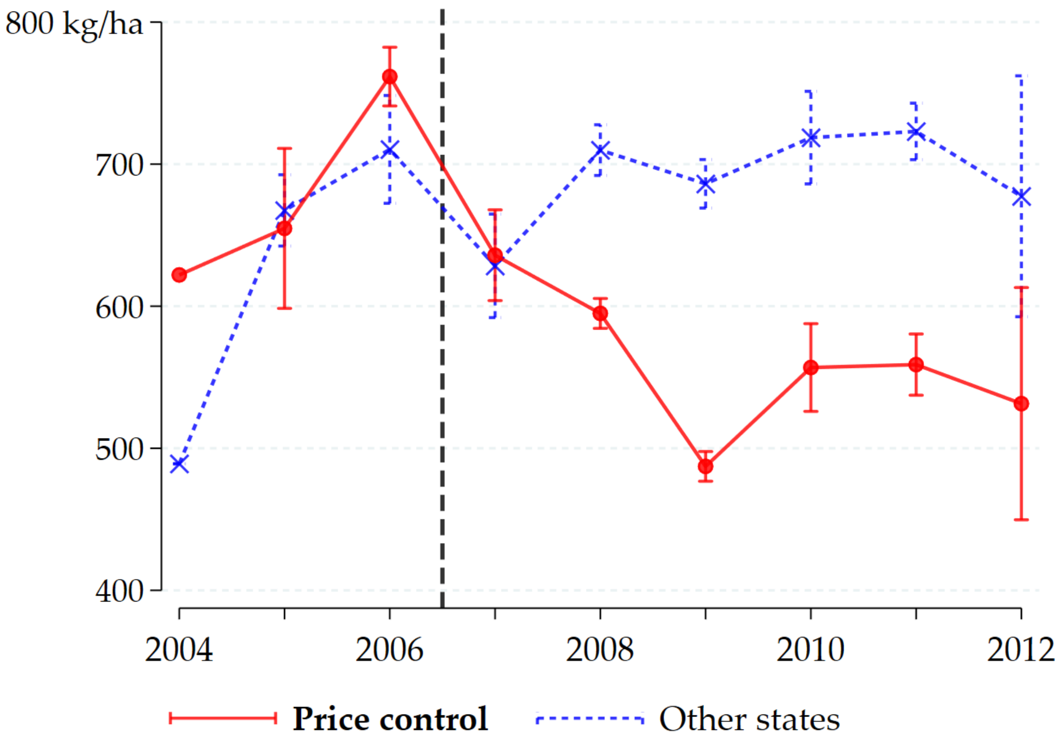
Notes: Data from the Genetic Engineering Appraisal Committee of the Ministry of Environment, Forest and Climate Change. The North zone includes the states of Haryana, Punjab, and Rajasthan; the Central zone includes the states of Gujarat, Madhya Pradesh, Maharashtra, and Odisha; the South zone includes the states of Andhra Pradesh, Karnataka, Tamil Nadu, and Telangana (formed in 2014).

Figure A9. Evolution of Seed Varietal Age at the Farm Level



Notes: "Seed Varietal Age" is defined as the number of years elapsed since the official market release of the seed variety planted in a certain plot. Data on the year of market release from the Genetic Engineering Appraisal Committee of the Ministry of Environment, Forest and Climate Change. Data on seed variety planted from ICRISAT's Village Dynamics in South Asia. Sample: Bt cotton farmers.

Figure A10. Means of Lint Yields from Agronomic Trials



Notes: Means estimates with 95 percent confidence intervals based on least-squares regressions with standard errors clustered at the company level. Unit of observation: seed variety × trial location. Data digitized from the Bt reports of the All India Coordinated Research Project on Cotton of the Indian Council of Agricultural Research. "**Price control**" includes all trial locations in states with price controls (i.e., Andhra Pradesh, Gujarat, Maharashtra, and Telangana).

Table A2. Excerpts from Monsanto's 10-K Annual Reports

Year	Statement
2004	"We experienced sales improvements in [...] cotton traits in the United States and India." (p. 24)
	"Fiscal year 2004 cotton trait revenues in India improved from the comparable period a year ago primarily because of increased acreage planted with *Bollgard* cotton traits. In the prior year, farmers realized the crop protection benefits of our cotton traits, which were approved in calendar year 2002. As a result, more farmers began using or increased the acreage they planted with *Bollgard* traits in 2004." (p. 34)
2005	"Net sales increased [...] as a result of [...] higher cotton trait revenues in Australia and India." (p. 21)
	"Sales of Bollgard traits in India improved in 2005 because of a significant increase in planted trait acres, increased penetration and new cotton hybrids. Increased acreage and penetration resulted from continued farmer experience and acceptance of our cotton traits." (p. 35)
2006	"Mahyco Monsanto Biotech Ltd. (MMB), a joint venture of our subsidiary Monsanto Holdings Private Limited and MAHYCO Seeds Limited, is currently defending complaints before the Monopoly and Restrictive Trade Practice Commission in India (MRTP), relating to the fees it charges on Bollgard technology. [...] On May 11, 2006, the MRTP concluded that MMB was in violation of law by engaging in restrictive trade practices by charging unreasonable trait fees, granted a temporary injunction and directed MMB not to charge Rupees 900 as a trait fee and to set a reasonable trait fee. Appeal was taken to India's Supreme Court. Pending determination of any appeal, MMB has complied with the directions of the order. MMB has also filed writs with the India Supreme Court challenging the state government orders." (p. 19)
	"Our India cotton business is currently operating under state governmental pricing directives which have increased our collection risk. We will continue to carefully monitor our Indian trade receivables in 2007." (p. 43)
2007	"Our international traits businesses, in particular, will probably continue to face unpredictable regulatory environments that may be highly politicized. We operate in volatile, and often difficult, economic environments. Although we see growth potential in our India cotton business with the ongoing conversion to new hybrids and Bollgard II, this business is currently operating under state governmental pricing directives that we believe limit near-term earnings growth." (p. 42)
2014	"Growth in India's cotton germplasm and traits business continues to be impacted by government controlled pricing and uncertainties in the regulatory approval process for new trait introductions." (p. 33)
2016	"Net sales decreased [...] in fiscal year 2016 compared to fiscal year 2015. (p. 19)
	"The net sales decrease [...] in cotton seed and traits was primarily due to lower average net selling price in India as a result of new government pricing policies. [...] Gross profit for cotton seed and traits decreased $126 million, or 31 percent, compared to the 16 percent decrease in net sales for cotton seed and traits primarily due to the effect on margins from the decline of the India business as a result of new government regulations coupled with higher costs in the United States." (p. 25)
	"India's cotton germplasm and traits business could continue to be significantly impacted by government policies, including controlled pricing and regulatory uncertainties, and we will continue to evaluate our cotton business in India." (p. 34)

Table A3. Excerpts from Interviews of Biotechnology Providers (first panel) and Other Press Coverage (second panel)

Source	Statement
'India: Biotech Companies Say Cotton Seed Price Cap Limiting Research', by Jacob P. Koshy, *Technology Networks*, May 11, 2010	Indian biotechnology companies that manufacture genetically modified cotton seeds say they are struggling to keep research activities afloat since the three top cotton-producing states – Andhra Pradesh, Maharashtra and Gujarat – fixed the prices at which cotton seeds could be sold to farmers four years ago. "I've spent over (Rs)25-30 crore in the last seven years on research and regulatory approvals around our Bt (Bacillus thuringiensis) genes, but with this price cap, I can't negotiate appropriate licensing fees with seed companies and I can't competitively price my seeds. So, we are bleeding," said K.K. Narayanan, managing director of Metahelix Life Sciences Pvt. Ltd, a Bangalore-based crop biotech firm. [...] In July 2009, Metahelix received approval to commercially launch a class of Bt cotton seeds. This year, it was planning to introduce a new class of Bt hybrids. But Narayanan said the price cap has forced them to put the plan on hold. [...] 'It's becoming harder and harder to convince investors of the viability of investing in new Bt seed technology in India, and that's largely because of states constantly meddling with the prices of cotton seeds,' said Narayanan."
Suresh and Rao (2009, p. 298)	"Among Indian seed companies, a few companies are struggling to attain technological capabilities [...]. The interference in the pricing of quality value-added hybrid seeds by state governments is creating a negative impact on indigenous seed companies in their efforts [to develop and commercialize biotech seeds]", M. Ramasami, founder of Rasi Seeds
Pray and Nagarajan (2010, pp. 305–306)	"During our interviews with multinational biotech firms in August 2010, it was clear they are wary of bringing in new GM traits such as drought tolerance or doing any research on traits for India-specific problems until the price-control situation is clarified. [...] Indian companies and Indian branches of multinationals say that as a result of the Bt cotton seed price controls, they are slowing down introduction of new technology, but it is too early to have any numbers to substantiate this claim."
"What's preventing India from achieving USD100 billion textile exports by 2030", by CD Mayee and Bhagirath Choudhary, *The Economic Times*, May 17, 2024	"By arbitrary fixing market price of Bt cotton seeds in favour of seed producer and thwarting technology fee, it has stymied R&D companies to innovate and develop new and high-yielding seed and biotech trait(s) for cotton sector in India. As a result, the cotton sector has not seen any new technology including next generation BG-IIRRF cotton.
"Cotton: Entangled in a web of double-edged MSPs", by CD Mayee and Bhagirath Choudhary, *The Hindu Businessline*, May 27, 2024	"As a result, the cotton sector has not seen any new technology, pest such as pink bollworm (PBW) has developed resistance to age-old Bt cotton, majority of tech related projects have been discontinued, and no new investment in research and integrated resistance management (IRM) in the country. The breeding of new varieties of cotton is in shamble with no improved genetic and trait in the offing."

B Directed Innovation in Cotton Breeding

The pervasive heterogeneity in geo-climatic conditions that span the Indian landscape requires seed development to be locally adapted. This is particularly true for cotton, a field crop that faces a broad spectrum of pest and disease threats across the country and that is grown on various soil types and under varying practices, including different irrigation systems and cropping sequences (Venugopalan et al., 2009; Blaise, 2021).

Historical Evolution of Cotton Varieties in India. Cotton (*Gossypium* species) has been cultivated in India since ancient times, with the earliest evidences of cotton fabric dating back to the Mohenjo-Daro and Harappa civilizations around 3,200 BCE (Gulati and Turner, 1929).[62] Throughout history, Indian cotton has been renowned not only for its significant role in the global market but also for its exceptional diversity. Witness to this, India is the only country that has cultivated all four domesticated *Gossypium* species of the *Malvaceae* family: the Old World diploids *Gossypium arboreum* and *Gossypium herbaceum* (Asian or *desi* cotton)[63] as well as the New World tetraploids *Gossypium barbadense* (Egyptian or Pima cotton) and *Gossypium hirsutum* (American or Upland cotton).[64] Moreover, the range of cotton produced in India includes various fiber qualities, from short to extra-long staple, and micronaire values, from coarse to fine. This diversity reflects the adaptation of cotton varieties to the country's varied growing conditions, agronomic practices, and market demands (Venugopalan et al., 2013).

During the colonial period, only indigenous *desi* cotton was grown with minimal external inputs and hardly any fertilizer application. This traditional practice was dramatically altered within a decade when the American Revolution (1765-1783) disrupted the supply of raw cotton fiber to the United Kingdom, forcing a shift toward cultivating American cotton, whose medium staple length and high fiber strength better suited industrial needs, arising from advancements in high-speed spinning machines. The British textile industry urged the government to promote the cultivation of this cotton species in India, starting with the introduction of the "Bourbon" cultivar from Malta and Mauritius in 1790. These efforts intensified in the early 1800s with extensive trials to adapt American cotton varieties to Indian conditions and to develop high-yielding strains (Sethi, 1960).

A further impetus for developing new cotton varieties came with World War I, which reduced the availability of cotton for the Lancashire textile industry. This led to the establishment of the Indian Central Cotton Committee (ICCC), which funded various research and development programs, resulting in the release of improved cotton cultivars with high yield potential (Sikka and Joshi, 1960). The ICCC was later replaced by the All India Coordinated Cotton Improvement Program (AICCIP – now

[62] For further historical accounts on cotton in India, see Santhanam and Hutchinson (1975), Santhanam and Sundaram (1997), and Santhanam et al. (2016).

[63] Regional *desi* cultivars are also known as Narasapur in Andhra Pradesh, Kala or Wagad cotton in Gujarat, Jayadhar in Karnataka, among many.

[64] Cotton is one of the earliest domesticated non-food crops. Archaeological evidence has revealed that cotton was domesticated independently in both the Old World and the New World: *Gossypium arboreum* in the Indus Valley of present-day Pakistan and India, *Gossypium herbaceum* in Arabia and Syria, *Gossypium barbadense* in Mesoamerica, and *Gossypium hirsutum* in coastal Peru (Brubaker et al., 1999; Zohary et al., 2012).

AICRP), which was responsible for significant historical milestones in hybridization for commercial cultivation, including the development of the extra-long staple cotton "Suvin", the world's first intra-hirsutum cotton hybrid "H4" or "Sankar-4", and the world's first inter-specific (*Gossypium hirsutum* × *Gossypium barbadense*) cotton hybrid "Varalaxmi" (Basu, 1983; Basu and Paroda, 1995).[65]

Despite the introduction of American cotton by the British, the cultivation of short and medium staple Asiatic cottons remained dominant prior to India's independence (Narayanan et al., 2014). Beginning in the 1950s, the share of American cotton rose significantly with the "Grow More Cotton" campaign, which provided incentives for farmers to apply chemical fertilizer to cotton, such as reduced prices and short-term credit, and cotton extension schemes. This rise was largely explained by the fact that American cotton hybrids were more responsive to fertilizer applications both at planting and as top dressing. However, the excessive use of synthetic pyrethroids in the 1990s generated insect pest resistance (Kranthi et al., 2001). As a main result, multiple pesticide applications were required to control bollworm infestations: "spray and pray", as the saying of those days goes.[66]

The decisive shift in species composition in favor of American cotton was brought about by the introduction and rapid adoption of Bt hybrids. By 2019, 97% of the cotton area was under intra-hirsutum hybrids, leaving less than 3% of the area under the Asiatic cotton varieties and confining Egyptian cotton to a limited area in South India (Blaise et al., 2014). This recent trend indicates that nearly all of the spatial variation in cotton cultivation and seed development today is driven by different hybrid varieties, which are the focus of our study.

The Advent of Private-Sector Breeding. The introduction of Bt cotton marked a significant turning point, not only for farmers' agronomic practices, but also for the focus of breeding programs. In this new context, private-sector breeding has come to play a major role in advancing technological improvements to meet the diverse needs of cotton farmers across the country.

The emergence of the private sector as a major contributor to investment in seed and pesticide research in India is a phenomenon that dates back to the 1980s. According to Murugkar et al. (2006), the expansion of private investment in cotton was a result of technological learning through experimentation using public-bred varieties as well as economic reforms that protected the discoveries (i.e., the new hybrids) made by private firms. A key reform for seed developers was the enactment of the Protection of Plant Varieties and Farmers' Rights Act in 2001 (Agrawal, 2019). The major change introduced by this law was that it allowed breeders to protect their developed varieties, granting them exclusive rights to produce and sell their protected varieties for a

[65] H4 was released by the celebrated cotton scientist Chandrakant T. Patel in 1970. Originating from the research station in Surat, Gujarat, it quickly spread across Gujarat, Maharashtra, Madhya Pradesh, Andhra Pradesh, and Karnataka. Inspired by the success of this hybrid, Bhimareddy Hanumareddy Katarki developed Varalaxmi at Dharwad Agricultural University, Karnataka, in 1972. Varalaxmi gained popularity for its superior fiber quality and was subsequently cultivated in Karnataka, Tamil Nadu, Andhra Pradesh, and Maharashtra.

[66] The over-reliance on insecticides and the associated evolution of resistance in insects and pests parallels the history of cotton in the US. This situation ultimately necessitated alternative solutions, culminating in the development and commercialization of genetically modified Bt cotton by Monsanto in the mid-1990s (Razaq et al., 2019).

specified amount of time. In order to be granted this protection, breeders must apply and prove that their varieties are distinctive from others in the market, stable, and novel.

Since then, private companies have largely taken the lead in hybrid development, while the public sector has focused on varietal development (Suresh et al., 2014).[67] Unlike a cotton "hybrid variety", which is created by crossing two different parent lines or varieties, a cotton "open-pollinated variety" (OPV) is a pure inbred line. Because it is self-pollinated and can reproduce true-to-type, farmers can re-use the seeds of an OPV for three to four seasons without any significant loss of vigor. This natural reusability, coupled with the longer time required to develop crop OPVs compared to hybrids, contributes to rendering varieties a less lucrative investment for private firms in this context.[68]

Evidence of Geography-Specific Seed Development. Faced with the many sources of heterogeneity in growing conditions mentioned above and a rising market size, cotton seed-producing companies in India tend to target their breeding programs to specific zones and segments, such as irrigated versus rain-fed areas, water storage versus no storage, and resistance to specific pests like whitefly versus pink bollworm. Despite the benefits of this segmentation, insights from qualitative interviews suggest that companies generally avoid focusing on too niche segments to ensure sufficient and consistent market demand.

We confirm this observation using our administrative data on seed companies and Bt varieties. Our descriptive analysis begins by documenting that cotton seed development is indeed geography-specific. Since the introduction of Bt hybrids, 55% of the varieties were approved for only one zone (Northern, Central or Southern), 38.8% were approved for two zones (typically, Central and Southern, which share important similarities in terms of cotton growing practices, such as being mostly rain-fed), and only 6.2% were approved for all three zones.

Interestingly, seed companies headquartered in a state s innovate more within their own state zone $z(s)$. We establish this second stylized fact by constructing a rectangularized dataset at the company-zone level, which records the number of seed varieties approved each year from 2002 to 2015. We then regress this count on an indicator variable that equals one if the company's zone matches the product's zone and zero otherwise, using a Poisson pseudo-maximum likelihood estimation procedure. The estimates in Table B1 show that a company headquartered in a certain zone is 21% more likely to release a seed variety in that same zone. This relationship is robust to the inclusion of year, company, and seed variety fixed effects. It is even stronger for Central and South zones, with companies being 47% and 37% more likely, respectively, to release a seed variety in their home zone. Conversely, the relationship does not hold for the North zone, which has a relatively smaller market and where fewer companies are headquartered.

From 2012 onward, the responsibility for approving new seed varieties for Bt cotton shifted to state governments. Therefore, we can replicate the previous results on

[67] Also see Kolady et al. (2012) on crops other than cotton.

[68] Given that OPVs became marginal in the context of the Indian cotton market and are supplied only by public institutions, our empirical analysis focuses almost exclusively on hybrids. Throughout the paper, references to cotton "varieties" generally serve as shorthand for "hybrid varieties".

Table B1. Company Location and Target Zone of Seed Varieties Released

	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}\left\{\text{Company zone = Product zone}\right\}$	0.213**	0.213**	0.214**			
	(0.100)	(0.100)	(0.092)			
$\times\,\mathbb{1}\left\{\text{zone = North}\right\}$				-0.143		
				(0.120)		
$\times\,\mathbb{1}\left\{\text{zone = Central}\right\}$					0.477***	
					(0.167)	
$\times\,\mathbb{1}\left\{\text{zone = South}\right\}$						0.370**
						(0.182)
Number of observations	1,764	1,764	1,764	1,764	1,764	1,764
Number of companies	49	49	49	49	49	49
Year fixed effects	✓	✓	✓	✓	✓	✓
Company fixed effects		✓	✓	✓	✓	✓
Product zone fixed effects			✓	✓	✓	✓

Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: company × zone × year. The outcome variable is the number of seed varieties approved. Data on seed varieties from the Genetic Engineering Appraisal Committee of the Ministry of Environment, Forest and Climate Change. Data on company headquarters from the Prowess database of the Centre for Monitoring Indian Economy. All regressions are Poisson pseudo-maximum likelihood regressions with multi-way fixed effects, as indicated in the last three rows of the table. Therefore, the coefficients approximate the percentage change in the count of the outcome variable. Standard errors clustered at the company level in parentheses.

this subset of approvals, using a rectangularized dataset at the company-state level. Table B2 confirms the importance of geography-specific innovation in this context: according to the most conservative specification in Column (4), controlling for year, company, and product-state fixed effects, seed companies headquartered in a state s are 31% more likely to release a seed variety in that same s.

Table B2. Company Location and Target Zone of Seed Varieties Released

	(1)	(2)	(3)	(4)
$\mathbb{1}\left\{ \text{Company state} = \text{Product state} \right\}$	0.732***	0.691***	0.577***	0.308***
	(0.142)	(0.124)	(0.100)	(0.118)
Number of observations	1,786	1,786	1,786	1,786
Number of companies	43	43	43	43
Year fixed effects	✓	✓	✓	✓
Company fixed effects		✓	✓	✓
Product zone fixed effects			✓	✓
Product state fixed effects				✓

Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: company × state × year. The outcome variable is the number of seed varieties approved. Data on seed varieties from the Genetic Engineering Appraisal Committee of the Ministry of Environment, Forest and Climate Change. Data on company headquarters from the Prowess database of the Centre for Monitoring Indian Economy. All regressions are Poisson pseudo-maximum likelihood regressions with multi-way fixed effects, as indicated in the last four rows of the table. Therefore, the coefficients approximate the percentage change in the count of the outcome variable. Standard errors clustered at the company level in parentheses.

As a result of this differential research effort and, possibly, of complementary investments in distribution and marketing, seed sales appear to be geography-specific too. Using a similar strategy as above, we rectangularize our four-state panel survey data at the farmer-company level. Then, we identify farmers who purchase their seeds from companies headquartered in the same state and regress the probability of purchasing a seed on this indicator variable using a linear probability model. Living in the same state as a certain company significantly increases the probability of a household purchasing a Bt cotton seed from such a company by 42% (Table B3 – top panel). A similar relationship is observed when focusing on non-Bt, conventional cotton seeds (Table B3 – bottom panel), confirming that this preference is a general phenomenon and is likely not due to the introduction of Bt.

Table B3. Company Location and Farmer Seed Choice

	(1)	(2)	(3)	(4)
Bt cotton seeds				
$\mathbb{1}\left\{\text{Farm state = Company state}\right\}$	0.027***	0.033***	0.031***	0.031***
	(0.005)	(0.006)	(0.007)	(0.007)
Sample mean of comparison group	0.065	0.065	0.065	0.065
Number of observations	13,482	13,482	13,482	13,482
Number of households	476	476	476	476
Number of companies	14	14	14	14
Conventional cotton seeds				
$\mathbb{1}\left\{\text{Farm state = Company state}\right\}$	0.037***	0.042***	0.047***	0.047***
	(0.009)	(0.011)	(0.008)	(0.008)
Sample mean of comparison group	0.065	0.065	0.065	0.065
Number of observations	7,228	7,228	7,228	7,228
Number of households	356	356	356	356
Number of companies	13	13	13	13
Survey wave fixed effects	✓	✓	✓	✓
Farm state fixed effects		✓	✓	✓
Company state fixed effects			✓	✓
Company fixed effects				✓

Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: farmer × company × survey wave. Data on seed choice from Kathage and Qaim (2012)'s panel survey. Data on company headquarters from the Prowess database of the Centre for Monitoring Indian Economy. All regressions are least squares with fixed effects, as indicated in the last four rows of the table. Standard errors clustered at the farmer level in parentheses. The outcome is a dummy variable, so that coefficients can be interpreted in terms of percentage points.

C Difference-in-Differences Estimates

C.1 Cluster-Robust Inference

Our empirical setting features panel data where observations can be naturally grouped into a certain number, G, of mutually independent clusters, g. Therefore, we consider statistical inference with regression model errors that are uncorrelated across clusters but arbitrarily correlated within clusters. Specifically, we assume that errors may be correlated across different time periods for a given unit while remaining uncorrelated for different units.

The importance of clustering is particularly salient in difference-in-differences studies with panel data (e.g., with state-by-year observations), where both the regressor of interest (e.g., a state-level policy) and the error term can be highly correlated across time (Bertrand et al., 2004). This issue is commonly addressed through a "cluster-robust" variance estimator, which averages across independent clusters (states in the example above) and provides a consistent estimate of the variance matrix as $G \to \infty$.[69] Such an estimator produces consistent standard errors and test statistics without imposing strong assumptions about the parametric structure of within-cluster error correlation. Yet the underlying asymptotic approximations require that the number of independent clusters sampled from an infinite super-population goes to infinity, rather than just the number of observations. In scenarios with few clusters, conventional cluster-robust standard errors tend to have poor properties, greatly overstate estimator precision, and over-reject true null hypotheses.

C.1.1 Determining the Appropriate Level of Clustering

Despite the need to control for clustered errors having been well established in the theoretical literature and appreciated in applied work, the structure of the clusters is often not known to the econometrician. For instance, in our empirical analysis of technological demand (Section 4.1), we observe individuals in different geographical locations, such as villages, districts, and states, and need to assume at what level the clustering occurs.

Some rules of thumb have become popular due to their wide applicability to the case of nested clusters, e.g., villages within a state (see, e.g., Roth et al., 2023 for the case of difference-in-differences). However, the decision on what to cluster over is not obvious a priori and can affect the validity of inference (MacKinnon et al., 2023). The fundamental trade-off boils down to the following: while clustering at a narrow level (e.g., household or village) can result in over-rejection, clustering too coarsely (e.g., state) can lead to complications arising from having few treated clusters, such as under-coverage of confidence intervals and tests with low power.[70] The latter issues arise because the researcher assumes less information than they actually have.

[69] A "robust" estimator, which naturally extends to balanced clusters, was initially derived by White (1980). It was then generalized to unbalanced clustered data and models with cluster-specific fixed effects Liang and Zeger (1986) and Arellano (1987), respectively. See Hansen and Lee (2019) for a recent attempt to provide a foundation to the large-sample theory of estimation and inference with clustered samples.

[70] These issues can be quite severe as the number of clusters becomes small, as demonstrated by MacKinnon and Webb (2017, 2018). Simulations in Abadie et al. (2023) show that cluster-robust covariance estimators based on excessively coarse clusters can be too large even in settings with many clusters.

To assess the validity of our chosen clusters, we use the modified randomization test developed by Cai (2023). This test particularly fits our application, given that it relies on asymptotics that take the number of coarse clusters (state, in our case) as fixed.[71] We define state as the coarse level of clustering in our data, villages and households as finer sub-clusters, and test them in our event-study regression in Equation 1.

Table C1 reports worst-case p-values from tests involving pairs of these three nested clusters for the two main outcomes of interest in Section 4.1: technology adoption and yields. For adoption, clustering at the village level is found to be appropriate at the 5 percent level, as the tests in Columns (1-2) reject the null hypotheses that the finer level of clustering, i.e., household, is correct against the alternative that coarser levels of clustering are better, while the test in Column (3) fails to reject that village clustering is more appropriate than state clustering. A similar conclusion is reached for yields: the high p-values suggest that state may be an unnecessarily coarse level, as finer clusters (i.e., households and villages) are approximately independent.

Table C1. Tests of Levels of Clustering

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome:	**Technology Adoption**			**Yields**		
Null hypothesis:	$i \to v$	$i \to s$	$v \to s$	$i \to v$	$i \to s$	$v \to s$
	0.003	0.004	0.504	0.963	0.600	0.351

Notes: p-values from worst-case randomization tests as in Cai (2023). The notation in the second column headers, e.g., $i \to v$, indicates that sub-cluster i is tested against coarse cluster v, where i, v, and s index household, village, and state, respectively. The null hypothesis is that sub-clusters i are uncorrelated within cluster v and, therefore, provide a more appropriate level of clustering. The regression specification is the one in Column (5) of Appendix Table C4 for technology adoption and Column (5) of Appendix Table C11 for yields. The treatment definition is the one from the first panel, i.e., using state-level price controls as treatment.

[71] The two other alternatives to test for the level of clustering (i.e., Ibragimov and Müller, 2016 and MacKinnon et al., 2023) both rely on an asymptotic framework with many clusters. Therefore, the test we implement is likely to be more conservative.

C.2 Auxiliary Results and Robustness

Table C2. Main Treatment Effects on Demand Side – Cluster-Robust Inference

	(1)	(2)	(3)	(4)	(5)
Cotton seed prices					
Price control × Post-2005	-0.460	-0.380	-0.378	-0.361	-0.399
	(0.106)***	(0.084)***	(0.084)***	(0.083)***	(0.093)***
	[0.155]*	[0.076]**	[0.083]**	[0.088]**	[0.092]**
p-values from small-sample adjustments					
Bell and McCaffrey (2002)	0.0769	0.0314	0.0381	0.0729	0.0474
Pustejovsky and Tipton (2018)	0.1271	0.0715	0.0809	0.1267	0.0997
p-values from wild-cluster bootstrap					
with Rademacher weights	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
with Webb (2023) weights	0.0580	0.0540	0.0540	0.0520	0.0730
p-values from cluster-robust *t*-statistic randomization inference	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
p-values from alternative non-standard inference methods					
Bertrand et al. (2004)$^{\neg}$		0.0099	0.0001	<0.0001	<0.0001
Donald and Lang (2007)		<0.0001	<0.0001	0.0007	0.0007
Ferman and Pinto (2019)$^{\neg}$		<0.0001	<0.0001	<0.0001	<0.0001
Number of observations	1,651	1,651	1,651	1,649	1,538
Number of villages	63	63	63	61	60
Number of states	4	4	4	4	4
Adjusted *R*-squared	0.524	0.571	0.572	0.685	0.663
Bt cotton adoption					
Price control × Post-2005	0.145	0.141	0.159	0.160	0.230
	(0.051)***	(0.051)***	(0.053)***	(0.054)***	(0.068)***
	[0.071]	[0.069]	[0.064]*	[0.048]**	[0.047]**
p-values from small-sample adjustments					
Bell and McCaffrey (2002)	0.2885	0.2931	0.2266	0.1460	0.0543
Pustejovsky and Tipton (2018)	0.3359	0.3413	0.2796	0.2044	0.1099
p-values from wild-cluster bootstrap					
with Rademacher weights	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
with Webb (2023) weights	0.0540	0.0670	0.0560	0.0590	0.0610
p-values from cluster-robust *t*-statistic randomization inference	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
p-values from alternative non-standard inference methods					
Bertrand et al. (2004)$^{\neg}$		<0.0001	<0.0001	<0.0001	<0.0001
Donald and Lang (2007)		0.0155	0.0082	0.0031	<0.0001
Ferman and Pinto (2019)$^{\neg}$		0.0200	<0.0001	<0.0001	<0.0001
Number of observations	1,681	1,681	1,681	1,679	1,577
Number of villages	63	63	63	61	60
Number of states	4	4	4	4	4
Adjusted *R*-squared	0.321	0.322	0.343	0.355	0.282
Wave fixed effects	✓	✓$^{\neg}$	✓$^{\neg}$	✓$^{\neg}$	✓$^{\neg}$
State fixed effects		✓	✓	✓	✓
District fixed effects			✓	✓	✓
Village fixed effects				✓	✓
Household fixed effects					✓

Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: household × plot × survey wave. Data from Kathage and Qaim (2012)'s panel survey. All regressions are least squares, as in Equation 1, with fixed effects, as indicated in the last four rows of the table, where, instead of interacting a dummy variable for each survey wave with the treatment, we pool all the waves after 2005 into *Post*. Non-standard inference methods tagged with the $^{\neg}$ superscript do not include wave fixed effects because of their aggregation procedure. Standard errors clustered at the village level in parentheses and clustered at the state level in brackets. All *p*-values below the estimate are for the case of state-level clustering.

Table C3. Treatment Effects on Cotton Seed Prices

	(1)	(2)	(3)	(4)	(5)
State-level price control					
Price control (2002-2003)	0.077	0.034	0.034	0.035	0.031
	(0.083)	(0.076)	(0.078)	(0.076)	(0.091)
	[0.043]	[0.016]	[0.015]	[0.011]*	[0.010]*
Price control (2006-2007)	-0.412	-0.370	-0.370	-0.357	-0.390
	(0.129)***	(0.118)***	(0.118)***	(0.113)***	(0.129)***
	[0.126]**	[0.087]**	[0.093]**	[0.091]**	[0.090]**
Price control (2008-2009)	-0.432	-0.357	-0.353	-0.333	-0.380
	(0.104)***	(0.087)***	(0.087)***	(0.085)***	(0.092)***
	[0.152]*	[0.069]**	[0.076]**	[0.089]**	[0.100]**
Number of observations	1,649	1,649	1,649	1,647	1,536
Number of villages	63	63	63	61	60
Number of states	4	4	4	4	4
Adjusted R-squared	0.524	0.570	0.571	0.684	0.662
Including spillover villages in the treatment group					
Price control (2002-2003)	0.081	0.029	0.031	0.024	-0.009
	(0.092)	(0.088)	(0.089)	(0.085)	(0.102)
	[0.042]	[0.021]	[0.019]	[0.008]*	[0.028]
Price control (2006-2007)	-0.519	-0.460	-0.462	-0.460	-0.482
	(0.132)***	(0.116)***	(0.116)***	(0.109)***	(0.134)***
	[0.125]**	[0.079]**	[0.087]**	[0.082]**	[0.076]***
Price control (2008-2009)	-0.513	-0.413	-0.409	-0.399	-0.436
	(0.110)***	(0.089)***	(0.088)***	(0.083)***	(0.096)***
	[0.178]*	[0.077]**	[0.085]**	[0.104]**	[0.104]**
Number of observations	1,649	1,649	1,649	1,647	1,536
Number of villages	63	63	63	61	60
Number of states	4	4	4	4	4
Adjusted R-squared	0.525	0.574	0.574	0.688	0.665
Dropping spillover villages from the estimation sample					
Price control (2002-2003)	0.083	0.032	0.033	0.024	-0.000
	(0.094)	(0.089)	(0.090)	(0.086)	(0.103)
	[0.040]	[0.020]	[0.018]	[0.012]	[0.019]
Price control (2006-2007)	-0.532	-0.471	-0.474	-0.465	-0.481
	(0.135)***	(0.118)***	(0.119)***	(0.110)***	(0.136)***
	[0.131]**	[0.091]**	[0.099]**	[0.090]**	[0.083]**
Price control (2008-2009)	-0.534	-0.433	-0.429	-0.413	-0.442
	(0.110)***	(0.088)***	(0.087)***	(0.082)***	(0.096)***
	[0.169]*	[0.074]***	[0.084]**	[0.099]**	[0.104]**
Number of observations	1,502	1,502	1,502	1,500	1,412
Number of villages	58	58	58	56	55
Number of states	4	4	4	4	4
Adjusted R-squared	0.524	0.572	0.572	0.691	0.662
Bt fixed effects	✓	✓	✓	✓	✓
Wave fixed effects	✓	✓	✓	✓	✓
State fixed effects		✓	✓	✓	✓
District fixed effects			✓	✓	✓
Village fixed effects				✓	✓
Household fixed effects					✓

Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: household × plot × survey wave. Data from Kathage and Qaim (2012)'s panel survey. All regressions are least squares, as in Equation 1, with fixed effects, as indicated in the last four rows of the table. Standard errors clustered at the village level in parentheses and clustered at the state level in brackets. The comparison period is 2004-2005, i.e., the pre-event survey wave. The outcome is expressed in natural logarithm, so that coefficients approximate percentage changes from that period. Treatment definition in the panel header. A graphical representation of the top-panel main estimates is displayed in Figure 1.

Table C4. Treatment Effects on Bt Cotton Adoption

	(1)	(2)	(3)	(4)	(5)
State-level price control					
Price control (2002-2003)	0.068	0.070	0.062	0.075	0.070
	(0.071)	(0.071)	(0.071)	(0.074)	(0.091)
	[0.051]	[0.050]	[0.046]	[0.055]	[0.053]
Price control (2006-2007)	0.210	0.209	0.220	0.226	0.287
	(0.061)***	(0.061)***	(0.062)***	(0.064)***	(0.080)***
	[0.040]**	[0.038]**	[0.036]***	[0.022]***	[0.026]***
Price control (2008-2009)	0.148	0.145	0.160	0.165	0.234
	(0.069)**	(0.070)**	(0.072)**	(0.075)**	(0.095)**
	[0.075]	[0.074]	[0.069]	[0.056]*	[0.051]**
Sample mean of outcome variable in comparison period	0.355	0.355	0.355	0.355	0.358
Number of observations	1,681	1,681	1,681	1,679	1,577
Number of villages	63	63	63	61	60
Number of states	4	4	4	4	4
Adjusted R-squared	0.321	0.322	0.343	0.356	0.281
Including spillover villages in the treatment group					
Price control (2002-2003)	0.069	0.073	0.061	0.055	0.021
	(0.084)	(0.083)	(0.083)	(0.086)	(0.103)
	[0.061]	[0.060]	[0.053]	[0.052]	[0.041]
Price control (2006-2007)	0.266	0.264	0.278	0.279	0.292
	(0.062)***	(0.063)***	(0.064)***	(0.067)***	(0.088)***
	[0.037]***	[0.034]***	[0.031]***	[0.019]***	[0.040]***
Price control (2008-2009)	0.203	0.199	0.218	0.219	0.257
	(0.078)**	(0.081)**	(0.083)**	(0.087)**	(0.112)**
	[0.071]*	[0.068]*	[0.063]**	[0.048]**	[0.049]**
Sample mean of outcome variable in comparison period	0.355	0.355	0.355	0.355	0.358
Number of observations	1,681	1,681	1,681	1,679	1,577
Number of villages	63	63	63	61	60
Number of states	4	4	4	4	4
Adjusted R-squared	0.324	0.325	0.346	0.359	0.282
Dropping spillover villages from the estimation sample					
Price control (2002-2003)	0.076	0.079	0.067	0.069	0.041
	(0.085)	(0.083)	(0.084)	(0.087)	(0.103)
	[0.065]	[0.065]	[0.057]	[0.061]	[0.049]
Price control (2006-2007)	0.271	0.268	0.283	0.286	0.308
	(0.064)***	(0.065)***	(0.066)***	(0.068)***	(0.089)***
	[0.034]***	[0.031]***	[0.028]***	[0.012]***	[0.025]***
Price control (2008-2009)	0.202	0.197	0.217	0.221	0.265
	(0.080)**	(0.082)**	(0.085)**	(0.088)**	(0.113)**
	[0.075]*	[0.073]*	[0.068]**	[0.053]**	[0.049]**
Sample mean of outcome variable in comparison period	0.353	0.353	0.353	0.353	0.353
Number of observations	1,534	1,534	1,534	1,532	1,453
Number of villages	58	58	58	56	55
Number of states	4	4	4	4	4
Adjusted R-squared	0.317	0.318	0.340	0.354	0.280
Wave fixed effects	✓	✓	✓	✓	✓
State fixed effects		✓	✓	✓	✓
District fixed effects			✓	✓	✓
Village fixed effects				✓	✓
Household fixed effects					✓

Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: household × plot × survey wave. Data from Kathage and Qaim (2012)'s panel survey. All regressions are least squares, as in Equation 1, with fixed effects, as indicated in the last four rows of the table. Standard errors clustered at the village level in parentheses and clustered at the state level in brackets. The comparison period is 2004-2005, i.e., the pre-event survey wave. Treatment definition in the panel header. A graphical representation of the top-panel main estimates is displayed in Figure 2.

Table C5. Treatment Effects on Seed Usage at the Intensive Margin

	(1)	(2)	(3)	(4)	(5)
State-level price control					
Price control (2002-2003)	5.6	38.8	36.3	45.6	25.5
	(73.1)	(62.8)	(63.9)	(64.5)	(67.0)
	[74.9]	[49.2]	[50.3]	[42.0]	[37.9]
Price control (2006-2007)	211.6	180.4	179.5	184.5	144.5
	(84.4)**	(74.2)**	(72.7)**	(78.0)**	(82.8)*
	[143.3]	[111.3]	[112.2]	[137.5]	[145.6]
Price control (2008-2009)	278.4	242.9	235.3	260.1	233.3
	(59.7)***	(55.8)***	(55.4)***	(79.4)***	(92.7)**
	[84.1]**	[45.0]**	[46.6]**	[115.9]	[154.8]
Sample mean of outcome variable in comparison period	605.468	605.468	605.468	605.468	592.708
Number of observations	1,652	1,652	1,652	1,650	1,538
Number of villages	63	63	63	61	60
Number of states	4	4	4	4	4
Adjusted R-squared	0.068	0.153	0.172	0.374	0.413
Including spillover villages in the treatment group					
Price control (2002-2003)	-129.1	-86.8	-88.5	-71.1	-55.0
	(68.5)*	(57.9)	(58.3)	(55.8)	(62.9)
	[46.9]*	[48.9]	[49.1]	[57.8]	[45.7]
Price control (2006-2007)	183.9	141.0	142.2	158.4	149.7
	(104.6)*	(86.4)	(86.5)	(97.4)	(98.5)
	[215.9]	[176.5]	[177.7]	[207.8]	[175.8]
Price control (2008-2009)	236.0	187.0	180.8	226.8	249.4
	(62.6)***	(55.1)***	(55.2)***	(96.3)**	(110.5)**
	[148.3]	[104.8]	[106.4]	[189.3]	[193.9]
Sample mean of outcome variable in comparison period	605.468	605.468	605.468	605.468	592.708
Number of observations	1,652	1,652	1,652	1,650	1,538
Number of villages	63	63	63	61	60
Number of states	4	4	4	4	4
Adjusted R-squared	0.060	0.159	0.172	0.375	0.419
Dropping spillover villages from the estimation sample					
Price control (2002-2003)	-90.7	-49.1	-51.1	-33.2	-30.8
	(67.0)	(56.2)	(56.8)	(53.2)	(62.1)
	[44.6]	[28.8]	[30.3]	[29.3]	[31.6]
Price control (2006-2007)	211.1	166.9	168.5	182.0	155.2
	(103.3)**	(85.2)*	(85.2)*	(96.7)*	(98.5)
	[195.1]	[156.7]	[157.4]	[189.9]	[172.3]
Price control (2008-2009)	272.8	223.6	217.1	259.3	260.3
	(58.8)***	(51.1)***	(51.2)***	(94.8)***	(110.0)**
	[117.3]	[72.1]*	[74.0]*	[163.3]	[185.7]
Sample mean of outcome variable in comparison period	580.181	580.181	580.181	580.181	576.519
Number of observations	1,506	1,506	1,506	1,504	1,415
Number of villages	58	58	58	56	55
Number of states	4	4	4	4	4
Adjusted R-squared	0.077	0.176	0.191	0.420	0.431
Wave fixed effects	✓	✓	✓	✓	✓
State fixed effects		✓	✓	✓	✓
District fixed effects			✓	✓	✓
Village fixed effects				✓	✓
Household fixed effects					✓

Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: household × plot × survey wave. Data from Kathage and Qaim (2012)'s panel survey. The outcome is measured in grams per acre. All regressions are least squares, as in Equation 1, with fixed effects, as indicated in the last four rows of the table. Standard errors clustered at the village level in parentheses and clustered at the state level in brackets. The comparison period is 2004-2005, i.e., the pre-event survey wave. Treatment definition in the panel header.

Table C6. Treatment Effects on Insecticide Expenditures

	(1)	(2)	(3)	(4)	(5)
State-level price control					
Price control (2002-2003)	-220.2	-264.0	-288.7	-310.5	-362.4
	(262.1)	(255.1)	(240.3)	(234.4)	(304.2)
	[520.7]	[535.7]	[516.1]	[515.5]	[523.6]
Price control (2006-2007)	-219.3	-237.3	-258.0	-198.7	-350.7
	(225.7)	(226.2)	(228.0)	(245.0)	(269.9)
	[375.3]	[340.8]	[333.3]	[319.9]	[337.4]
Price control (2008-2009)	-1,018.8	-981.6	-1,001.1	-905.7	-1,092.5
	(232.2)***	(229.2)***	(228.9)***	(229.2)***	(298.7)***
	[697.1]	[679.1]	[675.6]	[670.7]	[657.7]
Sample mean of outcome variable in comparison period	2115.1	2115.1	2115.1	2115.1	2127.4
Number of observations	1,681	1,681	1,681	1,679	1,577
Number of villages	63	63	63	61	60
Number of states	4	4	4	4	4
Adjusted R-squared	0.141	0.263	0.328	0.368	0.396
Including spillover villages in the treatment group					
Price control (2002-2003)	-650.0	-668.5	-670.2	-593.1	-573.4
	(210.0)***	(209.5)***	(205.4)***	(205.1)***	(256.5)**
	[432.8]	[439.1]	[424.5]	[440.8]	[467.3]
Price control (2006-2007)	-640.1	-671.6	-656.2	-632.7	-679.6
	(191.0)***	(188.2)***	(192.0)***	(203.5)***	(236.0)***
	[269.4]*	[265.4]*	[266.0]*	[252.0]*	[282.6]*
Price control (2008-2009)	-898.5	-881.2	-848.7	-799.3	-901.5
	(221.4)***	(225.9)***	(225.4)***	(219.2)***	(273.2)***
	[699.0]	[697.4]	[695.8]	[655.7]	[668.9]
Sample mean of outcome variable in comparison period	2115.1	2115.1	2115.1	2115.1	2127.4
Number of observations	1,681	1,681	1,681	1,679	1,577
Number of villages	63	63	63	61	60
Number of states	4	4	4	4	4
Adjusted R-squared	0.178	0.297	0.325	0.365	0.391
Dropping spillover villages from the estimation sample					
Price control (2002-2003)	-564.5	-586.5	-588.8	-549.1	-548.4
	(221.8)**	(220.0)***	(214.9)***	(214.3)**	(259.5)**
	[502.2]	[505.0]	[488.8]	[500.4]	[513.2]
Price control (2006-2007)	-553.2	-594.6	-576.2	-540.9	-609.9
	(201.7)***	(199.4)***	(203.2)***	(216.6)**	(243.6)**
	[323.8]	[314.1]	[315.2]	[296.9]	[317.7]
Price control (2008-2009)	-1,022.9	-1,008.9	-971.0	-905.5	-1,017.9
	(234.1)***	(236.5)***	(237.2)***	(233.4)***	(276.1)***
	[708.8]	[702.6]	[705.0]	[687.0]	[676.4]
Sample mean of outcome variable in comparison period	2039.6	2039.6	2039.6	2039.6	2059.3
Number of observations	1,534	1,534	1,534	1,532	1,453
Number of villages	58	58	58	56	55
Number of states	4	4	4	4	4
Adjusted R-squared	0.189	0.318	0.347	0.388	0.420
Wave fixed effects	✓	✓	✓	✓	✓
State fixed effects		✓	✓	✓	✓
District fixed effects			✓	✓	✓
Village fixed effects				✓	✓
Household fixed effects					✓

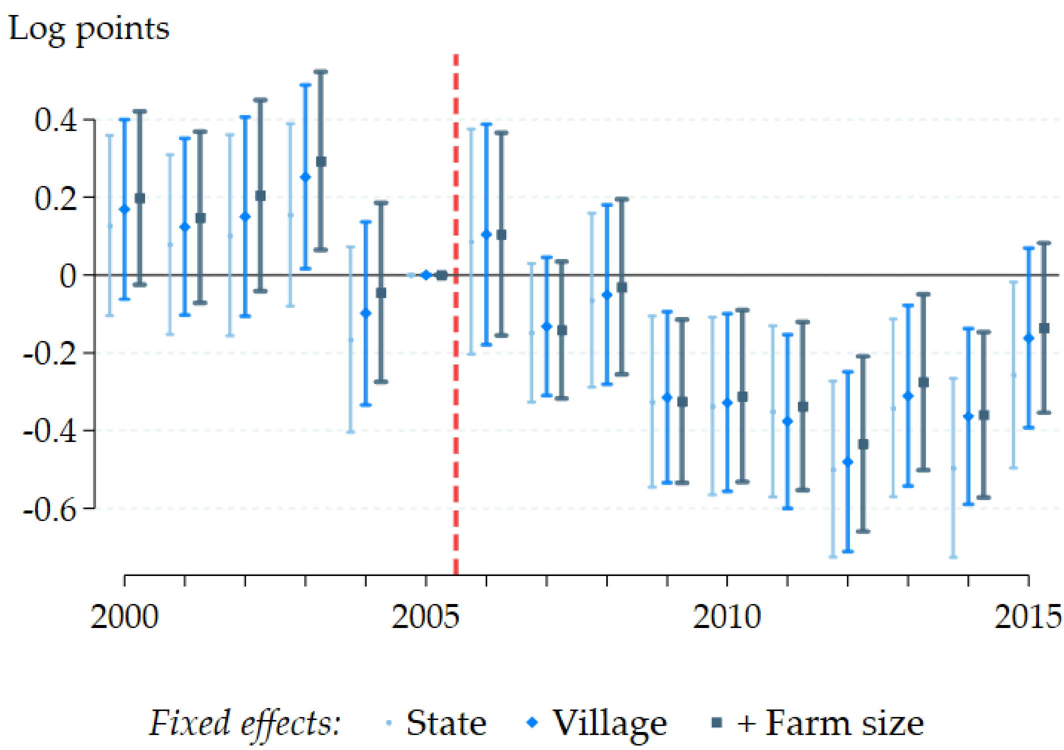
Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: household × plot × survey wave. Data from Kathage and Qaim (2012)'s panel survey. The outcome is measured in Indian rupees (₹). All regressions are least squares, as in Equation 1, with fixed effects, as indicated in the last four rows of the table. Standard errors clustered at the village level in parentheses and clustered at the state level in brackets. The comparison period is 2004-2005, i.e., the pre-event survey wave. Treatment definition in the panel header.

Table C7. Treatment Effects on Insecticide Use Against the Bollworm

	(1)	(2)	(3)	(4)	(5)
State-level price control					
Price control (2002-2003)	-0.133	-0.131	-0.132	-0.131	-0.101
	(0.062)**	(0.062)**	(0.064)**	(0.067)*	(0.079)
	[0.103]	[0.101]	[0.101]	[0.078]	[0.077]
Price control (2006-2007)	0.066	0.078	0.067	0.084	0.096
	(0.094)	(0.091)	(0.091)	(0.090)	(0.104)
	[0.117]	[0.106]	[0.101]	[0.089]	[0.092]
Price control (2008-2009)	-0.326	-0.298	-0.298	-0.279	-0.240
	(0.103)***	(0.097)***	(0.092)***	(0.092)***	(0.114)**
	[0.084]**	[0.067]**	[0.062]**	[0.046]***	[0.047]**
Sample mean of outcome variable in comparison period	0.836	0.836	0.836	0.839	0.843
Number of observations	1,385	1,385	1,385	1,382	1,261
Number of villages	61	61	61	58	57
Number of states	4	4	4	4	4
Adjusted R-squared	0.151	0.181	0.243	0.270	0.292
Including spillover villages in the treatment group					
Price control (2002-2003)	-0.166	-0.168	-0.173	-0.148	-0.123
	(0.077)**	(0.077)**	(0.078)**	(0.079)*	(0.093)
	[0.119]	[0.118]	[0.116]	[0.095]	[0.088]
Price control (2006-2007)	0.059	0.069	0.054	0.074	0.069
	(0.097)	(0.094)	(0.095)	(0.093)	(0.106)
	[0.117]	[0.109]	[0.102]	[0.085]	[0.087]
Price control (2008-2009)	-0.202	-0.181	-0.191	-0.181	-0.150
	(0.121)	(0.116)	(0.112)*	(0.112)	(0.128)
	[0.153]	[0.139]	[0.131]	[0.106]	[0.092]
Sample mean of outcome variable in comparison period	0.836	0.836	0.836	0.839	0.843
Number of observations	1,385	1,385	1,385	1,382	1,261
Number of villages	61	61	61	58	57
Number of states	4	4	4	4	4
Adjusted R-squared	0.157	0.192	0.230	0.259	0.279
Dropping spillover villages from the estimation sample					
Price control (2002-2003)	-0.167	-0.167	-0.181	-0.155	-0.125
	(0.078)**	(0.078)**	(0.079)**	(0.079)*	(0.093)
	[0.119]	[0.119]	[0.114]	[0.090]	[0.087]
Price control (2006-2007)	0.070	0.080	0.059	0.084	0.085
	(0.099)	(0.096)	(0.097)	(0.095)	(0.109)
	[0.129]	[0.120]	[0.112]	[0.096]	[0.097]
Price control (2008-2009)	-0.256	-0.233	-0.253	-0.230	-0.190
	(0.120)**	(0.115)**	(0.112)**	(0.112)**	(0.128)
	[0.107]*	[0.094]*	[0.085]*	[0.062]**	[0.059]**
Sample mean of outcome variable in comparison period	0.824	0.824	0.824	0.826	0.830
Number of observations	1,256	1,256	1,256	1,253	1,158
Number of villages	56	56	56	53	53
Number of states	4	4	4	4	4
Adjusted R-squared	0.165	0.190	0.232	0.259	0.290
Wave fixed effects	✓	✓	✓	✓	✓
State fixed effects		✓	✓	✓	✓
District fixed effects			✓	✓	✓
Village fixed effects				✓	✓
Household fixed effects					✓

Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: household × plot × survey wave. Data from Kathage and Qaim (2012)'s panel survey. All regressions are least squares, as in Equation 1, with fixed effects, as indicated in the last four rows of the table. Standard errors clustered at the village level in parentheses and clustered at the state level in brackets. The comparison period is 2004-2005, i.e., the pre-event survey wave. Treatment definition in the panel header.

Figure C1. Effects on Hired Labor Hours



Notes: Event-study estimates with 95 percent confidence intervals based on least-squares regressions as in Equation 1: year and season fixed effects are included in all the models; additional fixed effects are indicated in the legend below the graph. Standard errors clustered at the village level. Unit of observation: household × parcel × plot × season × year. Data from the Cost of Cultivation/Production Survey. The outcome is expressed in natural logarithm, so that coefficients approximate percentage changes from the pre-treatment comparison period (i.e., 2005). The vertical red line signals the treatment timing. Estimates pooling pre- and post-treatment periods are in Appendix Table C8.

Table C8. Treatment Effects on Labor Hours – Pooled Estimates

	(1)	(2)	(3)	(4)
Total Labor				
Price Control × Post-2005	-0.153	-0.133	-0.137	-0.143
	(0.032)***	(0.031)***	(0.031)***	(0.028)***
	[0.168]	[0.182]	[0.181]	[0.193]
Number of observations	24,542	24,542	24,542	24,541
Number of villages	655	655	655	655
Number of states	10	10	10	10
Adjusted R-squared	0.042	0.110	0.150	0.244
Household Labor				
Price Control × Post-2005	-0.044	-0.018	0.014	0.013
	(0.042)	(0.041)	(0.040)	(0.040)
	[0.117]	[0.113]	[0.127]	[0.125]
Number of observations	24,306	24,306	24,306	24,305
Number of villages	655	655	655	655
Number of states	10	10	10	10
Adjusted R-squared	0.026	0.168	0.227	0.230
Hired Labor				
Price Control × Post-2005	-0.410	-0.342	-0.387	-0.386
	(0.077)***	(0.074)***	(0.072)***	(0.064)***
	[0.316]	[0.300]	[0.290]	[0.311]
Number of observations	23,742	23,742	23,742	23,741
Number of villages	655	655	655	655
Number of states	10	10	10	10
Adjusted R-squared	0.097	0.197	0.217	0.349
Season fixed effects	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓
State fixed effects		✓	✓	✓
Agro-ecological zone fixed effects			✓	✓
Farm size fixed effects				✓

Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: household × parcel × plot × year. Data from the Cost of Cultivation/Production Survey. All regressions are least squares, as in Equation 1, with fixed effects, as indicated in the last four rows of the table, where, instead of interacting a dummy variable for each year with the treatment, we pool all the periods after 2005 into *Post*. Standard errors clustered at the village level in parentheses and clustered at the state level in brackets. The outcome is expressed in natural logarithm, so that coefficients approximate percentage changes from the pre-treatment comparison periods (i.e., 2000-2005). A graphical representation of the main dynamic estimates is displayed in Appendix Figure C1.

Table C9. Treatment Effects on Labor Expenses per Acre – Pooled Estimates

	(1)	(2)	(3)	(4)
Price Control × Post-2005	-0.332	-0.373	-0.415	-0.413
	(0.075)***	(0.073)***	(0.072)***	(0.064)***
	[0.250]	[0.235]	[0.229]	[0.248]
Number of observations	23,763	23,763	23,763	23,762
Number of villages	655	655	655	655
Number of states	10	10	10	10
Adjusted R-squared	0.247	0.319	0.338	0.450
Season fixed effects	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓
State fixed effects		✓	✓	✓
Agro-ecological zone fixed effects			✓	✓
Farm size fixed effects				✓

Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: household × parcel × plot × year. Data from the Cost of Cultivation/Production Survey. All regressions are least squares, as in Equation 1, with fixed effects, as indicated in the last four rows of the table, where, instead of interacting a dummy variable for each year with the treatment, we pool all the periods after 2005 into *Post*. Standard errors clustered at the village level in parentheses and clustered at the state level in brackets. The outcome is expressed in natural logarithm, so that coefficients approximate percentage changes from the pre-treatment comparison periods (i.e., 2000-2005).

Table C10. Treatment Effects on Cotton Costs Per Acre – Pooled Estimates

	(1)	(2)	(3)	(4)
Seeds				
Price control × Post-2005	-0.736	-0.833	-0.830	-0.838
	(0.045)***	(0.046)***	(0.045)***	(0.042)***
	[0.233]**	[0.247]***	[0.257]**	[0.272]**
Number of observations	24,539	24,539	24,539	24,538
Number of villages	655	655	655	655
Number of states	10	10	10	10
Adjusted R-squared	0.335	0.378	0.415	0.472
Total				
Price control × Post-2005	-0.118	-0.231	-0.235	-0.242
	(0.033)***	(0.032)***	(0.031)***	(0.028)***
	[0.135]	[0.143]	[0.146]	[0.159]
Number of observations	24,560	24,560	24,560	24,559
Number of villages	655	655	655	655
Number of states	10	10	10	10
Adjusted R-squared	0.343	0.400	0.433	0.506
Season fixed effects	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓
State fixed effects		✓	✓	✓
Agro-ecological zone fixed effects			✓	✓
Farm size fixed effects				✓

Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: household × parcel × plot × year. Data from the Cost of Cultivation/Production Survey. All regressions are least squares, as in Equation 1, with fixed effects, as indicated in the last five rows of the table, where, instead of interacting a dummy variable for each year with the treatment, we pool all the periods after 2005 into *Post*. Standard errors clustered at the village level in parentheses and clustered at the state level in brackets. The outcome is expressed in natural logarithm, so that coefficients approximate percentage changes from the pre-treatment comparison periods (i.e., 2000-2005). 'Other farm controls' are cultivated area and total value of capital (all logged). A graphical representation of the main dynamic estimates is displayed in Figure 3.

Table C11. Treatment Effects on Farm-Level Cotton Yields

	(1)	(2)	(3)	(4)	(5)
State-level price control					
Price control (2002-2003)	-37.5	-43.1	-51.8	16.4	56.6
	(84.9)	(86.0)	(80.4)	(72.9)	(73.4)
	[104.5]	[108.2]	[107.1]	[112.2]	[129.1]
Price control (2006-2007)	191.8	192.0	191.1	202.4	254.8
	(54.6)***	(53.2)***	(51.2)***	(49.5)***	(62.1)***
	[21.7]***	[28.4]***	[35.5]**	[32.9]***	[29.2]***
Price control (2008-2009)	-22.8	-13.9	-17.6	-19.4	1.3
	(59.6)	(61.2)	(60.9)	(57.5)	(75.1)
	[37.7]	[33.2]	[39.2]	[38.6]	[37.7]
Number of observations	1,655	1,655	1,655	1,653	1,542
Number of villages	63	63	63	61	60
Number of states	4	4	4	4	4
Adjusted R-squared	0.114	0.136	0.263	0.367	0.457
Including spillover villages in the treatment group					
Price control (2002-2003)	101.7	107.6	103.0	106.6	79.1
	(73.7)	(71.1)	(70.9)	(67.3)	(76.5)
	[208.7]	[209.2]	[207.3]	[169.2]	[157.2]
Price control (2006-2007)	226.5	218.6	229.8	228.3	271.0
	(46.1)***	(46.9)***	(44.9)***	(47.3)***	(60.7)***
	[33.6]***	[27.3]***	[31.5]***	[34.5]***	[35.6]***
Price control (2008-2009)	-24.4	-28.3	-20.4	-22.8	-20.4
	(63.5)	(67.0)	(68.3)	(66.4)	(87.6)
	[33.4]	[35.9]	[42.0]	[43.1]	[42.6]
Number of observations	1,655	1,655	1,655	1,653	1,542
Number of villages	63	63	63	61	60
Number of states	4	4	4	4	4
Adjusted R-squared	0.147	0.217	0.263	0.368	0.457
Dropping spillover villages from the estimation sample					
Price control (2002-2003)	58.5	62.2	57.7	84.7	77.4
	(71.7)	(70.6)	(70.4)	(68.5)	(78.1)
	[177.8]	[176.6]	[175.3]	[155.9]	[158.6]
Price control (2006-2007)	235.5	223.8	236.1	238.8	284.3
	(48.7)***	(50.2)***	(47.6)***	(50.1)***	(61.9)***
	[26.6]***	[23.1]***	[27.9]***	[28.3]***	[24.3]***
Price control (2008-2009)	-24.5	-29.4	-21.4	-24.8	-19.7
	(66.2)	(69.8)	(70.7)	(69.2)	(89.1)
	[35.3]	[38.1]	[45.2]	[44.9]	[44.1]
Number of observations	1,508	1,508	1,508	1,506	1,418
Number of villages	58	58	58	56	55
Number of states	4	4	4	4	4
Adjusted R-squared	0.177	0.204	0.262	0.326	0.403
Wave fixed effects	✓	✓	✓	✓	✓
State fixed effects		✓	✓	✓	✓
District fixed effects			✓	✓	✓
Village fixed effects				✓	✓
Household fixed effects					✓

Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: household × plot × survey wave. Data from Kathage and Qaim (2012)'s panel survey. All regressions are least squares, as in Equation 1, with fixed effects, as indicated in the last five rows of the table. Standard errors clustered at the village level in parentheses and clustered at the state level in brackets. The outcome is expressed in kilograms per hectare. The comparison period is 2004-2005, i.e., the pre-event survey wave. Treatment definition in the panel header. A graphical representation of the top-panel main estimates is displayed in Figure 4.

Table C12. Treatment Effects on Cotton Crop Production Statistics – Pooled Estimates

	(1)	(2)	(3)
Cotton acreage			
Price control × Post (2006-2008)	0.470	0.302	0.476
	(0.225)**	(0.183)	(0.121)***
	[0.343]	[0.191]	[0.199]**
Price control × Post (2009-2012)	0.491	0.503	0.781
	(0.212)**	(0.208)**	(0.142)***
	[0.295]	[0.299]	[0.223]***
Price control × Post (2013-2015)	-0.139	-0.283	0.455
	(0.288)	(0.262)	(0.218)**
	[0.674]	[0.649]	[0.279]
Number of observations	3,398	3,398	3,378
Number of villages	250	250	230
Number of states	11	11	10
Adjusted R-squared	0.239	0.299	0.898
Production			
Price control × Post (2006-2008)	0.370	0.338	0.540
	(0.236)	(0.181)*	(0.127)***
	[0.363]	[0.159]*	[0.209]**
Price control × Post (2009-2012)	0.462	0.498	0.685
	(0.220)**	(0.206)**	(0.146)***
	[0.333]	[0.320]	[0.241]**
Price control × Post (2013-2015)	0.045	-0.116	0.343
	(0.273)	(0.248)	(0.222)
	[0.618]	[0.530]	[0.310]
Number of observations	3,277	3,277	3,256
Number of villages	243	243	222
Number of states	11	11	10
Adjusted R-squared	0.226	0.337	0.879
Production per acre			
Price control × Post (2006-2008)	0.172	0.182	0.173
	(0.077)**	(0.059)***	(0.058)***
	[0.263]	[0.203]	[0.201]
Price control × Post (2009-2012)	-0.077	-0.080	-0.064
	(0.054)	(0.053)	(0.049)
	[0.138]	[0.131]	[0.138]
Price control × Post (2013-2015)	-0.175	-0.168	-0.147
	(0.056)***	(0.057)***	(0.060)**
	[0.124]	[0.142]	[0.153]
Number of observations	3,277	3,277	3,256
Number of villages	243	243	222
Number of states	11	11	10
Adjusted R-squared	0.166	0.468	0.574
Season fixed effects	✓	✓	✓
Year fixed effects	✓	✓	✓
State fixed effects		✓	✓
District fixed effects			✓

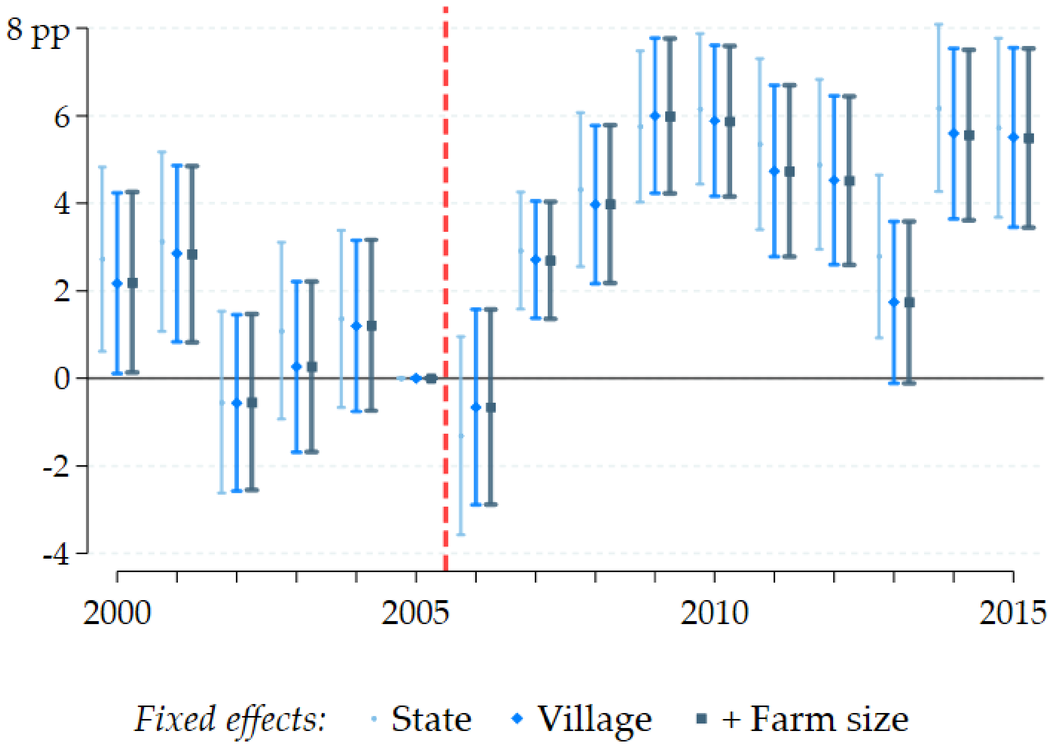
Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: district × year. Data from the Crop Production Statistics. Sample: cotton-growing states. All regressions are least squares, as in Equation 2, with fixed effects (indicated in the last four rows of the table) and standard errors clustered at the company level (in parentheses), where, instead of interacting a dummy variable for each year with the treatment, we pool the periods after 2005 into three *Post* time groups. Standard errors clustered at the village level in parentheses and clustered at the state level in brackets. The outcome is expressed in natural logarithm, so that coefficients approximate percentage changes from the pre-treatment comparison periods (i.e., 2000-2005). A graphical representation of the main estimates in the first and second panels is displayed in Figure 5a and 5b, respectively.

Table C13. Heterogeneous Treatment Effects on Cotton Crop Production Statistics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Outcome:	*Cotton acreage*			*Production*			*Production per acre*		
Below first tercile									
Price control × Post-2005	-0.063	-0.093	0.378	0.180	0.123	0.395	0.064	0.050	0.066
	(0.238)	(0.218)	(0.140)***	(0.240)	(0.210)	(0.154)**	(0.050)	(0.050)	(0.048)
	[0.371]	[0.348]	[0.139]**	[0.333]	[0.277]	[0.190]*	[0.104]	[0.118]	[0.124]
Price control × Post-2005 × Low cotton suitability	0.799	0.759	0.773	0.024	0.126	0.311	-0.608	-0.528	-0.504
	(0.412)*	(0.386)*	(0.527)	(0.418)	(0.380)	(0.551)	(0.088)***	(0.086)***	(0.084)***
	[0.459]	[0.350]*	[0.177]***	[0.420]	[0.278]	[0.253]	[0.080]***	[0.076]***	[0.078]***
Number of observations	3,217	3,217	3,199	3,096	3,096	3,077	3,096	3,096	3,077
Number of districts	239	239	221	232	232	213	232	232	213
Number of states	11	11	10	11	11	10	11	11	10
Adjusted R-squared	0.320	0.372	0.901	0.304	0.417	0.883	0.175	0.477	0.574
Below first quartile									
Price control × Post-2005	-0.073	-0.057	0.427	0.117	0.128	0.419	0.018	0.027	0.042
	(0.228)	(0.206)	(0.135)***	(0.227)	(0.196)	(0.144)***	(0.049)	(0.049)	(0.047)
	[0.390]	[0.364]	[0.119]***	[0.367]	[0.292]	[0.145]**	[0.111]	[0.134]	[0.147]
Price control × Post-2005 × Low cotton suitability	1.261	0.909	1.016	0.458	0.249	0.564	-0.603	-0.565	-0.514
	(0.528)**	(0.551)	(0.794)	(0.549)	(0.549)	(0.830)	(0.100)***	(0.099)***	(0.102)***
	[0.435]**	[0.392]**	[0.263]***	[0.437]	[0.309]	[0.329]	[0.101]***	[0.133]***	[0.163]**
Number of observations	3,217	3,217	3,199	3,096	3,096	3,077	3,096	3,096	3,077
Number of districts	239	239	221	232	232	213	232	232	213
Number of states	11	11	10	11	11	10	11	11	10
Adjusted R-squared	0.303	0.374	0.901	0.283	0.411	0.883	0.177	0.476	0.574
Season fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
State fixed effects		✓	✓		✓	✓		✓	✓
District fixed effects			✓			✓			✓

Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: district × season × year. Data on cotton acreage and output from the Crop Production Statistics; data on potential yields from the Global Agro-Ecological Zoning database, version 4, of the Food and Agriculture Organization of the United Nations. Sample: cotton-growing states. All regressions are least squares, as in Equation 1, with fixed effects, as indicated in the last four rows of the table, where, instead of interacting a dummy variable for each year with the treatment, we pool all the periods after 2005 into *Post*. Standard errors clustered at the district level in parentheses and clustered at the state level in brackets. The outcome is expressed in natural logarithm, so that coefficients approximate percentage changes from the pre-treatment comparison periods (i.e., 2000-2005). 'Low cotton suitability' is equal to one if the agro-climatic potential yield for cotton in the time period 1981-2010 with an available water content of 200 mm/m under irrigation conditions and high input level is below the sample statistics specified in the panel header.

Figure C2. Effects on Probability of Cultivating Cotton



Notes: Event-study estimates with 95 percent confidence intervals based on least-squares regressions as in Equation 1: year and season fixed effects are included in all the models; additional fixed effects are indicated in the legend below the graph. Standard errors clustered at the village level. Unit of observation: household × season × year. Data from the Cost of Cultivation/Production Survey. The outcome is a dummy variable, so that coefficients can be interpreted in terms of percentage points. The vertical red line signals the treatment timing. Estimates pooling pre- and post-treatment periods are in Appendix Table C14.

Table C14. Treatment Effects on Cotton Farming – Pooled Estimates

	(1)	(2)	(3)	(4)
Probability of cultivating cotton (0/1)				
Price control × Post-2005	0.048	0.034	0.034	0.034
	(0.006)***	(0.005)***	(0.005)***	(0.005)***
	[0.037]	[0.035]	[0.035]	[0.035]
Sample mean of outcome variable in comparison period	0.138	0.138	0.138	0.138
Number of observations	114,390	114,390	114,390	114,390
Number of villages	655	655	655	655
Number of states	10	10	10	10
Adjusted R-squared	0.163	0.203	0.301	0.302
Area cultivated with cotton (logged)				
Price control × Post-2005	0.079	0.026	0.023	0.023
	(0.037)**	(0.038)	(0.038)	(0.032)
	[0.117]	[0.111]	[0.108]	[0.130]
Number of observations	16,431	16,431	16,431	16,431
Number of villages	655	655	655	655
Number of states	10	10	10	10
Adjusted R-squared	0.031	0.075	0.114	0.435
Season fixed effects	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓
State fixed effects		✓	✓	✓
Agro-ecological zone fixed effects			✓	✓
Farm size fixed effects				✓

Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: household × parcel × plot × year. Data from the Cost of Cultivation/Production Survey. All regressions are least squares, as in Equation 1, with fixed effects (indicated in the last six rows of the table) and standard errors clustered at the household level (in parentheses), where, instead of interacting a dummy variable for each year with the treatment, we pool all the periods after 2005 into *Post*. The outcomes are expressed as follows: 'Probability of cultivating cotton' is a binary variable, while 'Area cultivated with cotton' is in natural logarithm Therefore, coefficients can be interpreted in terms of percentage points or percentage change from the pre-treatment comparison periods (i.e., 2000-2005), respectively. A graphical representation of the main dynamic estimates is displayed in Figure C2.

Table C15. Treatment Effects on Probability of Cultivating Cotton – Panel Survey Sample

	(1)	(2)	(3)	(4)	(5)
State-level price control					
Price control (2002-2003)	-0.052	-0.049	-0.047	-0.050	-0.044
	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)
	[0.038]	[0.036]	[0.037]	[0.036]	[0.039]
Price control (2006-2007)	0.221	0.221	0.221	0.212	0.226
	(0.080)***	(0.080)***	(0.080)***	(0.080)**	(0.078)***
	[0.048]**	[0.049]**	[0.049]**	[0.046]**	[0.043]**
Price control (2008-2009)	0.016	0.026	0.042	0.077	0.118
	(0.044)	(0.043)	(0.047)	(0.049)	(0.050)**
	[0.032]	[0.028]	[0.030]	[0.025]*	[0.023]**
Sample mean of outcome variable in comparison period	0.956	0.956	0.956	0.956	0.956
Number of observations	1,190	1,190	1,190	1,189	1,174
Number of villages	58	58	58	57	56
Number of states	4	4	4	4	4
Adjusted R-squared	0.141	0.155	0.172	0.244	0.219
Including spillover villages in the treatment group					
Price control (2002-2003)	-0.069	-0.066	-0.063	-0.062	-0.065
	(0.036)*	(0.036)*	(0.037)*	(0.036)*	(0.036)*
	[0.036]	[0.038]	[0.039]	[0.039]	[0.036]
Price control (2006-2007)	0.175	0.175	0.176	0.169	0.178
	(0.080)**	(0.081)**	(0.080)**	(0.080)**	(0.078)**
	[0.090]	[0.091]	[0.091]	[0.086]	[0.085]
Price control (2008-2009)	0.015	0.025	0.043	0.066	0.103
	(0.048)	(0.049)	(0.053)	(0.053)	(0.054)*
	[0.032]	[0.026]	[0.029]	[0.028]	[0.030]**
Sample mean of outcome variable in comparison period	0.956	0.956	0.956	0.956	0.956
Number of observations	1,190	1,190	1,190	1,189	1,174
Number of villages	58	58	58	57	56
Number of states	4	4	4	4	4
Adjusted R-squared	0.126	0.142	0.160	0.231	0.204
Dropping spillover villages from the estimation sample					
Price control (2002-2003)	-0.067	-0.066	-0.064	-0.063	-0.062
	(0.037)*	(0.037)*	(0.037)*	(0.037)*	(0.037)*
	[0.038]	[0.038]	[0.039]	[0.039]	[0.039]
Price control (2006-2007)	0.201	0.202	0.202	0.194	0.205
	(0.076)**	(0.076)**	(0.075)***	(0.076)**	(0.074)***
	[0.066]*	[0.067]*	[0.067]*	[0.063]*	[0.060]**
Price control (2008-2009)	0.015	0.028	0.046	0.073	0.113
	(0.049)	(0.049)	(0.053)	(0.052)	(0.053)**
	[0.033]	[0.028]	[0.031]	[0.026]*	[0.024]**
Sample mean of outcome variable in comparison period	0.953	0.953	0.953	0.953	0.953
Number of observations	1,115	1,115	1,115	1,114	1,110
Number of villages	54	54	54	53	53
Number of states	4	4	4	4	4
Adjusted R-squared	0.129	0.144	0.164	0.226	0.213
Wave fixed effects	✓	✓	✓	✓	✓
State fixed effects		✓	✓	✓	✓
District fixed effects			✓	✓	✓
Village fixed effects				✓	✓
Household fixed effects					✓

Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: household × survey wave. Data from Kathage and Qaim (2012)'s panel survey. Sample: households interviewed in Wave 1 (2002-2003). All regressions are least squares, as in Equation 1, with fixed effects, as indicated in the last five rows of the table. Standard errors clustered at the village level in parentheses and clustered at the state level in brackets. The comparison period is 2004-2005, i.e., the pre-event survey wave. Treatment definition in the panel header.

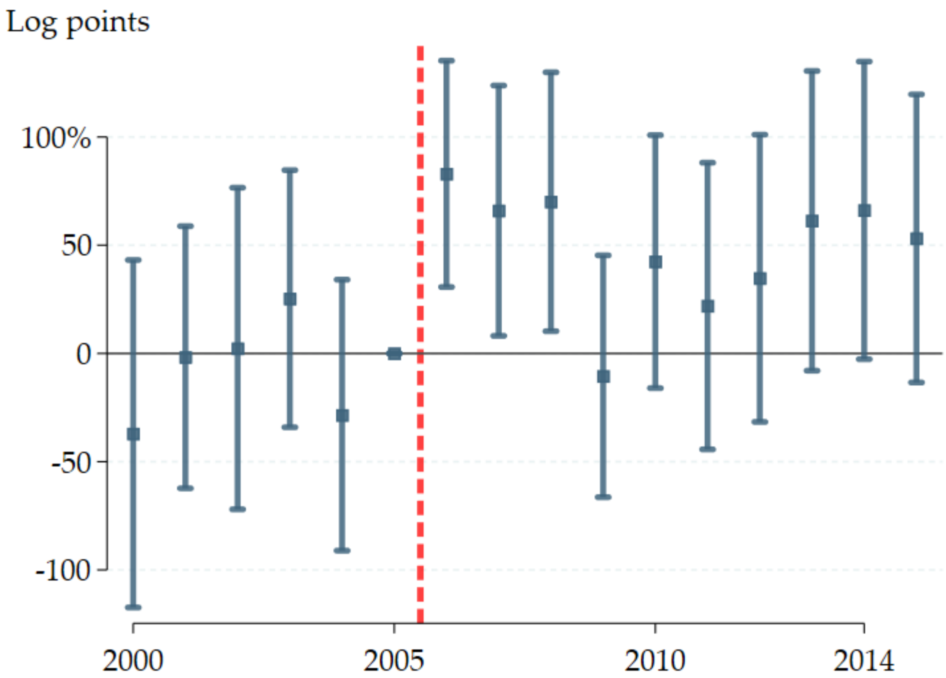
Table C16. Treatment Effects on Cotton Area – Panel Survey Sample

	(1)	(2)	(3)	(4)	(5)
State-level price control					
Price control (2002-2003)	-0.125	-0.121	-0.079	-0.063	-0.083
	(0.094)	(0.089)	(0.090)	(0.089)	(0.084)
	[0.141]	[0.141]	[0.143]	[0.140]	[0.134]
Price control (2006-2007)	-0.114	-0.059	-0.103	-0.059	-0.055
	(0.119)	(0.108)	(0.108)	(0.110)	(0.104)
	[0.181]	[0.164]	[0.165]	[0.170]	[0.144]
Price control (2008-2009)	-0.134	-0.077	-0.132	-0.021	0.031
	(0.133)	(0.115)	(0.114)	(0.126)	(0.125)
	[0.124]	[0.152]	[0.163]	[0.174]	[0.142]
Sample mean of outcome variable in comparison period	1.485	1.485	1.485	1.485	1.485
Number of observations	1,112	1,112	1,112	1,108	1,088
Number of villages	58	58	58	54	53
Number of states	4	4	4	4	4
Adjusted R-squared	0.045	0.143	0.213	0.322	0.604
Including spillover villages in the treatment group					
Price control (2002-2003)	-0.058	-0.089	-0.067	-0.054	-0.125
	(0.092)	(0.091)	(0.097)	(0.096)	(0.085)
	[0.133]	[0.127]	[0.131]	[0.129]	[0.122]
Price control (2006-2007)	-0.111	-0.035	-0.064	-0.038	-0.076
	(0.126)	(0.113)	(0.113)	(0.115)	(0.108)
	[0.175]	[0.158]	[0.157]	[0.158]	[0.135]
Price control (2008-2009)	-0.149	-0.063	-0.096	0.008	0.035
	(0.147)	(0.127)	(0.122)	(0.138)	(0.139)
	[0.113]	[0.148]	[0.155]	[0.172]	[0.152]
Sample mean of outcome variable in comparison period	1.485	1.485	1.485	1.485	1.485
Number of observations	1,112	1,112	1,112	1,108	1,088
Number of villages	58	58	58	54	53
Number of states	4	4	4	4	4
Adjusted R-squared	0.070	0.155	0.212	0.322	0.605
Dropping spillover villages from the estimation sample					
Price control (2002-2003)	-0.080	-0.097	-0.076	-0.061	-0.118
	(0.094)	(0.095)	(0.100)	(0.099)	(0.088)
	[0.141]	[0.141]	[0.144]	[0.141]	[0.134]
Price control (2006-2007)	-0.118	-0.051	-0.080	-0.047	-0.073
	(0.129)	(0.116)	(0.115)	(0.118)	(0.110)
	[0.186]	[0.165]	[0.165]	[0.170]	[0.144]
Price control (2008-2009)	-0.152	-0.079	-0.112	-0.001	0.035
	(0.149)	(0.129)	(0.124)	(0.139)	(0.139)
	[0.127]	[0.162]	[0.169]	[0.183]	[0.156]
Sample mean of outcome variable in comparison period	1.465	1.465	1.465	1.465	1.465
Number of observations	1,043	1,043	1,043	1,039	1,030
Number of villages	54	54	54	50	50
Number of states	4	4	4	4	4
Adjusted R-squared	0.068	0.152	0.213	0.327	0.621
Wave fixed effects	✓	✓	✓	✓	✓
State fixed effects		✓	✓	✓	✓
District fixed effects			✓	✓	✓
Village fixed effects				✓	✓
Household fixed effects					✓

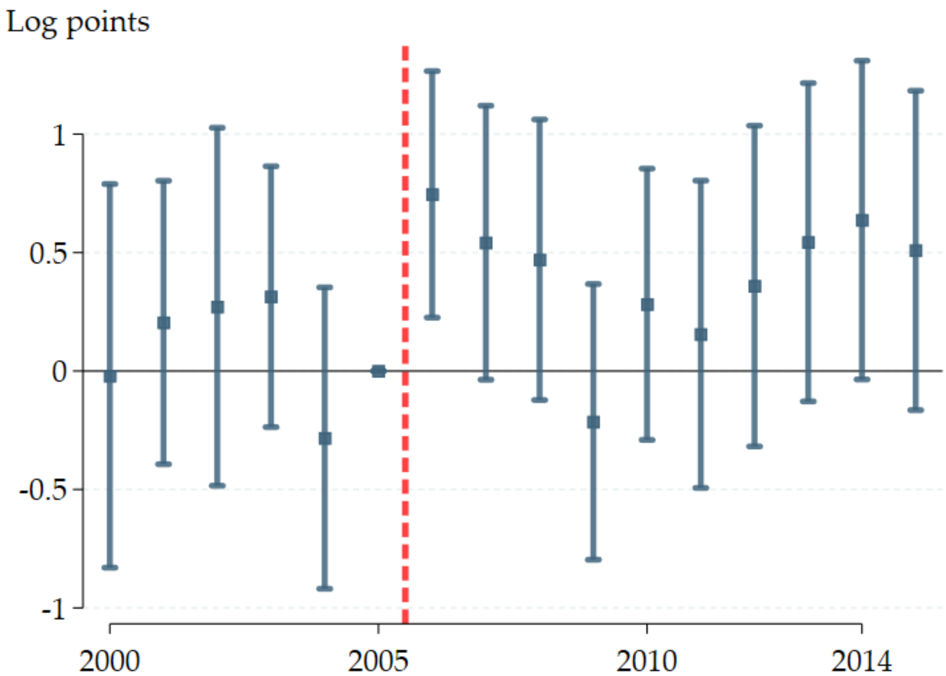
Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: household × survey wave. Data from Kathage and Qaim (2012)'s panel survey. Sample: households interviewed in Wave 1 (2002-2003). All regressions are least squares, as in Equation 1, with fixed effects (indicated in the last four rows of the table) and standard errors clustered at the household level (in parentheses). The comparison period is 2004-2005, i.e., the pre-event survey wave. Treatment definition in the panel header.

Figure C3. Effects on Cotton Seed Sales

(a) Compared to Other Agricultural Inputs



(b) Compared to Other Seeds



Notes: Event-study estimates with 95 percent confidence intervals based on least-squares regressions with company and year fixed effects as in Equation 2. The comparison group is in the subfigure caption. Standard errors clustered at the company level. Unit of observation: company × product × year. Data from the Prowess database of the Centre for Monitoring Indian Economy. The outcome is expressed in natural logarithm, so that coefficients approximate percentage changes from the pre-treatment comparison period (i.e., 2005). The vertical red line signals the treatment timing. Estimates pooling pre- and post-treatment periods are in Appendix Table C17.

Table C17. Treatment Effects on Cotton Seed Sales – Pooled Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome:	*Sale values*		*Quantities*		*Sale values* (sample with non-missing *quantities*)	
Compared to other agricultural inputs						
Cotton seeds × Post (2006-2009)	0.592**	0.663***	0.779**	0.994***	0.903***	0.773***
	(0.235)	(0.207)	(0.355)	(0.286)	(0.331)	(0.262)
Cotton seeds × Post (2010-2012)	0.087	0.040	0.241	0.468	0.107	0.138
	(0.217)	(0.178)	(0.357)	(0.312)	(0.381)	(0.296)
Cotton seeds × Post (2013-2015)	0.586**	0.439**	0.093	-0.351	0.251	-0.378
	(0.247)	(0.219)	(0.613)	(0.616)	(0.648)	(0.620)
Number of observations	6,795	6,795	2,359	2,359	2,222	2,222
Effective number of observations	6,795	6,516	2,359	2,205	2,222	2,072
Number of clusters	1,293	1,014	562	408	530	380
Adjusted R-squared	0.051	0.765	0.004	0.682	0.046	0.755
Compared to other seeds						
Cotton seeds × Post (2006-2009)	0.453*	0.471**	0.800*	0.983***	0.848**	0.756**
	(0.244)	(0.212)	(0.413)	(0.316)	(0.355)	(0.292)
Cotton seeds × Post (2010-2012)	-0.038	-0.066	0.091	0.463	-0.245	0.024
	(0.226)	(0.185)	(0.393)	(0.323)	(0.407)	(0.310)
Cotton seeds × Post (2013-2015)	0.504**	0.406*	-0.320	0.010	-0.120	-0.172
	(0.254)	(0.221)	(0.604)	(0.461)	(0.700)	(0.516)
Number of observations	4,885	4,885	1,477	1,477	1,382	1,382
Effective number of observations	4,885	4,663	1,477	1,349	1,382	1,261
Number of clusters	989	767	409	281	382	261
Adjusted R-squared	0.054	0.723	0.016	0.664	0.035	0.726
Year fixed effects	✓	✓	✓	✓	✓	✓
Company fixed effects		✓		✓		✓
Company × year fixed effects						

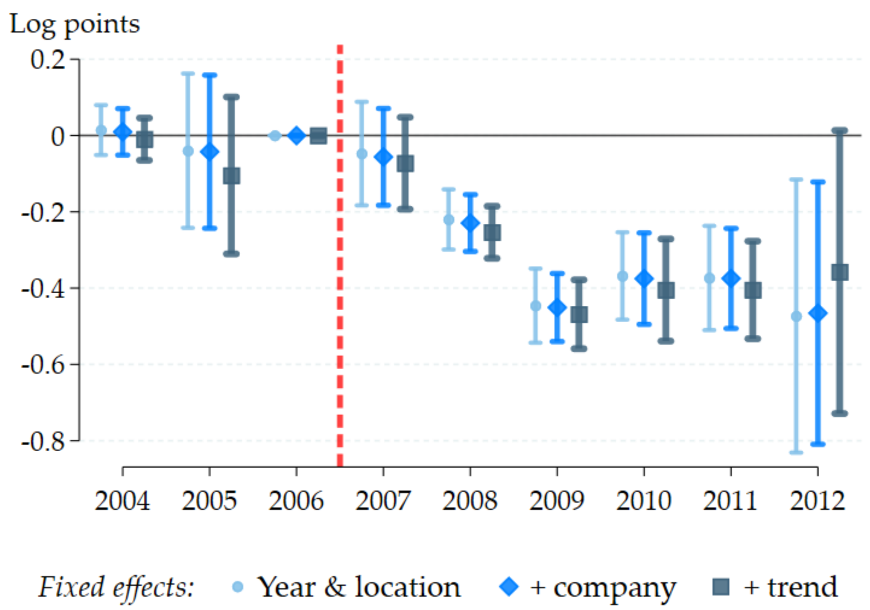
Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: company × product × year. Data from the Prowess database of the Centre for Monitoring Indian Economy. Estimation sample in the panel header. All regressions are least squares, as in Equation 2, with fixed effects (indicated in the last three rows of the table) and standard errors clustered at the company level (in parentheses), where, instead of interacting a dummy variable for each year with the treatment, we pool the periods after 2005 into three *Post* time groups. The outcome is expressed in natural logarithm, so that coefficients approximate percentage changes from the pre-treatment comparison periods (i.e., 2000-2005). A graphical representation of the dynamic estimates from Column (2) of the first and second panels is displayed in Figure C3.

Table C18. Treatment Effects on Agronomic-Trial Yields of Cotton Varieties – Pooled Estimates and Cluster-Robust Inference

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Price control × Post short-term (2007-2008)	-124.2	-129.3	-155.6	-108.0	-114.8	-130.0	-130.1	-138.1
	(38.5)***	(31.7)***	(30.9)***	(31.8)***	(29.9)***	(32.3)***	(32.4)***	(20.5)***
	[102.7]	[103.5]	[106.4]	[104.6]	[100.6]	[98.3]	[98.1]	[54.7]**
Price control × Post long-term (2009-2012)	-226.5	-222.6	-250.4	-205.6	-208.8	-218.0	-218.2	-222.7
	(30.6)***	(24.3)***	(22.3)***	(27.3)***	(24.2)***	(27.9)***	(28.2)***	(18.7)***
	[69.5]***	[82.1]**	[82.6]**	[78.2]**	[77.3]**	[77.4]**	[76.8]**	[43.3]***
p-values from small-sample adjustments								
Bell and McCaffrey (2002)	0.0141	0.0421	0.0819	0.0457	0.0413	0.0374	0.0368	0.0020
Pustejovsky and Tipton (2018)	0.0275	0.0623	0.1070	0.0693	0.0643	0.0598	0.0590	0.0484
p-values from wild-cluster bootstrap								
with Rademacher weights	0.0310	0.0300	0.1170	0.0610	0.0590	0.0520	0.0580	0.1080
with Webb (2023) weights	0.0220	0.0270	0.1100	0.0600	0.0530	0.0520	0.0530	0.1030
p-values from cluster-robust *t*-statistic randomization inference	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Number of observations	2,524	2,524	2,524	2,523	2,523	2,523	2,523	2,523
Number of companies	34	34	34	34	34	34	34	34
Number of states	10	10	10	10	10	10	10	10
Adjusted *R*-squared	0.073	0.211	0.317	0.508	0.518	0.538	0.538	0.566
Year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Variety zone fixed effects		✓	✓	✓	✓	✓	✓	✓
Trial state fixed effects			✓	✓	✓	✓	✓	✓
Trial location fixed effects				✓	✓	✓	✓	✓
Company fixed effects					✓	✓	✓	✓
Company × year fixed effects						✓	✓	✓
Controlling for varietal age at release							✓	✓
Variety fixed effects								✓

Notes: [*]Significant at 10%. [**]Significant at 5%. [***]Significant at 1%. All regressions are least squares, as in Equation 4, with fixed effects (indicated in the last six rows of the table), where, instead of interacting a dummy variable for each year with the treatment, we pool the periods after 2006 into *Post short-term* (2007 and 2008) and *Post long-term* (2009 onwards). Standard errors clustered at the company level in parentheses and clustered at the state level in brackets. The outcome is expressed in kilograms per hectare. A graphical representation of the main dynamic estimates from the first panel is displayed in Figure 6.

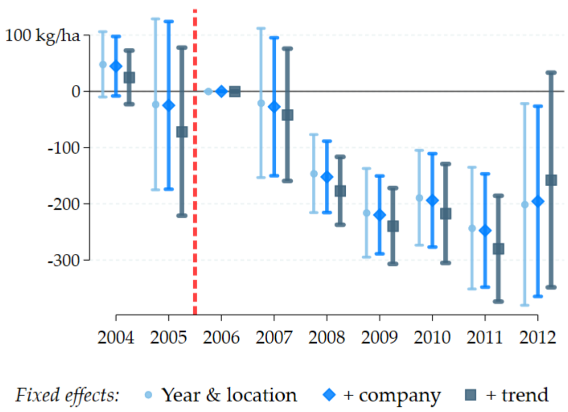
Figure C4. Effects on Agronomic-Trial Yields of Cotton Varieties (Logged)



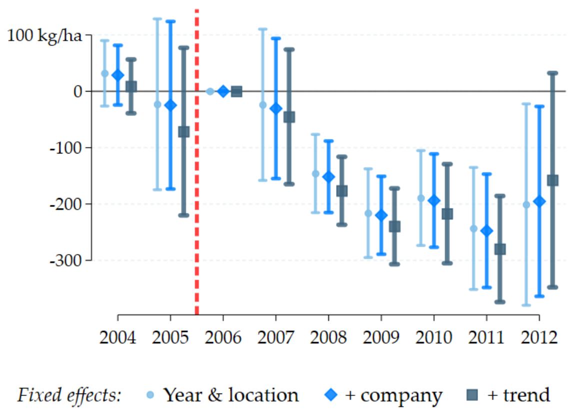
Notes: Event-study estimates with 95 percent confidence intervals based on least-squares regressions as in Equation 4: fixed effects are indicated in the legend below the graph. Standard errors clustered at the company level. Unit of observation: seed variety × company × trial location. Data digitized from the Bt reports of the All India Coordinated Research Project on Cotton of the Indian Council of Agricultural Research. The outcome is expressed in natural logarithm, so that coefficients approximate percentage changes from the pre-treatment comparison period (i.e., 2006). The vertical red line signals the treatment timing. Estimates on the untransformed outcome (in kilograms per hectare) are in Figure 6.

Figure C5. Effects on Agronomic-Trial Yields of Cotton Varieties – Robustness to Distribution Moments

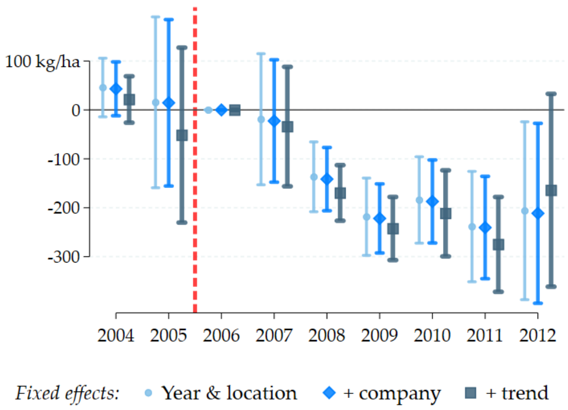
(a) Average (same as Figure 6)



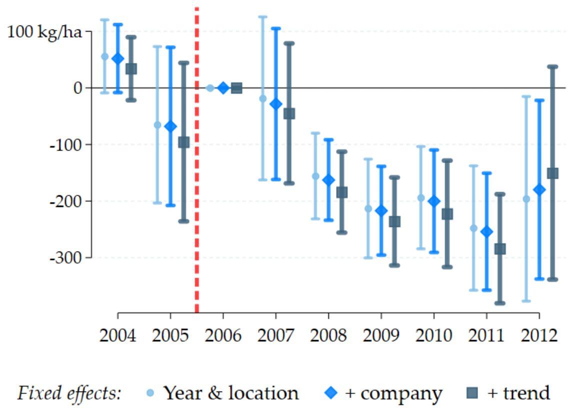
(b) Median



(c) Minimum



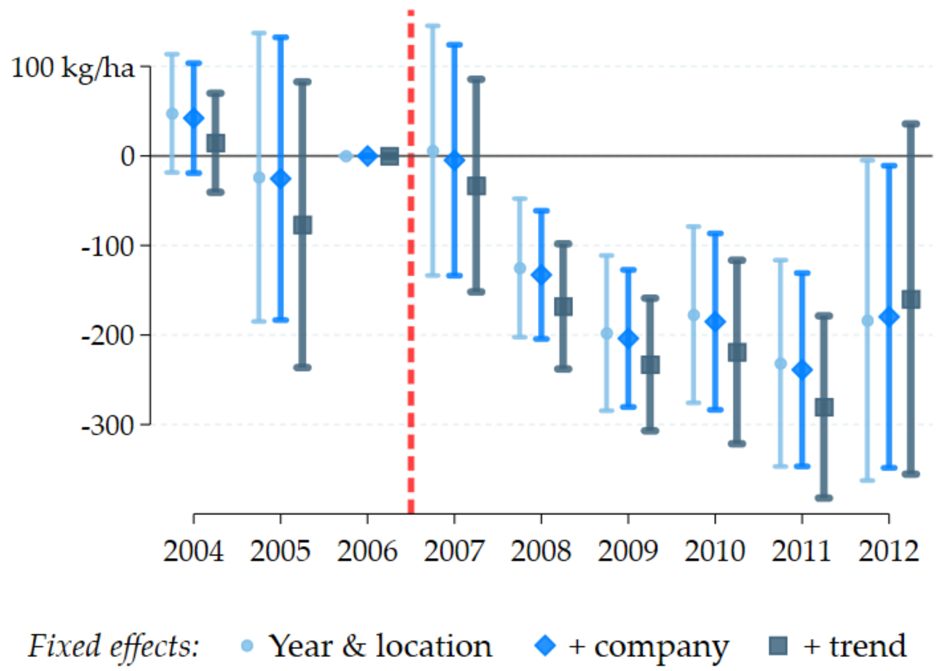
(d) Maximum



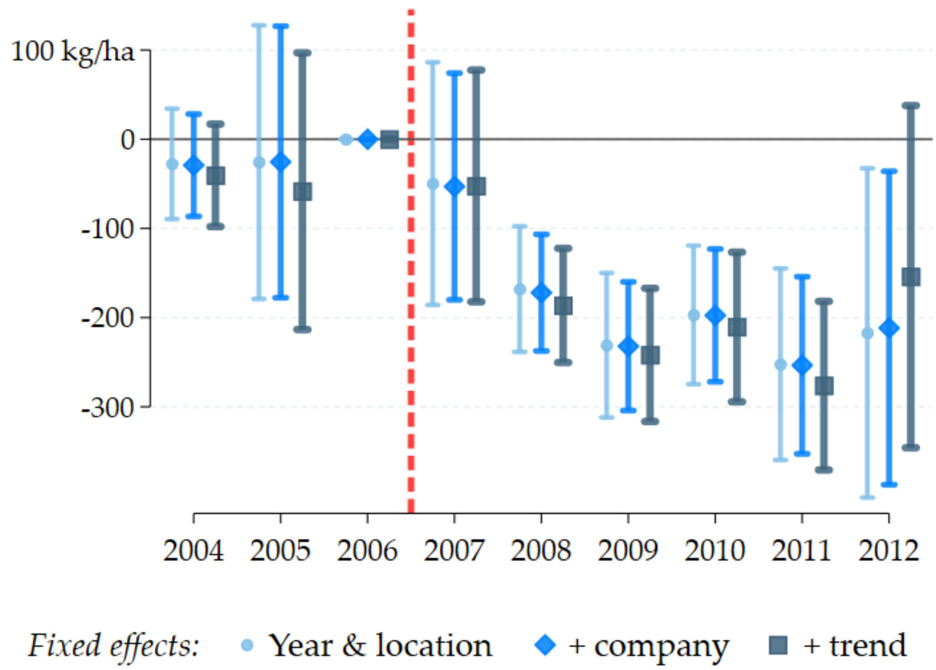
Notes: Event-study estimates with 95 percent confidence intervals based on least-squares regressions as in Equation 4: seed variety release year fixed effects are included in all the models; additional fixed effects are indicated in the legend below the graph. Standard errors clustered at the company level. Unit of observation: seed variety × company × trial location. Data digitized from the Bt reports of the All India Coordinated Research Project on Cotton of the Indian Council of Agricultural Research. The outcome is expressed in kilograms per hectare. The vertical red line signals the treatment timing.

Figure C6. Effects on Agronomic-Trial Yields of Cotton Varieties – Robustness to Time of Trial

(a) First Year of Trials



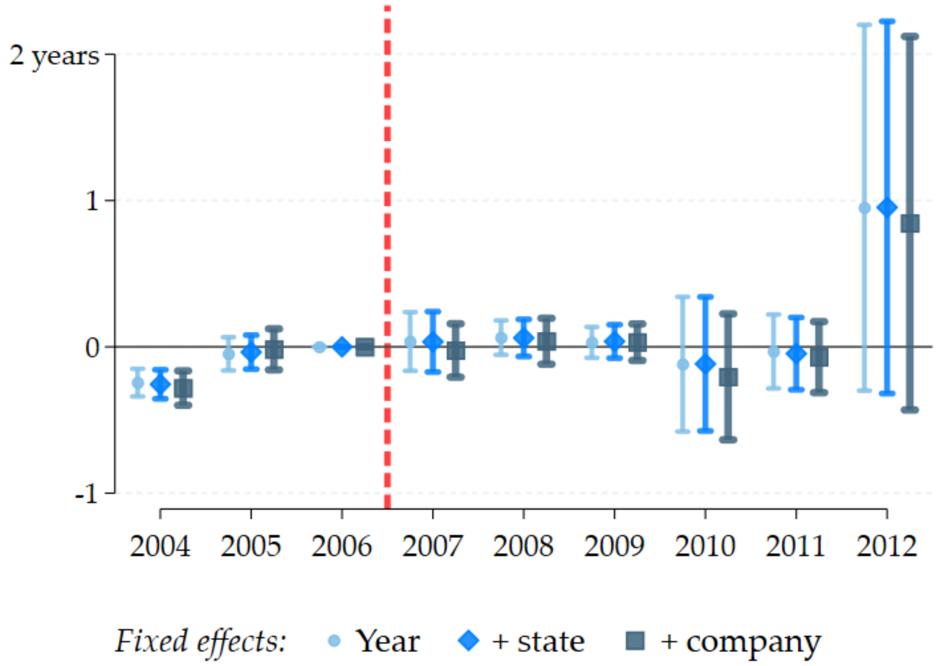
(b) Last Year of Trials



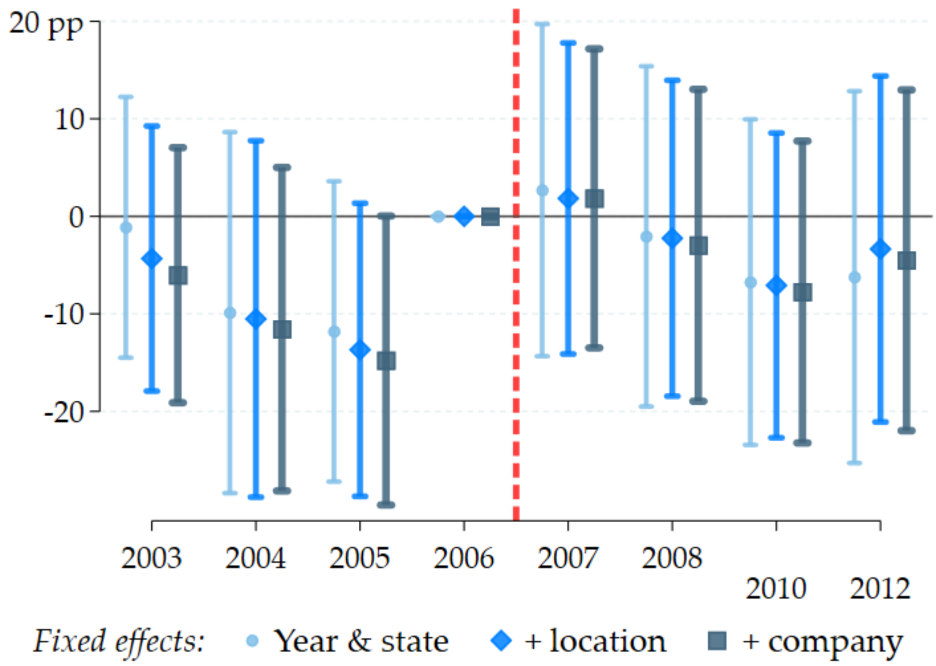
Notes: Event-study estimates with 95 percent confidence intervals based on least-squares regressions as in Equation 4: seed variety release year fixed effects are included in all the models; additional fixed effects are indicated in the legend below the graph. Standard errors clustered at the company level. Unit of observation: seed variety × company × trial location. Data digitized from the Bt reports of the All India Coordinated Research Project on Cotton of the Indian Council of Agricultural Research. The outcome is expressed in kilograms per hectare. The vertical red line signals the treatment timing.

Figure C7. Balance in Characteristics of Cotton Varieties Tested in Agronomic Trials

(a) Varietal Age at Release



(b) Probability of Approval



Notes: Event-study estimates with 95 percent confidence intervals based on least-squares regressions as in Equation 4: fixed effects are indicated in the legend below the graph. Standard errors clustered at the company level. Unit of observation: seed variety × company × trial state. Data digitized from the Bt reports of the All India Coordinated Research Project on Cotton of the Indian Council of Agricultural Research. The outcome variable of panel (a) is defined as the number of years elapsed from the first agronomic trial we observe in the data to the official market release of the seed variety. The outcome of panel (b) is a binary variable equal to one if the seed variety was approved and to zero otherwise. The vertical red line signals the treatment timing.

Table C19. Heterogeneous Treatment Effects on Agronomic-Trial Yields of Cotton Varieties (Logged) – Pooled Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Firm is headquartered in price-controlled state								
Price control × Post-2006	-0.148	-0.166	-0.210	-0.133	-0.142	-0.216	-0.216	-0.111
	(0.098)	(0.086)*	(0.027)***	(0.038)***	(0.035)***	(0.037)***	(0.037)***	(0.021)***
	[0.139]	[0.138]	[0.144]	[0.131]	[0.137]	[0.152]	[0.151]	[0.080]
[...] × Heterogeneity margin	-0.190	-0.165	-0.183	-0.178	-0.162	-0.085	-0.085	-0.121
	(0.108)*	(0.091)*	(0.046)***	(0.050)***	(0.049)***	(0.065)	(0.064)	(0.033)***
	[0.079]**	[0.084]*	[0.080]**	[0.075]**	[0.066]**	[0.076]	[0.078]	[0.098]
Number of observations	2,524	2,524	2,524	2,523	2,523	2,523	2,523	2,523
Number of companies	34	34	34	34	34	34	34	34
Number of states	10	10	10	10	10	10	10	10
Adjusted R-squared	0.053	0.213	0.324	0.501	0.516	0.548	0.546	0.570
Firm size is below sample median								
Price control × Post-2006	-0.247	-0.259	-0.315	-0.240	-0.235	-0.258	-0.258	-0.197
	(0.048)***	(0.037)***	(0.036)***	(0.041)***	(0.039)***	(0.043)***	(0.045)***	(0.029)***
	[0.155]	[0.160]	[0.164]*	[0.140]	[0.141]	[0.151]	[0.150]	[0.069]**
[...] × Heterogeneity margin	-0.272	-0.189	-0.203	-0.181	-0.194	-0.173	-0.174	-0.072
	(0.082)***	(0.072)**	(0.064)***	(0.069)**	(0.068)***	(0.077)**	(0.079)**	(0.074)
	[0.132]*	[0.124]	[0.109]*	[0.102]	[0.110]	[0.106]	[0.106]	[0.073]
Number of observations	2,524	2,524	2,524	2,523	2,523	2,523	2,523	2,523
Number of companies	34	34	34	34	34	34	34	34
Number of states	10	10	10	10	10	10	10	10
Adjusted R-squared	0.055	0.214	0.324	0.502	0.517	0.548	0.547	0.569
Year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Variety zone fixed effects		✓	✓	✓	✓	✓	✓	✓
Trial state fixed effects			✓	✓	✓	✓	✓	✓
Trial location fixed effects				✓	✓	✓	✓	✓
Company fixed effects					✓	✓	✓	✓
Company × year fixed effects						✓	✓	✓
Controlling for varietal age at release							✓	✓
Variety fixed effects								✓

Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. All regressions are least squares, as in Equation 4, with fixed effects (indicated in the last seven rows of the table) and standard errors clustered at the household level (in parentheses), where, instead of interacting a dummy variable for each year with the treatment, we pool all the periods after 2006 into *Post*. The outcome is expressed in natural logarithm, so that coefficients approximate percentage changes from the pre-treatment comparison period (i.e., 2006).

Table C20. Productivity Gains from New Seed Varieties

	(1)	(2)	(3)	(4)	(5)	(6)
New variety	0.195**	0.182**	0.175**	0.161**	0.165**	0.209***
	(0.084)	(0.080)	(0.079)	(0.072)	(0.073)	(0.080)
Number of observations	6,145	6,145	6,145	6,145	6,145	6,145
Number of clusters	607	607	607	607	607	607
Adjusted R-squared	0.087	0.163	0.193	0.274	0.301	0.316
Year fixed effects	✓	✓	✓	✓	✓	✓
Variety zone fixed effects		✓	✓	✓	✓	✓
Trial state fixed effects			✓	✓	✓	✓
Trial location fixed effects				✓	✓	✓
Company fixed effects					✓	✓
Company × year fixed effects						✓

Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: seed variety × year × trial location. Data digitized from the Bt reports of the All India Coordinated Research Project on Cotton of the Indian Council of Agricultural Research. All regressions are least squares, as in Equation 5 with fixed effects (indicated in the last six rows of the table) and standard errors clustered at the seed variety level (in parentheses). The outcome, i.e., lint yields, is expressed in natural logarithm, so that coefficients approximate percentage changes.

D Triple Difference-in-Differences Estimates

The Cost of Cultivation/Production Survey (CCS) covers the principal crops cultivated in India and so allows us to extend our empirical design in Section 4.1 to a triple difference-in-differences. We compare the evolution of cotton outcomes with that of other crops, such as rice, wheat, etc.[72] The event-study model becomes

$$Y_{i,c,t} = \alpha_{s(i)} + \alpha_c + \alpha_t + \sum_{\tau \neq -1} \beta_\tau \cdot PriceCap_{s(i)} \cdot \mathbb{1}\{c = \text{cotton}\} \cdot \mathbb{1}\{t = \tau\} + \varepsilon_{i,c,t} \quad \text{(D1)}$$

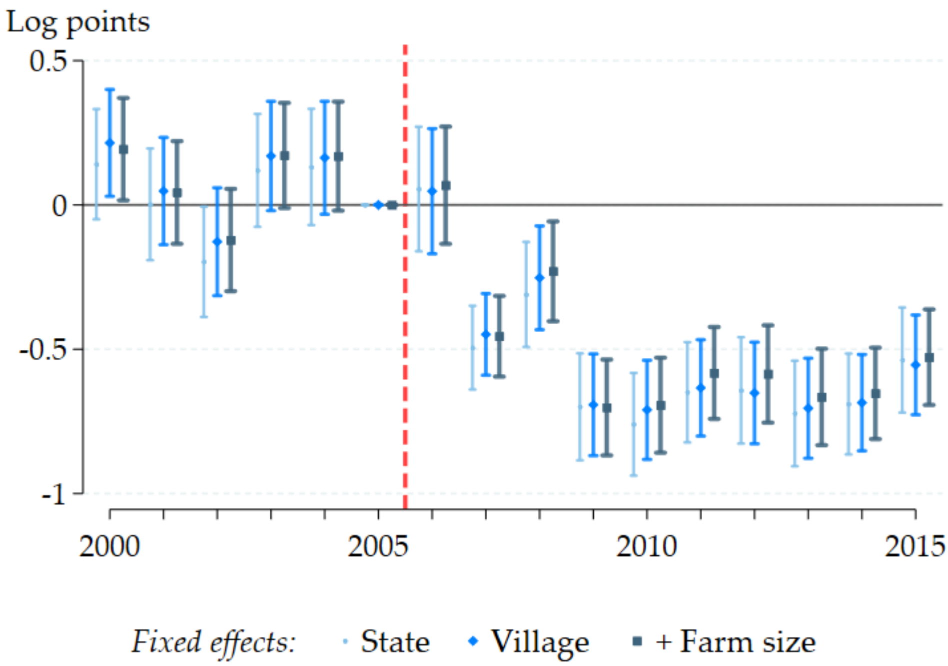
where the notation is the same as for Equation 1 with the addition of index c that refers to the crop cultivated by a certain household i in state s and time period t. The identifying assumption of parallel trends now applies to the evolution of crop-level counterfactual outcomes. In particular, β_τ identifies the average treatment effect on the treated (ATT) if cotton quantities had grown at the same rate across states, *as compared to other crops*, in the absence of the treatment.

The results on seed and total production costs, presented in Appendix Figure D1, decisively confirm the baseline findings in the paper (see Figure 3 for comparison). The event-study point estimates are slightly smaller but maintain both a negative sign and statistical significance at the 0.1 percent level for the vast majority of post-treatment periods. We interpret this as evidence that state-level policy changes or economic shocks, other than cotton seed price controls, are not responsible for our results.

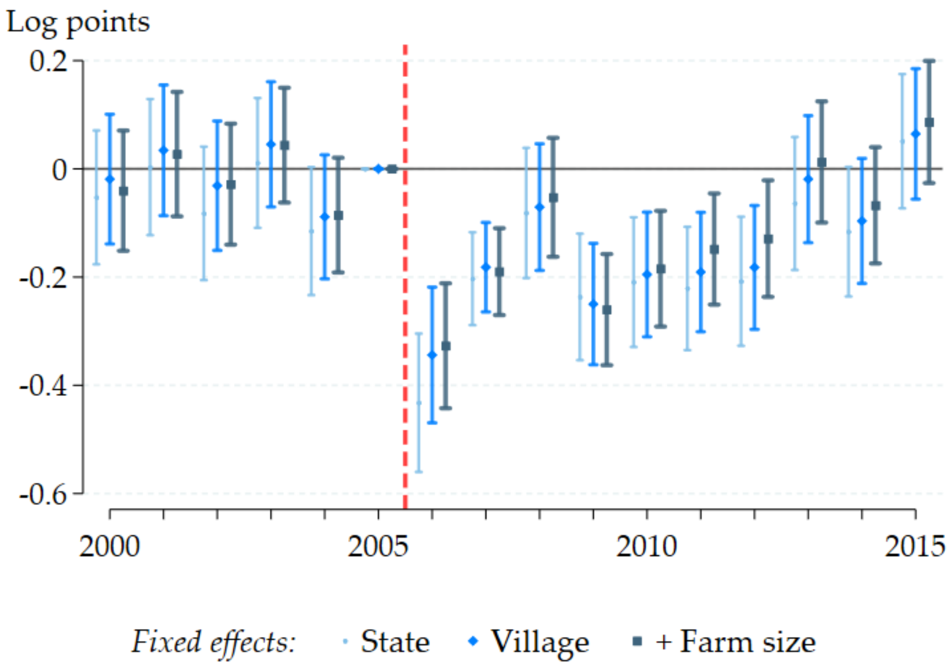
[72] In the estimation sample used in this appendix, we have 499,302 observations, of which 37.6% refer to rice cultivation, 16.4% to wheat, 4.9% to cotton, 4.2% to maize, 4.1% to mustard or rapeseed, 3.5% to pearl millet, 3.4% to sugarcane, and 3.1% to soybeans. All the other crops are below 3%.

Figure D1. Differential Treatment Effects on Costs per Acre

(a) Seeds



(b) Total



Notes: Event-study estimates with 95 percent confidence intervals based on least-squares regressions as in Equation D1: year and season fixed effects are included in all the models; additional fixed effects are indicated in the legend below the graph. Standard errors clustered at the village level. Unit of observation: household × parcel × plot × crop × season × year. Data from the Cost of Cultivation/Production Survey. The outcome is expressed in natural logarithm, so that coefficients approximate percentage changes from the pre-treatment comparison period (i.e., 2005). The vertical red line signals the treatment timing. Estimates pooling pre- and post-treatment periods are in Appendix Table D1.

Table D1. Differential Treatment Effects on Costs per Acre – Pooled Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
Seeds						
Price Control × Post-2005 × Cotton	-0.534***	-0.531***	-0.664***	-0.692***	-0.663***	-0.598***
	(0.054)	(0.052)	(0.052)	(0.050)	(0.048)	(0.040)
Number of observations	477,653	477,653	477,653	477,652	477,649	477,649
Number of clusters	48,037	48,037	48,037	48,037	48,037	48,037
Adjusted R-squared	0.442	0.469	0.566	0.585	0.640	0.830
Total						
Price Control × Post-2005 × Cotton	0.081**	0.088**	-0.126***	-0.134***	-0.109***	-0.066***
	(0.039)	(0.038)	(0.035)	(0.034)	(0.031)	(0.020)
Number of observations	499,338	499,338	499,338	499,337	499,334	499,334
Number of clusters	48,293	48,293	48,293	48,293	48,293	48,293
Adjusted R-squared	0.309	0.365	0.512	0.533	0.606	0.857
Season fixed effects	✓	✓	✓	✓	✓	✓
Crop fixed effects	✓	✓	✓	✓	✓	✓
Year fixed effects		✓	✓	✓	✓	✓
State fixed effects			✓	✓	✓	✓
Village fixed effects				✓	✓	✓
Farm size fixed effects					✓	✓
Other farm controls						✓

Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: household × parcel × plot × crop × season × year. Data from the Cost of Cultivation/Production Survey. All regressions are least squares, as in Equation D1, with fixed effects (indicated in the last seven rows of the table) and standard errors clustered at the household level (in parentheses), where, instead of interacting a dummy variable for each year with the treatment and cotton indicator, we pool all the periods after 2005 into *Post*. The outcome is expressed in natural logarithm, so that coefficients approximate percentage changes from the pre-treatment comparison periods (i.e., 2000-2005). 'Other farm controls' are cultivated area, hours of attached labor, and total value of capital (all logged). A graphical representation of the main dynamic estimates is displayed in Figure D1.

E Structural Model: Details

E.1 Data Construction

Our main source of data to estimate the structural parameters from our model of the Indian cotton seed market is the Cotton Crop Track by Francis Kanoi Marketing Research, a market research and consultancy company based in Chennai, Tamil Nadu. Our estimation procedure, further detailed in Section E.2, requires two main datasets: a product-level and a consumer-level dataset. This subsection reviews and motivates the decisions made to construct both of these, while keeping the model tractable for estimation.

Product Data. To begin with, we assume farmers choose among *brands* (i.e., seed-producing firms) rather than specific *varieties*. This is a result of both the detail of survey responses, which always include the brand but do not record the specific variety in a consistent and systematic way across data collection waves, and of our sample size, where some varieties are planted by very few farmers.[73] As a consequence, in our model, a *differentiated product* represents a portfolio of varieties offered by a seed firm in a market. Even though seed firms may sell more than one variety in the same market (according to our data, an average of 3.29 and a median of 2 varieties), our definition of a seed brand captures key attributes that matter most to farmers, such as the nature of the germplasm and the presence of genetically engineered (GE) traits.

We restrict our attention to the 14 largest Bt brands in our data. We collapse all other Bt varieties into a residual brand that we define as "Other Bt" and collapse all non-Bt varieties into two categories: "*Desi* public varieties" and "Non-Bt hybrids". Table E1 presents the shares of each brand by year: our 14 main brands of Bt cotton cover more than 90% of the Bt market. Two prominent patterns emerge: between 2002 and 2005, the adoption of GE varieties was low, with only three brands offering varieties containing this genetic trait (*Mahyco* in both 2002 and 2005, *Navabharat* and *Rasi* only in 2005). During the early years of Bt, non-GE hybrid varieties were the most popular choice among cotton farmers with around 20% of market share for public varieties (which can be saved for more than one cropping season) and 70% for hybrids (which have to be re-purchased every season to avoid steep productivity losses). After price controls were enacted by state governments, both the number of companies selling Bt cotton and the share of Bt adoption went up significantly, the latter passing 95% by 2013. This major shift aligns with the reduced-form evidence presented in Section 4 and illustrates the rich variation in seed choice in our data, covering both the early years of Bt in India and the subsequent development of a more mature market. Finally, we notice that the Herfindahl-Hirschman Index (HHI) is equal to 1,000, on average, and relatively stable over time. According to the Merger Guidelines of the United States Department of Justice and Federal Trade Commission, this level of HHI is indicative of a "moderately concentrated" market.

We construct choice sets at the market level. We define a *market* as a district-year pair and assume that any product (i.e., brand) that was bought by at least one farmer in that district-year is available to all farmers in the same district and year. We make

[73] There were more than 1,000 Bt varieties approved between 2002 and 2014, i.e., our sample period for the structural analysis. In addition, qualitative interviews highlight that varietal turnover was generally high during this period, partly due to the proliferation of hybrid cotton varieties in the aftermath of Bt introduction in India (Stone, 2007; Ramaswami et al., 2009).

Table E1. Individual Product Choice: Data Frequency and Shares

	Year →									
	2002/2003		**2004/2005**		**2008/2009**		**2013/2014**		**Total**	
Brand ↓	N	%	N	%	N	%	N	%	N	%
Ajeet					934	3.75	2,512	14.02	3,446	4.20
Ankur					1,428	5.73	1,422	7.93	2,850	3.48
Bio-seeds / Sri Ram					985	3.95	1,087	6.07	2,072	2.53
Desi/Public varieties	4,491	23.42	4,181	20.97	269	1.08	128	0.71	9,069	11.06
JK Seeds					381	1.53	200	1.12	581	0.71
Kaveri					203	0.81	2,337	13.04	2,540	3.10
Krishidhan					550	2.21	211	1.18	761	0.93
Mahyco	1,042	5.43	1,718	8.62	2,132	8.55	2,061	11.50	6,953	8.48
Monsanto / Paras					1,230	4.93	947	5.28	2,177	2.66
Navabharat			436	2.19	359	1.44			795	0.97
Non-Bt hybrids	13,647	71.15	12,275	61.56	326	1.31	278	1.55	26,526	32.36
Nuziveedu					5,831	23.39	2,637	14.71	8,468	10.33
Other Bt					2,554	10.25	1,541	8.60	4,095	5
Rasi			1,331	6.67	4,781	19.18	1,378	7.69	7,490	9.14
Tulasi					1,546	6.20	394	2.20	1,940	2.37
Vibha Agrotech					912	3.66	65	0.36	977	1.19
Vikram					506	2.03	724	4.04	1,230	1.50
Total	19,180	100	19,941	100	24,927	100	17,922	100	81,970	100
HHI	1,480		966		1,422		1,029		1,000	

Notes: Unit of observation: household × plot. Data from Francis Kanoi Marketing Research's Cotton Crop Track. All brands other than 'Desi/Public varieties' are hybrids; all brands other than "Desi/Public varieties' and 'Non-Bt hybrids' are Bt hybrids. Even columns report the count of observations of a certain brand in a certain agricultural season. Odd columns report the brand percentage, relative to the total number of observations in that season. Empty cells indicate frequencies equal to zero. The HHI in the last row is computed as the sum of squared market shares of private *companies* (rather than *brands*, after dropping public varieties and missing values); the index is scaled from 0 to 10,000.

this assumption because this is the smallest level of aggregation at which we can compute our product-level variables on cotton, allowing us to capture substantial heterogeneity in product characteristics that farmers take into account for their seed choices.

Our key product characteristics are price and yield. We input the average price paid for all varieties offered by a brand in a particular district-year pair as the price for that brand. An obvious choice for measuring yields would be the agronomic yield data used in the reduced-form analysis of technological innovation (Section 4.4). As noted in Footnote 20, these data come from experimental field trials, as carried out by agricultural researchers in testing centers, and so are not affected by other endogenous decisions by farmers, which would take place after the observed seed choice. However, they are only available for a subset of market-brand pairs (namely, 296 out of 1,980 observations) and are not available for non-Bt varieties. A brand's yield is therefore computed as the maximum yield in our survey data across all farmers in a market, a proxy of "expected yield at purchase".[74] We keep this measure in physical

[74] Unlike prices, information on yields is not officially released to buyers, e.g., through advertising on

units (i.e., quintals of unginned cotton per hectare) in order to capture expected production without conflating it with changes in expected output prices. Variation in the latter is likely to affect farmers' decisions to cultivate cotton, rather than substitution patterns across different brands. We model this extensive-margin choice through a market-varying utility from the outside option (Equation 7).

Product-level variables from complementary data sources are taken at the state-year level due to the fact that not all districts are sampled for such data. In particular, we construct a measure of outside option profits per hectare, which aims to reflect the profitability from cultivating crops other than cotton.

Summary statistics on product characteristics are reported in Table E2.

Table E2. Product Characteristics: Summary Statistics

	(1) Mean	(2) Std. Dev.	(3) Median	(4) 25th pct.	(5) 75th pct.
Seed price ('00 ₹)	7.93	3.00	7.64	7.12	9.30
Physical yield ('00 kg/ha)	10.48	4.17	10.00	7.50	13.50
Outside option profits ('00 ₹/ha)	171.84	140.84	112.40	80.64	225.98
Number of rival varieties	27.79	12.84	28.00	18.00	37.00
Number of own varieties	3.29	4.06	2.00	1.00	4.00
Number of products	1,918				
Number of markets	240				
Number of brands	17				

Notes: Unit of observation: brand × market (i.e., district × year). Data on inside goods from Francis Kanoi Marketing Research's Cotton Crop Track. Data on outside goods sampled with replacement from the Cost of Cultivation/Production Survey to match shares from the Crop Production Statistics (CPS).

Consumer Data. We include all observations in the four waves of our farmers' microdata. Importantly, our individual choice data only includes observations from cotton farmers. To incorporate extensive-margin choices (i.e., growing crops other than cotton), we augment the sample using draws from a nationally representative survey of farmers growing any crop.[75] We sample farmers with replacement from the Cost of Cultivation/Production Survey (CCS) in a proportion that makes the share of observations choosing inside goods (i.e., cotton seeds) match the share of total acreage

seed packets, given the wide heterogeneity in productivity by growing conditions and practices; however, at the moment of purchase, farmers tend to discuss expected yields with input agrodealers, who are in turn informed about previous trials by seed companies. Given that our model does not feature any market intermediaries, we are implicitly assuming that input agrodealers collectively help farmers choose the product with price and quality that maximizes their utilities. On the supply side, we do not model retail pricing decisions either. Our implicit assumption is that retailers charge constant markups, which are not affected by price controls.

75 Following the recommendations of Nevo (2000), we check the sensitivity of our results to the market definition. In particular, to compute the outside good shares, we restrict our attention to the main alternative crops to cotton in the Indian setting, i.e., sorghum, pearl millet, groundnut, soybeans, sunflower, rice, maize, green gram, and pigeon peas. The demand estimates using the latter approach are very similar to our baseline specification.

planted under cotton in the districts where the main farmers' survey was carried out. We obtain these shares from our district-level acreage measures obtained from the Crop Production Statistics (CPS) of the Directorate of Economics and Statistics in the Ministry of Agriculture and Farmers Welfare. Table E3 presents the frequencies and shares of inside and outside goods. The share of cotton farmers ranges between 9% in 2013 and 15% in 2004. As a result, the final sample for our demand estimation is of 628,143 micro-consumers.

Table E3. Inside and Outside Goods: Frequencies and Shares

	Year →									
	2002/2003		**2004/2005**		**2008/2009**		**2013/2014**		**Total**	
Chosen option ↓	N	%	N	%	N	%	N	%	N	%
Inside good: cotton	18,282	10.94	19,398	15.15	17,734	13.84	17,922	8.75	73,336	11.68
Outside option	148,826	89.06	108,655	84.85	110,407	86.16	186,919	91.25	554,807	88.32
Total	167,108	100.00	128,053	100.00	128,141	100.00	204,841	100.00	628,143	100.00

Notes: Unit of observation: household × plot. Inside good observations from Francis Kanoi Marketing Research's Cotton Crop Track. Outside good observations sampled with replacement from the Cost of Cultivation/Production Survey to match shares from the Crop Production Statistics. Even columns report the count of observations in a certain agricultural season. Odd columns report the percentage, relative to the total number of observations in that season.

For all observations (from the cotton farmers' data or the non-cotton CCS draws) we also include a measure of plot size, i.e., hectares under cultivation, to allow preferences to be heterogeneous across farmers. Among our sample of cotton farmers, we have additional demographics: summary statistics on these demographics can be found in Table E4 (Columns 1 to 5). Plot size is positively correlated with education, asset ownership, total landholdings, and mechanization (Column 7). Therefore, we view plot size as a proxy of scale of cultivation, use of complementary technologies, and wealth.

Table E4. Demographics of Cotton Farmers

	(1) Mean	(2) Std. Dev.	(3) Median	(4) 25th pct.	(5) 75th pct.	(6) No. of obs.	(7) Correlation with plot size
Age	40.28	12.51	40	31	48	20,464	0.421
Education							
No education	0.137	0.344	0	0	0	20,464	-0.043
Has primary education	0.217	0.412	0	0	0	20,464	-0.038
Has middle education	0.197	0.398	0	0	0	20,464	0.027
Has secondary education	0.221	0.415	0	0	0	20,464	0.012
Has senior secondary education	0.114	0.318	0	0	0	20,464	0.008
Has undergraduate degree	0.049	0.216	0	0	0	20,464	0.031
Has graduate or professional degree	0.020	0.141	0	0	0	20,464	0.024
Durable assets							
Owns a telephone	0.140	0.347	0	0	0	14,888	0.118
Owns a television	0.481	0.500	0	0	1	14,888	0.110
Owns a refrigerator	0.132	0.338	0	0	0	14,888	0.212
Owns a moped	0.308	0.462	0	0	1	14,888	0.150
Owns a motorcycle	0.018	0.133	0	0	0	14,888	0.040
Owns a car	0.011	0.104	0	0	0	14,888	0.142
Owns a truck	0.002	0.042	0	0	0	14,888	0.041
Landholding							
Under cotton cultivation	3.63	12.36	1.00	3.00	4.00	20,464	0.421
Total	7.26	41.16	5.00	2.50	8.00	13,950	0.148
Mechanization							
Owns a tractor	0.099	0.299	0	0	0	18,140	0.283
Used a tractor	0.568	0.495	1	0	1	11,101	0.283

Notes: Unit of observation: household for Columns (1) to (6), household × plot for Column (7). Data from Francis Kanoi Marketing Research's Cotton Crop Track. Durable assets and total landholding were not asked in the 2002/2003 wave, tractor usage was not asked in the 2009/2010 and 2013/2014 waves, explaining the missing values compared to the other variables; other missings are due to the interviewee not answering the question. All the variables under the '**Education**' and '**Durable assets**' panels are binary variables (0/1). 'Primary education' indicates up to Grade 6, 'middle' up to Grade 8, 'secondary' up to Grade 10, and 'senior secondary' up to Grade 12.

E.2 Demand: Details

For notational simplicity, we express all the components of the utility of inside goods as differences relative to the utility of the outside option. Thus, $\delta_{bm} \equiv \alpha \cdot p_{bm} + \gamma \cdot y_{bm} + \xi_m + \xi_{bm} - \beta_1 \cdot \Pi_{0m} - \beta_2 \cdot t_m$. This can be expressed in matrix form as $\vec{\delta} = \mathbb{X}_{bm}^T \vec{\beta}$. Further, let $\mu_{fbm}^z = \mu_{bm}^{z_{fm}}(\theta^z) - \mu_{0m}^{z_{fm}}(\theta^z)$ and $\vec{\theta} = \left[\theta^z, \theta^v\right]$. Then, our parameters of interest become $\left(\hat{\beta}, \hat{\theta}, \hat{\delta}\right)$, which are the solution to

$$\left(\hat{\beta}, \hat{\theta}, \hat{\delta}\right) = \arg\min_{\beta,\theta,\delta} \left(- \log \hat{L}^{\text{micro}}(\theta, \delta) - \log \hat{L}^{\text{MACRO}}(\theta, \delta) + \hat{\Pi}(\beta, \delta) \right)$$

where $\hat{L}$ denotes likelihood functions and $\hat{\Pi}$ is a vector of product-level moment conditions. The first, 'micro' term is the log-likelihood of the sample of farmer individual choices, which follows the mixed logit. The second, 'MACRO' term is the log-likelihood of the market shares: it integrates over the distribution of farmer characteristics in the population. The third term directly incorporates information from the product-level exogeneity restrictions, which are additional assumptions on the data-generating process.

Let $d_{fbm} = 1$ if a farmer f chooses brand b in market m and $d_{fbm} = 0$ otherwise. Under the assumption that preference shocks ε in Equation 6 are independent and identically distributed (iid) and follow a type 1 extreme value (Gumbel) distribution, we can express choice probabilities for farmer f selecting brand b given their plot size z and product characteristics $X := (p, y, \xi)$ for any set of parameters (θ, δ). Formally,

$$\pi_{bm}^{z_{fm}} \equiv \mathbb{P}\mathrm{r}\{d_{fbm} = 1 \,|\, z_{fm}, X_m;\, \theta, \delta\} = \int \frac{\exp(\delta_{bm} + \mu^{z}_{fbm} - \mu^{\nu}_{f0m})}{1 + \sum_{\ell=1}^{J_m} \exp(\delta_{\ell m} + \mu^{z}_{f\ell m} - \mu^{\nu}_{f0m})} \, \mathrm{d}F_m(\nu)$$

Unconditional choice probabilities, or expected market shares, are obtained in a similar fashion by integrating $\pi_{bm}^{z_{fm}}$ with respect to the distribution of farmer plot sizes

$$\pi_{bm} \equiv \mathbb{P}\mathrm{r}\{d_{fbm} = 1 \,|\, x_m \; \theta, \delta\} = \int \pi_{bm}^{z_{fm}}(\theta, \delta)\, dG_m(z)$$

where the G distribution is taken from the data. Now, let S_{fm} be equal to 1 if a farmer f is in the 'micro' sample of a certain market m and to 0 otherwise. Then, using our model choice probabilities, we can write the mixed-data likelihood as

$$\begin{aligned}
\log \hat{L}(\theta, \delta) &= \sum_{m=1}^{M} \sum_{b=0}^{B_m} \sum_{f=1}^{N_m} d_{fbm} \left(S_{fm} \log\left(\pi_{bm}^{z_{fm}}(\theta, \delta)\right) + (1 - S_{fm}) \log\left(\pi_{bm}(\theta, \delta)\right) \right) \\
&= \sum_{m=1}^{M} \sum_{b=0}^{B_m} \sum_{f=1}^{N_m} S_{fm} d_{fbm} \log\left(\frac{\pi_{bm}^{z_{fm}}(\theta, \delta)}{\pi_{bm}(\theta, \delta)}\right) + \sum_{m=1}^{M} N_m \sum_{b=0}^{B_m} \frac{\sum_{i=1}^{N_m} d_{fbm}}{N_m} \log \pi_{bm}(\theta, \delta)
\end{aligned}$$

This expression clearly separates the contribution of the farmer-level data, from the first term, and the market-level data, from the second term. In order to avoid farmers observed in the micro-sample are double counted in the second MACRO-term, the estimator can be rewritten as

$$\log \hat{L}(\theta, \delta) = \sum_{m=1}^{M} \sum_{b=0}^{B_m} \sum_{f=1}^{N_m} S_{fm} d_{fbm} \log \pi_{bm}^{z_{fm}} + \sum_{m=1}^{M} \sum_{b=0}^{B_m} \left(N_m s_{bm} - \sum_{f=1}^{N_m} S_{fm} d_{fbm} \right) \log \pi_{bm}(\theta, \delta)$$

where the first term captures the contribution of consumer-level 'micro' data, while the second term captures the contribution of product-level 'MACRO' data. $\pi_{0m}^{z_{fm}}$ is the model choice probability of the outside good, $b = 0$, and π_{0m} its market share.

The second part of the objective function of our estimator arises from imposing moment conditions on the unobserved utility, or structural error, ξ_{bm}. These involve exogeneity restrictions of the form

$$\mathbb{E}\left[\xi_{bm} \,\middle|\, W_{bm}^{\text{demand}}\right] = 0$$

where W^{demand} is a vector of instruments. These exogeneity restrictions are needed to identify the mean product utility parameters $\vec{\beta}$. Formally, the penalty function is

$$\hat{\Pi}(\beta, \delta) = \frac{1}{2}\hat{m}^T(\beta, \delta)\,\hat{\mathcal{W}}\,\hat{m}(\beta, \delta)$$

where

$$\hat{m}(\beta, \delta) = \sum_{m=1}^{M}\sum_{b=1}^{B_m} W_{bm}^{demand}(\delta_{bm} - \beta^T x_{bm})$$

and $\hat{\mathcal{W}}$ is the optimal GMM weighting matrix converging to the inverse of $\mathbb{V}(W_{bm}^{demand}\,\xi_{bm})$.

As mentioned in Grieco et al. (2025), when the model is exactly identified (i.e., if the dimension of β equals that of W^{demand}), the estimator is equivalent to a two-step procedure that estimates $\hat{\theta}$ and $\hat{\delta}$ through maximum likelihood estimation and then estimates $\hat{\beta}$ through GMM using $\hat{\delta}$. When the model is over-identified, both $\hat{L}$ and $\hat{\Pi}$ contribute to estimating θ and δ.

E.3 Supply: Details

Aggregation. As mentioned above, a market is defined at the district-year level to better capture local variation in product characteristics and consumers' choice set. Firms typically target a larger consumer base when designing a seed variety and incur common costs at a larger scale. Therefore, we aggregate product characteristics and demand estimates (namely, price and yield derivatives and elasticities) at the state-by-year level. We do so by taking their weighted average at the district-by-year level and using conditional choice probabilities within each state-year pair. This results in a sample of 323 observations and 35 relevant "macro" markets for the supply-side analysis.

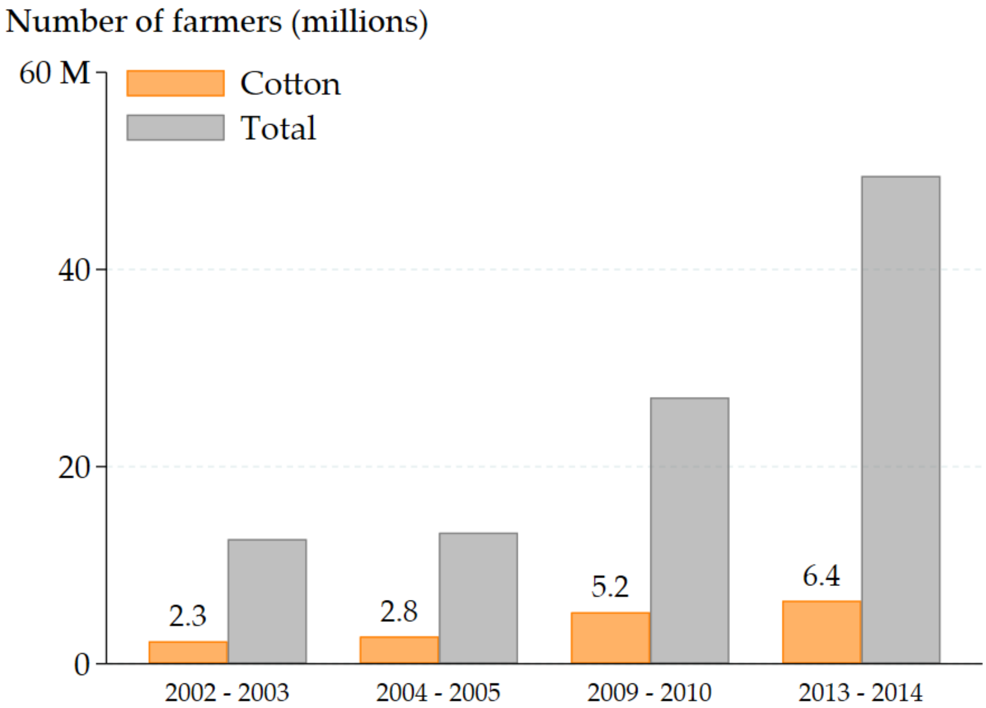
We calculate *market size* in the following three steps. First, in our main dataset, we observe the distribution of plot sizes and the total acreage under cotton in each market across four strata: up to 2 acres, over 2 and up to 5 acres, over 5 and up to 10 acres, over 10 acres. By combining this information, we can approximate the number of cotton farmers in each market-strata pair and obtain an estimate of total cotton farmers in India. Second, we use the CPS data to calculate the ratio of non-cotton farmers to cotton farmers. Finally, to get at total farmers, we scale this ratio by our estimate of cotton farmers and add the result to the latter. Figure E1 plots the resulting number of farmers over time. This procedure provides a close match with national estimates of the farming population during the study period.[76]

Given that cotton farmers typically purchase more than one seed packet (2.9, on average in our sample), we multiply the number of farmers by the average number

[76] For comparison, the 70th round of the National Sample Survey by the Ministry of Statistics and Programme Implementation estimated 90.2 million agricultural households in India, of which 33.6 million in the nine cotton states we consider in our analysis (NSSO, 2014).

of seed packets purchased at the district-year level to obtain our final measure of potential market size. The resulting estimate of total cotton seed packets aligns well with the one by (Pray and Nagarajan, 2010, p. 303). Average seed packets purchased will also be used to rescale per-packet surplus (i.e., total surplus shown in Column 5 in Table 4 divided by total seed packets) to per-farmer surplus (Column 6).

Figure E1. Market Size by Year



Notes: Unit of observation: year. Cotton farmers are calculated in each market-strata pair using plot size distributions and total acreage data from Francis Kanoi Marketing Research's Cotton Crop Track. Total farmers are calculated by scaling the latter to match cotton shares from the Crop Production Statistics.

Marginal Cost Estimation. This paragraph details the estimation procedure to recover marginal costs in constrained markets using our supply-side estimates from unconstrained markets. We assume that marginal costs can be decomposed into the expression in Equation 9, i.e., a linear function of yields, a brand component, a Bt-specific trend, and an unobserved cost shock that is brand-market specific. To estimate these components in unconstrained markets (i.e., all markets before 2006 and the non-price-controlled states from 2006 onward), we follow the three-step procedure outlined in Section 5.2. Provided that these three components are identified under our assumptions in Footnote 56, the remaining ν^{mc} is mean-zero. We can then predict marginal costs for any market, including constrained markets (i.e., price-controlled states after 2006), as

$$\widehat{mc}_{bm} = \widehat{\omega} \cdot y_{jt} + \widehat{\xi}_b + \widehat{\kappa} \cdot t_{\mathbb{1}\{b \in \mathcal{B}t\}}$$

This step implicitly assumes that the cost structure estimated from unregulated markets – including both the fixed components and the relationship between yields and costs – can be extrapolated to regulated markets.

E.4 Counterfactuals: Algorithm

Consider a market, m, with price cap policy, such that $\mathcal{P} = (0, \check{p}]$, where $\check{p}$ is the price cap and $\check{p} = +\infty$ in the absence of such policy. Following Barwick et al. (2024), and their matrix notation, an equilibrium outcome is defined as the set of prices $p \in \mathcal{P}$ and yields $y \in \mathbb{R}^{++}$ that satisfy firms' first-order conditions (FOCs), conditional on profits Π being positive.

Let

$$dM_y = \begin{pmatrix} \frac{\partial mc_1}{\partial y_1} & 0 & \cdots & 0 \\ 0 & \frac{\partial mc_2}{\partial y_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\partial mc_J}{\partial y_J} \end{pmatrix} \quad \text{and} \quad dFC_x = \begin{pmatrix} \frac{\partial FC_1}{\partial y_1} \\ \frac{\partial FC_2}{\partial y_2} \\ \vdots \\ \frac{\partial FC_J}{\partial y_J} \end{pmatrix}$$

Then, the unconstrained FOCs are:

$$\begin{aligned} d\Pi_p &:= \quad Q + \Omega \otimes \Delta_p(p - r - mc) = 0 \\ d\Pi_y &:= \quad -dM_y Q + \Omega \otimes \Delta_y(p - r - mc) = dFC_y \end{aligned}$$

where $d\Pi_p$ and $d\Pi_y$ represent the derivatives of total profit with respect to a firm's choice variables, i.e., price and yield, respectively. Q is quantity sold, r are trait fees, and mc are marginal costs. See Equations 11 and 12 in the main body of the paper.

The iterative Algorithm 1 jointly updates (the vectors of) product characteristics until changes become trivial, using the following quasi-Newton method.[77]

Price Caps. To capture price caps, we add a slack parameter, λ, in regulated markets, such that the actual maximization problem of price-controlled firms becomes

$$\max_{\substack{p_{bm} \in \mathbb{R}^{++} \\ y_{bm} \in \mathbb{R}^{++}}} \Pi_{bm}(\vec{p}_m, \vec{y}_m) - \lambda \cdot \mathbb{1}\{p_{bm} > \check{p}_m\}(p_{bm} - \check{p}_m)^2$$

The first term is the same as in Equation 8. The second (and new) term applies a penalty for pricing above the mandated cap, $\{p_{bm} > \check{p}\}$: firms have to pay a "monetary fine" that is convex in the magnitude of their violation, $(p_{bm} - \check{p})$. The resulting price FOC explicitly features a "numerically friendly" version of the cap, resembling the intended firm behavior:

$$[p_{bm}] : \quad Q_{bm} + \Big(p_{bm} - r_m - mc_{bm}(y_{bm})\Big)\left(\frac{\partial Q_{bm}}{\partial p_{bm}}\right) = 2\lambda \cdot \mathbb{1}\{p_{bm} > \check{p}\}(p_{bm} - \check{p})$$

[77] To guide optimization, we compute the FOC violations at the beginning of each iteration and apply a diagonal Hessian approximation, where updates are obtained as $dp = \text{FOC}_p \oslash \left|\frac{\partial Q}{\partial p}\right|, dy = \text{FOC}_y \oslash \left|\frac{\partial Q}{\partial y}\right|$; $\oslash$ denotes element-wise division. If the preferred simultaneous price-yield update fails to reduce the overall FOC violation (defined as the maximum of the normalized L2-norms of price and yield FOCs), we fall back to: (i) sequential price-then-yield updates, (ii) targeted updates focusing on the three worst-violating products, (iii) collective price reductions for all products at the regulatory cap (see the next paragraph for details on how we account for price caps), and (iv) individual price adjustments for each capped product.

Algorithm 1 Estimating Counterfactual Outcomes

for each market $m \in \mathcal{M}$ **do**

Initialize $p^{\text{old}} = p^{\text{initial}}$, $y^{\text{old}} = y^{\text{initial}}$

while $\max\left\{\frac{\|dp\|^2}{\|p^{\text{old}}\|^2}, \frac{\|dy\|^2}{\|y^{\text{old}}\|^2}\right\} \geq 1\text{e-}4$ **do**

Compute sales Q, marginal costs mc, and demand derivatives (Δ_p, Δ_y)

Solve price FOC: $d\Pi_p + \left[\Delta_p + (\Omega \otimes \Delta_p)^T\right] dp = 0$

Solve yield FOC: $d\Pi_y - \left[M_y\Delta_y + (\Omega \otimes \Delta_y)^T M_y\right] dy = dFC_y$

Update prices and yields: $p^{\text{new}} = p^{\text{old}} + dp$, $y^{\text{new}} = y^{\text{old}} + dy$

Update sales Q, marginal costs mc, and demand derivatives (Δ_p, Δ_y)

Reset prices and yields for next iteration: $p^{\text{old}} = p^{\text{new}}$, $y^{\text{old}} = y^{\text{new}}$

Return updated equilibrium prices and yields for market m

$p^\star = p^{\text{new}}$, $y^\star = y^{\text{new}}$

We therefore adjust the calculation of the price FOC in Algorithm 1 as follows: (i) for prices above the cap, i.e., $p_{bm} > \check{p}$, we set λ to a very large value so as to approximate a discontinuous constraint in the profit maximization problem; (ii) when prices are at the cap, i.e., $p_{bm} = \check{p}$, and the unconstrained price FOC is negative, we set the latter to zero so that $p^{\text{new}}_{bm} = p^{\text{old}}_{bm} = \check{p}$; (iii) for the remaining cases, i.e., if the price cap does not affect firms' decision, the penalty term becomes inactive, so we proceed as for unregulated markets and just use the unconstrained FOC.

Farm Input Subsidy. Input subsidies affect the FOCs in the following two ways. First, farmers choose a brand based on the vector of subsidized prices, $p^{\text{sub}\rightarrowtail f} = (1-\tau)\cdot p$ Second, using the chain rule, price derivatives for Bt brands, $b \in \mathcal{B}t$, become:

$$\frac{\partial Q_{bm}}{\partial p_{bm}} = \frac{\partial Q_{bm}}{\partial p^{\text{sub}\rightarrowtail f}_{bm}} \cdot \frac{\partial p^{\text{sub}\rightarrowtail f}_{bm}}{\partial p_{bm}} = \frac{\partial Q_{bm}}{\partial p^{\text{sub}\rightarrowtail f}_{bm}} \cdot (1-\tau)$$

For non-Bt brands, the price derivative and resulting FOC apply unchanged since $\tau = 0$ and thus $(1-\tau) = 1$.

The total government expenditure, B, in market m, is calculated by multiplying the chosen subsidy rate by the equilibrium price and quantity of subsidized products sold. Namely,

$$B^{\text{sub}\rightarrowtail f}_m = \sum_{b \in \mathcal{B}t} \tau \cdot p_{bm} \cdot Q_{bm}$$

This calculation makes it clear how the budgetary cost of input subsidies depends not only on the subsidy rate applied, but also on the equilibrium market response through quantity demanded and the endogenous price adjustments by firms, potentially amplifying such cost beyond the mechanical effect of the subsidy alone.

Firm Innovation Subsidy. We consider a certain innovation budget, $B^{\text{sub}\rightarrowtail j}_m$, for each market and allocate it proportionally to firms that operate in that market based on their yield performance. Specifically, each product bm receives a grant of $G_{bm}(\vec{y}_m) = B_m^{\text{sub}\rightarrowtail j} \cdot \frac{y_{bm}}{\sum_b y_{bm}}$. This performance-based allocation ensures that higher-yielding products capture larger shares of the innovation budget; e.g., if product A achieves twice the yield of product B, then A receives exactly twice the grant allocation of B. The grant payment is then used to reduce the slope of the fixed cost in the yield FOC (Equation 12), so that $\frac{\partial FC_{bm}^{\text{sub}\rightarrowtail j}}{\partial y_{bm}} = \phi' + \phi'' \cdot y_{bm} - G_{bm}(\vec{y}_m)$, generating incentives for quality competition and yield improvement. The price FOC remains the same as in the standard case given that farmers continue to pay the full price.

E.5 Auxiliary Results and Robustness

Table E5. First Stage of Demand Instrumental Variables

	(1)	(2)
Number of rival varieties	-0.007	-0.010*
	(0.006)	(0.006)
Number of own varieties	-0.184***	-0.243***
	(0.014)	(0.018)
Own yields ('00 kg/ha)		0.134***
		(0.020)
Number of observations	1,918	1,918
Number of markets	240	240
F-statistic	96.0	64.1

Notes: *Significant at 10%. ***Significant at 1%. Unit of observation: brand × market (i.e., district × year). The outcome variable is average prices in Indian rupees (₹) per 450-gram bag, the typical size of a hybrid cotton seed package (enough to plant an acre of land) in India. Data from Francis Kanoi Marketing Research's Cotton Crop Track. All regressions are least squares with clustered standard errors at the market level (in parentheses). The estimates of the second stage are in Table 2.

Table E6. First Stage of Supply Instrumental Variables

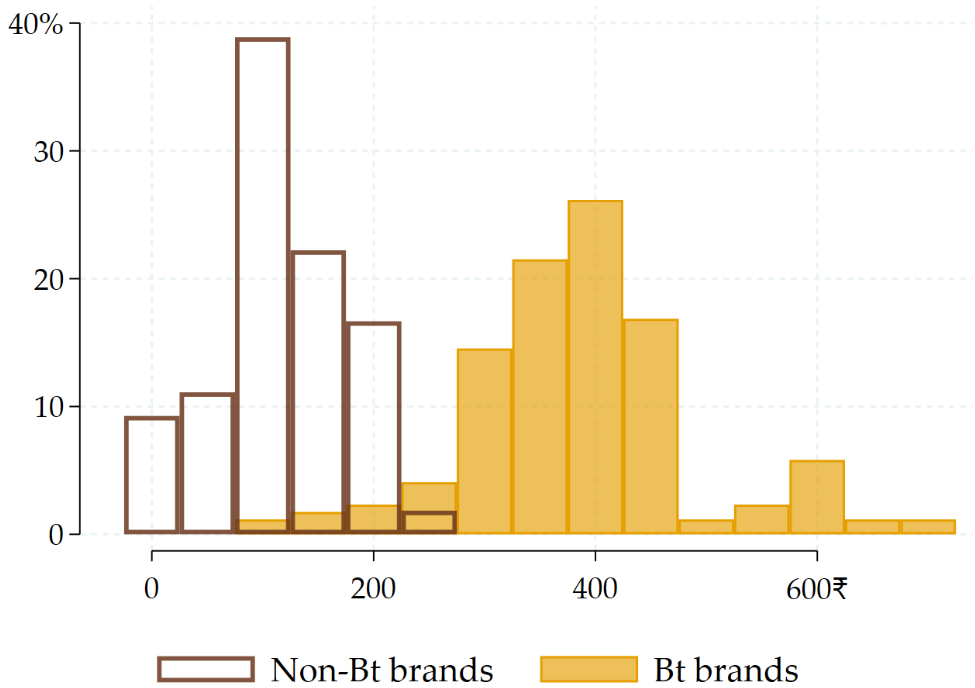
	(1)	(2)	(3)
Plot size (ha)	1.432**	1.439**	1.494***
	(0.598)	(0.603)	(0.565)
Outside option log-profits-per-ha	-3.059***	-2.828***	-5.065***
	(0.492)	(0.545)	(0.953)
Interaction (centered)	4.270***	4.204***	4.173***
	(1.126)	(1.128)	(1.122)
Bt linear time trend		-0.000	0.431***
		(0.000)	(0.134)
Number of observations	226	226	226
Effective *F*-statistic	17.1	13.0	13.2
Product fixed effects			✓

Notes: **Significant at 5%. ***Significant at 1%. Unit of observation: brand × state × year. The outcome variable is yields (in 100 kilograms per hectare). Data on inside good from Francis Kanoi Marketing Research's Cotton Crop Track. Data on outside good sampled with replacement from the Cost of Cultivation/Production Survey to match shares from the Crop Production Statistics. All regressions are least squares with robust standard errors (in parentheses). The effective *F*-statistic is based on the weak instrument test of Montiel Olea and Pflueger (2013). The second stage is in Table 3.

Figure E2. Distribution of Price and Yield Elasticities

Notes: The two histograms plot price (in red) and quality (in green) elasticities implied by the structural estimates of demand in Table 2. Unit of observation: market × product. The vertical axis indicates the share of observations that have a value of elasticity within the bin in the horizontal axis.

Figure E3. Distribution of Marginal Cost Estimates



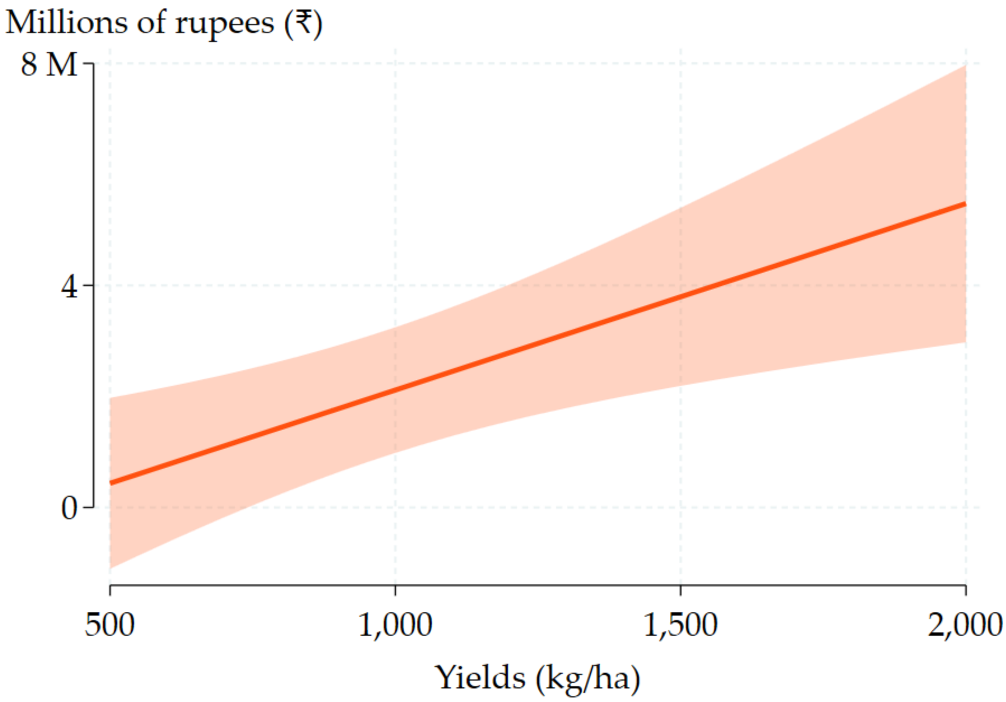
Notes: The two histograms plot marginal costs of non-Bt brands (i.e., *desi* public varieties and non-Bt hybrids) and of Bt brands. Marginal costs are recovered through the estimation procedure outlined in Section 5.2 and Appendix Section E.3. Unit of observation: brand × state × year. The vertical axis indicates the share of observations that have a value of marginal costs within the bin in the horizontal axis.

Table E7. Treatment Effects on Structural Model Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
	Prices			Yields of Bt products		
Price control × Post-2005	-0.602***	-0.585***	-0.483**	-0.243*	-0.241*	-0.141*
	(0.175)	(0.143)	(0.147)	(0.116)	(0.115)	(0.071)
Number of observations	321	321	321	245	245	245
Number of clusters	9	9	9	9	9	9
Year fixed effects		✓	✓		✓	✓
State fixed effects			✓			✓

Notes: *Significant at 10%. **Significant at 5%. ***Significant at 1%. Unit of observation: brand × state × year. All regressions are least squares with fixed effects (indicated in the last two rows of the table), weighted by brand share within each state-year, and clustered standard errors at the state level (in parentheses). Equilibrium outcomes on the right hand side are consistent with our structural model in Section 5 and estimated through the quasi-Newton algorithm described in Appendix E.4. They are expressed in natural logarithms, so that coefficients approximate percentage changes.

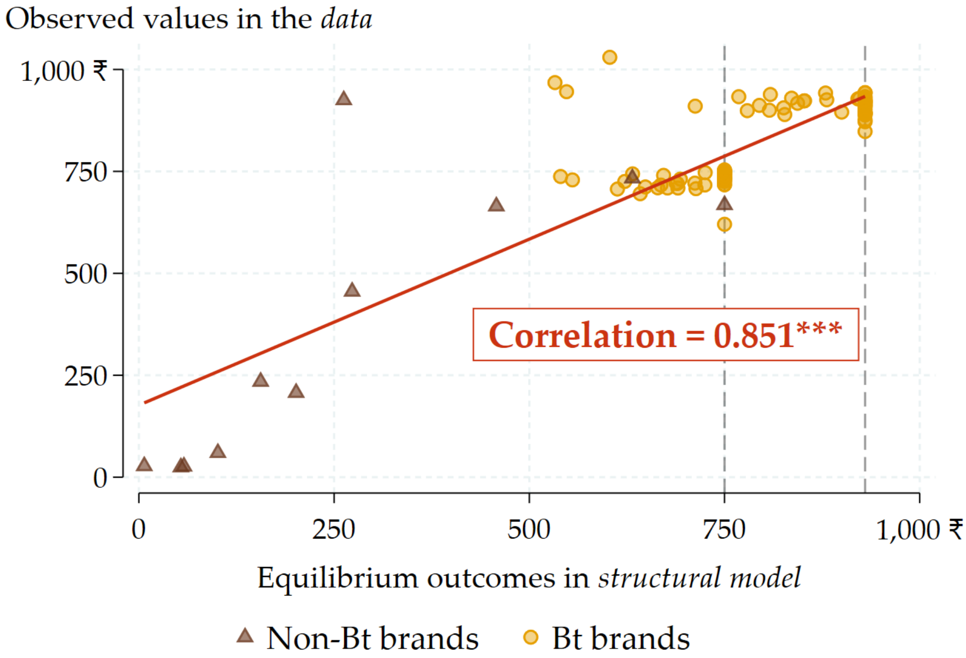
Figure E4. Slope of Fixed Costs



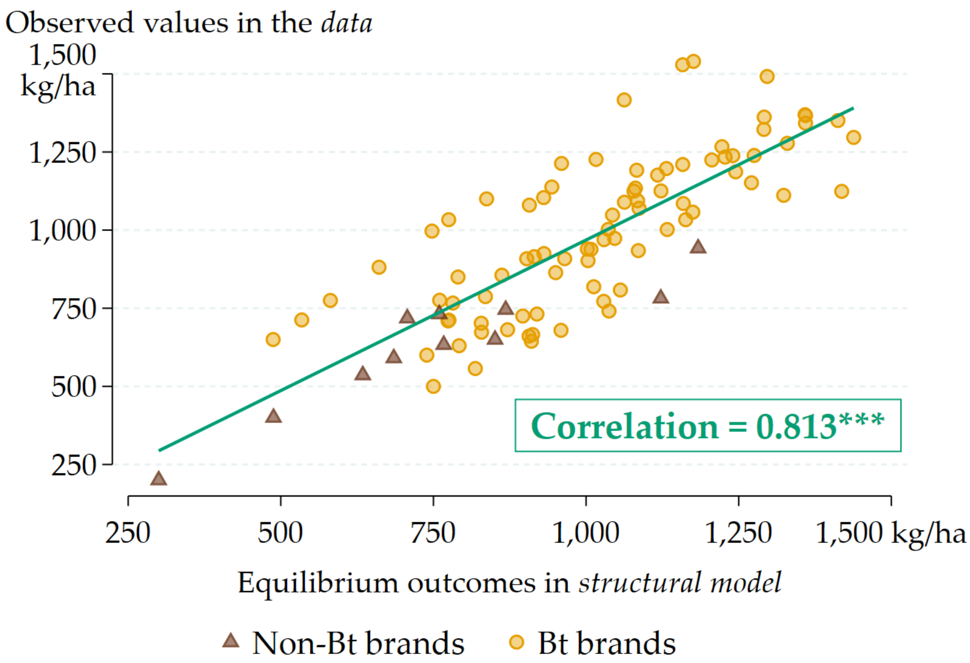
Notes: The solid orange line plots the slope of fixed costs, as defined in Equation 10, using GMM estimates from Column (3) of Table 3. 95 percent confidence intervals, based on robust standard errors, are shaded.

Figure E5. Goodness of Fit between Data and Model

(a) Prices



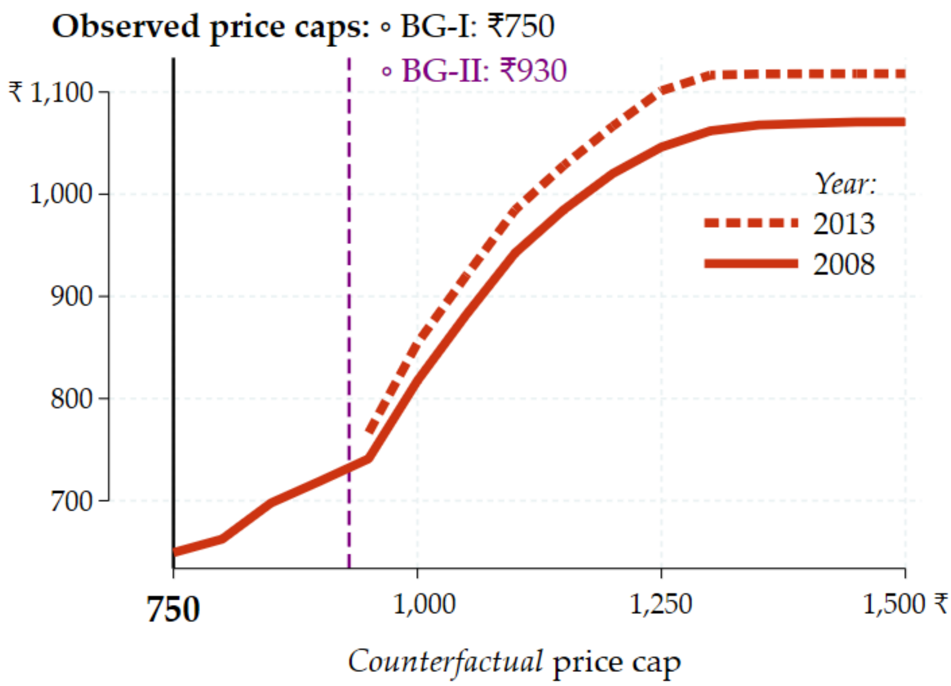
(b) Yields



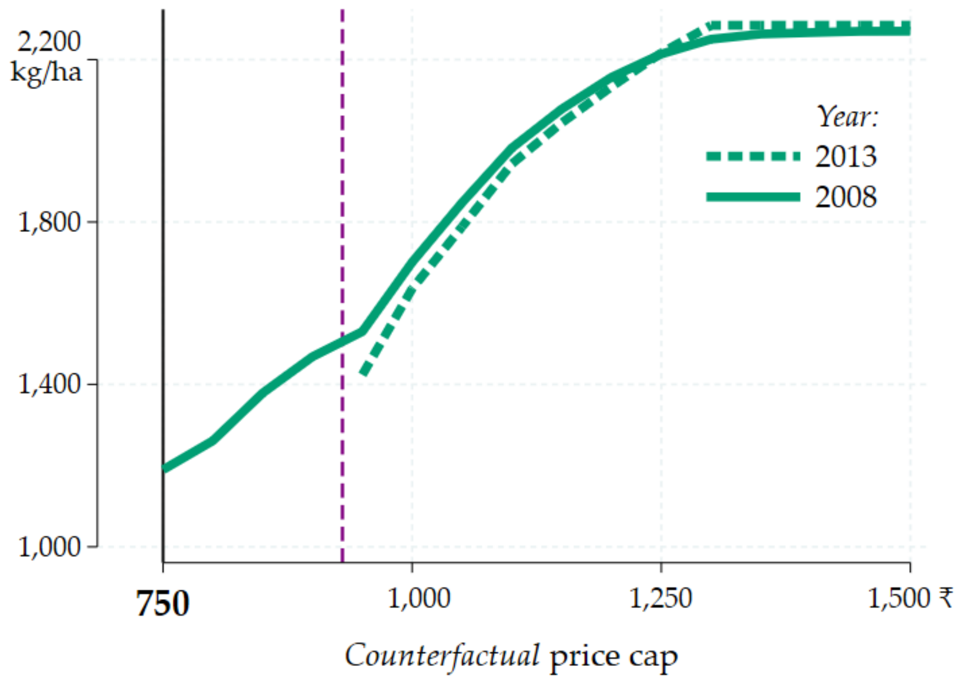
Notes: Unit of observation: brand × state × year. Data on prices and yields (on the y-axis) are from Francis Kanoi Marketing Research's COTTON CROP TRACK. Equilibrium outcomes are consistent with our structural model in Section 5 and estimated through the quasi-Newton algorithm described in Appendix E.4.

Figure E6. Average Product Characteristics by Counterfactual Price Caps

(a) Prices



(b) Yields

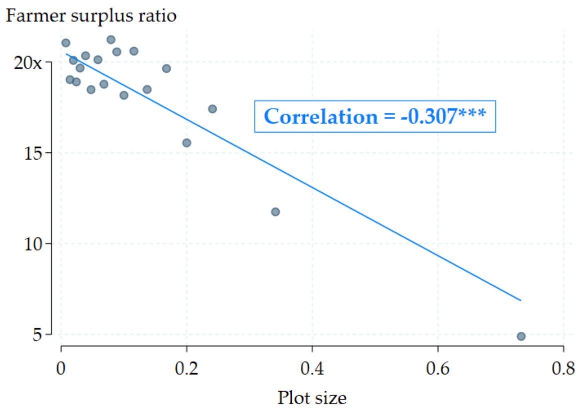


Notes: ***Significant at 1%. Unit of observation: brand × state × year. Average equilibrium outcomes (on the y-axis) are consistent with our structural model in Section 5, estimated through the quasi-Newton algorithm described in Appendix E.4, and weighted by product market share. The lines start at the observed price cap: ₹750 in 2008 and ₹930 in 2013.

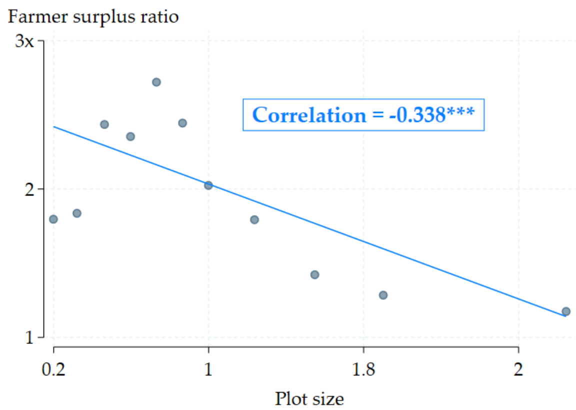
Figure E7. Distribution of Welfare Gains across Plot Size

Observed Policy versus Benchmark

(a) Full Population of Farmers

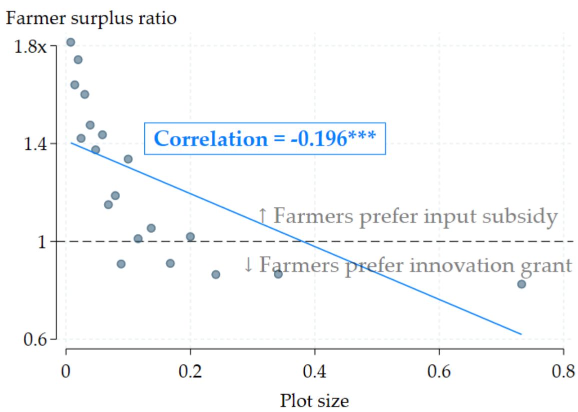


(b) Sub-Population of Cotton Farmers under Benchmark Scenario of No Policy

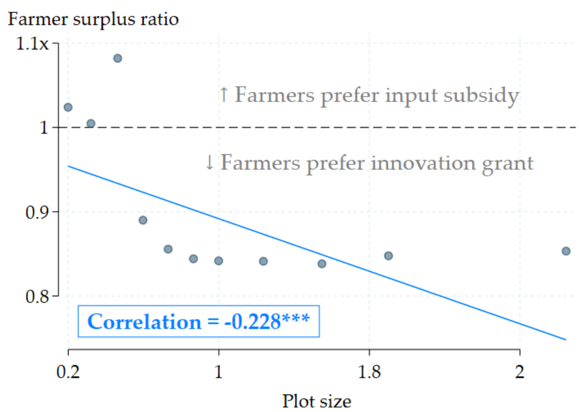


Input Subsidy versus Innovation Grant

(c) Full Population of Farmers



(d) Sub-Population of Cotton Farmers under Benchmark Scenario of No Policy



Notes: ***Significant at 1%. Unit of observation: household × plot. Estimation sample in the figure caption. Binned scatterplots are obtained by grouping farmer surplus ratio (on the y-axis) and plot size (on the x-axis) into twenty equal-sized bins. The solid line plots a linear fit, i.e., the prediction for farmer surplus ratio on plot size. Farmer surplus under each scenario is based on the logit formula in Equation 13. The ratios are equal to farmer surplus under price controls ("observed policy") divided by farmer surplus under no policy ("benchmark") in the upper panel and to farmer surplus under a 54.9% linear farm subsidy ("input subsidy") divided by farmer surplus under an aggregate-welfare-equivalent firm subsidy ("innovation grant") in the bottom panel. Data on plot size for cotton farmers are from Francis Kanoi Marketing Research's Cotton Crop Track and for non-cotton farmers are sampled with replacement from the Cost of Cultivation/Production Survey to match shares from the Crop Production Statistics.

Additional References Cited in the Appendices

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge (2023) "When Should You Adjust Standard Errors for Clustering?" *The Quarterly Journal of Economics*, 138 (1), 1–35, 10.1093/qje/qjac038.

Agrawal, RC (2019) "Opportunities and Challenges Created by the Plant Variety Protection and Farmers' Rights Act in India," in Adhikari, Kamalesh and David Jefferson eds. *Intellectual Property Law and Plant Protection*, Chap. 5, 86–103: Routledge, 10.4324/9780429059520.

Arellano, Manuel (1987) "Computing Robust Standard Errors for Within-Groups Estimators," *Oxford Bulletin of Economics and Statistics*, 49 (4), 431–434, 10.1111/j.1468-0084.1987.mp49004006.x.

Basu, A.K. (1983) "Spectrum of Hybrid Cotton in India," *Indian Farming*, 33 (9), 16–22.

Basu, A.K. and R.S. Paroda (1995) *Hybrid Cotton in India: A Success Story*, Bangkok, Thailand: Asia-Pacific Association of Agricultural Research Institutions.

Bell, Robert M and Daniel F McCaffrey (2002) "Bias Reduction in Standard Errors for Linear Regression with Multi-Stage Samples," *Survey Methodology*, 28 (2), 169–182.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (2004) "How Much Should We Trust Differences-in-Differences Estimates?" *The Quarterly Journal of Economics*, 119 (1), 249–275, 10.1162/003355304772839588.

Blaise, Desouza (2021) "Cotton (*Gossypium Species*) Production Systems of India: Historical Perspective, Achievements and Challenges," *Indian Journal of Agronomy*, 66 (2), 119–128.

Blaise, Desouza, MV Venugopalan, and AR Raju (2014) "Introduction of Bt Cotton Hybrids in India: Did It Change the Agronomy?" *Indian Journal of Agronomy*, 59 (1), 1–20.

Brubaker, Curt L, Fred M Bourland, and Jonathan F Wendel (1999) "The Origin and Domestication of Cotton," in Smith, C Wayne and Joe Tom Cothren eds. *Cotton: Origin, History, Technology, and Production*, 4, 3–31, New Work City, New York: John Wiley & Sons.

Donald, Stephen G and Kevin Lang (2007) "Inference with Difference-in-Differences and Other Panel Data," *The Review of Economics and Statistics*, 89 (2), 221–233, 10.1162/rest.89.2.221.

Ferman, Bruno and Cristine Pinto (2019) "Inference in Differences-in-Differences with Few Treated Groups and Heteroskedasticity," *Review of Economics and Statistics*, 101 (3), 452–467, 10.1162/rest_a_00759.

Gulati, A.N. and A.J. Turner (1929) "A Note on the Early History of Cotton," *Journal of the Textile Institute Transactions*, 20, 1–9.

Hansen, Bruce E and Seojeong Lee (2019) "Asymptotic Theory for Clustered Samples," *Journal of Econometrics*, 210 (2), 268–290, 10.1016/j.jeconom.2019.02.001.

Ibragimov, Rustam and Ulrich K Müller (2016) "Inference with Few Heterogeneous Clusters," *Review of Economics and Statistics*, 98 (1), 83–96, 10.1162/REST_a_00545.

Kolady, Deepthi Elizabeth, David J Spielman, and Anthony Cavalieri (2012) "The Impact of Seed Policy Reforms and Intellectual Property Rights on Crop Productivity in India," *Journal of Agricultural Economics*, 63 (2), 361–384, 10.1111/j.1477-9552.2012.00335.x.

Kranthi, Keshav Raj, Deepak Jadhav, Ravindra Wanjari, Sandhya Kranthi, and Derek Russell (2001) "Pyrethroid Resistance and Mechanisms of Resistance in Field Strains of *Helicoverpa Armigera* (Lepidoptera: Noctuidae)," *Journal of Economic Entomology*, 94 (1), 253–263, 10.1603/0022-0493-94.1.253.

Liang, Kung-Yee and Scott L Zeger (1986) "Longitudinal Data Analysis Using Generalized Linear Models," *Biometrika*, 73 (1), 13–22, 10.2307/2336267.

MacKinnon, James G, Morten Ørregaard Nielsen, and Matthew D Webb (2023) "Testing for the Appropriate Level of Clustering in Kinear Regression Models," *Journal of Econometrics*, 235 (2), 2027–2056, 10.1016/j.jeconom.2023.03.005.

MacKinnon, James G and Matthew D Webb (2017) "Wild Bootstrap Inference for Wildly Different Cluster Sizes," *Journal of Applied Econometrics*, 32 (2), 233–254, 10.1002/jae.2508.

——— (2018) "The Wild Bootstrap for Few (Treated) Clusters," *The Econometrics Journal*, 21 (2), 114–135, 10.1111/ectj.12107.

Montiel Olea, José Luis and Carolin Pflueger (2013) "A Robust Test for Weak Instruments," *Journal of Business & Economic Statistics*, 31 (3), 358–369, 10.1080/00401706.2013.806694.

Narayanan, S.S., P. Vidyasagar, and K.S. Babu (2014) "Cotton Germplasm in India – New Trends," in Abdurakhmonov, Ibrokhim ed. *World Cotton Germplasm Resources*, 97–118: InTech Open, 10.5772/58622.

Nevo, Aviv (2000) "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand," *Journal of Economics & Management Strategy*, 9 (4), 513–548, 10.1111/j.1430-9134.2000.00513.x.

NSSO (2014) "Income, Expenditure, Productive Assets and Indebtedness of Agricultural Households in India," National Sample Survey Office, Ministry of Statistics and Programme Implementation, Government of India, `https://mospi.gov.in/sites/default/files/publication_reports/nss_rep_576.pdf`.

Pustejovsky, James E and Elizabeth Tipton (2018) "Small-Sample Methods for Cluster-Robust Variance Estimation and Hypothesis Testing in Fixed Effects Models," *Journal of Business & Economic Statistics*, 36 (4), 672–683, 10.1080/07350015.2016.1247004.

Ramaswami, Bharat, Milind Murugkar, and Mahesh Shelar (2009) "Product Proliferation in India's Cotton Seed Market: Are There Too Many Varieties?" *Journal of Agricultural & Food Industrial Organization*, 7 (1), 10.2202/1542-0485.1256.

Razaq, Muhammad, Robert Mensah, and Habib-ur-Rehman Athar (2019) "Insect Pest Management in Cotton," in Jabran, Khawar and Bhagirath Singh Chauhan eds. *Cotton Production*, Chap. 5, 85–107, New Work City, New York: John Wiley & Sons, 10.1002/9781119385523.ch5.

Roth, Jonathan, Pedro HC Sant'Anna, Alyssa Bilinski, and John Poe (2023) "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature," *Journal of Econometrics*, 235 (2), 2218–2244, 10.1016/j.jeconom.2023.03.008.

Santhanam, Vaidhyanathaswamy and Sir Joseph Burtt Hutchinson (1975) "Cotton," in *Evolutionary Studies in World Crops*, 89–100, Cambridge, UK: Cambridge University Press.

Santhanam, Vaidhyanathaswamy and V Sundaram (1997) "Agri-History of Cotton in India: an Overview," *Asian Agri-History*, 1, 235–251.

Santhanam, Vaidhyanathaswamy, V Sundaram, and SS Narayanan (2016) "Historical Perspective of Cotton in India," *Asian Agri-History*, 7, 1–11.

Sethi, M. L. (1960) "History of Cotton," in *Cotton in India, a Monograph*, 1, 1–39, Mumbai, India: Indian Central Cotton Committee.

Sikka, S.M. and A.B. Joshi (1960) "Breeding," in *Cotton in India, a Monograph*, 1, 137–335, Mumbai, India: Indian Central Cotton Committee.

Stone, Glenn Davis (2007) "Agricultural Deskilling and the Spread of Genetically Modified Cotton in Warangal," *Current Anthropology*, 48 (1), 67–103, 10.1086/508689.

Suresh, A, P Ramasundaram, Josily Samuel, and Shwetal Wankhade (2014) "Cotton Cultivation in India since the Green Revolution: Technology, Policy, and Performance," *Review of Agrarian Studies*, 4 (2).

Suresh, Narayanan and Ch Srinivas Rao (2009) "Profiles of Four Top Biotech Companies in India. Rasi Seeds: Changing the face of India's bioagriculture," *Biotechnology Journal: Healthcare Nutrition Technology*, 4 (3), 295–300, 10.1002/biot.200900028.

Venugopalan, M. V., Desouza Blaise, and Sandhya Kranthi (2013) "Cotton," in Prasad, Rajendra ed. *Text Book of Field Crops Production, Vol. 2: Commercial Crops*, 305–344, New Delhi, India: Indian Council of Agricultural Research.

Venugopalan, MV, K Sankaranarayanan, Desouza Blaise, P Nalayini, CS Prahraj, and B Gangaiah (2009) "Bt Cotton (*Gossypium sp.*) in India and its Agronomic Requirements – A Review," *Indian Journal of Agronomy*, 54 (4), 343–360.

Webb, Matthew D (2023) "Reworking Wild Bootstrap-Based Inference for Clustered Errors," *Canadian Journal of Economics/Revue canadienne d'économique*, 56 (3), 839–858, 10.1111/caje.12661.

White, Halbert (1980) "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, 817–838, 10.2307/1912934.

Zohary, Daniel, Maria Hopf, and Ehud Weiss (2012) *Domestication of Plants in the Old World: The Origin and Spread of Domesticated Plants in Southwest Asia, Europe, and the Mediterranean Basin*, Oxford, UK: Oxford University Press, 4th edition, 10.1093/acprof:osobl/9780199549061.001.0001.