# SKIMMING OFF THE TOP: THE UNINTENDED CONSEQUENCES OF MARKET EXPANSION IN THE INDIAN DAIRY INDUSTRY

EMILY BREZA†, ARUN G. CHANDRASEKHAR‡, ASHISH SHENOY§,
AND (PRELIMINARY AND INCOMPLETE)

ABSTRACT. The dairy sector in India is orgnaized into village cooperatives, in which many indivudals pour milk together for sale to the regional market. In the last decade the Karnataka Milk Federation, the largest organizer of cooperatives in the Indian state of Karnataka, has invested heavily in bulk milk chillers (BMCs) that drastically lower the time between production and refrigeration. These chillers, by lowering the perceived risk of spoilage, both raise the potential returns to high quality milk and increase the temptation to engage in unsavory practices such as milk dilution. We investigate the effects of village access to a BMC on the production process through a difference-in-difference approach using village-level data from the district of Kolar. We find that production quantity increases with access to a chiller but average production quality decreases, as does the likelihood of being punished for low quality. The results are consistent with a story in which villagers increase their use of dishonest practices such as dilution after being connected to a BMC because they face less risk of being punished. The effect size varies with village social characteristics, indicating that it is driven in part by a village's ability to manipulate the behavior of BMC officers in a manner not possible at central processing plants. We propose an instrumental variables strategy to supplement our initial analysis and evaluate the impact of BMC access on broader village-level economic and political outcomes. The instrument is based on the optimal placement of chilling centers, as computed by a facility location algorithm inspired by work in organizational engineering.

JEL CLASSIFICATION CODES: D23, D73, L23, Q13

NEUDC CLASSIFICATION: Land and Agriculture - Agriculture

## 1. INTRODUCTION

In India, dairy production is a key source of income for approximately 20% of rural households. Each producer operates at an extremely small scale, with the average household owning fewer than three cows. Bringing milk to market and producing value-added

*Date*: August 2012.

We thank Manaswini Rao, Vikas Dimble, Madhumitha Hebbar, and Ben Marx for excellent research support. Also, Abhijit Banerjee, Tavneet Suri, and Esther Duflo for helpful comments and Adam Sacarny for coding advice.

†Columbia Business School.

‡Microsoft Research New England and Stanford University.

§Massachusetts Institute of Technology.

dairy products presents the considerable challenge challenge of maintaining profitablility despite high fixed costs.

The dairy sector in India is generally organized into cooperatives, through which milk producers in a village join a Dairy Cooperative Society (DCS). Milk from each village DCS is collected and sent to a production facility owned by the cooperative's umbrella organization. In the Indian state of Karnataka, the Karnataka Milk Federation (KMF) comprises of over 2 million members from more than 11,000 DCSs statewide and procures over 4 million kg of milk per day[1].

Villagers in Karnataka have limited ability to produce milk for the general market until the KMF decides to charter a new DCS in their village; they are generally otherwise restricted to selling selling their product exclusively wothon the village. Once a DCS is formed, any milk in excess of local demand is integrated into the broader retail market for dairy products. Producers integrated into the state system earn a higher price per liter both due to high demand from urban centers with little local production and due to the production of value-added goods such as yoghurt and cheese. One key set of factors limiting market integration is the cost and spoilage risk in transporting unrefrigerated milk from villages with poor roads to central processing facilities, which are generally found in large towns up to 80 km away.

Over the past decade, the KMF has invested heavily in reducing transportation risks by commissioning and installing thousands of bulk milk chillers (BMCs), providing refrigeration at the point of village milk collection. One BMC allows access to refrigeration facilities to up to 4 neighboring villages. Furthermore, the KMF installs the chillers at no cost to the DCSs that benefit, providing a large free public good. We seek to estimate the effect of building a BMC on the behavior and production habits of individual villagers.

While reducing the spoilage risk borne by each village and increasing the sale price given to each farmer are unequivocally positive developments for any DCS,[2] the total value of constructing each BMC may not be so clear-cut. Because milk from several villages is combined at the point of refrigeration, the KMF is unable to provide village-level quality testing at the point of delivery, and instead must rely on the measurements taken at the site of the BMC. These measurements form the basis for each village's payment, so villages may find it difficult to self-monitor. Thus, installing new BMCs has the potential to reduce production risks and increase profitability in remote villages, but also introduces incentive problems into the dairy production process. In this paper, we seek to understand how producers respond to the installation of a BMC in or near their village. Namely, what are the magnitudes and relative importances of these opposing effects?

[1]http://www.kmfnandini.coop

[2]In general, the per liter price offered to farmers by the KMF is substantially higher than the price available on the local market.

Using administrative records for more than 1,600 DCSs in 2 districts of Karnataka, we use a differences-in-differences approach to estimate the effects of BMC infrastructure investment on milk quality, quantity, and household investment. The 11 years covered by the data span a rapid period of BMC growth, and we exploit the timing of the actual construction across the districts to identify the desired parameters. Furthermore, we decompose the effects of becoming connected along pre-period quality dimensions. Are benefits and/or costs accrued to the previously high-quality villages or to low-quality villages? In addition to the differences-in-differences model, we outline an instrumental variables approach that will be used to corroborate the reduced form results and explore further outcomes as data becomes available.

We find that while production quantity improves following the installation of a chiller, there are declines in average quality. The results show that while average quality declines, the number of days per month with low quality penalty payments decreases. While chillers may reduce uncertainty for producers, this result is consistent with the effects of reduced monitoring, leading to perverse behavior by dairy producers. We also find some evidence of disinvestment in quality on the part of milk producers. The fraction of the village's cow herd that is a modern cross-breed decreases with becoming connected.

There has been a growing literature estimating the impacts of infrastructure investment in a variety of settings, using an expanding set of empirical strategies. Duflo and Pande (2007) and Lipscomb et al. (2011) evaluate the incidence of the benefits of dams and hydroelectric power in India and Brazil. Transportation infrastructure such as rails and roads is evaluated by Banerjee et al. (2012), Donaldson (2010), Donaldson and Hornbeck (2012) and Datta (2011). Our project is also related to the empirical trade literature studying market integration such as Costinot and Donaldson (2012) and Michaels (2008).

The analysis also has connections with political economy and contract theory questions. Our paper follows the of the analysis of the PE of sugar cooperatives by Banerjee et al. (2001) and that of public goods allocation by Banerjee and Somanathan (2007). Finally, a set of papers including Banerjee et al. (2008) and Glewwe et al. (2010) describes the pitfalls of decentralizing incentives in sectors such as education and health. These studies show that individuals are quite apt at gaming incentive systems.

**Structure of the Paper.** The remainder of the paper is organized as follows. In Section 2, we describe the experimental subjects, network and survey data sources and the experimental design. Section 3 discusses the reduced form empirical approach. In section 4 we present the results. Section 5 details our IV procedure and predicted placement algorithm, while 6 concludes.

## 2. Institutional Setting and Data

### 2.1. **Setting.**

*The Karnataka Milk Federation.* The KMF has used the same model for milk procurement and governance across the state of Karnataka since its inception in the 1970's . In viable milk-producing villages, farmers are invited to join a dairy cooperative socirty (DCS). Once a new village DCS is chartered, each member becomes a shareholder in the statewide institution, earning voting rights in cooperative elections and a share of annual profits. Each DCS collects milk twice a day in both the early morning and evening. Producers bring their milk to the village office, where a nominal quality test if performed.[3]

In some villages, the milk cans are loaded onto a truck and delivered directly to one of the four district processing plants. There, full-time KMF employees test the milk's quality, measured by fat and solid non-fat (SNF) levels, and inspect for evidence of dilution or adulteration. In other villages, the milk cans are loaded into trucks and delivered to a nearby chilling facility, or bulk milk chiller (BMC). (See Figure 1 for a picture of a BMC.) Milk delivered to BMCs is tested by local DCS officers where the BMC is located. The milk is then combined with the milk from other villages and chilled. Once per day, a refrigerated tanker truck delivers milk from the chillers to one of the four production facilities. The average contents of the refrigerated truck are tested by the full-time KMF employees, but measurements cannot be traced back to individual villages. Figure 2 details this procurement process in a flow chart.

*Bulk Milk Chiller (BMC) Expansion.* Transportation costs and milk spoilage are significant barriers to expansion for the KMF. As a result, with the help of the Government of India, the organization has invested heavily in bringing refrigeration technology to remote villages. Each new BMC constructed produces 5 villages that are connected to refrigerated technology. With the chillers, milk only needs to be collected once daily, further reducing costs. In the past decade, the KMF has built more than 100 BMCs in the two districts we study. Figure 3 shows the frequency of new BMCs over time. We seek to estimate the value in terms of milk quality and farmer co-investment of a village being connected to a BMC.

It is important to note that the selection criteria for receiving a BMC are not random. Some of the determinants are minimum levels of daily milk procurement, the presence of other producing villages nearby, ownership of a structure that could accommodate a BMC, proximity to a road where tanker trucks can pass, and reliable power supply. The KMF banks the biggest gain from installing BMCs farther away from the processing plants where sspoilage risk en route to the processor is high.

*Incentives for Quality and Milk Pricing.* The processing center pays each DCS a prt liter rate based on milk quality. The procurement price is increasing in both quality dimensions, fat and SNF. If either the fat or SNF levels fall below some pre-specified threshold, then

[3]The DCS secretary measures the CLR, or corrected lacto-meter reading. This is a temperature-adjusted density measure. However, the field test is highly manipulable.

the DCS is punished with a discretely lower payment. Low payment or no payment is also given if milk is spoiled, though this occurs extremely infrequently in practice. Milk that is nearly spoiled can only be used for cheap retail products such as highly pasturized shelf-stable milk packets, and therefore lowers the state organization's annual profitability. However, individual villages are rarely penalized for such occurrences. Qualitative surveys suggest that farmers believe spoilage to pose a large threat to their income, perhaps due to past payment schemes before the availability of high-quality pasturization technology, despite the low prevalence of reported spoilage in the current data.

While village-level milk prices are increasing in quality, two problems limit the power of these incentives. First, villagers are only paid directly for quantity. They may receive year-end bonuses if the average village quality is high, but individual incentives are weak. Furthermore, the power of the quality incentives are quite low in terms of the marginal price for quality. A 2 standard deviation increase in quality is only accompanied by a small increase in price, on the order of magnitude of 2%. The incentive is much steeper when quality falls below a certain threshold, resulting in a payment decrease of 50 or even 100%. Thus, there are hgh returns to producing milk that meets a certain standard, but weak incentives to exceed the standard.

Milk quality and yield are determined by several factors including breed of cow, feed type, health and vaccination record of the animals, and water availability. Notably, producers may also choose to dilute their milk to increase their supplied quantity. Because thinning the milk decreases both fat and SNF, other adulterants may be added with the water such as milk powder, butter, salt, sugar, urea or even shampoo to avoid detection. Adulterant testing is costly, so only a small subsample of pooled DCS milk is tested, with no payment given for milk found to be adulterated. Anecdotal evidence suggests that milk dilution and adulteration is relatively commonm but rarely punished. External audits suggest that in Karnataka, approximately 20% of samples contain adulterants.[4] Determining the effects of BMC expansion on quality, quantity, and production behavior is a key goal of this paper.

Each DCS's total earnings comprise the difference between the price received from the KMF and the price paid to the farmer, plus a year-end bonus based on the KMF's annual profits. A portion of these earnings go to DCS building maintenance and staff payments at the DCS president's discretion. The remainder are returned to farmers on a per-liter basis, again independent of individual quality.

When a village is connected to a BMC, the monitoring of quality is transferred from the processing plant to the village where the BMC is located. Thus, villages connected to a BMC are paid based on measurements taken at the BMC by local officers while villages that deliver directly to processing centers are paid based on measurements taken by central

[4]Times of India, 01/12/2012

staff. This creates additional incentive problems. It is possible that village personnel help their own members and the members of contributing DCSs by inflating certain quality parameters. It is also possible that DCS officials dilute the milk to achieve the highest possible volume meeting still meeting minimum standard for normal payment. While the bulk milk chiller facilitates market expansion and consolidates transportation costs, it also decentralizes the monitoring process.

2.2. **Data.** We use four main data sources in our analysis.

*KMF Administrative Records.* The KMF has generously shared administrative records with us for the districts of Kolar and Chikballapur. These districts, formerly a single district until 2009, are both managed by the Kolar Milk Union; thus the same policies and prices apply to all villages in the sample. The administrative records detail village-level quantity and quality for each morning and evening collection from April 1, 2000 to March 31, 2011. We restrict analysis to a balanced panel of DCSs that report data in every month of the study period, consisting of 842 villages. The reported quality characteristics include fat and SNF as well as the per liter price paid to each DCSs milk on a twice-daily basis, including penalry payments for low quality or spoilage. The KMF also provided us with a list of all of the villages with BMCs as of July 2011.

*Survey of BMC Villages.* Using the list of BMCs provided by the KMF, we surveyed the DCS secretary in each of the 100 villages. From these personnel, we collected the number of members at the time of the survey, the date of commissioning of the BMC, the date of installation of the BMC, and the names of the other DCSs that contribute milk to their BMC.

*Department of Animal Husbandry Records.* To measure the composition of each DCS's herd of cattle, we obtained livestock census data from the Karnataka Department of Animal Husbandry. The organization records detailed information at the village-level on types and breeds of cows and buffaloes, along with animal husbandry participation rates by households in the village. The censuses are collected every 5 years, and we use the 2002 and 2007 data in our analysis.

*Census of India.* Finally, we use data from the 2001 Village Census of India. The key variables available from the census are GPS coordinates (used to calculate distances between villages), population, number of scheduled castes and scheduled tribes members, total land area, total cultivated land, and total irrigated land. We are waiting for the 2011 village census data to become available so that we can use other variables in our diff-in-diff analysis. We supplement the geographical data with GPS coordinates of the four district milk processing plants read from Google Maps. Census and livesrock census data are easily matched using the national census code. These are then matched to the

KMF production data by village name and location. We are able to match 761 out of 842 from the KMF records to the census.

2.3. **Descriptive Statistics.** Table 1 displays an overview of our final data set. The first column contains means of the pertinent variables for villages that receive a BMC by 2011, the second for villages eventually connected to another village's BMC, and the third for never-connected villages. Note that in this sample, 104 villages ever receive a BMC, 185 villages become connected to another village's BMC, while 553 villages remain unconnected. The average fat levels tend to hover around 4.10 with 8.45 SNF in all categories of village. The high variance in rate paid stems from rate chart adjustments over time (with the average payment increasing from Rs. 10/ltr. to Rs. 18/ltr. over this period) rather than differences between villages. There are substantial differences between the composition of villages which receive a BMC and that of the other two categories. BMC villages tend to be bigger with smaller scheduled tribes populations. This is not surprising in light of KMF's selection criteria.

## 3. Empirical Strategy: Differences-in-Differences

We first evaluate the village-level production response to chilling centers using regression analysis. All regressions are run on a balanced panel of village-month observations ranging from April, 2000, to March, 2011, with errors clustered at the village level. In all of our reduced-form analysis, we employ a differences-in-differences approach. Our key regression of interest is

$$y_{v,t} = \alpha_v + \alpha_t + \beta connected_{v,t-9} + \delta connected_{v,t+9} + \varepsilon_{v,t}$$

$v$ indexes the village or DCS and $t$ indexes the month. $connected_{v,t}$ is an indicator for whether village $v$ in month $t$ is connected to a BMC (either has a BMC in the village or delivers to a nearby village with one). We also include an extra term, $\beta connected_{v,t+9}$ to capture anticipatory effects, as each BMC is commissioned approximately 9 months before the actual equipment is installed. At this stage, village officials may yet be worried that the BMC placement decision may be altered based on village performance. However, there is not yet any difference in the technology available.

The coefficient $\delta$is the effect of having a BMC on a given outcome, $y_{v,t}$, relative to the commissioning period. The full effect of receiving a BMC compared to the period before commissioning is captured by the sum of coefficients, $\beta + \delta$. Time fixed effects, $\alpha_t$, partial out any time-generated variation, including overall trends and seasonal variation. Village fixed effects, $\alpha_t$, partial out level differences between villages. The remaining identifying assumption is that villages that become connected to a BMC would have followed the same production trend as unconnected villages were they not connected to a BMC.

3.1. **Framework.** Village incentives change after installation in multiple ways. First, BMC installation lowers the travel time between production and chilling, which significantly shortens the period in which milk may spoil. In practice, reported spoilage is extremely low, accounting for less than 0.04% of milk, and does not significantly change after the installation of a BMC. However, anecdotal evidence suggests there is a very pervasive belief that transportation time is strongly linked to dairy payments.[5]

BMC installation also incurs a large fixed cost, making it very unlikely that a BMC is subsequently removed or replaced. In addition, milk delivered to a BMC is tested by local village officers rather than central staff at a processing center. As a result, DCS officials may have more scope to manipulate the readings to avoid low payment outcomes.

3.2. **Milk Production Outcomes.** We are primarily interested in the effect of the commissioning and the installation of a BMC on milk production and the subsequent investment response of producers. Village-level outcomes include average milk fat, SNFs, monthly volume produced, days in which some milk is flagged as low quality, and total portion of milk for which villages receive low payment. The first two outcomes represent village average milk quality, and the third total quantity. The fourth comprises the portion of days in which some milk receives low payment, and the fifth the total portion of milk for which low payment is received. It is never the case that a village receives low payment for its entire milk production in a day. Low payment is only meted to those cans from which low measurements are taken; the remaining cans receive full payment. Low payment very rarely stems from spoilage; the vast majority of low payment instances are caused by low quality, as measured by fat and SNF content.

In a story of virtuous BMC effects, we might expect quality (both fat and SNF) to increase due to less uncertainty about spoilage or rejection ;eading to an increase in effort and decrease in adulteration. Quantity may increase due to more investment, but if farmers adulterate less following BMC expansion, then quantity may decrease. Similarly, low payments should decrease with the increase in quality.

Alternatively, if the monitoring effects dominate, then we should expect higher payments and higher quantities, with no observable increase in the fat or SNF levels, and potentially an overall decline in average quality. These outcomes would correspond to a situation in which villagers expend minimal effort and dilute milk down to the minimum threshold, and then monitors bump up any low readings that fall below the regular payment threshold.

3.3. **Livestock Outcomes.** We supplement our analysis of village-level outcomes using village livestock censuses. In our period of study, censuses were conducted in 2002 and 2007. Using these, we implement a simple difference-in-difference taking villages that had a BMC installed in the interim period as the treatment group and those that were

[5]It also might be the case that changes in producer incentives offset the virtuous effect on spoilage.

unconnected to a BMC in 2007 as control. If villages respond to chilling facilities by improving production, we would expect to see a shift in herd composition from indigenous cow varieties to crossbreeds, which require greater investment in feed and care but also provide higher quality milk. Inversely, if BMCs degrade the ability for the KMF to monitor quality, we would expect a decline in livestock investment.

3.4. **Heterogeneous Effects Outcomes.** We predict that effects may be heterogeneous along two margins. First, villages connected to BMCs may respond differently based on the initial quality of their production. Second, there may be a differential response to BMC connection based on the relationship between villagers and the officers responsible for measuring the quality of their milk.

Villages with low initial quality are presumably those that have the most trouble internally solving the collective action problem. As a result, we may expect that they are the most likely to continue in a dishonest equilibrium with high shirking and cheating after BMC installation. On the other hand, the presence of a chiller could increase the perceived returns to high quality milk, facilitating a shift to a more honest equilibrium where villagers put in high effort to produce high quality, undiluted milk.

Similarly, villages with high initial quality are presumably those initially in a more honest equilibrium, meaning they may be more likely to continue thusly and enjoy the higher perceived returns to high quality production. However, these villages also have the most to gain from dilution, as their high quality gives more buffer to dilute before hitting the low payment threshold. If villagers believe that diluting milk with unsanitary water raises the chance of spoilage, then the perceived low spoilage risk due to chilling may make dilution appear profitable when it previously was not.

We also test for differential responses by social distance. The temptation to dilute milk is tied to the risk of being punished for it. Sources of punishment include low payment due to low quality measurements and spoilage due to bacteria in added water. The perceived probability of being punished likely falls with BMC installation, both due to lower spoilage risk and the possibility to manipulate measurements. Quality measurements at chilling centers are taken by local village officers who are subject to more social pressure from milk producers than central staff.

The social pressure placed on local officers is a function of their social distance from the producers they evaluate. Therefore, we expect there to be greater scope for manipulation of measurement among villages that deliver to socially close BMC relative to those that deliver to a socially distant BMC. This would result in both lower average quality and fewer instances of low payment. We employ three proxies for social distance: geographic distance, scheduled caste composition, and scheduled tribe composition.

## 4. Results: Differences-in-Differences

4.1. **Graphical Analysis.** Figures 4 through 7 present our main results in graphical form. These graphs represent our differences in differences estimators. DCS and time effects are partialled out of each outcome variable, and the graphs are centered at the period at which the "treatment" villages become connected. Villages never conneced to a BMC are centered so that 0 falls at the midpoint of the period of study. The x-axis represents months since becoming connected. All outcome data is adjusted so that villages that never become connected are normalized to zero. Finally, a vertical bar is placed at 9 months prior to the installation of a BMC to indicate anticipation effects. Recall that village DCS officials may make sure not to jeopardize the arrival of the BMC before it is installed. Once the chiller is installed, it is very hard to remove.

Figure 4 shows the effects of BMC connections on days with low quality. The pattern is quite clear. Villages have far less low quality/low payment milk once they become connected. This result alone could either be driven by improved quality or lax monitoring. Figure 5 shows the same type of specification but using quantity as the outcome variable. It appears that quantity increases over time once the villages are connected. Note that there are no strong pretrends in either picture in the period before the commissioning of the BMC.

We also present our preliminary investigation into heterogeneous effects in Figures 6 and 7. In these regressions, the "treament effect" is split between villages that were in the bottom quartile of the fat distribution in the pre-period and villages in the top quartile. Figure 6 shows the heterogeneous effects of becoming connected on fat. The results are quite striking. The worst villages see moderate quality improvement after the connection, but the previous best performers experience a reduction in average fat content. In Figure 7, it also appears that previously high quality villages also increase production quantity by more after BMC introduction.

4.2. **Regression Results.** The patterns in the graphical analysis carry over to the regression results. Table 2 displays the preliminary OLS and Differences-in-Differences results. The top panel does not include any time or DCS fixed effects, thus the parameter estimates should be thought of as simple correlations. Villages that become connected tend to have higher production quantities and lower fat content. In panel B, we add time fixed effects, and the patterns look similar. In panel C, we run the full Differences-in-differences specification. We find that average fat increases during the commissioning time, but the gains are erased after BMC installation. We also find that connections to a BMC slightly decrease SNF, significantly increase quantity, and also significantly decrease the number of low payments received by the DCS. The increase in quantity is on the order of adding 2 new farmers to the DCS.

The results presented in Table 3 suggest that the change in quantity is likely not driven by the addition of new members or livestock to the DCS. If anything, the number of cows in the village declines after BMC installation. In addition, the portion of cross-breed cattle, which produce more milk but require greater investment, declines with BMC installation. These changes are consistent with the prediction that farmers take advantage of the lax monitoring by BMC staff to substitute low effor and dilution for high effort in milk production.

In Table 4 we show the same production outcomes, but the regressions are broken down by pre-period fat quartile. In terms of quality, there are large commissioning date responses by the low quartile villages. These changes persist after installation as well. In contrast, the top quality performers exhibit an economically and statistically significant decline. We also find that low quality villages are less likely to receive low payments from the KMF.

Finally, Table 5 shows additional heterogeneous results. The notable effect is that villages connected but far from the BMC exhibit the strongest declines in fat content, but still manage to reduce their reports of low payment quality levels.

Taken together, the evidence suggests that lax monitoring contributes significantly to the value and costs faced by the KMF. Villages (especially those with previously well-functioning DCSs) most likely exert influence over measurements taken by local staff. Officers at BMC centers are likely pressured to make reports such that neighboring villages rarely fall below regular payment threshold.

## 5. Empirical Strategy: Facility Placement Problem

While the differences-in-differences results are quite striking, we propose an alternate estimation procedure that generates an instrument for BMC placement. Such an approach has the benefit that we can estimate the effects of becoming connected on data without needing to rely on pre-period information. Further, while our diff-in-diff estimates of fat content and quantity effects are probably not susceptible to differences in potential trends in connected vs. unconnected villages, other outcomes of interest may be less plausibly excluded. As a result, when completed, our instrumental variables approach will serve as a robustness check and will also provide estimates on supplemental outcomes such as political affiliations.

In our IV estimation procedure, we first estimate a model of optimal BMC placement based on the desirability of each village as a potential home to a BMC. As we mentioned in Section 2, the biggest benefits are accrued from building BMCs in villages far from the processing plant and close to other DCSs. Furthermore, the fixed capacity of each chiller places an upper limit on the number of DCSs which can contribute to any given BMC. Using data from the first 2 years, we estimate the parameters of the model, which include transportation and placement costs. Once we have those parameters, we simulate the BMC placement in the following 8 years. We simulate the model and then calculate each village's propensity both of receiving a BMC and becoming connected in an optimal placement scheme. We will use these propensities as instruments for our IV regressions.

Choosing the optimal placement for the 100 BMCs out of a network of possible villages is an application of the "facility location problem." It is a well-known NP-hard problem in the computer science and operations research literature. Thus, any any algorithm must approximate the optimum. Several papers provide useful algorithms for this kind of problem such as Singhtaun and Charnsethikul (2007), Guha and Khuller (1998), and Shmoys et al. (1997). These types of algorithms have also begun being applied to development work such as Athreya and Somanathan (2008) in their examination of public goods allocation as well as Rahman et al. (1999).[6]

Our algorithm uses the following heuristic:

- Step 1: Global Stochastic Greedy Algorithm
  - Place one out of the 5 most valuable chiller first, chosen randomly
  - Recalculate the values of placing a chiller in remaining unconnected villages
  - Place the next chiller as before
  - Recalculate values and continue until all 100 chillers are placed
- Step 2: Local Optimization
  - Break the full geographic region in to polygons
  - Given the number allocated to each polygon in Step 1, re-optimize using a local greedy algorithm, but only within that area.

We plan to complete the IV estimation by November. We will then use the instrument to corroborate our difference in difference results as well as estimate the effect of access to chilling facilities on economic and political outcomes. In particular, given the political nature of cooperatives in India, we will perform IV using election data to evaluate whether chilling facilities may be used to influcene local voting behavior.

## 6. Conclusion

We document potential perverse effects associated with infrastructure investment. Connected villages experience quantity increases with access to a chiller, but average production quality decreases, as does the likelihood of being punished for low quality. From a policy perspective, it is necessary to keep in mind the potential perverse effects associated with any changes in monitoring or incentives that accompany public investment and infrastructure expansion.

[6]Other recent papers generally using spatial networks or predicted placement algorithms include Dell (2012) and Lipscomb et al. (2011).

## References

Athreya, S. and R. Somanathan (2008): "Quantifying spatial misallocation in centrally provided public goods," *Economics Letters*, 98, 201–206.

Banerjee, A., E. Duflo, and R. Glennerster (2008): "Putting a Band-Aid on a corpse: Incentives for nurses in the Indian public health care system," *Journal of the European Economic Association*, 6, 487–500.

Banerjee, A., E. Duflo, and N. Qian (2012): "On the road: Access to transportation infrastructure and economic growth in China," *NBER Working Paper*.

Banerjee, A., D. Mookherjee, K. Munshi, and D. Ray (2001): "Inequality, control rights, and rent seeking: sugar cooperatives in Maharashtra," *Journal of Political Economy*, 109, 138–190.

Banerjee, A. and R. Somanathan (2007): "The political economy of public goods: Some evidence from India," *Journal of Development Economics*, 82, 287–314.

Costinot, A. and D. Donaldson (2012): "How Large are the Gains from Economic Integration? Theory and Evidence from US Agriculture, 1840-2002," *Working Paper*.

Datta, S. (2011): "The impact of improved highways on Indian firms," *Journal of Development Economics*.

Dell, M. (2012): "Trafficking Networks and the Mexican Drug War," *Working Paper*.

Donaldson, D. (2010): "Railroads of the Raj: Estimating the impact of transportation infrastructure," Tech. rep.

Donaldson, D. and R. Hornbeck (2012): "Railroads and American Economic Growth: A Market Access Approach," *Working Paper*.

Duflo, E. and R. Pande (2007): "Dams*," *The Quarterly Journal of Economics*, 122, 601–646.

Glewwe, P., N. Ilias, and M. Kremer (2010): "Teacher Incentives," *American Economic Journal: Applied Economics*, 2, 205–227.

Guha, S. and S. Khuller (1998): "Greedy strikes back: Improved facility location algorithms," in *Proceedings of the ninth annual ACM-SIAM symposium on Discrete algorithms*, Society for Industrial and Applied Mathematics, 649–657.

Lipscomb, M., A. Mobarak, and T. Barham (2011): "Development Effects of Electrification: Evidence from the Geologic Placement of Hydropower Plants in Brazil," *Working Paper*.

Michaels, G. (2008): "The effect of trade on the demand for skill: evidence from the Interstate Highway System," *The Review of Economics and Statistics*, 90, 683–701.

Rahman, S., D. Smith, et al. (1999): "Deployment of rural health facilities in a developing country," *Journal of the Operational Research Society*, 50, 892–902.

Shmoys, D., É. Tardos, and K. Aardal (1997): "Approximation algorithms for facility location problems," in *Proceedings of the twenty-ninth annual ACM symposium on*

*Theory of computing*, ACM, 265–274.

SINGHTAUN, C. AND P. CHARNSETHIKUL (2007): "An efficient algorithm for capacitated multifacility location problems," *Journal of Computer Science*, 3, 583–591.

## Figures

Figure 1. Bulk Milk Chiller

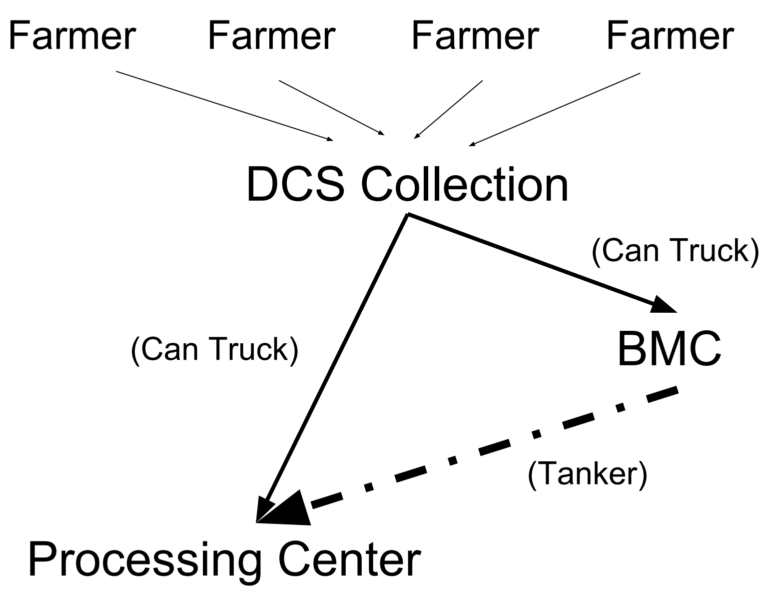


Figure 2. Dairy Collection and Transport Procedure

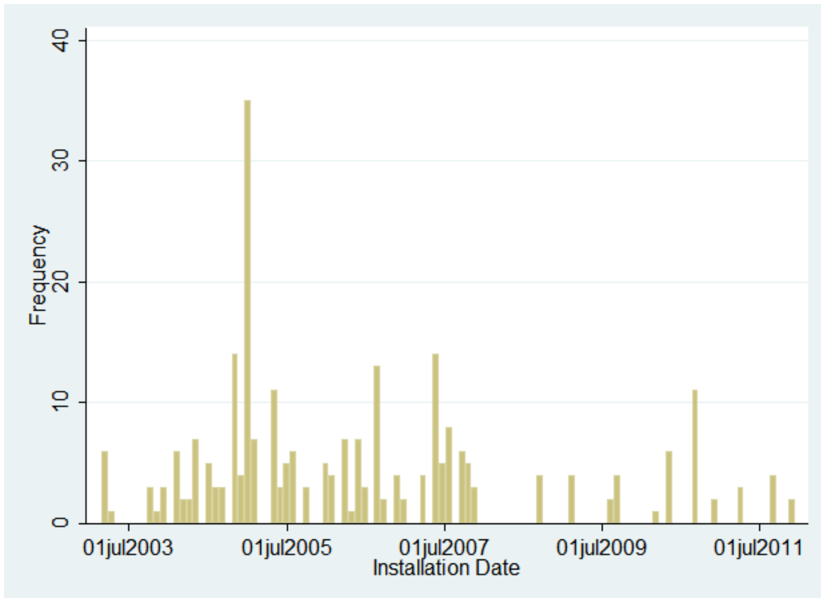


FIGURE 3. Frequency of BMC Expansion over Time

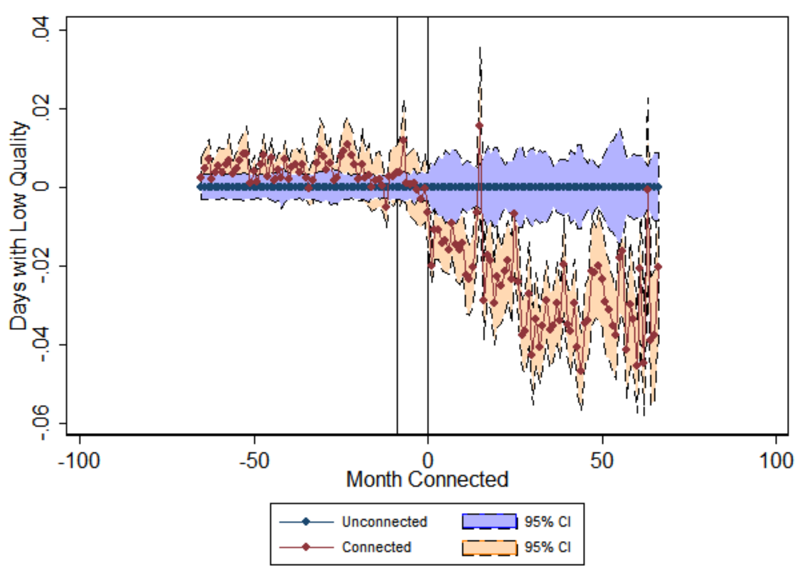


FIGURE 4. Effect of BMC Connection on Days with Low Payment

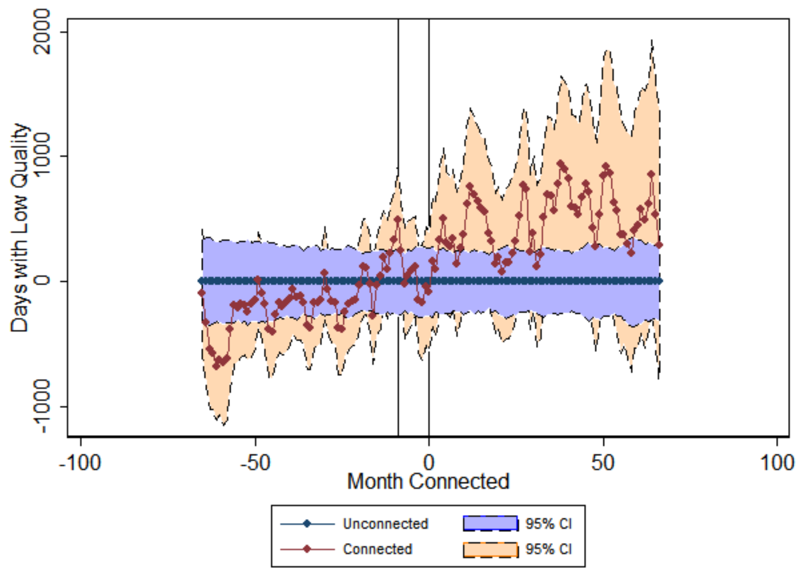


Figure 5. Effect of BMC Connection on Quantity

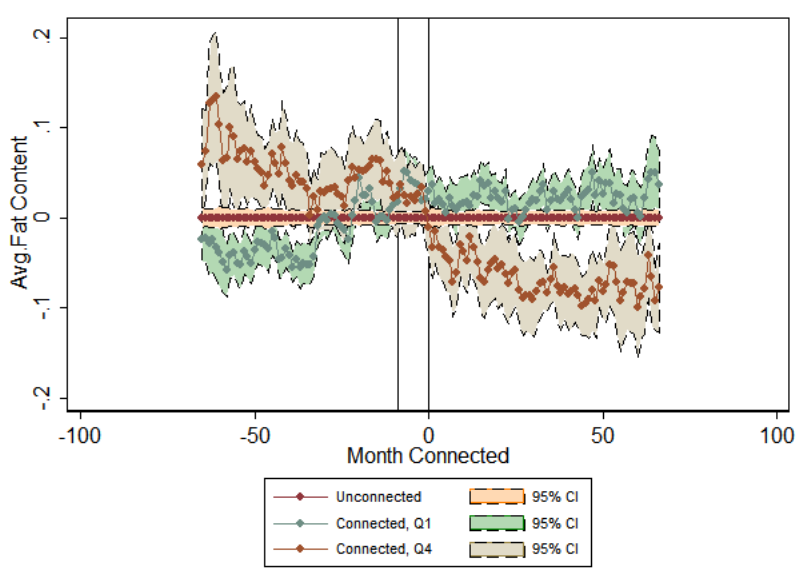


Figure 6. Effect of BMC Connection on Fat Content: Top and Bottom Quartiles

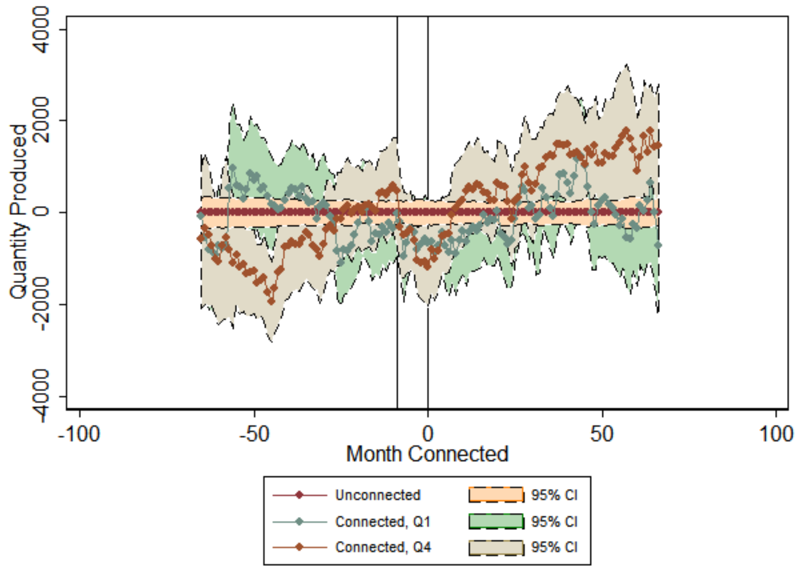


FIGURE 7. Effect of BMC Connection on Quantity: Top and Bottom Quartiles

# Tables

Table 1. Village Summary Statistics

| | With BMC | Connected to BMC | Direct Delivery |
|---|---|---|---|
| Avg. Fat | 4.105 | 4.010 | 4.123 |
| | (0.160) | (0.178) | (0.191) |
| Avg. SNF | 8.455 | 8.447 | 8.448 |
| | (0.085) | (0.092) | (0.081) |
| Avg. Monthly Quantity (1,000 ltr.) | 22.339 | 11.995 | 12.945 |
| | (11.018) | (6.127) | (7.980) |
| Avg. Rs./Ltr. | 10.354 | 10.312 | 10.313 |
| | (2.491) | (2.477) | (2.478) |
| Days w/ Spoiled Milk (%) | 0.102 | 0.071 | 0.061 |
| | (1.157) | (0.872) | (0.790) |
| Days w/ Milk of Low Quality (%) | 2.219 | 3.001 | 0.461 |
| | (7.103) | (8.513) | (10.780) |
| Portion of Milk w/ Low Payment (%) | 0.469 | 0.557 | 0.614 |
| | (1.673) | (1.750) | (1.652) |
| 2010 Status (Number) | 104 | 185 | 553 |
| Matched to 2001 Census/2002 Livestock Census | | | |
| Population | 1525.02 | 829.14 | 995.15 |
| | (1213.33) | (679.91) | (861.60) |
| Portion SC (%) | 24.35 | 24.65 | 26.22 |
| | (10.89) | (15.05) | (14.74) |
| Portion ST (%) | 6.15 | 7.89 | 9.24 |
| | (7.22) | (13.47) | (12.42) |
| Number of HH | 308.77 | 161.47 | 201.31 |
| | (255.10) | (132.95) | (186.68) |
| Portion Owning Livestock (%) | 60.94 | 66.03 | 65.59 |
| | (17.60) | (17.19) | (17.67) |
| Portion of livestock in dairy (%) | 18.82 | 17.14 | 16.86 |
| | (9.18) | (8.83) | (8.46) |
| Milking Cows | 158.82 | 94.00 | 112.23 |
| | (109.69) | (50.37) | (77.35) |
| Portion Hybrid (%) | 70.26 | 64.46 | 57.01 |
| | (20.47) | (22.28) | (21.21) |
| Cultivated Land | 229.23 | 181.16 | 179.78 |
| | (193.32) | (136.73) | (343.83) |
| Portion Irrigated | 31.58 | 32.61 | 32.94 |
| | (23.97) | (25.19) | (25.03) |
| Number of Villages | 100 | 164 | 497 |

Mean values reported with standard deviations in parentheses.

Table 2. OLS, Fixed Effects and Basic Diff-in-Diff Regressions

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| VARIABLES | Avg. Fat | Avg. SNF | Ltrs./Month | Low Quality Days | Portion Low Payment |
| Panel A: OLS | | | | | |
| commissioned | -0.0183** | 0.00113 | 2,692*** | -0.00820* | -0.000689 |
| | (0.00825) | (0.00429) | (533.5) | (0.00440) | (0.000647) |
| connected | -0.0707*** | -0.0583*** | 618.2* | 0.00296 | 0.00349*** |
| | (0.00776) | (0.00400) | (345.9) | (0.00449) | (0.000757) |
| R-squared | 0.030 | 0.060 | 0.021 | 0.000 | 0.004 |
| Panel B: Time FEs | | | | | |
| commissioned | -0.0105 | 0.00538* | 2,388*** | -0.0114*** | -0.00101* |
| | (0.00786) | (0.00307) | (539.5) | (0.00401) | (0.000585) |
| connected | -0.0353*** | -0.00816*** | 1,397*** | -0.0270*** | -0.00119 |
| | (0.00706) | (0.00295) | (369.5) | (0.00442) | (0.000726) |
| R-squared | 0.264 | 0.576 | 0.043 | 0.162 | 0.144 |
| Panel C: Village and Time FEs (Diff in Diff) | | | | | |
| commissioned | 0.0202*** | 0.00107 | 232.4 | -0.00799** | -0.000695 |
| | (0.00634) | (0.00263) | (246.3) | (0.00357) | (0.000564) |
| connected | -0.0209*** | -0.00613** | 780.0*** | -0.0291*** | -0.00131* |
| | (0.00697) | (0.00289) | (262.2) | (0.00409) | (0.000703) |
| R-squared | 0.575 | 0.726 | 0.795 | 0.292 | 0.257 |
| Observations | 111,144 | 111,144 | 111,144 | 111,144 | 111,144 |

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 3. Difference-in-Difference of Livestock Composition

| VARIABLES | (1)<br>Milking Cows | (2)<br>Portion Cross-Breed | (3)<br>Total Livestock | (4)<br>Total Poultry | (5)<br>Total Animals |
|---|---|---|---|---|---|
| Time | 10.94*** | 0.0657*** | 64.70*** | -379.7*** | -315.0*** |
| | (2.028) | (0.00622) | (12.45) | (96.55) | (97.26) |
| Treatment | 21.38*** | 0.0605*** | 67.93* | -89.59 | -21.66 |
| | (7.122) | (0.0183) | (40.20) | (169.9) | (179.7) |
| Time*Treat | -6.389 | -0.0440*** | -27.15 | 154.3 | 127.1 |
| | (5.568) | (0.0160) | (30.95) | (170.5) | (176.8) |
| Constant | 95.28*** | 0.735*** | 652.1*** | 693.5*** | 1,346*** |
| | (2.179) | (0.00777) | (14.12) | (96.40) | (98.13) |
| Observations | 2,456 | 2,451 | 2,456 | 2,456 | 2,456 |
| R-squared | 0.011 | 0.022 | 0.005 | 0.007 | 0.005 |

Treatment group was connected to a BMC between 20002 and 2007; control had no BMC access in 2007.
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 4. Heterogeneous Effects by Pre-Period Quality

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| VARIABLES | Avg. Fat | Avf.SNF | Ltrs./Month | Low Quality Days | Portion Low Payment |
| commissioned | 0.0698*** | 0.0126** | -1,079** | -0.00508 | -0.000625 |
| | (0.0106) | (0.00492) | (497.2) | (0.00757) | (0.00104) |
| com. x Q2 | -0.0518*** | -0.0208*** | 983.8 | 0.0118 | 0.00250 |
| | (0.0134) | (0.00640) | (624.0) | (0.0107) | (0.00161) |
| com. x Q3 | -0.0564*** | -0.0155** | 2,123*** | -0.0101 | -0.00117 |
| | (0.0153) | (0.00689) | (649.4) | (0.00840) | (0.00119) |
| com. x Q4 | -0.117*** | -0.0321*** | 1,229* | -0.00542 | -0.000690 |
| | (0.0179) | (0.00784) | (686.9) | (0.00964) | (0.00139) |
| connected | -0.00853 | -0.00865 | 486.4 | -0.0209** | 0.00156 |
| | (0.0124) | (0.00563) | (466.4) | (0.00942) | (0.00150) |
| con. x Q2 | 0.00640 | 0.0123* | 217.8 | -0.0126 | -0.00329* |
| | (0.0162) | (0.00709) | (639.7) | (0.0116) | (0.00195) |
| con. x Q3 | -0.0316* | -0.00377 | -62.84 | 0.000105 | -0.00235 |
| | (0.0171) | (0.00782) | (712.6) | (0.0111) | (0.00196) |
| con. x Q4 | -0.0763*** | -0.0176* | 1,061 | -0.0126 | -0.00476*** |
| | (0.0225) | (0.00942) | (678.9) | (0.0110) | (0.00180) |
| Observations | 136,356 | 136,356 | 136,356 | 136,356 | 136,356 |
| R-squared | 0.581 | 0.710 | 0.792 | 0.291 | 0.254 |

Quartiles based on avg. fat content in April, 2000 - December, 2002.
Q2 = 25th-50th percentile; Q3 = 50-75; Q4 = 75-100; Omitted: Q1(0-25)
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 5. Heterogeneous Effects by Village Ties to BMC

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| VARIABLES | Avg. Fat | Avf.SNF | Ltrs./Month | Low Quality Days | Portion Low Payment |
| commissioned | 0.0223*** | -0.00183 | 190.9 | -0.00886* | -0.000465 |
| | (0.00842) | (0.00366) | (350.6) | (0.00481) | (0.000761) |
| com. x SC dist. | -0.0120 | 0.00664 | -795.0 | 0.0559 | 0.00504 |
| | (0.0599) | (0.0225) | (1,962) | (0.0559) | (0.00691) |
| com. x ST dist. | 0.00726 | 0.000409 | 3,013 | 0.0285 | 0.00361 |
| | (0.0483) | (0.0375) | (2,048) | (0.0434) | (0.00583) |
| com. x Geog. dist. | -0.0226 | -0.0959 | -9,068 | -0.113 | -0.0217 |
| | (0.233) | (0.0900) | (6,303) | (0.0852) | (0.0133) |
| connected | -0.00824 | -0.00404 | 890.7** | -0.0199*** | -0.000160 |
| | (0.00936) | (0.00419) | (397.5) | (0.00518) | (0.000932) |
| con. x SC dist. | 0.0282 | 0.00837 | -1,079 | -0.0977 | -0.0115 |
| | (0.0617) | (0.0252) | (1,996) | (0.0658) | (0.00883) |
| con. x ST dist. | -0.0747 | -0.0241 | -918.1 | -0.0334 | -0.00549 |
| | (0.0602) | (0.0287) | (2,919) | (0.0403) | (0.00647) |
| con. x Geog. dist. | -0.821*** | -0.204** | -7,738 | 0.104 | 0.0167 |
| | (0.223) | (0.0846) | (5,627) | (0.102) | (0.0200) |
| Observations | 133,056 | 133,056 | 133,056 | 133,056 | 133,056 |
| R-squared | 0.575 | 0.707 | 0.792 | 0.291 | 0.253 |

SC dist. = absolute difference in portion Scheduled Caste between village and BMC location.
SC dist. = absolute difference in portion Scheduled Tribe between village and BMC location.
Geog. dist. = physical distance between village and BMC location.
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1