Does Firm Ownership Structure Matter? Evidence from Sugar Mills in India*

Sandip Sukhtankar†
Dartmouth College

February 26, 2016

Abstract

This paper examines the effect of firm ownership structure on firm behavior and the economic outcomes of upstream suppliers by comparing privately owned sugar mills to cooperatives and public mills in India. In this setting, government support for cooperative and public mills allows meaningful variation in ownership structure to be observed, while the "command area" zoning system - whereby mills are given monopsony power to operate within an assigned area - helps tackle the identification challenge. The borders of command areas allow for a geographic regression discontinuity design, where underlying soil, weather, and institutional characteristics are exactly the same but ownership structure changes across boundaries. Using satellite images overlaid on digital maps to measure sugarcane grown along the borders, as well as a survey to determine the effects of crop choices on farmer welfare, I find that private mills encourage sugarcane production. Greater cane cultivation is tied to better credit provided by private mills, and it results in higher income and consumption for farmers.

JEL codes: D22, H19, L23, L33, O25, Q13

Keywords: ownership, firms, cooperatives, sugar, agricultural markets, remote sensing, satellite data

*I am extremely grateful to Sendhil Mullainathan for collaboration on previous versions of this paper, to Jonathan Chipman for his patient and generous assistance with remote sensing, and Kevin Mpwenda for help with GIS analysis. I also thank Daniel Keniston, Gabrielle Kruks-Wisner, Rocco Macchiavello, Ulrike Malmendier, Dilip Mookherjee, Karthik Muralidharan, Paul Niehaus, Rohini Pande, Andrei Shleifer, Monica Singhal, two anonymous referees, and seminar and workshop participants at Dartmouth, NEUDC (Yale), Econometric Society Meetings 2012, Harvard Kennedy School, NBER Organizational Economics Workshop 2014, and the Indian Statistical Institute-Delhi for comments, and Aravind Moorthy, Anand Kothari, Kelly Tapani, Bernard Rump, Deepa Ranganathan, Edward Lescher, Kevin Li, Jung Kim, Brendan Norris, Marco Martinez del Angel, Shivam Chadha, and Zheng-Yi Yang for excellent research assistance. I thank Dominik Bulla, Sujatha Suresh, and Vidhya Muthuram for managing field operations at CIFD and CDF, and K Sridevi and KPR Menon at the National Remote Sensing Centre (NRSC) for assistance with the satellite images. I gratefully acknowledge funding from the International Growth Center (IGC).

†BREAD, JPAL, and the Department of Economics, Dartmouth College, 326 Rockefeller Hall, Hanover, NH 03755. sandip.sukhtankar@dartmouth.edu.

1 Introduction

How does organizational form affect firm behavior and performance? While this question has received much theoretical attention, empirical analysis is more challenging.[1] In addition to the usual identification challenge, observing meaningful variation in organizational form in practice is difficult. If theory suggests that one organizational form dominates, then one should only observe that particular type of organization in the data; if, on the other hand, we observe variation in organization form, this likely indicates that organizational form is unimportant. Answering the question, "How does the performance of a firm that adopted a particular arrangement compare with how *that same firm* would have performed had it adopted an alternative?" (Masten, 2002) is, therefore, a formidable task.

This paper examines how variation along one dimension of organizational form – ownership structure – affects firm behavior and the economic outcomes of upstream suppliers. The sugar industry in India provides us with a useful context in which government regulation helps resolve empirical challenges. First, the "observation" challenge is overcome since government support – financial and otherwise – ensures the existence of cooperative and government-owned sugar mills operating in parallel with privately owned mills. Second, regulations governing the operations of sugar mills address the econometric challenge. Mills are subject to a zoning system wherein they are assigned monopsony power to purchase sugarcane from farmers in a specific "command area"; these areas are historically fixed and clearly delineated, and their borders can be considered to be randomly placed. Command area boundaries provide a regression discontinuity design, since farmers on either side of the boundary must sell to mills of different ownership types – cooperative, private, and public – even though other factors such as weather, soil quality, institutions, etc. are constant across the borders. Thus, any differences in farmer outcomes can be attributed to differences in ownership structure right at the border.[2]

The question of ownership structure is particularly important in agricultural markets, in which issues of hold-up and economies of scale often compel governments to nationalize firms or set up cooperatives. For example, raw produce takes a long time to grow but must be processed immediately after harvest, while processing plants require large-scale investments; the resulting threat of hold-up and monopoly power motivates government intervention. Non-governmental organizations also attempt to intervene: for example, the fair trade movement has become synonymous with small farmer cooperatives, in the process channeling large amounts of funds and technical assistance to these associations.[3] The idea of cooperatives has generated enormous interest; UN

[1] Shleifer (1998) and Megginson and Netter (2001) review the vast theoretical and empirical literature.

[2] There may well be selection into who becomes a farmer or grows sugarcane, but this could also be attributed to differences in ownership structure. Inframarginal characteristics of the command area may matter too. I discuss these and other qualifications below.

[3] See, for example, the review by Dragusanu et al. (2014), and the websites of organizations promoting fair trade: (`http://www.cdf.coop/`), `http://www.globalexchange.org/fairtrade/coffee/cooperatives`.

Secretary General Ban Ki-moon, in declaring 2012 the "International Year of Cooperatives", said that "cooperatives are a reminder to the international community that it is possible to pursue both economic viability and social responsibility."[4] Whether governments should subsidize and promote cooperatives is therefore an empirical question, one that assumes significant importance given the scarce public resources directed towards cooperatives in developing countries.

This paper captures one aspect of performance – the effect on upstream suppliers – which is important in this context, given the presumption that cooperatives maximize farmer welfare. In order to examine the effects of ownership structure, this study uses several unique sets of data on farmer outcomes. The first set of data comprises multi-spectral satellite images that allow me to directly observe sugarcane grown along the borders. The second set is a more conventional survey to determine the effects of crop choices on farmer welfare, which includes questions about income and mill interactions. The third and final set of data is based on soil tests from plots along borders to test whether results are driven simply by variation in soil quality.

Before I analyze whether ownership structure affects any of these outcomes, however, I check whether the empirical strategy is valid. Figure 1 shows the standard regression discontinuity figure, and it is clear that a stark discontinuity exists at the border. Meanwhile, other characteristics are continuous as we move across borders. Importantly, soil testing confirms that there are no differences in soil quality (Table 1). Differences in intrinsic farmer characteristics such as literacy can also be ruled out. On the other hand, farmers on the private side of the border seem to have more land, and are situated farther from private mills. The former could potentially be an outcome of private ownership, and hence, I show estimations with and without acreage as a control; the latter, meanwhile, I control for flexibly in all estimations. Finally, to ensure that conditions away from the border are not affecting mill operations and hence outcomes at the border, I also control for the amount of cultivable area available, mill crushing capacity, mill age, soil types, and elevation/ ruggedness in the command area (i.e., conditions that are not a result of mill performance).

Examining first the question of whether hold-up issues lead to undersupply of cane to private mills, I find the exact opposite result: private mills encourage more sugarcane production than co-operative mills. Overlaying satellite images on maps of command areas, I determine that the sides of the borders owned by private mills are actually planted with a greater proportion of sugarcane (by about 2 percentage points, or 4 percent) than those owned by cooperative or government mills. This result is mirrored in the surveys of farmers with plots close to the borders, with farmers on the private side about 5 percentage points (22 percent) more likely to be growing cane.

How does the growing of extra cane translate into farmer welfare? It appears that total income, farm income, and consumption are actually higher for farmers on the private side of the border. In addition, all these variables are particularly higher for land-poor farmers on the private side of the

[4]See `http://www.un.org/en/events/coopsyear/`

border. These results are in contrast to claims that cooperatives particularly help poor farmers (see, for example, the UN website on Year of the Cooperative: `http://www.un.org/en/events/coopsyear/`). Note that farmers with similar amounts of land may have different unobservable characteristics across the borders, given selection into farming or other changes induced by longer term effects of ownership. To the extent that this is true, the results on land-poor farmers must be interpreted with caution.

Do private and cooperative mills differ in the interactions they have with their suppliers? While these results are not very robust, I find that private mills appear to provide more loans to land-poor sugarcane farmers, and also pay higher prices to these farmers,[5] possibly encouraging them to cultivate cane. Since sugarcane is a lumpy crop, with harvest coming a year to eighteen months after planting, advance credit is critical to cane farmers (see also Macchiavello and Morjaria (2014) on coffee farmers).

These results suggests that in identical conditions at the border, private and cooperative mills perform differently as captured by farmer outcomes. One caveat is that it is possible that conditions away from the border affect how firms perform at the border - for example, if private mills are consistently located on better quality land - and the controls used may not capture unobservables. However, for this caveat to affect the interpretation of the results, the following two conditions must both be true (i) inframarginal characteristics of command areas are correlated with ownership structure in a way that favors private mills, and (ii) there are other constraints, such as credit constraints, which affect the ability of cooperative mills to function. For example, even if inframarginal characteristics of command areas favor private mills, if there are no credit constraints then both mills should invest in border areas until marginal cost equals marginal returns, which should be the same since the areas are identical. If capital constraints are on the other hand the only issue, then this does not affect the interpretation of the results as firm organizational structure matters; rather it just highlights the fact that organizational structure matters because of access to credit. Of course, it is also possible that both conditions are met. The paper attempts to show, using the best available data, that condition (i) is likely not true; but results must be viewed with caution given that it is impossible to conclusively show this.

The discussion above suggests that access to capital may be one mechanism for these results: for example, cost of capital may be lower for private mills (Rey and Tirole, 2007; Hart and Moore, 1998; Kremer, 1997), which helps them provide loans and other inputs to farmers. Unfortunately, lack of data on interactions between mills and farmers, as well as data on mill finances, limits the ability to say anything categorically on mechanisms, although there are several possibilities. First, it is possible that long-run profit maximization may not be incompatible with keeping suppliers happy; the problems seen as characteristic of private firms – monopsony power, hold-up, etc – could be mitigated by repeated interactions between these firms and farmers. The

[5]While state governments regulate cane prices by setting a price floor, prices are tied to the relative efficiency of the mill, and mills are free to pay above those prices. I discuss cane pricing in more detail in the Section 2.

results are consistent with the simplest models of monopsony purchase, under which more efficient private mills both purchase more inputs and pay higher prices. Second, the objective function of cooperative management might not be the same as that of its members. For example, Banerjee et al. (2001) find that richer farmers – who control the cooperative – prefer that the cooperative not distribute profits as higher prices for cane. Given this and the political objectives of cooperative chairmen (Sukhtankar, 2012), interference in management is likely.[6]

This paper contributes to, and sits at the intersection of, two literatures: (1) the impact of organizational form on economic outcomes, and (2) the industrial organization of agricultural markets in developing countries. Although a vast theoretical literature exists within the first area,[7] clean empirical estimates of the impact of organizational form on economic outcomes are not as common. In the particular case of firm ownership, endogeneity concerns plague existing estimates (for example, selection – governments may choose to sell the worst-performing units – or corruption – the value of state units may be deliberately suppressed) (Megginson and Netter, 2001; Masten, 2002).

Within the second set of literature, this paper is directly concerned with work on understanding rural cooperatives, which have gained prominence as potential pathways to assist poor rural farmers, although empirical work on the impact of cooperative versus private ownership has so far been missing. This paper is also related to research on interlinked transactions, which has a long history in development economics. Early theoretical work by Bardhan (1983) and Mukherjee and Ray (1995) has been complemented by more recent empirical work by Blouin and Macchiavello (2013) on coffee mills and foreign clients; by Casaburi and Macchiavello (2014) on loans to dairy producers in Kenya; and, most closely related to the current paper, work by Macchiavello and Morjaria (2014) on inputs and loans provided to coffee farmers in Rwanda. The command area system exploited for identification in my paper also relates to work on agricultural marketing arrangements and market structure. Such systems are not uncommon in the agricultural sector in developing countries where contracts are unenforceable (Macchiavello and Morjaria (2014)).

Finally, this paper also introduces an empirical innovation. The combined use of satellite and survey data to observe sugarcane provides a methodological proof-of-concept that can be extended to gathering data on other crops. While economists have previously used multi-spectral satellite images to measure forest cover (Foster and Rosenzweig, 2003), I use higher resolution (23.5m) data and actual field measurements to calibrate and measure the extent of sugarcane grown, thus conducting the first such analysis (to my knowledge) in the economics literature.[8]

[6]Note that lack of access to technology cannot explain the results: the machinery to produce sugar is easily available for purchase, and scope for research and development on this front is limited. Moreover, the impacts we see are on *inputs*, not *outputs*; it is difficult to imagine that sugarcane production is affected by mill hardware.

[7]See, for example, Hart et al. (1997); Boycko et al. (1996); Hart (2003); Laffont and Tirole (1991); Stiglitz (1994), amongst others.

[8]The World Bank's Development Impact Blog recently featured a post that suggested such techniques will become easier to implement as satellite data becomes more widely available: see `http://blogs.worldbank.org/impactevaluations/node/1225`, accessed March 20, 2015. Published papers in economics are as yet unknown.

The rest of the paper proceeds as follows. Section 2 provides background on the zoning system and the sugar industry in Tamil Nadu. Section 3 describes the empirical strategy, while section 4 presents the sample selection procedure and summary statistics. Section 5 provides checks of the empirical strategy. Section 6 presents the results on farmer outcomes; section 7 discusses mechanisms; while 8 concludes.

2 Background

The sugar industry in India offers several attractive features as an empirical setting to explore the impact of cooperative versus private ownership. Sugarcane is one of the biggest cash crops in India, and the sugar industry employs a substantial share of the rural population. Many industry features are common to sugarcane and other important cash crops such as coffee and cocoa, including institutional and contractual arrangements and zoning regulations (see, for example, Macchiavello and Morjaria (2014) on coffee). The next few sections describe sugar production, the sugar industry, and the ways that differences in ownership may matter practically in this context, and in the command area system.

2.1 The Sugar Production Process

Sugarcane is a cash crop grown in large parts of India, from the semi-arid tropics of the south to the sub-tropical plains of the north. It is a water- and fertilizer-intensive crop that takes a year to grow. Irrigation is usually necessary, although rainfall is also important, since it reduces irrigation costs. Sugar is made by crushing sugarcane via massive rollers to extract sucrose-rich juice. Lime is then added to the juice to balance pH and clump together impurities; sulfur is bubbled through to bleach the juice; and finally, it is boiled and refined to make processed crystalline sugar.

The intrinsic sucrose and water content of sugarcane help determine the potential amount of sugar that can be extracted from it. In addition, the efficiency and organization of the mill itself can play a large role in the quality of the crop. Once cane is harvested, it dries out rapidly, and hence must be crushed within hours of cutting. Given the generally poor transportation infrastructure in rural India, this means that farms cannot be located more than 15-20 kilometers from the factory. The coordination and efficiency of the mill determine how much sugar is obtained per ton of cane crushed. Mills need to coordinate cane harvesting in order to run the factory at precise capacity every day. If too much cane arrives at the factory gates daily, some of it cannot be crushed, and it dries out. If too little cane arrives, recovery is also lower due to the fixed width between the rollers. Moreover, keeping the rollers running is costly, so it may not be cost effective to run the machinery for small quantities of cane. Machinery breakdowns are also extremely costly, since the cane at the factory starts drying out, and the harvesting schedule must be readjusted.

Since prices paid to the farmer are per ton of cane (regardless of quality), drier cane means less

earnings for farmers. In order to ensure a regular supply of good quality cane, mills provide seeds, loans, and agricultural extension services to farmers. Each factory pays its farmers a unique price per metric ton of cane. A single price for sugarcane is paid per year on the basis of weight alone. Usually, a price is announced just before the beginning of the season (in September/October), and adjustments (upwards only) are made at the end of the season. Sugar prices and rainfall affect cane prices, as does the recovery rate of the mill, as statutory prices are tied to this recovery rate.

2.2 The Sugar Industry and Ownership Structure

The sugar industry emerged in north India after sugar tariffs were imposed in the 1930s, with the establishment of private British and Indian sugar-producing factories in Uttar Pradesh and Bihar. After Independence, the federal government, as well as state governments, made their way into sugar production. In the 1950s, the cooperative sector burgeoned in the western state of Maharashtra, and from there, it spread to other states (Baru, 1990).

Historically, cooperatives were a response to the government's distrust of powerful landowners and private industry. Public funds were (and still are) used to set up mills, provide bailouts when the mills faced threats of bankruptcy, provide subsidized loans for operation, and provide state-guaranteed loans for many other purposes. In addition to funding cooperatives, both State and Central governments have also heavily regulated the sugar industry. Table A.1 presents the list of operating mills in Tamil Nadu as of 2010, along with their ownership structure; note that none of these mills has ever experienced a change in ownership structure.

How might ownership structure matter in the sugar industry? First, as has been alluded to above, the objective function of private versus cooperative mills might be different. Certainly there is a strong perception that cooperative mills maximize farmer welfare, while private mills may exploit monopsony power and hold-up issues. As described above, sugarcane must be crushed as soon as it is harvested, and farmers cannot sell their cane to mills that are far away. Combined with the fact that there are economies of scale in crushing sugarcane, mills have local monopsony power and the opportunity to hold-up farmers ex post. Farmers may anticipate these problems and undersupply cane, and one might expect this problem to affect private mills more than it does cooperatives or public mills.

Second, the management of cooperatives and private mills may differ along several dimensions. Most importantly, chairmen of cooperatives must be elected by member-shareholders from within their group. Given the political economy of rural sugarcane farming, these chairmen tend to be active in rural politics (Baru, 1990; Sukhtankar, 2012), and consequently have different objectives for the mill than other members. While these chairmen may not be active in day-to-day management of the mill, they are responsible for hiring operating officers, and may prefer to hire managers who maximize the chairmen's objective function rather than profits or farmer welfare. Compensation for managers and other mill employees may also be restricted in terms of the incentives they may

earn for performance, as cooperative mill by-laws are influenced by government hiring practices with their attendant compression of wage structures.

Finally, private firms may raise financing better than cooperatives do (Rey and Tirole, 2007; Hart and Moore, 1998; Kremer, 1997). This may be reflected in the fact that these mills tend to be larger, although this could also reflect constraints on management and the ability to attract more member-shareholders. Cooperatives, moreover, also have access to subsidized credit from the government.

2.3 The Command Area System

The constraints imposed by the fact that cane must be crushed immediately after harvest means that sugar factories cannot bring in cane from large distances. Moreover, the technology of cane crushing has large economies of scale, and thus gains to be had from building large factories. Finally, unlike in Brazil, where cane is grown in plantations owned by the sugar factories themselves, cane in India is grown by a large number of individual farmers. Combined, these factors mean that ensuring adequate supplies of cane is a first-order problem for sugar mills in India.

Adequate supply for processing plants is a pervasive problem in agricultural markets in developing countries for cash crops such as sugarcane, coffee, and cocoa. Blouin and Macchiavello (2013) and Macchiavello and Morjaria (2014) describe a similar problem in the market for coffee cherries in Rwanda. The problem is exacerbated by problems with contract enforcement in these contexts.

The government's solution to this problem was to designate reserved sugarcane zones for each mill, thereby limiting competition between mills for cane and providing incentives for the mill to assist in cane development within their zones.[9] The policing of the command areas is left to the mills, who have strong incentives to ensure that farmers do not sell their cane to other mills.[10] In practice, the complex relationship that a cane farmer needs to have with the mill to procure seed, fertilizer, credit, pesticide, etc effectively binds her to her current mill. The agricultural extension officers that mills send to assist farmers with growing cane also help the mill monitor farmers; moreover, because the cane must be crushed immediately after harvesting farmers cannot simply harvest their cane and show up at another mills door to sell it.[11] In order to control supply of

[9]This was an old idea; in a meeting of the Sugar Committee in 1933, a Mr. Noel Deerr noted that "With the adoption of a zone system, that is to say, with an area given over to the miller to develop in sympathy with the small holder, there should follow at once an association of agriculture and manufacture for the common benefit of both interests. It will be the object of the mill to reduce the price of the raw material and this can best be done by increasing the production per acre, and with an increment in the yield the net income of the small holder will increase even with a decrease in the rate paid per unit of raw material. (as cited in Baru (1990), p 33)"

[10]The Essential Commodities Act (1955, amended 2003) governs sugar regulation in India, along with various Sugarcane Control Orders. These acts and ordinances allow national and state governments to restrict entry in the sugar industry, including via the creation of command areas. States implement entry restrictions differently; while Tamil Nadu assigns specific command areas to mills, Maharashtra simply prohibits mills from opening within 15 kilometers of existing mills.

[11]This is in contrast to the coffee and dairy markets described in Macchiavello and Morjaria (2014) and Casaburi and Macchiavello (2014); there, the ease of transporting the raw material, and in the case of dairy daily harvests,

cane arriving at the factory, mills assign particular "cutting dates" to farmers.

In order to protect farmers from the monopsony power thus created, the government would set a floor for the price of cane to be paid by each mill. Currently, cane prices are regulated by state governments, which set a State Advisory Price (SAP); mills are, of course, free to pay above these floors, and more efficient mills will pay higher prices for cane as the SAP is tied to the recovery rate of sugar from cane. Sale of processed sugar is also restricted, with a certain proportion (which varies over the years, and is currently 10%) to be sold at a low rate ("levy price") to the Central Government, and the rest on the open market (at the "free price").

While some states have now abolished the command area system, replacing it with looser rules that require new factories to be built at least a certain distance (20-25km) away from existing factories, the system still exists in the state of Tamil Nadu. Most of the boundaries of the command areas of the 39 operating sugar mills in the state were historically set. Some followed natural geographical features, like rivers, canals, or hills. Others were set to equate the number of villages that neighboring mills had in their command areas. Anyone who wished to establish a new mill had to obtain permission from the sugar commissioner, proving that she had the potential to obtain adequate supplies of cane from a heretofore undesignated command area, or that existing mills were not using cane from their currently assigned areas.[12] It is, of course, possible and even likely that these command areas differ in various characteristics: however, as the section below explains, what is important for the empirical strategy is that the areas close to borders between private and cooperative/government sugar mills are not different from each other.

3 Empirical Framework

3.1 Empirical Strategy

The approach to estimating the effect of ownership structures on farmer outcomes uses regression discontinuity, similar to that followed by Black (1999). This approach takes advantage of a discontinuity in ownership structure at the border, while other characteristics – such as weather, soil quality, pest exposure, the institutional environment, etc – are continuous. The advantage of this approach over that of simply comparing farmers' outcomes in areas served by private and cooperative mills respectively is that it is difficult to control for all pertinent characteristics that may affect these outcomes. While average characteristics may differ across command areas, characteristics at the border should be continuous. Thus, instead of estimating:

$$Y_{ij} = \alpha + X'_{ij}\beta + A'_j\gamma + \delta Private_j + \epsilon_{ij} \tag{3.1}$$

make defection by farmers easy.

[12]Notes from August 2009 meeting with then Tamil Nadu Sugar Commissioner Mr. Sandeep Saxena and Tamil Nadu Sugar Corporation's Chief Cane Development Officer Dr. A. Sekar.

where Y is an outcome of interest for farmer i in area j, X are individual farmer characteristics, and A are area characteristics, and the outcome of interest is coefficient δ on a dummy variable indicating whether the area is served by a private mill, I estimate:

$$Y_{ib} = \alpha + X'_{ib}\beta + \sum_{1}^{B} \gamma_b + \delta Private_b + \epsilon_{ib} \tag{3.2}$$

where b is a particular border and a series of indicator variables γ control for characteristics that vary at the border. This approach makes sense when comparing the entirety of border areas, which I do in the satellite data analysis. However, the survey is based on sampling a few pairs of villages that are directly across from each other on different sides of the border; some borders may be very long, and there may be significant differences in characteristics on different parts of the border. In order to account for these differences, instead of including indicator variables for the border, I use indicator variables for the village pairs, and estimate:

$$Y_{ipb} = \alpha + X'_{ipb}\beta + \sum_{1}^{P} \nu_p + \delta Private_b + \epsilon_{ipb} \tag{3.3}$$

where p refers to the village pair.

In all estimations, I cluster standard errors at the mill-border level. While in principle the satellite data analysis covers the entire population of border villages, and hence there is no sampling error, other sources of error – such as differences in atmospheric conditions affecting the satellite images and bleeding of pixels across plots – remain. Hence, I present conservative estimates in the satellite data analysis: I first aggregate all data to the village level, rather than using each pixel as a separate observation, and continue to cluster standard errors by mill-border.

One potential concern with this strategy is that conditions far from the border but correlated with private or cooperative ownership may affect outcomes at the border. For example, a mill assigned a relatively less fertile area on average may perform worse than a mill assigned a relatively more fertile area, even though the border areas are equally fertile and all outcomes are measured at these border areas. However, note that for such external conditions to matter, one must assume that (i) external inputs matter for cane production at border areas and (ii) there are constraints on providing these external inputs that prevent mills from operating each area as an independent unit. Of the main inputs into production – land, labor, and capital – land and labor are locally provided, and there are no differences in quality of these inputs at the border.[13] Capital, in the form of loans to farmers for example, is more likely to be externally provided. Again, however, assuming no differences in conditions at the border, the returns to capital should be the same,

[13]One might imagine that management is an external input that would matter for production; however, management is also likely to be mobile, so it would be difficult to imagine a situation in which characteristics of command areas away from the borders determine the availability of management. As discussed above, differences in management are more likely to stem directly from ownership structure.

and there is no reason to expect differences in capital provision on either side of the border. Thus external conditions will only affect border areas if we expect there to be constraints on capital availability to the mills. To the extent that such constraints are possible, I control for external conditions in certain specifications as described below (section 5.3).

3.2 Threats to Discontinuity Design

Regression discontinuity designs that include geographical discontinuities must carefully consider three sets of issues: the process of boundary creation, the endogenous sorting of economic actors across boundaries, and the differences between regions other than the treatment of interest (Lee and Lemieux, 2010). I next explicitly consider these threats to internal validity and explain how this project deals with them. In addition, I also consider a common criticism of regression discontinuity-type designs, namely the external validity of the results.

1. **Process of Boundary Creation** As described above, the boundaries of command areas were historically set and are clearly delineated. I will also directly test observable characteristics to ensure that they do not vary across borders. Moreover, as is standard in these analyses, I will exclude any parts of boundaries that follow natural borders such as lakes, rivers, hills, etc. Finally, all decisions about which parts to include or exclude are transparent and available to anyone using Google Earth. Figure 3 presents a sample *taluk* (sub-district) split between two mills, showing how it is basically split down the middle into two mills' command areas. Figure 4 shows the distribution of mill border areas across Tamil Nadu.

2. **Endogenous Location of Farmers** Given that the boundaries have been historically set, it is possible that farmers move selectively across borders by purchasing land. For example, farmers who work harder might move to mills that reward effort. This is not a threat, however, to the validity of estimates but rather an interpretational issue. If farmers move because certain mills reward effort, this can still be interpreted as the causal impact of ownership structure, although due to selection rather than other mechanisms. Moreover, this kind of mobility can be measured to some extent by posing questions to farmers. While the survey did not directly ask about migration, it did ask about land sales, and the vast majority (75%) of farmers note that the land they farm was simply inherited rather than purchased. Thus, it is unlikely that endogenous movement of farmers drives the results.

3. **Other Differences between Regions** I directly test other relevant characteristics to ensure that they do not jump discontinuously across borders. The most obvious characteristic is soil, and we can directly measure soil traits such as granularity and chemical content that would affect crop choices and yields. Some of these characteristics – for example, the mineral content of the soil– might be affected by farmer effort such as application of fertilizer and indeed by

ownership structure. However, other characteristics such as the nature of soil and the size of soil grains are not affected by farmer effort. A remaining possible difference is that one side of the border is farther away from its mill than the other side; the estimates control non-parametrically for distance, and results are not different when restricted to borders located at similar distances from mills.

4. **External Validity** Since regression discontinuity estimates relate to observations close to the discontinuity, one concern is that they have limited external validity. Certainly in some contexts where the marginal complier is questionable or different from the rest of the population – for example, a student in an ability distribution with high variance where the cutoff is some score – this concern is valid. However, in the sugarcane farmers' context, it is difficult to imagine that farmers close to the border are systematically different from those who are not. It is possible that mills treat farmers who are close to the border in a different way than they treat other farmers, perhaps due to competition across the border; results from a small and hence admittedly underpowered survey of farmers at various distances from the mill do not show any differences in agricultural extension services provided by mills based on distance. Finally, these results from the sugarcane industry are applicable to various similar industries in India and elsewhere; for example, dairy and coffee (Macchiavello and Morjaria, 2014; Casaburi and Macchiavello, 2014).

4 Sample Selection and Data Description

Table A.1 presents the list of sugar mills that were operating in the state of Tamil Nadu in 2010. From the universe of potential borders between these mills, I did not consider those borders that were along a river, or separated by large geographic features like canals or mountains where the two sides are likely to be very different. I further considered only borders which did not overlap district/sub-district borders, since this would mean that the two sides are in different administrative divisions. In addition, I also collected soil samples from a subset of farmers, and tested these samples for various physical and chemical characteristics. Finally, the National Remote Sensing Centre (NRSC) of India provided satellite imagery in order to determine how much sugarcane was grown on either side of the border.

For the survey, I sampled pairs of villages, located across from each other and along command area boundaries, that did not overlap any major administrative divisions. This yielded 26 village pairs (52 villages) along 14 mill pair borders. Within these villages, I compiled a list of all plots that were within a kilometer of the border by obtaining land records from the Village Administrative Officer (VAO). The VAOs also denoted whether or not the plots were farmed with sugarcane. Based on this information, I picked a stratified random sample of sugarcane growers and non-growers, aiming to survey 25 sugarcane farmers and 15 non-sugarcane farmers in each village. All

regressions are weighted to account for these differential sampling probabilities, so that results are representative of the entire 2 kilometer strip (1 kilometer on each side) along the selected borders.

Table A.2 presents summary statistics separately for private and non-private mills. In general, the different areas appear to be balanced. The only obvious differences are in the average amount of land owned, and loans provided by mills; both these outcomes are discussed further in the results section below.

For a subsection of the surveyed farmers – approximately 3 per village – I collected soil samples from their fields. The samples were collected according to the procedures set forth by the Tamil Nadu Agricultural University on the following website: `http://agritech.tnau.ac.in/agriculture/agri_soil_sampling.html` (originally accessed October 2010, confirmed still active as of March 2015). The same institution conducted the analysis on the samples, providing us with data on the texture, type of soil, available amounts of nitrogen, phosphorus, and potassium, as well as the electrical conductivity and ph of the soil samples.

In addition, I obtained multi-spectral satellite images of the state of Tamil Nadu from the National Remote Sensing Centre (NRSC). In order to digitally distinguish vegetation as well as separate sugarcane from other crops, I make use of the fact that different crops will have different digital spectral signatures.[14] Overlaying the calibrated images over GIS maps of the borders allows us to determine how much sugarcane is growing on each side. I included only villages that were on the borders of command areas, and calculated the number of pixels in these villages that were crops in general and sugarcane, and then the proportion of crops that were sugarcane. Appendix A provides more details on the procedure.

5 Checks of Empirical Strategy

5.1 First Stage

This empirical strategy is valid if there is actually a discontinuity in ownership structure at the border and continuities in other characteristics. Whereas the law says that farmers must sell to the mill whose command areas their land is located in, it is possible that this law is flouted in practice. Some flexibility in this law may also be possible in case of cane shortages or overages on different sides of the border. I first check that a discontinuity does indeed exist at the border; that is, farmers on one side of the border sell to the mill on their own side and not the other side. Moreover, I also check that other variables do not display a discontinuity at the border.

Data for these checks come from a small survey of 80 households implemented prior to the main survey. Sugarcane growers at various distances from the border (at a set of different borders) were asked about which mills they had sold sugarcane to in the last five years, their yields, and their land ownership and rental details. Not a single respondent claimed to sell sugarcane *regularly* to

[14] For more on remote sensing of vegetation, see Jensen (2007).

the mill on the other side of the border. There are, however, farmers who have sold cane to the mill on the other side of the border occasionally over the last five years.

Figure 1 presents these results by designating the cooperative mill as Mill A and the private mill (on the other side of the border) as Mill B. This figure is conservative and biased against demonstrating a discontinuity, since it shows the proportion of respondents with land in the command area of Mill A who have *ever* sold cane to Mill B on the left hand side of the graph, compared against those who exclusively sell to Mill B on the right hand side of the graph. Despite this bias, however, the discontinuity at the border is clear. Since no one on the side of Mill A sells exclusively to Mill B, there will clearly be an even sharper discontinuity at the border if this metric were used instead. Meanwhile, other variables such as cane yields do not show this discontinuity (Figure 2). Of course, farmer yields are outcomes rather than underlying characteristics of the areas. The following section discusses testing one important underlying characteristic: soil quality.

5.2 Soil testing

I next check that the soil quality was indeed the same across either side of the border. Table 1 presents the results, which show no significant differences between private and cooperative/public mills. The coefficient on any of the soil characteristics is smaller than 5% of the standard deviation of one of these variables. Another way to gauge the magnitude of this coefficient is to project what it means to income, by multiplying it with the coefficient on the regression of farm income on the given characteristic. For example, soil on the private side of the border has 20 kg/hectare more nitrogen. An additional kg/hectare of nitrogen is associated with a Rs. 31.8 increase in annual farm income. Thus, the difference between the private and cooperative soil samples corresponds to about Rs. 500 in annual farm income, which is only 1% of mean annual farm income.

5.3 Farmer and Mill Characteristics

Finally, I check whether there are any differences in characteristics of farmers at command area borders. Since I do not have any information on the landless, it is difficult to separate out the compositional effects of people drawn into farming. Note, however, that this is an issue of interpretation: if certain types of people are drawn into farming because certain mills reward effort, this can still be interpreted as the causal impact of ownership structure. Table 2 presents results on farmer characteristics, separated into all farmers and cane farmers.

In general, farmers appear to be similar on both sides of the border, yet there are nevertheless some important differences. The first difference is that private mills seem to be located farther from farmer's plots. This could be a result of private mills having larger command areas, or simply because mill command areas are irregularly shaped. To the extent that mills find it easier to provide services to farmers that are close by, this would bias us against finding results in favor

of private mills overall. In any case, all regressions control non-parametrically for distance to the border. Second, farmers on the private side of the border have more land than those on the cooperative side. This may very well be a long-term consequence of working with private mills. Given this difference, however, I also present a specification that controls for acreage as well as an interaction of acreage with the indicator for private, in order to check for differential effects on land-poor and land-rich farmers. Note that land-poor farmers across the borders may well have different observable and unobservable characteristics, precisely due to selection or longer term consequences of ownership, and hence these comparisons must be viewed with caution.

Finally, it is also possible that conditions away from the border affect mill operations at the border. Note that it is important to control for only conditions that are *not* a result of mill performance, that is to not control for mill outcomes such as utilization or recovery rates of sugar from cane. I present specifications controlling for plausibly exogenous conditions: the amount of cultivable area available to the mill within its command area, mill crushing capacity,[15] the age category of the mill,[16] the elevation and ruggedness of the terrain in the command area, and finally, the area of the command area under various soil types (alluvium residual, marine alluvial, granite-red, aeolian-alluvial, laterite, alluvial, and granite-mixed).

6 Results on Farmer Outcomes

The perception amongst governments that fund cooperatives is that incentives for private mills to hold up farmers are higher and will hence result in an undersupply of cane to these mills. However, I find exactly the opposite result: private mills seem to encourage production of cane. Both satellite and survey data are consistent in this regard.

The satellite analysis suggests that villages on the private side of the border have a higher proportion of cane planted on all vegetated land, by about 2.3 percentage points, or 3.8% (Table 3). Moreover, using the same technique as that used for identifying sugarcane, I distinguish land planted with any crops by simply observing NDVI value ranges of all observed crops by image; it appears as though more of the land on the private side of the border is planted with any crops (1.6 percentage points, or about 2%). Finally, farmers on the private side of the border also plant more cane as a proportion of all planted crops (outcome in column 1 divided by outcome in column 2), by 1.6 percentage points, or about 2.4%. While these magnitudes might appear small, it is important to note that the two areas are practically identical in underlying characteristics.

The satellite data analysis is corroborated by the survey, with similar observed magnitudes. While these results are not robust to the addition of mill controls, the survey data suggests that

[15]Note that capacity could also be considered a mill outcome; removing it does not affect results qualitatively, largely because it is highly correlated with cultivable area.

[16]I was unable to determine exact age of entry for all mills, given confusion over when plant construction started, when it finished, and when mill operations started. However, it is clear that mill entry happened in clumps, and hence I designate mills by category as more than fifty years old, between 2-49, and less than 2 years old.

farmers on the private side of the border plant about 0.3 additional acres of sugarcane, corresponding to 5% more sugarcane on owned or rented land (Table 4). This number compares reasonably to the 3.8% additional sugarcane on all vegetated land observed via satellite, particularly since the survey excluded areas where no sugarcane was grown on either side of the border.

The extensive margin results mirror these intensive margin results. Farmers are 6-28 percentage points more likely to have cultivated sugarcane in the past five years on the private side of the border, and 5-24 percentage points more likely to be growing sugarcane at the time of the survey. These results suggest that private mills are more efficient at using a given unit of command area assigned to them, and more successful at convincing farmers to grow and supply them with sugarcane.

The coefficients are somewhat sensitive to the addition of mill-level controls, although the sign and significance level does not change. In order to assess robustness, I provide results with the various mill controls – age, elevation, and soil proportion – added in separately. The Appendix tables (A.3-A.7) show these results, in addition to results with more flexible controls for mill capacity. As is clear, the results are not qualitatively different.

Finally, I consider the effects on overall welfare of farmers. Sugarcane is an extremely lucrative cash crop. Farmers may choose not to plant it if they have no source of irrigation, or if they are liquidity constrained and cannot afford the upfront costs of seed and fertilizer, or if they fear that sugar mills may not purchase their cane or hold them up ex post. Therefore, if poorer farmers are indeed able to plant cane, this could have significant effects on their overall income and consumption.

The fact that poorer farmers grow sugarcane on the private side does appear to have some effect on their finances (Table 5). Farm income is significantly higher, by about 13%. The addition of mill controls makes the coefficients on overall income and consumption positive and strongly significant. Moreover, land-poor farmers on the private side of the border seem to be better off in terms of all three outcomes than their counterparts on the cooperative side. This set of results is particularly important, since proponents of the cooperative model highlight its benefits for poorer farmers. Again, while the coefficients are somewhat sensitive to the addition of mill controls, the broad story is consistent, as seen in the Appendix tables.

7 Mechanisms and Discussion

What do private mills do differently that encourages farmers to grow sugarcane? Sugarcane is a lumpy crop, and farmers often require credit in order to pay for seeds and fertilizer. Overall, the amount of credit assistance provided by both types of mills is very similar: however, controlling for acreage, it appears as though land-poor farmers receive more loans from private mills (Table 6). The story is similar for the cane price: while overall there are few differences, once I control for

acreage, land-poor farmers seem to get slightly higher prices from private mills. This is consistent with stories in which cooperative mills are captured by richer farmers.[17] However, given the fact that selection into who is a sugarcane farmer cannot be ruled out, precise mechanisms for these results cannot be defined.

While precise mechanisms may be difficult to pin down, it is clear that farmers seem to fare better under private mills rather than cooperatives, and the simple characterization that cooperative mills maximize farmer welfare may not be true. When we refer back to the ways in which cooperatives and private mills may operate differently in agricultural markets, we see that there may be at least three non-mutually exclusive explanations for these results. Note that a lack of data on interactions between mills and farmers (asked on the survey but a large number of missing observations), as well as data on mill finances (not collected; and in any case, empirical identification strategy would not work), limits the ability to directly test for mechanisms.

First, given repeated interactions between farmers and mills, holding-up farmers as in a one-shot game may not be profit maximizing in the long run. Hence, it is not clear that the long-run interests of mills, which include ensuring a steady supply of sugarcane for their factories (which require large upfront investments), are incompatible with those of farmers. In this context, the results that suggest more sugarcane grown and higher prices paid (at least to land-poor farmers) are consistent with the models of monopsony purchase in which more efficient private mills both purchase more inputs and pay higher prices.

This explanation does not fully account for the fact that land-poor farmers do better under private mills; in other words, why do cooperative mills discriminate against the worst off? To explain this, one must turn to management of mills: in particular, the suggestion that the objective function of cooperative management might not be the same as that of the worst-off farmers. Richer farmers, who control the cooperative, may be more interested in using the mill for political gain (Sukhtankar, 2012), or in distributing profits not as higher prices for cane, but rather on "public goods," such as temples and engineering colleges, from which they benefit (Banerjee et al., 2001).[18]

A third potential explanation involves differences in resources available to mills, such as access to technology and capital. For example, the cost of capital may be lower for private mills, which helps them provide loans and other inputs such as extension services to farmers. This is a possibility that cannot be ruled out with the available data. However, note that cooperative mills also have access to very large sources of subsidized credit through both government banks and "priority sector" lending by private banks (Cole, 2009). On the other hand, it is easier to rule out technology as the driver of the results here. The rationale for ruling out technology relies on the production technology for sugar. The actual production of sugar from cane is a simple process: large machines

[17]Note that I did attempt to collect data on other aspects of mill performance which might encourage sugarcane production – such as paying on time and delays from optimal harvesting dates for farmers – but a large number of missing observations (over 50%) precludes analysis of these data.

[18]Note that both the Banerjee et al. (2001) and Sukhtankar (2012) evidence comes from a different state (Maharashtra), but anecdotal evidence suggests that the rural political economy is similar in Tamil Nadu.

crush cane to extract juice, and this juice is boiled to extract sugar. These machines can be purchased on the open market easily, and do not rely on specialized research and development done by individual mills. Hence, lack of access to technology cannot explain the results. Next, the impacts we see are on *inputs*, not *outputs*; if the results involved quality of sugar produced, for example, the technology of the machinery might be a plausible explanation. Here, it is difficult to imagine that *sugarcane production* is affected by mill machinery. To encourage farmers to grow sugarcane, the mill (1) provides outreach through mill officers, which it hires, trains, and incentivizes; and (2) ensures smooth processing of cane through managing cutting dates, paying on time, and ensuring no factory stoppages. These are tasks more likely to be affected by management rather than technology.

8 Conclusion

Does organizational form matter for firm behavior? I examine this question in the context of sugar mills in India, relying on government policies to overcome endogeneity and variation challenges. The uniqueness and simplicity of the context – where we see dissimilarly governed firms performing the same economically significant, yet simple, activity in the same place at the same time – allows us the opportunity to answer this question.

I find evidence consistent with the importance of ownership structure for economic outcomes in this context: private mills encourage sugarcane production. Farmers are more likely to have cultivated sugarcane on the private side of the border. Income and consumption is also higher for farmers living on the private side of the border, and particularly so for land-poor farmers. Meanwhile, soil and other local conditions are equivalent, and the results are robust to controlling for mill characteristics outside the border as well as farmer acreage.

The reasons for higher sugarcane production appear to be loans and prices. Sugarcane has a yearly harvest, hence the income stream of its farmers is lumpy, and providing loans can ameliorate cash flow constraints and encourage productive activities. However, private mills seem to be just as good at making these loans as cooperate and public mills, while offering higher prices and loans to land-poor farmers. One caveat to these results is that inframarginal characteristics of command areas, combined with credit constraints, may influence mill behavior at the border; to the extent these conditions cannot be controlled for adequately, the results must be viewed with caution.

Why are private mills more adept at providing credit and higher prices to farmers? Data availability constraints preclude answering this question precisely. Potential explanations include the fact that long-term relationships between mills and farmers may help solve hold-up issues. Evidence from other studies suggests managerial quality matters and may be constrained in developing countries (Bloom et al., 2012). Managerial quality may be worse in cooperative and public mills because of elite capture, and because mistakes by management go unpunished. Recent history

suggests some recognition that cooperatives are mismanaged: financial support for cooperatives is being withdrawn even in the state of Maharashtra, previously synonymous with cooperative sugar mills (Damodaran, 2014). In Tamil Nadu, no new cooperative mills have opened since 1997, while at least six new private mills have opened in that period. Meanwhile, pure access to technology can be ruled out as an explanation, particularly given the simple technology of sugar production.

The lessons from this study are applicable to various other realms where governments feel forced to intervene in agricultural markets in developing countries due to the threat of market failure. These interventions are costly, and the benefits of the intervention may be captured by special interests. Concrete empirical evidence on the productivity or equity gains of these interventions is essential before they are funded.

References

Banerjee, Abhijit, Dilip Mookherjee, Kaivan Munshi, and Debraj Ray, "Inequality, Control Rights, and Rent Seeking: Sugar Cooperatives in Maharashtra," *Journal of Political Economy*, 2001, *109* (1), 138–90.

Bardhan, Pranab K., "Labor-tying in a Poor Agrarian Economy: A Theoretical and Empirical Analysis," *The Quarterly Journal of Economics*, 1983, *98* (3), 501–514.

Baru, Sanjaya, *The Political Economy of Indian Sugar: State Intervention and Structural Change*, Delhi: Oxford University Press, 1990.

Black, Sandra, "Do better schools matter? Parental Valuation of Elementary Education," *Quarterly Journal of Economics*, 1999, *114* (2), 577–599.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts, "Does Management Matter? Evidence from India," *The Quarterly Journal of Economics*, 2012.

Blouin, Arthur and Rocco Macchiavello, "Tropical Lending: International Prices, Strategic Default and Credit Constraints among Coffee Washing Stations," Working Paper, Warwick University 2013.

Boycko, M., A. Shleifer, and R.W. Vishny, "A Theory of Privatization," *Economic Journal*, 1996, *106*, 309–319.

Casaburi, Lorenzo and Rocco Macchiavello, "Loyalty, Exit and Enforcement: Evidence from a Kenya Dairy Cooperative," Working Paper, Warwick University 2014.

Cole, Shawn, "Fixing Market Failures or Fixing Elections: Agricultural Credit in India," *American Economic Journal: Applied Economics*, 2009, *1* (1), 219–250.

Damodaran, Harish, "Another Maharashtra verdict: Sugar goes private," *The Indian Express*, October 2014.

Dragusanu, Raluca, Daniele Giovannucci, and Nathan Nunn, "The Economics of Fair Trade," *The Journal of Economic Perspectives*, 2014, *28* (3), 217–236.

Foster, Andrew and Mark Rosenzweig, "Economic Growth and the Rise of Forests," *Quarterly Journal of Economics*, 2003, *118*, 601–637.

Hart, Oliver, "Incomplete Contracts and Public Ownership: Remarks, and an Application to Public-Private Partnership," *Economic Journal*, 2003, *113*, C69–C76.

_ **and John Moore**, "Cooperatives vs. Outside Ownership," NBER Working Paper 6421, National Bureau of Economic Research, Inc 1998.

_ **, Andrei Shleifer, and Robert W. Vishny**, "The Proper Scope of Government: Theory and an Application to Prisons," *Quarterly Journal of Economics*, 1997, *112*, 1127–61.

Jensen, John, *Remote Sensing of the Environment: An Earth Resource Perspective*, Upper Saddle River, NJ: Pearson Prentice Hall, 2007.

Kremer, Michael, "Why are Worker Cooperatives so Rare?," NBER Working Paper 6118, National Bureau of Economic Research, Inc 1997.

Laffont, Jean-Jacques and Jean Tirole, "Privatization and Incentives," *Journal of Law, Economics, and Organization*, 1991, *7*.

Lee, David and Thomas Lemieux, "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, June 2010, *48* (2), 281–355.

Macchiavello, Rocco and Ameet Morjaria, "Competition and Relational Contracts: Evidence from Rwanda's Coffee Mills," Working Paper, Warwick University 2014.

Masten, Scott E, "Modern evidence on the firm," *American Economic Review Papers and Proceedings*, 2002, *92* (2), 428–432.

Megginson, William L. and Jeffry M. Netter, "From State to Market: A Survey of Empirical Studies on Privatization," *Journal of Economic Literature*, 2001, *39*, 321–389.

Mukherjee, Anindita and Debraj Ray, "Labor tying," *Journal of Development Economics*, 1995, *47* (2), 207 – 239.

Rehman, Md. Rejaur, A.H.M. Hedayatul Islam, and Md. Ataur Rehman, "NDVI Derived Sugarcane Area Identification and Crop Condition Assessment," *Planplus*, 2004, *2*.

Rey, Patrick and Jean Tirole, "Financing and access in cooperatives," *International Journal of Industrial Organization*, 2007, *25* (5), 1061 – 1088.

Shleifer, Andrei, "State versus Private Ownership," *Journal of Economics Perspectives*, 1998, *12*, 133–50.

Stiglitz, Joseph, *Whither Socialism*, Cambridge: MIT Press, 1994.

Sukhtankar, Sandip, "Sweetening the Deal? Political Connections and Sugar Mills in India," *American Economic Journal: Applied Economics*, July 2012, *4* (3), 43–63.

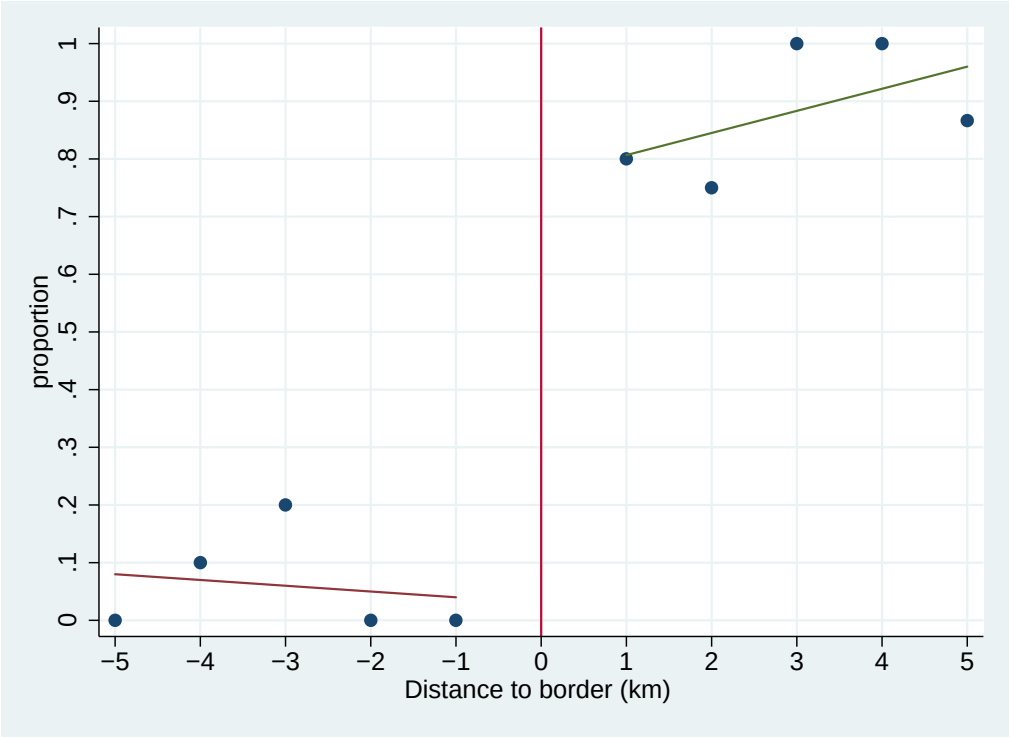

Figure 1: Proportion of Farmers Selling to Mill B

This figure shows how the proportion of farmers selling to a given private mill (Mill B) changes as we get closer to and then cross the border of that mill's command area with a cooperative mill (Mill A). Note that the question in the survey asks whether farmers have "ever sold" cane to Mill B, so is more conservative than a figure that would show the mill farmers are currently selling to, which is in almost all cases the mill whose command area the farmer's land lies in.

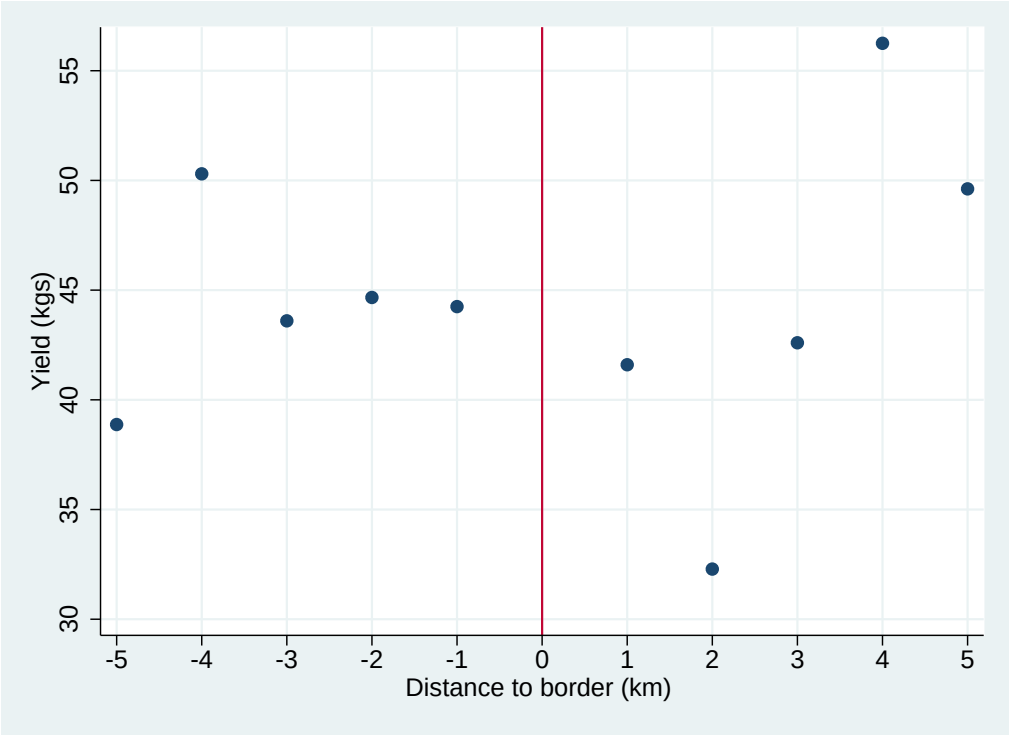

Figure 2: Yield

This figure shows sugarcane yield at various distances as we get closer to and cross borders between cooperative (left) and private (right) mills.

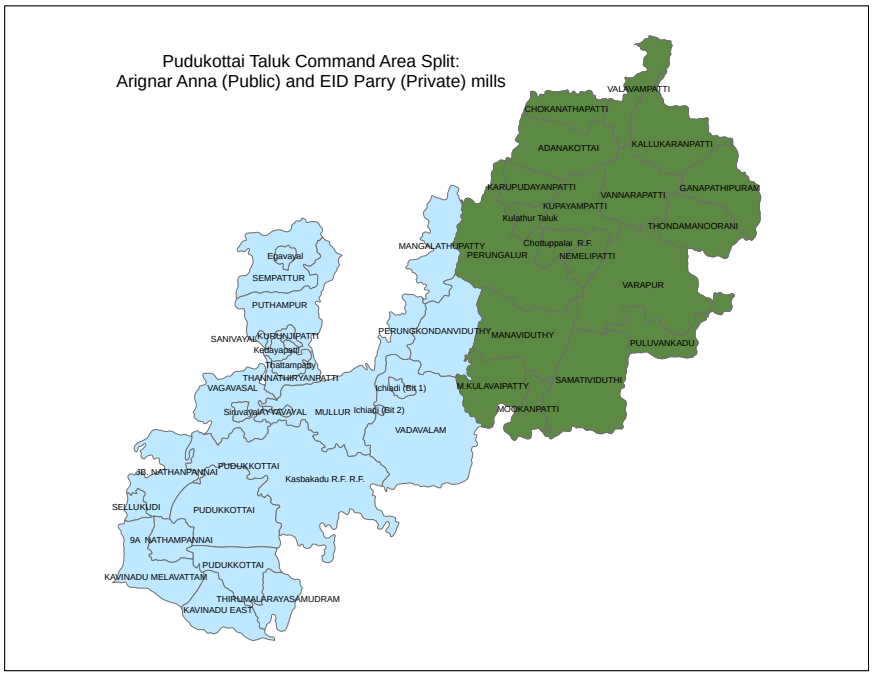

Figure 3: Sample Border Area within Taluk (sub-district)

This figure shows a sample border area in a sub-district divided between Arignar Anna, a public mill (dark green or dark gray in grayscale), and EID Parry Pudukottai, a private mill (light blue or light gray in grayscale). This split roughly halfway through the subdistrict is illustrative of borders in the sample of borders used in the analysis, which lie entire with sub-districts.

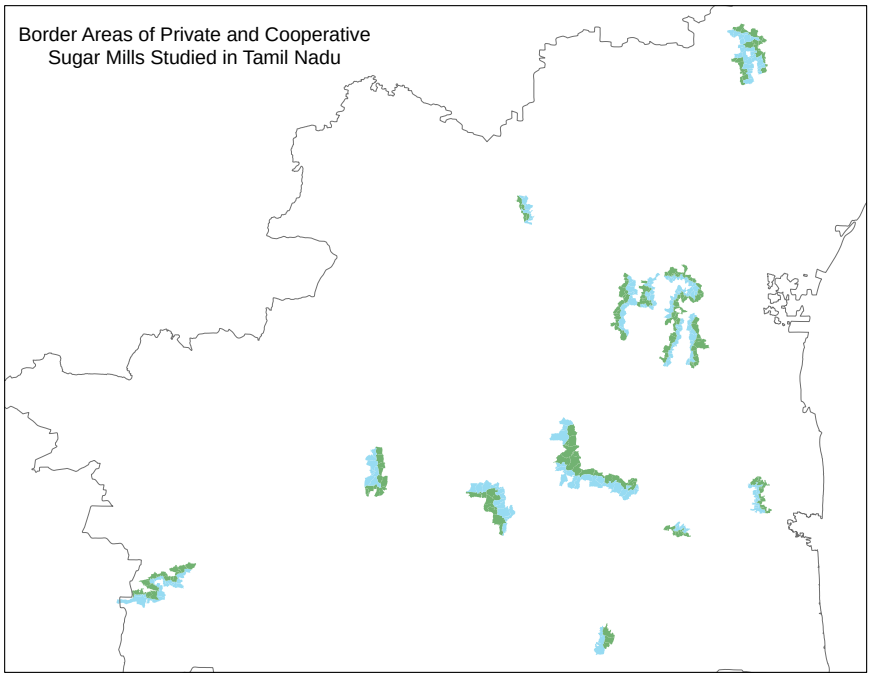

Figure 4: Border Areas in Tamil Nadu

This figure shows the entire set of 14 border areas used in the analysis of this paper. The sample is spread across the state of Tamil Nadu, and is representative of all major sugarcane growing regions in the state. Areas in dark green (dark gray in grayscale printing) belong to cooperative/government mills, while those in light blue (light gray in grayscale printing) belong to private mills.

Table 1: Soil Quality

	(1) Texture	(2) Conductivity	(3) Nitrogen	(4) Phosphorus	(5) Potassium	(6) ph
Private	.154	-.00397	20.1	8.54	17.4	-.00285
	(.183)	(.0258)	(30.1)	(8.5)	(28.9)	(.091)
Observations	136	136	136	136	136	136
R-squared	.493	.465	.134	.7	.256	.476
Outcome Mean	2.48	.168	272	53.6	213	7.75

This table presents regressions of various indicators of soil quality on an indicator for being on the private side of the border ("Private"). "Texture" refers to the size of the grain of soil. "Conductivity" is the electrical conductivity measured in deci-Siemens/meter; range is .01-1.39. "Nitrogen" is the kg/hectare content of nitrogen; range is 70-1989. "Phosphorus" is the kg/hectare content of phosphorus; range is 8-455. "Potassium" is the kg/hectare content of potassium; range is 35-1456. "Ph" measures acidity/alkalinity; it ranges from 1-14. All regressions include indicators for village pairs. Standard errors clustered at the mill-border level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2: Farmer Characteristics

	Acreage		Irrigation		Distance to Mill		Literacy		Land Value	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Private	2.29***	7.16**	-.0456	.00926	5.5**	10.6**	-.0173	.038	1,020	49,137
	(.803)	(2.64)	(.069)	(.0425)	(2.54)	(4.29)	(.0311)	(.0747)	(16,642)	(78,925)
Observations	1,091	407	1,091	407	1,091	407	1,081	405	1,055	141
R-squared	.0805	.0989	.667	.587	.757	.58	.0933	.117	.26	.262
Outcome Mean	6.22	7.87	.348	.309	43.7	30.5	.7	.676	78,530	494,233
Farmer Sample	All	Cane	All	Cane	All	Cane	All	Cane		
Include Zeros									Yes	No

This table presents regressions of farmer characteristics on an indicator for being on the private side of the border ("Private"). For columns 1-8, odd-numbered columns include the full sample of households, while odd-numbered columns only include farmers who grew sugarcane in the last five years. "Acreage" is the sum of the amount of land owned, rented or sharecropped by the household. "Irrigation" refers to whether the sampled plot was on irrigated land. "Distance to mill" is the distance from the farmers' plots to the sugar mill. "Literacy" refers to whether the respondent can read. "Land value" is the average value per acre of owned land; column 9 shows results where values were recorded as 0, while column 10 excludes these values. These results are only shown for the full sample of farmers, since there are only 42 non-missing and non-zero observations for sugarcane farmers. All regressions include indicators for village pairs. Standard errors clustered at the mill-border level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Crops Planted as Viewed from Satellites

	Cane		Planted		Cane Proportion	
	(1)	(2)	(3)	(4)	(5)	(6)
Private	.0226*	.0766***	.016*	.00697***	.016*	.0731***
	(.0128)	(3.54e-14)	(.00893)	(2.24e-14)	(.00845)	(1.06e-13)
Observations	306	306	306	306	304	304
R-squared	.907	.913	.72	.729	.927	.931
Outcome Mean	.6	.6	.839	.839	.684	.684
Mill Controls	No	Yes	No	Yes	No	Yes

The dependent variable in columns 1-2 is the proportion of land which corresponds to NDVI values representative of sugarcane. In columns 3-4, it is the proportion of land which corresponds to NDVI values of all observed crops. In column 5-6, it is the amount of land corresponding to NDVI values of sugarcane divided by the amount of land which corresponds to NDVI values of all observed crops. The main independent variable is an indicator for being on the private side of the border ("Private"). All regressions include indicators for mill border pairs, as well as indicators for each different satellite image. Mill controls include indicators for age, log crushing capacity, log cultivable area, mean and standard deviation of elevation in the operational area, and proportion area under various types of soil. Standard errors clustered at the mill-border level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Sugarcane Production

	Grew Cane Recently			Grows Cane Now			Cane Acreage	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Private	.0643***	.275***	.278***	.0495***	.227***	.239***	.363	.0352
	(.0216)	(.0359)	(.0463)	(.0169)	(.0486)	(.0518)	(.267)	(.36)
Acreage			.0047			.00355		
			(.00285)			(.00282)		
Private*Acreage			-.00114			-.00129		
			(.00392)			(.00326)		
Observations	1,078	1,078	1,078	1,057	1,057	1,057	1,089	1,089
R-squared	.408	.422	.436	.413	.424	.429	.0674	.0708
Outcome Mean	.253	.253	.253	.229	.229	.229	.502	.502
Mill Controls	No	Yes	Yes	No	Yes	Yes	No	Yes
Acreage	No	No	Yes	No	No	Yes	No	No

Columns 1-3 are linear probability estimations of whether respondent has ever grown sugarcane in the last 5 years, columns 4-6 are linear probability estimations of whether the respondent is currently growing sugarcane, and columns 7-8 are estimations of the amount of land devoted to sugarcane, all on an indicator for being on the private side of the border ("Private"). "Acreage" refers to the number of acres owned or rented. All regressions include indicators for village pairs, as well as non-parametric controls for distance from the mill. Mill controls include indicators for age, log crushing capacity, log cultivable area, mean and standard deviation of elevation in the operational area, and proportion area under various types of soil. Standard errors clustered at the mill-border level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Farmer Welfare

(a) All Farmers

	Ln Income			Ln Farm Income			Ln Consumption		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Private	.0111	.395***	.516***	.131**	.706***	.86***	.0192	.276***	.331***
	(.0576)	(.0935)	(.107)	(.0573)	(.111)	(.165)	(.0372)	(.0699)	(.0841)
Acreage			.0118			.0207			.00901
			(.011)			(.0153)			(.00646)
Private*Acreage			-.0188			-.0188			-.0103
			(.0111)			(.0163)			(.00651)
Observations	1,074	1,074	1,074	986	986	986	1,066	1,066	1,066
R-squared	.153	.163	.195	.186	.198	.237	.136	.144	.17
Outcome Mean	12.3	12.3	12.3	11	11	11	10.9	10.9	10.9
Mill Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Acreage	No	No	Yes	No	No	Yes	No	No	Yes

(b) Sugarcane Farmers

	Ln Income			Ln Farm Income			Ln Consumption		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Private	.276**	.301	.411**	.166	.45	.566*	.337***	.306**	.386***
	(.124)	(.2)	(.197)	(.172)	(.312)	(.307)	(.0849)	(.127)	(.128)
Acreage			.00733			.00841			.00607
			(.00846)			(.00957)			(.00461)
Private*Acreage			-.015			-.0162			-.0113**
			(.00893)			(.00999)			(.00449)
Observations	404	404	404	395	395	395	400	400	400
R-squared	.224	.236	.249	.221	.245	.256	.226	.243	.265
Outcome Mean	12.3	12.3	12.3	11	11	11	10.9	10.9	10.9
Mill Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Acreage	No	No	Yes	No	No	Yes	No	No	Yes

This table presents regressions of overall farmer outcomes on an indicator for being on the private side of the border ("Private"). Panel (a) presents results for all farmers, while panel (b) presents results for sugarcane farmers only. "Ln income" is log overall income over the previous year. "Ln farm income" is log income from crop harvests over the last year. "Ln Consumption" refers to total regular consumption over the last year. All regressions include indicators for village pairs, as well as non-parametric controls for distance from the mill. Mill controls include indicators for age, log crushing capacity, log cultivable area, mean and standard deviation of elevation in the operational area, and indicators for various types of soil. Standard errors clustered at the mill-border level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6: Mill Assistance

	Mill Loans			Cane Price		
	(1)	(2)	(3)	(4)	(5)	(6)
Private	1,556	-9,976	-1,630	-115	-56.3	21.4
	(6,797)	(10,856)	(11,594)	(109)	(398)	(401)
Acreage			1,675**			4.41
			(711)			(2.69)
Private*Acreage			-1,567**			-10.1***
			(729)			(2.41)
Observations	396	396	396	158	158	158
R-squared	.214	.252	.296	.366	.377	.39
Outcome Mean	10,581	10,581	10,581	1,558	1,558	1,558
Mill Controls	No	Yes	Yes	No	Yes	Yes
Acreage	No	No	Yes	No	No	Yes
Farmer Sample	Cane	Cane	Cane	Cane	Cane	Cane

This table presents regressions of mill financial relationships with farmers on an indicator for being on the private side of the border ("Private"). "Mill loans" refer to the total amount lent by the sugar mill. "Cane price" is the price per ton of cane paid by the mill. All regressions include indicators for village pairs, as well as non-parametric controls for distance from the mill. Mill controls include indicators for age, log crushing capacity, log cultivable area, mean and standard deviation of elevation in the operational area, and proportion area under various types of soil. Standard errors clustered at the mill-border level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

A Details of Satellite Analysis Procedure

I obtained multi-spectral satellite images of the state of Tamil Nadu from the National Remote Sensing Centre (NRSC). These images were of 23.5m resolution, which corresponds to 1/8th of an acre on the ground. For comparison, the average number of acres owned or rented in the survey was about 6, and even if this land were to be divided into 5 plots, each pixel of resolution would amount to about 1/10th of these plots, allowing us to precisely identify sugarcane through the satellite images. The images were all captured by satellite IRS-P6 in October 2010. The particular month was chosen since a) all sugarcane that will be crushed in the season has been planted and is growing, but not yet harvested, by October; and b) field teams were on the ground at the time, allowing us to match crops on the ground with the satellite data.

More broadly, we can take advantage of the fact that chlorophyll in vegetation absorbs visible light – especially light in the red frequency – for photosynthesis but does not absorb near-infrared light (since the energy in near-infrared light would destroy proteins in the leaf). The near-infrared light is then reflected or transmitted, with denser canopies of vegetation reflecting more light since light that is transmitted by one leaf might be reflected by the leaf below it, and captured by satellite sensors. The sensor captures the strength of the electromagnetic radiation within each wavelength band, with values ranging from 0-255. An index called the Normalized Difference Vegetation Index (NDVI) transforms the near-infrared (NIR) and red wavelengths of the satellite images into a single dimension ranging from -1 to 1 according to the simple formula below:

$$NDVI = \frac{NIR - Red}{NIR + Red} \tag{A.1}$$

NDVI values in general above 0 represent vegetation, since vegetation has low red reflectance but high NIR reflectance. Moreover, different crops correspond to different ranges of NDVI, with denser vegetation, such as thick tree canopies, having higher values. This fact allows us to distinguish sugarcane from rice, the other main crop in sugarcane growing areas. Figure A.3 illustrates how a sugarcane plant's canopy is much denser and higher off the ground than a rice plant. In order to identify the exact NDVI thresholds for sugarcane, I follow standard procedures in remote sensing (see, for example, Rehman et al. (2004)) that involve calibrating NDVI values by individual image by referencing coordinates of sample fields. I obtained GPS coordinates of over 200 fields in Tamil Nadu in October-December 2010 – at the same time the satellite images were captured – and calibrated NDVI values for sugarcane fields by image. Each image will have slightly different NDVI ranges for a crop due to differences in atmospheric conditions that scatter and reflect light differentially over time and space. Since we are comparing very localized areas, this does not pose a problem; I only compare border pairs within images.

Below is the step-by-step description of the procedure used to determine the proportion of sugarcane grown in border areas. Python scripts available on request.

1. Convert Red and Near-Infrared Band satellite image into vegetation Index (NDVI): As noted above, the Normalized Difference Vegetation Index (NDVI) transforms the near-infrared (NIR) and red wavelengths of the satellite images into a single dimension ranging from -1 to 1 according to the simple formula given above

   I use this formula to convert the following six images from Satellite IRS-P6 taken in October 2010 (first number denotes flight path, second number denotes image row) – 100-66, 101-65,

101-66, 102-64, 102-65, 102-66 – over Tamil Nadu.

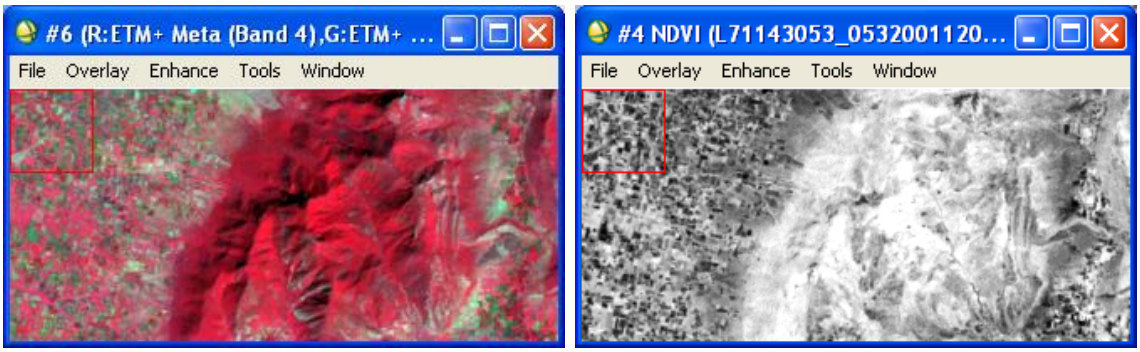

Figure A.1: Raw image converted to NDVI

2. Calibrate NDVI values of sugarcane using GPS coordinates of actual fields: I captured the GPS coordinates and noted the current growing crop of 203 fields in border areas. Overlaying these fields on the NDVI images, I determined the NDVI ranges, by image, of fields growing sugarcane. I repeated this exercise for all observed crops.

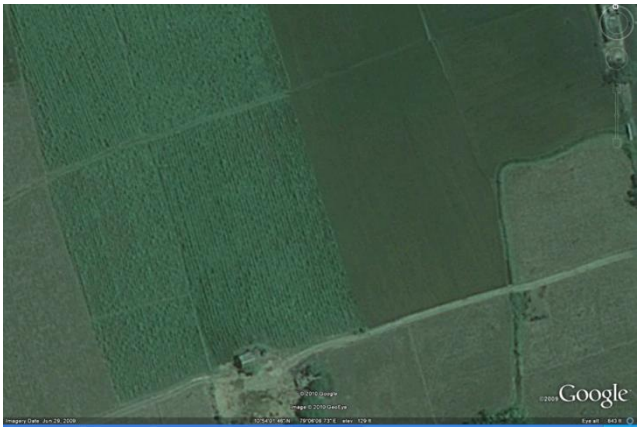

Figure A.2: Photo Showing Distinct Sugarcane Field on Right

3. Classify NDVI images into sugarcane/non-sugarcane based on these values: Using the NDVI range for sugarcane for each image observed above, I classified the images into pixels that represented sugarcane and those that did not. I also repeated this exercise for each observed crop.

4. Restrict images to positive NDVI values for vegetated areas: Restricting the coverage of images to positive NDVI values, as noted above, determines land that is covered with vegetation. In addition, this restriction also automatically excludes cloud cover and water bodies, since these have negative NDVI values.

5. Overlay border areas on classified sugarcane and vegetated area images to determine proportion cane by village: Finally, I overlay the GIS maps of border areas on the classified sugarcane

Figure A.3: Sugarcane vs Rice plants

The two photos on top show sugarcane plants in a field, while the two at the bottom show rice plants..

and vegetation images, and count the number of pixels per village that are classified as sugarcane, other planted crops, and all vegetation respective. Note that the same border area may be covered by multiple images, since there is some overlap between the vertical paths that the satellite travels on. I included all observations as long as the image covered both sides of the border area in its entirety, and included image fixed effects since the NDVI ranges will differ by image due to atmospheric variance.

Table A.1: List of Sugar Mills in Operating in Tamil Nadu (2009-11)

Mill	Type	In Sample?	New mill?	Capacity
Amaravathy (Krishnapuram)	Coop	Yes	No	1250
Ambur (Vadapudupattu)	Coop	No	No	1400
Chengalrayan (Periyasevalai)	Coop	Yes	No	3000
Cheyyar (Anakavoor)	Coop	No	No	2500
Dharmapuri	Coop	No	No	2000
Kallakurichi I (Moongilthuraipattu)	Coop	Yes	No	2500
Kallakurichi II (Kachirapalayam)	Coop	Yes	No	2500
KRR Ramasamy (Thalaignairu)	Coop	Yes	No	3500
MRK (Sethiathope)	Coop	No	No	2500
National (B. Mettupatti)	Coop	No	No	2500
Salem (Mohanur)	Coop	Yes	No	2500
Subramania Siva (Gopalapuram)	Coop	Yes	No	2500
Tirupattur (Kethandapatti)	Coop	Yes	No	1250
Tiruttani (Tiruvalangadu)	Coop	Yes	No	2500
Vellore (Ammundi)	Coop	Yes	No	2500
Arignar Anna (Kurungulam)	Public	Yes	No	2500
Perambalur (Eraiyur)	Public	Yes	No	3000
Banniyariamman (Kolunthampattu)	Private	Yes	Yes	5000
Banniyariamman (Sathiyamangalam)	Private	Yes	No	4000
Dhanalakshmi Srinivasan	Private	Yes	Yes	3500
Dharani Unit I	Private	No	No	2500
Dharani Unit II (Polur)	Private	No	No	5000
Dharani Unit III	Private	Yes	Yes	3500
EID Parry (Nellikuppam)	Private	No	No	5000
EID Parry (Pettavaithalai)	Private	No	No	2500
EID Parry (Pudukkottai)	Private	Yes	No	3500
EID Parry (Pugalur)	Private	Yes	No	4000
Kothari Sugars I	Private	No	No	2900
Kothari Sugars II	Private	No	Yes	3000
Ponni (Odapalli)	Private	Yes	No	2500
Rajshree Unit I (Varadaraj Nagar)	Private	No	No	2500
Rajshree Unit II (Mundiyampakkam)	Private	No	No	4000
Rajshree Unit III	Private	No	Yes	5000
S.V. Sugars (Palayaseevaram)	Private	Yes	No	3500
Sakthi (Sakthinagar)	Private	No	No	9000
Sakthi (Sivaganga)	Private	No	No	4000
Shree Ambika (Pennadam)	Private	No	No	7500
Thiruarooran (A.Chithoor)	Private	Yes	No	3500
Thiruarooran (Thirumandangudi)	Private	Yes	No	6000

Table A.2: Summary Statistics

	sumStats					
	Mean	SD	N	Mean	SD	N
Texture	2.36	1.31	72	2.63	1.33	64
Conductivity	.192	.234	72	.138	.136	64
Nitrogen	269	84.6	72	275	226	64
Phosphorus	49.4	84.9	72	58.7	97.7	64
Potassium	191	102	72	239	215	64
Mean elevation in 20km radius	257	198	547	154	100	544
Standard deviation of elevation in 20km radius	120	131	547	60.6	70.9	544
Area alluvium residual soil in 20km radius	213	185	547	167	169	544
Area marine alluvial soil in 20km radius	51.7	158	547	.16	.753	544
Area granite-red soil in 20km radius	608	356	547	605	450	544
Area aeolian-alluvial soil in 20km radius	66.6	107	547	18.6	56.6	544
Area laterite soil in 20km radius	43.6	179	547	38.4	180	544
Area alluvial soil in 20km radius	150	267	547	231	330	544
Area granite-mixed soil in 20km radius	118	269	547	191	413	544
Acreage	5.3	8.17	547	7.28	16	544
Land value	76,351	262,194	532	81,110	291,368	523
Irrigated	.371	.483	547	.322	.468	544
Mill distance	37.7	17.9	547	50.5	34.9	544
Literacy	.732	.443	546	.695	.461	536
Grew cane recently	.262	.44	543	.243	.43	535
Grow cane now	.228	.42	533	.23	.421	524
Cane acreage	.332	1.07	547	.7	5.88	542
Cane price	1,555	355	97	1,562	338	67
Total loans	97,703	184,929	547	86,729	144,005	544
Mill loans	11,714	44,909	540	9,278	39,985	540
Ln income	12.4	.891	540	12.2	.887	534
Ln consumption	10.9	.562	533	10.9	.551	533
Ln farm income	11	1.11	496	10.9	1.09	490

This table presents summary statistics on the main outcomes from the soil testing data as well as farmer survey, split by private and non-private mills. "Texture" refers to the size of the grain of soil. "Conductivity" is the electrical conductivity measured in deci-Siemens/meter; range is .01-1.39. "Nitrogen" is the kg/hectare content of nitrogen; range is 70-1989. "Phosphorus" is the kg/hectare content of phosphorus; range is 8-425. "Potassium" is the kg/hectare content of potassium; range is 35-1456. "Ph" measures acidity/alkalinity; it ranges from 1-14. "Acreage" is the sum of "acres owned" - the total amount of land owned by the household - and "acres rented" - total amount of land rented or sharecropped. "Land value" is the average value per acre of owned land. "Irrigated" refers to whether the sampled plot was on irrigated land. "Distance to mill" is the distance from the farmers' plots to the sugar mill. "Literacy" refers to whether the respondent can read. "Grew cane recently" is an indicator for whether respondent has ever grown sugarcane in the last 5 years. "Grow cane now" is an indicator for whether the respondent is currently growing sugarcane. "Cane acreage" is the amount of respondent's land devoted to sugarcane. "Total loans" is the household's total debt. "Mill loans" refer to the total amount lent by the sugar mill. "Cane price" is the price per ton of cane received by the respondent. "Ln income" is log overall income over the previous year. "Ln consumption" refers to total regular consumption over the last year. "Ln farm income" is log income from crop harvests over the last year.

Table A.3: Outcomes with Only Age Controls Added

	Grew cane recently	Grows cane	Cane acreage	Income	Farm income	Consumption	Mill loans	Cane price
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Private	.104***	.0723***	.542*	.0874	.162*	-.00086	-2,808	50.2
	(.0228)	(.0192)	(.291)	(.0677)	(.082)	(.0452)	(5,080)	(138)
Observations	1,078	1,057	1,089	1,074	986	1,066	396	158
R-squared	.412	.413	.0679	.154	.187	.137	.226	.371
Outcome Mean	.253	.229	.502	12.3	11	10.9	34,328	1,579
Mill Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Farmer Sample	All	All	All	All	All	All	Cane	Cane

This table presents regressions of farmer outcomes on an indicator for being on the private side of the border ("Private"). "Grew cane recently" is an indicator for whether the farmer grew sugarcane in the last 5 years. "Grows cane now" is an indicator for whether the respondent is currently growing sugarcane. "Cane acreage is the amount of land devoted to sugarcane. "Mill loans" refer to the total amount lent by the sugar mill. "Cane price" is the price per ton of cane paid by the mill. All regressions include indicators for village pairs, as well as non-parametric controls for distance from the mill. "Ln income" is log overall income over the previous year. "Ln farm income" is log income from crop harvests over the last year. "Ln Consumption" refers to total regular consumption over the last year. *Mill controls only include indicators for age.* Standard errors clustered at the mill-border level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.4: Outcomes with Only Elevation Controls Added

	Grew cane recently	Grows cane	Cane acreage	Income	Farm income	Consumption	Mill loans	Cane price
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Private	.115***	.0721***	.57*	.111*	.176**	.00628	-12,150*	42.5
	(.0223)	(.0191)	(.284)	(.0632)	(.0798)	(.0438)	(6,623)	(117)
Observations	1,078	1,057	1,089	1,074	986	1,066	396	158
R-squared	.413	.415	.0683	.156	.186	.138	.25	.377
Outcome Mean	.253	.229	.502	12.3	11	10.9	34,328	1,579
Mill Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Farmer Sample	All	All	All	All	All	All	Cane	Cane

This table presents regressions of farmer outcomes on an indicator for being on the private side of the border ("Private"). "Grew cane recently" is an indicator for whether the farmer grew sugarcane in the last 5 years. "Grows cane now" is an indicator for whether the respondent is currently growing sugarcane. "Cane acreage is the amount of land devoted to sugarcane. "Mill loans" refer to the total amount lent by the sugar mill. "Cane price" is the price per ton of cane paid by the mill. All regressions include indicators for village pairs, as well as non-parametric controls for distance from the mill. "Ln income" is log overall income over the previous year. "Ln farm income" is log income from crop harvests over the last year. "Ln Consumption" refers to total regular consumption over the last year. *Mill controls only include mean and standard deviation of elevation in the operational area.* Standard errors clustered at the mill-border level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.5: Outcomes with Only Soil Controls Added

	Grew cane recently	Grows cane	Cane acreage	Income	Farm income	Consumption	Mill loans	Cane price
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Private	.0796*	.0259	-.189*	.176*	.224*	.309***	2,814,843***	403**
	(.0453)	(.0492)	(.0986)	(.0923)	(.12)	(.0696)	(511,999)	(175)
Observations	1,078	1,057	1,089	1,074	986	1,066	396	158
R-squared	.416	.417	.0704	.159	.191	.143	.252	.377
Outcome Mean	.253	.229	.502	12.3	11	10.9	34,328	1,579
Mill Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Farmer Sample	All	All	All	All	All	All	Cane	Cane

This table presents regressions of farmer outcomes on an indicator for being on the private side of the border ("Private"). "Grew cane recently" is an indicator for whether the farmer grew sugarcane in the last 5 years. "Grows cane now" is an indicator for whether the respondent is currently growing sugarcane. "Cane acreage is the amount of land devoted to sugarcane. "Mill loans" refer to the total amount lent by the sugar mill. "Cane price" is the price per ton of cane paid by the mill. All regressions include indicators for village pairs, as well as non-parametric controls for distance from the mill. "Ln income" is log overall income over the previous year. "Ln farm income" is log income from crop harvests over the last year. "Ln Consumption" refers to total regular consumption over the last year. *Mill controls only include proportion area under various types of soil.* Standard errors clustered at the mill-border level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.6: Linear Control for Capacity

	Grew cane recently	Grows cane	Cane acreage	Income	Farm income	Consumption	Mill loans	Cane price
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Private	.297***	.249***	.0889	.43***	.745***	.277***	-6,033	698***
	(.0389)	(.0529)	(.41)	(.101)	(.121)	(.0759)	(8,804)	(213)
Observations	1,078	1,057	1,089	1,074	986	1,066	396	158
R-squared	.422	.424	.0708	.163	.198	.144	.252	.377
Outcome Mean	.253	.229	.502	12.3	11	10.9	34,328	1,579
Mill Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Farmer Sample	All	All	All	All	All	All	Cane	Cane

This table presents regressions of farmer outcomes on an indicator for being on the private side of the border ("Private"). "Grew cane recently" is an indicator for whether the farmer grew sugarcane in the last 5 years. "Grows cane now" is an indicator for whether the respondent is currently growing sugarcane. "Cane acreage is the amount of land devoted to sugarcane. "Mill loans" refer to the total amount lent by the sugar mill. "Cane price" is the price per ton of cane paid by the mill. All regressions include indicators for village pairs, as well as non-parametric controls for distance from the mill. "Ln income" is log overall income over the previous year. "Ln farm income" is log income from crop harvests over the last year. "Ln Consumption" refers to total regular consumption over the last year. Mill controls include indicators for age, *a linear control for crushing capacity*, log cultivable area, mean and standard deviation of elevation in the operational area, and proportion area under various types of soil. Standard errors clustered at the mill-border level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.7: Quadratic Controls for Capacity

	Grew cane recently	Grows cane	Cane acreage	Income	Farm income	Consumption	Mill loans	Cane price
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Private	.325***	.274***	.136	.509***	.844***	.286***	31,515*	793***
	(.0426)	(.0582)	(.447)	(.111)	(.128)	(.0809)	(15,972)	(234)
Observations	1,078	1,057	1,089	1,074	986	1,066	396	158
R-squared	.422	.424	.0708	.163	.198	.144	.252	.377
Outcome Mean	.253	.229	.502	12.3	11	10.9	34,328	1,579
Mill Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Farmer Sample	All	All	All	All	All	All	Cane	Cane

This table presents regressions of farmer outcomes on an indicator for being on the private side of the border ("Private"). "Grew cane recently" is an indicator for whether the farmer grew sugarcane in the last 5 years. "Grows cane now" is an indicator for whether the respondent is currently growing sugarcane. "Cane acreage is the amount of land devoted to sugarcane. "Mill loans" refer to the total amount lent by the sugar mill. "Cane price" is the price per ton of cane paid by the mill. All regressions include indicators for village pairs, as well as non-parametric controls for distance from the mill. "Ln income" is log overall income over the previous year. "Ln farm income" is log income from crop harvests over the last year. "Ln Consumption" refers to total regular consumption over the last year. Mill controls include indicators for age, *linear and quadratic controls for crushing capacity*, log cultivable area, mean and standard deviation of elevation in the operational area, and proportion area under various types of soil. Standard errors clustered at the mill-border level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.8: Total Loans

	Mill Loans		
	(1)	(2)	(3)
Private	-1,388	102,700***	125,322***
	(7,938)	(14,940)	(21,237)
Acreage			3,448
			(2,124)
Private*Acreage			-3,341
			(2,148)
Observations	1,091	1,091	1,091
R-squared	.0701	.0762	.13
Outcome Mean	92,617	92,617	92,617
Mill Controls	No	Yes	Yes
Acreage	No	No	Yes
Farmer Sample	All	All	All

This table presents regressions of mill financial relationships with farmers on an indicator for being on the private side of the border ("Private"). "Total loans" is total household debt. Mill controls include indicators for age, log crushing capacity, log cultivable area, mean and standard deviation of elevation in the operational area, and proportion area under various types of soil. Standard errors clustered at the mill-border level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.9: Acreage Interactions

	Grew cane recently	Grows cane	Cane acreage	Income	Farm income	Consumption	Mill loans	Cane price
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Private	.144	.0908	-1.28	.258	.501	.235**	-25,407	603
	(.0943)	(.107)	(1.36)	(.211)	(.3)	(.0901)	(16,520)	(405)
Poorest x private	.021	-.00494	.117	-.0775	-.0548	-.121	-23,147	-570**
	(.0905)	(.117)	(.168)	(.166)	(.214)	(.102)	(17,331)	(212)
Poor x private	.00188	.0878	.44	-.067	-.164	-.0629	12,037	-527**
	(.108)	(.114)	(.523)	(.179)	(.268)	(.115)	(15,098)	(214)
Middle x private	.106	.125	2.72	-.0706	-.125	-.00692	-14,910	-587**
	(.139)	(.127)	(2.58)	(.217)	(.291)	(.124)	(19,876)	(263)
Observations	1,078	1,057	1,089	1,074	986	1,066	396	158
R-squared	.482	.481	.113	.265	.406	.226	.318	.4
Outcome Mean	.253	.229	.502	12.3	11	10.9	34,328	1,579
Mill Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Acreage	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Farmer Sample								

This table presents regressions of farmer outcomes on an indicator for being on the private side of the border ("Private"), along with interactions with various classes of acreage. "Grew cane recently" is an indicator for whether the farmer grew sugarcane in the last 5 years. "Grows cane now" is an indicator for whether the respondent is currently growing sugarcane. "Cane acreage is the amount of land devoted to sugarcane. "Mill loans" refer to the total amount lent by the sugar mill. "Cane price" is the price per ton of cane paid by the mill. All regressions include indicators for village pairs, as well as non-parametric controls for distance from the mill. "Ln income" is log overall income over the previous year. "Ln farm income" is log income from crop harvests over the last year. "Ln Consumption" refers to total regular consumption over the last year. Mill controls include indicators for age, log crushing capacity, log cultivable area, mean and standard deviation of elevation in the operational area, and proportion area under various types of soil. Standard errors clustered at the mill-border level in parentheses. *** p<0.01, ** p<0.05, * p<0.1