ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Jayaraman, Rajshri; Ray, Debraj; de Véricourt, Francis

Working Paper

Productivity Response to a Contract Change

CESifo Working Paper, No. 4679

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Jayaraman, Rajshri; Ray, Debraj; de Véricourt, Francis (2014) : Productivity Response to a Contract Change, CESifo Working Paper, No. 4679, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:
https://hdl.handle.net/10419/93420

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

The international platform of Ludwig-Maximilians University's Center for Economic Studies and the Ifo Institute

CESifo Working Papers

www.cesifo.org/wp

Productivity Response to a Contract Change

Rajshri Jayaraman
Debraj Ray
Francis de Véricourt

CESIFO WORKING PAPER NO. 4679
CATEGORY 13: BEHAVIOURAL ECONOMICS
MARCH 2014

An electronic version of the paper may be downloaded
- *from the SSRN website: www.SSRN.com*
- *from the RePEc website: www.RePEc.org*
- *from the CESifo website: www.CESifo-group.org/wp*

CESifo
Center for Economic Studies & Ifo Institute

CESifo *Working Paper No. 4679*

Productivity Response to a Contract Change

Abstract

This paper studies the productivity impact of a contract change for tea pluckers in an Indian plantation. The contract, implemented at the end of a three-year cycle in which contracts are generally revised, was (a) the joint outcome of negotiations between twenty unions and plantations, (b) mandated to respect a state government notification stipulating a new minimum wage for plantation workers statewide, and (c) applicable equally to all the plantations in the local region. The contract raised the baseline wage by 30% but lowered marginal incentives, by shifting the existing piece rates to higher minimum thresholds and eliminating an existing penalty per unit for low output. In the one month following the contract change, output increased by a factor between 30-60%, the exact number depending on the choice of counterfactual and the set of controls applied. This large and contrarian response to a flattening of marginal incentives is at odds with the standard model, including one that incorporates dynamic incentives, and it can only be partly accounted for by higher supervisory effort. We conclude that the increase is a "behavioral" response. Yet in subsequent months, the increase is comprehensively reversed. In fact, an entirely standard model with no behavioral or dynamic features that we estimate off the pre-change data, fits the observations four months after the contract change remarkably well. While not an unequivocal indictment of the recent emphasis on "behavioral economics," the findings suggest that non-standard responses may be ephemeral, especially in employment contexts in which the baseline relationship is delineated by financial considerations in the first place. From an empirical perspective, therefore, it is ideal to examine responses to a contract change over an substantial period of time.

JEL-Code: J430.

Keywords: incentive pay, personnel economics, productivity, contracts.

Rajshri Jayaraman
European School of Management and Technology
Berlin / Germany
jayaraman@esmt.org

Debraj Ray
New York University
New York / NY / USA
debraj.ray@nyu.edu

Francis de Véricourt
European School of Management and Technology / Berlin / Germany
francis.devericourt@esmt.org

January 2014
We are grateful to seminar and conference participants at the DIW and the Behavioral Economics Workshop in Berlin, Goethe University Frankfurt, NEUDC, CESifo Behavioral Economics Conference, University of Namur, and University of Heidelberg for useful comments and suggestions. Ray acknowledges funding from the National Science Foundation under Grant SES-1261560.

1. Introduction

This paper studies the productivity impact of a contract change. The setting is a tea plantation in India. The activity in question is tea-plucking, the output from which is measurable and contractible. Payments to pluckers consist of a baseline wage, along with piece rates per kilogram of output that vary over different production intervals.

In 2008 a contract change was instituted in the plantation, as part of a customary practice of renegotiation once every three years. This plantation represented only a fraction of the parties active in the negotiations, which involved representatives from all the unions and large plantations operating in the tea-growing region. Around 20 unions and employers participated, with no dominant players on either side of the table. Negotiations commenced a few months prior to the expiration of the going contract, which had been in place since 2005, and went through several rounds.

The replacement contract significantly increased the baseline wage, by approximately 30%. Almost all of this increase in the baseline was mandated by state government legislation. In June 2008, the government announced a preliminary notification under the Minimum Wages Act, 1948, setting the minimum wage payable to plantation workers everywhere in the state at Rs.101.52. The baseline in force in our plantation was, at the time, Rs. 77.55.[2] The new wage contract, signed and instituted on September 2, 2008 set a fixed daily wage of Rs. 103.76. It covered over 10,000 pluckers from different plantations and unions in the same tea-growing area. From the point of view of an individual worker, therefore, both the timing and the structure of the contract change can be plausibly viewed as exogenous.

An important feature of the new contract is that it also altered marginal incentives. It did away with an existing penalty per kilogram for outputs below a minimum "standard," which could be viewed as just another piece rate. It left the remaining piece rates unchanged, but they now came into play only when higher output thresholds were crossed. The contract therefore effectively flattened the piece rate structure, and with it, presumably, incentives. We will presently have more to say about the contract change, but this description is a good starting point.

Classical principal-agent theory yields an unambiguous prediction for worker productivity. In the absence of dynamic incentives, the change in the contract cannot increase worker effort: productivity must weakly decline. In Section 4.1, we confirm the classical prediction in a general model that includes both minimum standards and a multiplicity of piece rates. But classical theory does not appear to explain what happened following the contract change. Just afterward, there was a large and significant increase in worker productivity. Below, we document the increase in some detail, but here is a summary.

Average output increased by almost 60%, from 35 kg. per person per day to 55 kg. between the last week of August, just prior to the contract change, and September of 2008. As a comparison, the same calendar period for the plantation in the previous year, 2007, showed an increase from 25 to 34 kg. A "control plantation" in a neighboring region where no contract change occurred (it was within its three-year contract cycle) exhibited a contemporaneous increase from 35 to 39 kg. With rainfall and other controls accounted for, it is possible to temper the increase to some extent, but its substantial

[2] Indeed, the effective base wage was even lower if a penalty clause were to be netted out; see below.

existence is undeniable, lying somewhere between 33% and 44% — depending on the exact controls used, or relative to the trends in the two counterfactuals just mentioned.

This is interesting, because it is apparently opposed to a standard model of incentives. Three possibilities suggest themselves, and in the analysis we will attempt to distinguish between them. First, a static framework may be the wrong one: dynamic incentives based on contract non-renewal or termination may now have been heightened. Second, the increase may be driven by enhanced supervisory effort. Finally, notions of gratitude, reciprocity and gift-giving — "behavioral responses," in short — might enter in a fundamental way, as opposed to the colder calculus of moral hazard and its attendant payoff-maximizing response to a changed incentive structure.

We argue that the first two explanations are far from satisfactory. Briefly — but in more detail below — both permanent workers (who cannot be fired) and temporary workers (who are hired seasonally) exhibit the same magnitude of increase, and there is no significant difference across the two categories. We also show that heightened supervisory effort accounts for a quarter to a third of increased productivity. Hence, the presence of a large "behavioral" response appears undeniable.

That would be the end of the story, were it not for the fact that later observations in the months following the contract change tell a different tale. The output jump occurred in Month 1 (September), but in the months thereafter, a decline set in, until by the end of month four, output had returned to pre-change levels.

To assess the extent of this reversal, we estimate a simple structural model from the data *before* the contract change. In this model, each individual maximizes a "classical" payoff function defined on income and effort, purged of any behavioral effects. There are stochastic shocks, and there is a minimum standard below which an additional payoff loss is presumed to occur to the plucker, possibly non-monetary but expressed simply in monetary equivalent, and presumably stemming from direct supervision. Detailed data at the plucker level over several months permits us — under some functional form assumptions that we describe below — to obtain estimates of individual and aggregate heterogeneity, a parameter controlling the curvature of effort disutility, and the additional payoff loss arising per kilogram drop below the minimum standard. This estimation exercise may be of independent interest because the underlying plucker problem is non-convex: the fact that piece rates increase on higher production slabs implies that global no-deviation constraints need to be respected, in addition to the local first-order conditions.

We then apply the estimates to the data *after* the contract change. Not surprisingly, the predicted distribution of output is far removed from the distribution observed just after the contract change. This confirms what we've already noted: that output in the immediate aftermath of the new contract is significantly out of line from any prediction of a standard model. But in subsequent months, there is significant and steady improvement in the fit of the standard model, and the majority of the output increase (relative to the "standard prediction") vanishes by the end of Month 4. In fact, at that time, which is the last period for which we have data, the standard model works remarkably well.

Our study documents that a contract change in a field setting has significant "behavioral effects" in the short run, which, however, dissipate just as significantly as time goes by.[3] These findings echo

[3]More generally, our study of contractual change in a natural employment setting belongs to a growing empirical literature on the implications of the contract form for effort across individuals, teams or within firms in natural employment contexts; see, e,g., Shaban (1987), Lazear (2000), Ackerberg and Botticini (2002), Shearer (2004), Bellemare and Shearer

those of Gneezy and List (2006) who, in a experimental casual employment setting, find that positive productivity responses to gift exchange erode in a matter of hours. Our findings have implications that are both of intrinsic and methodological interest. We are fortunate to have access to a unique data set that records outcomes not just immediately following the contract change, but for a considerable time period afterward (a longer time period would have been still more preferable). It appears that behavioral effects stemming from reciprocity or gratitude, or perhaps other sources that we discuss below, played an important role, but those effects were short-lived, "ultimately" giving way to outcomes that are by and large consistent with the responses predicted by classical incentive theory.

Methodologically, our findings serve as a warning against the claim of dominant behavioral responses by pointing to outcomes just following a change. Once the euphoria of change dies down, such effects may vanish in a matter of weeks or months. In particular, we warn against the use of regression discontinuity-type methods that are designed to emphasize just what happens at the point of the event. To be sure, *enough* time-averaging both before and after the change will dampen or eliminate the shorter-lived effects, but then we run up against the other side of the coin, which is that behavioral responses do appear to be strong in the short run. Both sets of effects are important, and ideally, we would like to identify both. The way to do that is to track them over time.

In this sense, our paper fits into a research program outlined by Gneezy, Meier, and Rey-Biel (2011), who wish to examine "some general aspects of how extrinsic incentives may come into conflict with other motivations." They write:

> "…[Monetary] incentives might have the desired effects in the short term, but they still weaken intrinsic motivations. Thus, once the incentives are removed, people may pursue the desired outcome less eagerly. To put it in concrete terms, an incentive for a child to read more might achieve that goal in the short term, but then be counterproductive as an incentive for students to enjoy reading and seek it out over their lifetimes."

The authors clearly emphasize not just the interplay between intrinsic and extrinsic motivations, but the *intertemporal* evolution of those motivations following a policy change. They emphasize how the introduction of the "extrinsic" (money) might erode the "intrinsic" (reciprocity, gratitude or fair play). That is, in situations in which the fundamental *baseline* interaction is non-monetary — say blood donations or social work — financial incentives can cut into and displace such intrinsic motivations, sometimes (but not always) to the point where the net effect is negative; see, e.g., Gneezy and Rustichini (2000), Lacetera, Macis, and Slonim (2012), the several references to the literature in Gneezy, Meier, and Rey-Biel (2011) and Frey and Jegen (2001) for a review of empirical evidence.[4]

Our study complements these findings from a different perspective. The baseline is different: unlike social activities, paid employment represents a starting point in which *monetary* incentives are intrinsic. Gratitude, reciprocity and pro-social behavior are extrinsic in this context, because the entire interaction is mediated by money to begin with. The contract was presented and perhaps

(2009), and Fehr and Goette (2007). Some of this research, most notably, the contributions of Bandiera, Barankay, and Rasul (2005, 2007, 2009, and 2010) are based on experiments in a field setting that deliberately changed the contract form, often with implications that combine monetary and social motivations.

[4]It is fair to point out that sometimes there is no evidence that crowding-out occurs; see, e.g., Ashraf, Bandiera, and Jack (2012). Theoretical research that bears on these issues includes Bénabou and Tirole (2003) and Besley and Ghatak (2005) and is reviewed in Kőszegi (2013).

also perceived as an act of economic generosity, as conversations with the workers in this plantation appear to indicate. Indeed, this pro-social ramification of the contractual change led to a large change in output. Nevertheless, monetary incentives ultimately held sway, despite a possibly "behavioral" response in the shorter term. While reiterating the somewhat depressing lessons from the earlier literature, the more general point is that all evaluations of "intrinsic versus extrinsic" need to be based on a dynamic evaluation that examines both short- and longer-term responses.

Section 2 describes the economic setting, the contract change and the data structure. The short run response to the contract change and possible drivers for this response are analyzed in Sections 3 and 4. Section 5 studies the long run response and Section 6 presents a structural approach to understanding post-contract output. Section 7 concludes.

2. Contract Change in a Tea Plantation

2.1. **Setting.** We study a tea plantation, owned and operated by a large producer in India, in a state where tea is an important source of employment. This plantation is one of several operating in the local tea-growing region. Plantations in the local region offer identical wage contracts to all their pluckers, and are the dominant source of employment in the locality.

The plantation is made up of a number of fields on which tea bushes grow in rows. The production of green tea leaf — fresh, unprocessed leaves — is labor-intensive. Flushes, each comprising "three leaves and a bud," are plucked from tea bushes manually, either by hand or with metal shears. For the purpose of our study, this is the final output. Around 87% of the pluckers in our plantation are female. Approximately 73% of the pluckers are permanently employed and have a median tenure of approximately 21 years. They are entitled to work on all workdays. They cannot be fired. Moreover, the Plantation Labour Act of 1951 stipulates generous non-pecuniary benefits for all permanent workers, including free housing, health care and children's education. The remaining 27% of workers are hired on a temporary basis season by season. Temporary workers are not entitled to the same non-pecuniary benefits. However, *all* workers, temporary or permanent, receive identical wage contracts. According to management, absenteeism is idiosyncratic (i.e., not in the form of coordinated shutdowns or strikes) and driven largely by illness or family obligations.

Workers are meant to work full days. They are permanently assigned to labor gangs, composed of 20–40 members. Each workday, gangs are assigned to plucking or non-plucking duties, which include pruning, weeding, and the spraying of pesticides. Job duties (plucking, pruning, etc.) and field assignment are decided centrally, by managers. So is the decision of whether a given field is to be plucked using hand or shears. All of this is based on an annually predetermined schedule that is conditioned on the season as well as the field life cycle.

A gang will typically pluck one or two fields a day, returning to the same field within a week. On a given field, workers are always assigned to the same row in order to internalize dynamic considerations, since over- or under-plucking can damage the long-term yield of a plant. There are no direct complementarities in production.

Tea pluckers collect their leaves in a bag carried on their backs. These bags are weighed and the amount of green leaf plucked is recorded for each plucker on each day. Each gang has a supervisor, who is paid a fixed wage. She closely monitors this small group of workers. The supervisor records each gang member's daily activity, which captures whether a worker was present or absent for work,

and in the former case, either the non-plucking task to which the worker was assigned, or the number of kilograms of tea plucked for those workers assigned to plucking duty. There were no changes in gang or supervisor assignment following the contract change.

All told, tea plucking is a pretty routine task. Tea pluckers (or "workers") engage in this activity on a daily basis over the course of years under close monitoring, contracts are written on an objective measure of productivity, and the company has a considerable history in this industry. It is therefore reasonable to assert that the employment relationship is not characterized by asymmetric information regarding worker ability, the difficulty of the task, or the suitability of the worker. The absence of both asymmetric information and task complexity means that complications to standard contract theory settings — such as multitasking (Holstrom and Milgrom (1991)), ratchet effects (Gibbons (1987)) or career concerns (Gibbons and Murphy (1992)) — are unlikely to be central to our economic environment. Finally, permanent workers cannot be dismissed due to legal restrictions, which makes the threat of firing an ineffectual instrument in inducing extra effort (Dutta, Ray, and Sengupta (1989)). On the other hand, temporary workers can be fired, and this distinction will inform part of the analysis below.

2.2. **Contract.** Workers are paid a fixed baseline wage on days when they work. Otherwise they receive no payment. In addition, they receive affine incentives – piece-rates per kilogram of plucked green leaf. The piece rates vary across different slabs of output, and the thresholds at which the different rates kick in (though, as it happens, not the rates themselves) vary by estate-climate type (we observe 2 such types in our data) and whether the plucking is done by hand or shears. Essentially, in estates with higher overall month-end yields per hectare, and on days where workers are assigned to pluck with shears, the thresholds at which larger piece rates kick in are higher. In all cases, piece rate incomes are calculated on the basis of daily output, and overall wages are paid at month-end. In addition, the contract stipulates a "base output", which indicates a minimum standard that workers are expected to maintain.

The solid line in Figure 1 depicts a typical contract structure before September 2, 2008. The minimum standard is shown by s and the fixed daily wage by F. There are three piece rates; vertical lines denote the cutoffs for "slabs" under which different piece rates apply. It is of interest to note that the contract has no kink at s: the minimum standard is located "within" the first piece rate slab, and the piece rate to the "left" of s is framed as a monetary penalty. In total, we observe 4 such contracts in our data, with cutoffs as well as the minimum standards being higher for plucking by shears rather than by hand, and for the high relative to the low month-end yields.

This contract, in place since 2005, was renegotiated in 2008 as part of customary practice in the industry to write a new contract every three years. The negotiations involved representatives from all the unions and large plantations operating in the tea-growing region. There were approximately 20 unions and employers in these negotiations, with no dominant players on either side of the table. In keeping with industry norms, negotiations began 3–4 months prior to contract expiration and went through several rounds. Furthermore, state government legislation effectively set a lower bound for an increase in the baseline wage. In June 2008, the state government confirmed a preliminary notification under the Minimum Wages Act, 1948, setting the minimum wage payable to plantation workers in

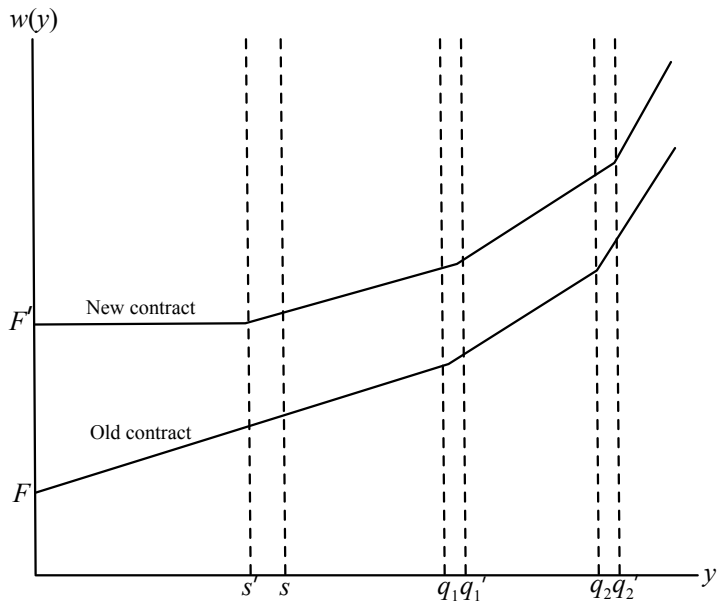


Figure 1. Wage contract. *Notes.* This figure depicts the daily wage as a function of output y for a typical contract before and after the contract change. F and F' are the fixed daily wage, paid conditional on participation; q and q' denote output thresholds corresponding to higher piece rates. Positive piece rates correspond to Rs. 0.4, Rs. 0.55 and Rs. 0.85, respectively. The wage loss below the standard s in the old contract was worded as a penalty.

the state at Rs.101.52. The daily wage being paid at the study plantation was, at the time, Rs. 77.55.[5] Writ petitions by planters seeking a stay on the minimum wage notification were dismissed by the state's high court on August 27, 2008. The new wage contract, signed by all parties and instituted on September 2, 2008, set a fixed daily wage, F', of Rs. 103.76.

Additionally, the new contract flattened the piece rate structure and hence reduced incentives at the margin. As we've already noted, the baseline wage increased by approximately 30%. The effective increase was even larger than this, because the penalty provisions for dropping below the standard were also done away with. In addition, the announced minimum standard was lowered somewhat, by 1–2 kg., and the piece rate for the first slab up to the standard was now zero, because of the absence of a pecuniary penalty. The piece rates for the remaining slabs were unchanged at Rs. 0.45, 0.55 and 0.85, respectively. However, the cutoffs shifted to the right. Writ petitions and on-going press reports clearly indicated that owners objected to the new baseline wages. In conversation, while emphasizing their commitment to the use of piece rates to elicit effort, planters explained the flatter incentive structure as an effort to offset higher labor costs imposed by the substantial baseline increase.

Over 10,000 pluckers from different plantations were covered by the wage agreement (our plantation had around 2000 pluckers). From the point of view of an individual worker, therefore, not only the timing but also the structure of the contract change can be reasonably treated as exogenous. The

[5]This does not account for the monetary penalty that would accrue if a worker came in and plucked no output, which would be between Rs 16 and 28, depending on the yield class and plucking method. To be sure, there were presumably other non-pecuniary penalties, such as supervisory pressure, which we will account for in our estimation exercise.

	Month 0	Month 1	Month 2	Month 3	Month 4
			Study plantation: 2008[a]		
No. Days	25	25	26	25	24
No. unique workers	1992	1977	1958	2047	2131
No. Obs.	30,783	34,369	31,986	32,022	31,136
			Study plantation: 2007[b]		
No. Days	24	25	25	25	25
No. unique workers	2318	2064	2114	2249	2155
No. Obs.	37,131	32,730	35,202	33,682	28,729
			Control plantation: 2008[c]		
No. Days	25	26	25	25	25
No. unique workers	1397	1499	1636	1461	1391
No. Obs.	21,951	24,765	23,018	23,630	22,636

Table 1. Data Structure. *Notes.* This table describes the number of observations for the geographic locations and time periods for which we have data. The contract change was instituted at the beginning of Month 1 in the study plantation in 2008. *a*. Study plantation in the year of the contract change, *b*. Study plantation in the year prior to the contract change, *c*. Control plantation in 2008, where no contract change occurred.

management claims there were no strikes or slowdowns before or after contract change. Nevertheless, as we will see later, the data seems to indicate abnormally low output in the first three weeks of August, 2008, directly before the contract change, and we will deal with this by considering different counterfactuals.

The new contract is shown by the dashed line in Figure 1. For each yield class and plucking method, the two contracts compare as shown in the Figure: the new contract represented an upward shift and "flattening" of payments. In what follows, we attempt to assess the implications of this contract change for labor productivity.

2.3. **Output Data.** Our data are obtained from personnel records, drawn from supervisor entries, comprising daily observations on all workers employed on the plantation. Our unit of analysis is the worker-day. For days on which workers participate and are assigned to plucking duties, the observations include the number of kilograms of green leaf tea plucked by each worker. This one number is our measure of productivity.

We restrict attention to days on which pluckers participated *and* were assigned to plucking duty. The reason for our focus on actual plucking days is twofold. First, on days when the worker is absent, we do not have any means of knowing to what activity she *would* have been assigned. This means that we are compelled to assign a missing value for productivity, rather than a zero, to such an observation. Second, when workers are assigned to non-plucking activities, there is no comparable measure of productivity.

Table 1 describes the basic data structure. For 2008, we have daily observations for roughly 2000 workers, observed for 5 months: one month prior to the contract change and 4 months following the contract change. In this and all future tables and figures, Month 0 refers to the month before the (date of the) contract change and Months 1-4 refer to the corresponding months following the contract change. Note that the number of worker-day observations is not necessarily the product of the number of work days and the number of workers because workers may have been absent or been assigned to non-plucking duties.

We exploit two useful counterfactuals. The first is our plantation observed over the same 5 months in 2007, the year before the contract change; 82% of the workers in the study plantation sample are observed in both 2007 and 2008, and within each year, roughly 85% of workers are observed both in Month 0 and thereafter. Hence, participation does not change dramatically over this period. In fact, we will show below that it cannot account for the productivity changes we observe. Moreover, our results are robust to the inclusion of worker fixed effects.

The second counterfactual is a plantation owned by the same company and located in a hill station at a 100 km. great circle distance north of the study plantation, in the same state. This "control" plantation is observed contemporaneously in 2008, but it belongs to another employer's association, which is on a different 3-year renegotiation cycle. Hence it did not experience a contract change over the observation period.

Workers are permanently assigned to work gangs with an average of approximately 30 members. Gangs are assigned to pluck fields using either hands or shears based on a predetermined schedule set by management. Hence, there is no endogenous choice on the part of workers or supervisors regarding plucking method, field, or task assignment. This allows us to control for field and plucking method fixed effects, as well as the time-varying intensity with which a field was recently plucked, measured by a quadratic term capturing the number of days in the previous week that a field has been plucked using hands and (separately) shears. To account for additional time-varying weather related conditions, we correct for rainfall with an appropriate lag structure discussed later, constructed from daily Tropical Rainfall Measuring Mission (TRMM) grid cells whose centroid is within 20 km. of 4 villages near which the plantation is located.[6] For the control plantation, we do not have field, plucking method, or gang information, but we do have rainfall data.

3. Productivity Response in the Short Term

We begin by examining average daily output in Month 0 (August) and Month 1 (September). The time series in Figure 2 shows that output increased sharply in the treatment plantation in the month after the contract change. Despite an upward blip in week 2 after the date of the contract change, no corresponding increase is visible in the control plantation over the same time period. At the same time, it is evident that output in the treatment plantation in the first 3 weeks of Month 0 in 2008 is unusually low, both relative to the control plantation in 2008, and relative to Month 0 in the treatment plantation in 2007, denoted by the horizontal line. As we show later, this abnormally low output persists even when we correct for other (potentially time-varying) variables. In order to avoid the possibility that our estimates for the productivity response to the contract change are simply capturing a potential work slow-down leading up to the contract change, in what follows we restrict our "before

[6] We are grateful to Eric Stroebl for these data.

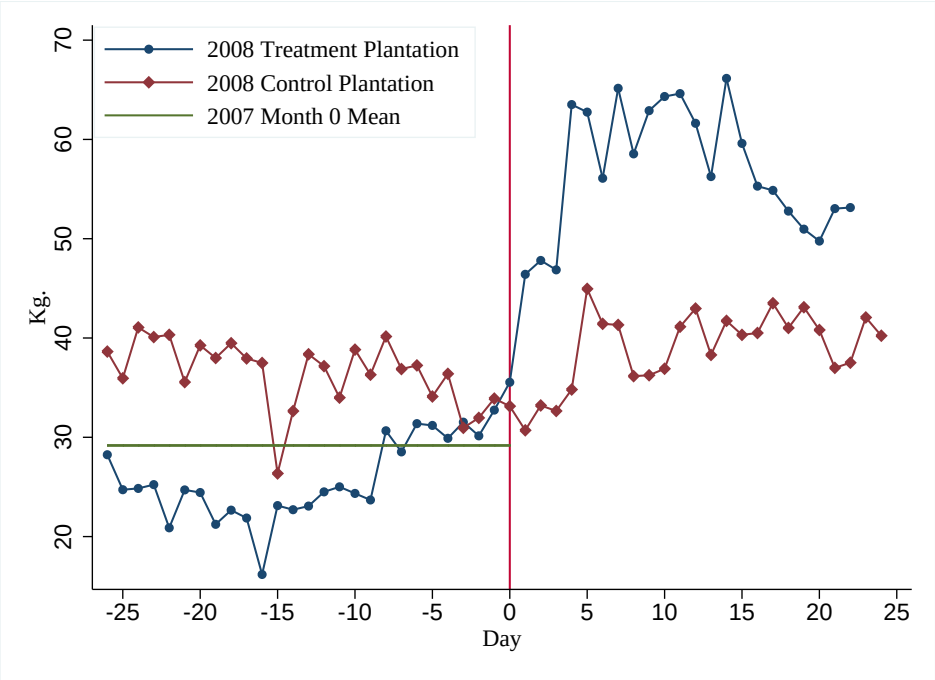
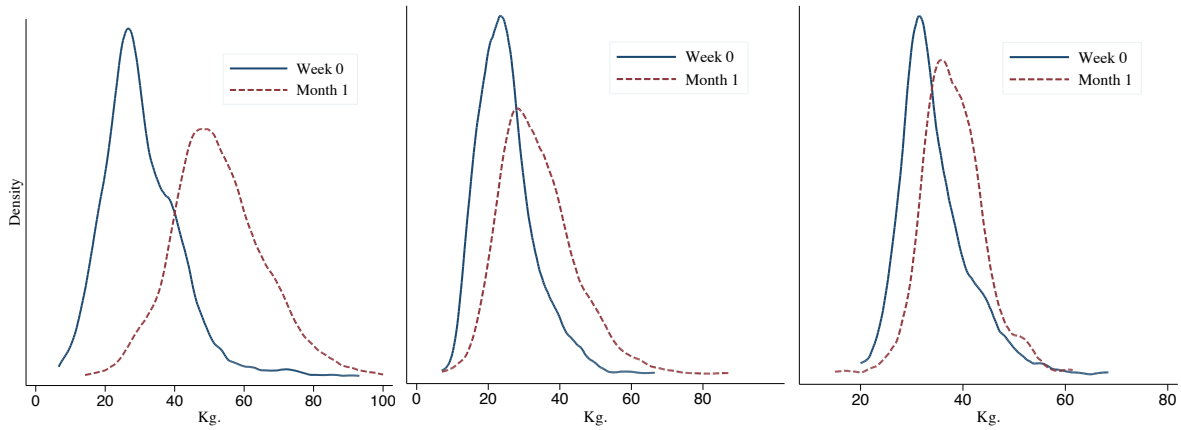


Figure 2. Time Series: Average Daily Output. *Notes.* The time series depicted in this graph describe average daily output in the treatment plantation in 2008 and in the control plantation in 2008 for two months on either side of Day 0, the date of the contract change. The horizontal line denotes Month 0 average output in the treatment plantation in 2007.

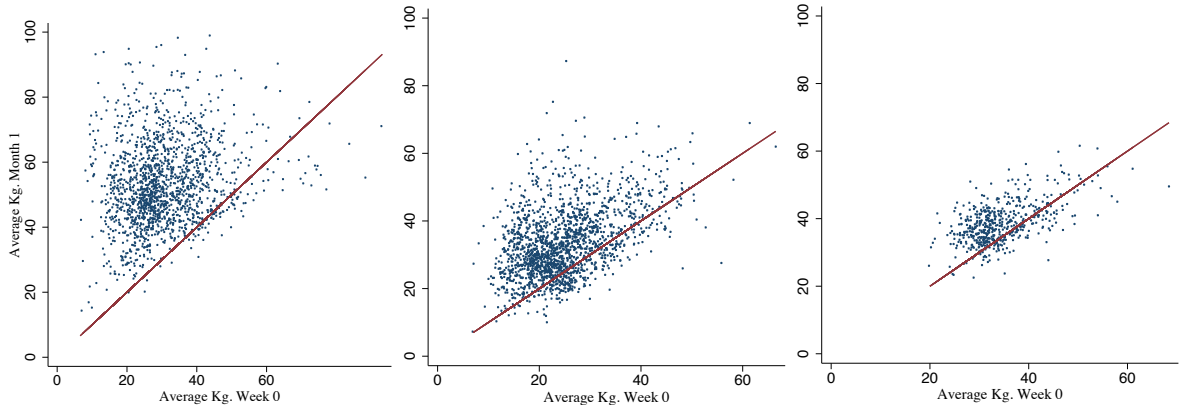
the contract change" sample to the work week, i.e. 6-day period, directly preceding the date of the contract change. We refer to this as Week 0. In the Appendix, we present analogous results using data from all of Month 0.

Figure 3 depicts kernel densities of daily individual output for Week 0 and Month 1, in three cases. The first panel depicts the situation of interest, which is output in the study plantation in the last week of August and in the month of September, 2008. The middle panel depicts the study plantation again, but this time over the same periods in 2007, when there was no contract change. The right panel depicts output in the control plantation in the contemporaneous period in 2008 (again, there was no change in contract). There is a mild shift in the distribution to the right for both the treatment plantation in 2007 and in the control plantation in 2008 (it is close to the start of the plucking season), but the rightward shift in the treatment plantation in 2008 is dramatic. Average output increased by 75%, from 31 to 54 kg. between Week 0 and Month 1 in the treatment plantation in 2008 compared to a 38% increase from 25 to 34 kg. in 2007 and a contemporaneous 11% increase from 35 to 39 kg. in the control plantation. Figure A1 in the Appendix shows that both the relative and absolute increases are only magnified when one uses all of Month 0 rather than Week 0 as the baseline. Average output more than doubled from 25 to 54 kg. between Month 0 and Month 1 in the treatment plantation in 2008 compared to a 17% increase from 29 to 34 kg. in 2007 and a contemporaneous 8% increase from 36 to 39 kg. in the control plantation.



(a) 2008 Treatment Plantation **(b)** 2007 Treatment Plantation **(c)** 2008 Control Plantation

Figure 3. Kernel Density: Average Daily Output. *Notes.* This figure depicts kernel density estimates for the average daily output of workers in Week 0 (solid line) and Month 1 (dashed line). Panel (a) corresponds to the study plantation in the year of the contract change, Panel (b) to the study plantation in the year before the contract change, and Panel (c) to the control plantation in the year the contract change took place. Densities are calculated using an Epanechnikov kernel



(a) 2008 Treatment Plantation **(b)** 2007 Treatment Plantation **(c)** 2008 Control Plantation

Figure 4. Scatter Plot: Average Daily Output. *Notes.* This scatter plot depicts average daily output per worker in Week 0 (x-axis) and Month 1 (y-axis). Each dot represents an individual worker. The solid line is drawn at 45°. Panel (a) corresponds to the study plantation in the year of the contract change, Panel (b) to the study plantation in the year before the contract change, and Panel (c) to the control plantation in the year the contract change took place.

Plots of the densities before and after do not capture what happens at the level of an individual worker. Figure 4 takes this up by looking at within-worker changes in productivity, measured in average daily kilograms of tea plucked in Week 0 (the horizontal axis) and Month 1 (the vertical axis). Each observation in the scatter plot corresponds to an individual worker. In 2008, the left panel, 94% of observations are located above the 45^0 line: almost *all* the workers are more productive after the

contract change. While in 2007, 82% lie above the 45^0 line, and 75% in the control plantation in 2008, the increase is clearly more muted. Again, this upward shift in the treatment plantation is more dramatic both in relative and absolute terms, when one includes data from all of Month 0 (see Figure A2 in the Appendix).

This is a remarkable increase, and we will need to account for various sources that might bear on it. In this section, we address some of the preliminary considerations.

Start of the Plucking Season. August is the first month of the plucking season, and there is a natural tendency for output to grow over this period as the season picks up. Different estates on the plantation are more likely to fall under a lower yield class (yield class 2) in August, and a higher yield class (yield class 3) in September. This is true of our plantation both in 2007 and 2008. All the estates in the treatment plantation were under yield class 2 in Month 0, and under yield class 3 in Month 1. We can form a rough estimate of how much output is expected to grow over the two yield classes by looking at the change in the minimum standard: which is from 23 kg. to 28 kg. for hands, and 28 kg. to 33 kg. for shears. This is an increase of 18–22%, and it roughly explains the growth in output in our plantation in 2007, and on the control plantation in 2008. But that is nowhere close to what happened on the treatment plantation in 2008.

Hands Versus Shears. The higher output could reflect an increase in use of shears in the month(s) following the contract change, since output is higher with shears than with hand plucking. However, Figure 5 shows that the change we observe does not simply represent a technological shift, as output in the treatment plantation increased both for hands and shears. Additional graphical support — kernel densities for hands and shears, and accompanying scatter plots at the worker level — is available in Figures A3 and A4 in the Appendix.

At the same time, we need to correct for the possibility of a compositional change from hands to shears, which could confound a proper estimate of the output increase following the contract change. We take care of this below; see Table 3.

Participation Rates. The scatter plots presented in Figure 4 demonstrate within-worker increase in output, conditional on workforce participation. But it is also worth noting that there was no overall change in participation following the contract change. When we estimate a probit model (not reported) for all working days with a binary dependent variable indicating work participation on the left hand side and a dummy variable equal to 1 in the period after the contract change on the right hand side, the coefficient on the dummy variable is statistically insignificant and close to zero.

Rainfall. We do need to control for weather patterns. The southwest monsoon is relevant for the period under consideration. In the tea-growing region we study, it would generally last through the summer months until the end of August, contributing to a growth in productivity into September. (As we shall see below, rainfall is beneficial to plucking, but with a lag to account for increased growth of the tea bushes.)

The analysis that follows not only takes care of weather patterns, it does so in conjunction with the possibility that higher output can be attributed to increased use of shears, as well as the plucking of different and more productive fields. We account for all of these together by estimating the residuals of an OLS regression controlling for field fixed effects, a quadratic term for the number of days in the previous week a field was plucked by hand and (separately) by shears, as well as time-varying weather patterns.

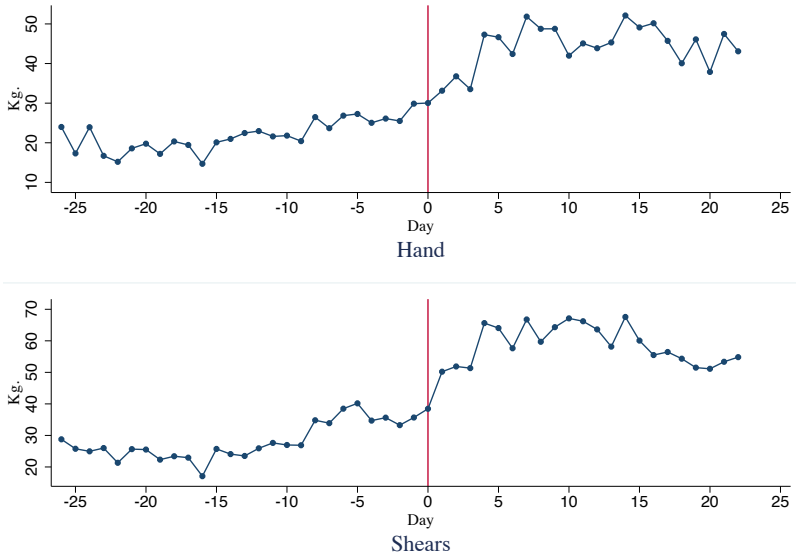
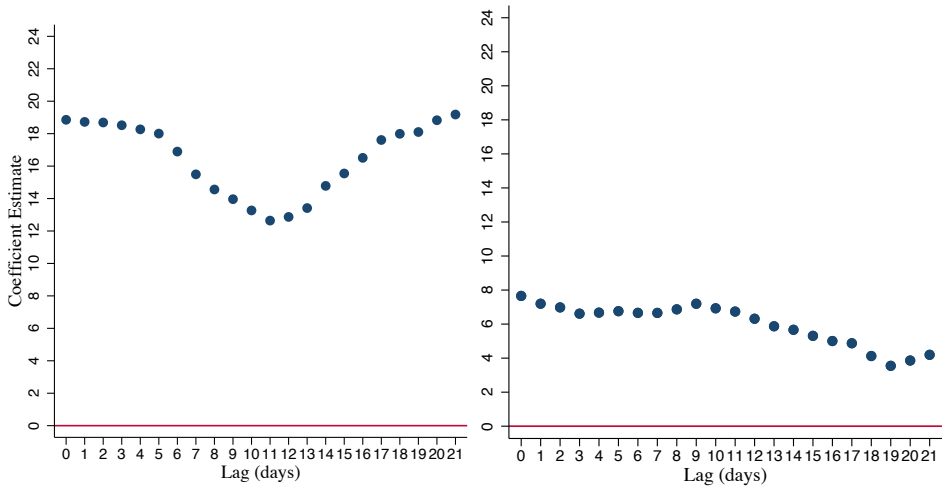


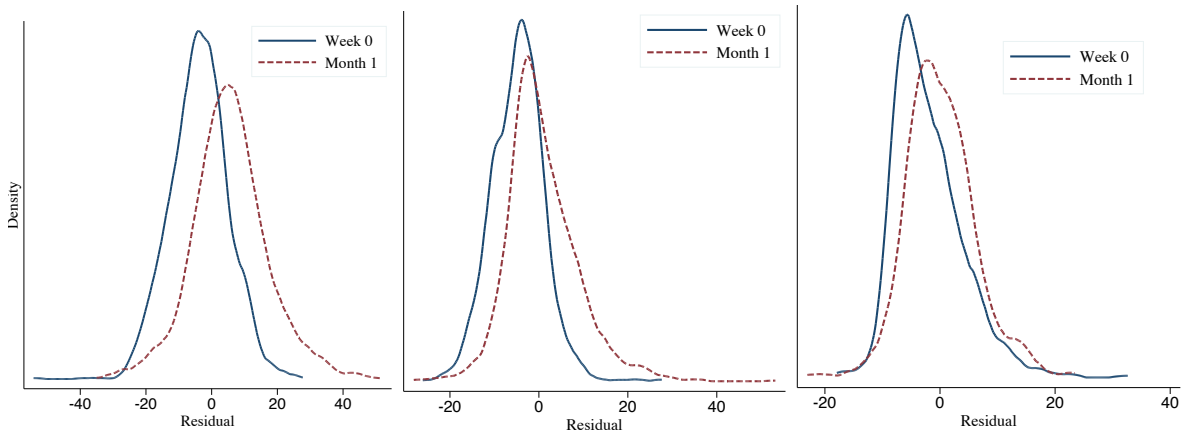
Figure 5. Time Series: Average Daily Output Disaggregated by Hand and Shears. *Notes.* This figure depicts average daily output in the treatment plantation in 2008. The top panel averages over plucker-days assigned to hand plucking and the bottom panel to plucker-days assigned to shears plucking.



(a) 2008 First Difference **(b)** Double Difference

Figure 6. Productivity Change with Different Rainfall Lags. *Notes.* This figure depicts the first difference (panel a) and the double difference (panel b) estimates for the residuals using for different lags of a seven-day moving average of rainfall. Notice that the minimum value of the first difference estimate is reached at around day 11.

Figure 6 depicts the first difference and the double difference estimates for Week 0 and Month 1, for various lags of a 7-day moving average of local rainfall. The left panel does this for the treatment plantation in 2008. It reports OLS coefficients for the first difference estimate τ_1 from the following



(a) 2008 Treatment Plantation **(b)** 2007 Treatment Plantation **(c)** 2008 Control Plantation

Figure 7. Kernel Density: Average Daily Residual. *Notes.* This figure depicts kernel density estimates analogous to those in Figure 3, but for residuals obtained after controlling for rainfall, plucking method and field type.

regression (this is the estimate of output change after the new contract):

$$\text{Output}_{it} = \alpha + \tau_1 \text{ After}_t + \psi \text{ Controls}_{i(t)} + \rho \text{ Rainfall}_t + \varepsilon_{it}, \tag{1}$$

where "After" equals 0 in Week 0 and 1 in Month 1, and "Controls" in this and all future residual estimates include time-varying field plucking intensity mentioned earlier, as well as method and field fixed effects. The variable "Rainfall" is entered with a multitude of lags, ranging from contemporaneous to three weeks; in each case, the estimate τ_1 is recorded on the vertical axes of Figure 6.

The right panel of that Figure presents the OLS double difference estimate τ_2 from the following regression. It measures how productivity in the treatment plantation in 2008 changed relative to the analogous change in 2007. We have:

$$\text{Output}_{it} = \alpha + \tau_0 \text{ Treat}_{i(t)} + \tau_1 \text{ After}_t + \tau_2(\text{ After}_t \times \text{Treat}_{i(t)}) + \beta \text{ Controls}_{i(t)} + \rho \text{ Rainfall}_t + \varepsilon_{it} \tag{2}$$

where $\text{Treat}_{i(t)}$ is a dummy variable equal to 1 in 2008 and the interaction term is a dummy taking value 1 after the contract change in 2008. The right panel of Figure 6 presents the double difference estimate τ_2 from (2) using observations from Week 0 and Month 1.

Figure 6 indicates that regardless of the choice of lag structure, the change in output both for the first difference in 2008, as well as the double difference, remains firmly positive following the contract change. The very lowest increase in output (in the left panel) after accounting for all the controls is around 12 kg. per person, which represents an increase of almost 40% (see Table 6 for a more precise estimate). Figure 6 also shows that an 11-day lagged rainfall structure gives us the most conservative estimate of the productivity increase. (Using the full Month 0 data also yields an 11-day lagged rainfall structure as our most conservative estimate.)

In all future estimates controlling for unobserved heterogeneity in the treatment plantation, we calculate residuals of an OLS regression with output on the left hand side, field fixed effects, plucking method fixed effects, quadratic terms for the number of days a field has been plucked using hand and

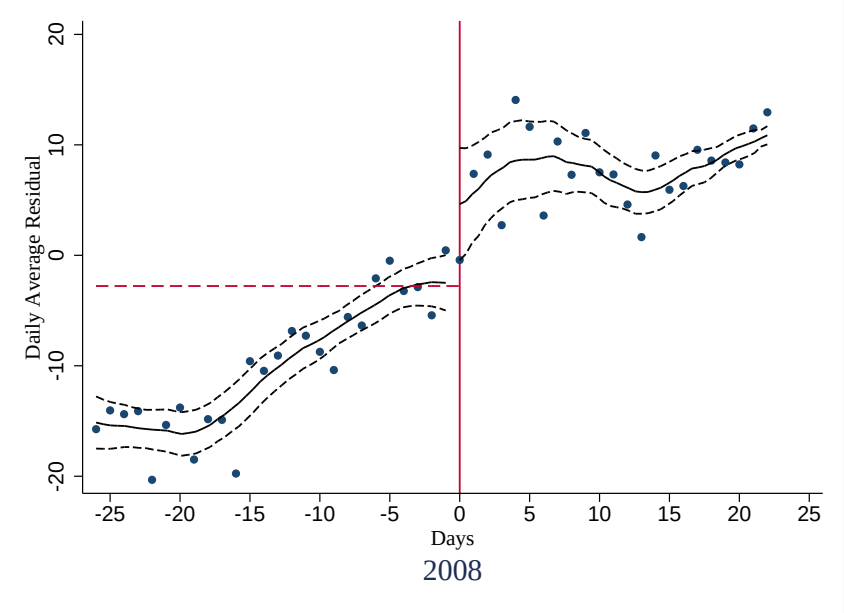


Figure 8. Time Series: Average Daily Residual. *Notes.* This figure depicts average daily residual in the treatment plantation in 2008. Each dot in the figure represents a different day. The figure is overlaid with a local polynomial smoothed using an Epanechnikov kernel (solid curve). Dashed curves denote the 95% confidence interval. The horizontal line denotes the 2007 Month 0 average residual in the treatment plantation.

shears in the last week, and an 11-day lagged 7-day moving average of local rainfall. In the control plantation, our residuals only account for rainfall since we do not have field-level data.

Figure 7 is the analogue of Figure 3 for residuals thus estimated, rather than the raw outputs themselves. The productivity increase, while somewhat muted relative to the raw output estimates, is still highly pronounced in the treatment plantation in 2008. There remains a small shift of the residual distributions to the right in the treatment plantation in 2007 and the control plantation in 2008, but this shift is much smaller relative to the treatment plantation in 2008. (As usual, this increase is far more pronounced in both relative and absolute terms when all of Month 0 is used; see Appendix Figure A5.) The scatter plot in Figure A6 in the Appendix (this is the residuals analogue of Figure 4) confirms that the increase holds not just for the overall density but at the individual level: 80% of individuals lie above the 45^0 line in 2008, compared to 72% of individuals in 2007 and 75% of individuals in the control plantation in 2008. Moreover, the per-individual increase is far more pronounced for the treatment plantation in 2008.

Figure 8 plots the time series of daily average residuals for Months 0 and 1 in 2008, fitted with a local polynomial using an Epanechikov kernel. As before, there is a depression in the first 3 weeks of Month 0, but by the week directly preceding the contract change – Week 0 – the residuals are approximately 0, which coincides with the Month 0 average residual in the treatment plantation in 2007 (marked with the horizontal line).

The figure displays a sharp discontinuity at the date of the contract change. Such a discontinuity is absent in the residual time series for both the treatment plantation in 2007 and the control plantation

in 2008 (Figures A8 and A9 in the Appendix). Table 2 puts a number on this increase by presenting regression discontinuity (RD) estimates of τ_1 using Month 0 and Month 1 data. At an optimal bandwidth of 3 days for the full sample, these estimates are insensitive to the seemingly depressed output in the first few weeks of observation. The RD estimates from column 1 of Table 2 show that just in the region of the discontinuity, there is a sharp increase of between 5–7 kilograms, somewhere in the vicinity of 20% in a matter of days.

	All	Shears[a]	Hand[b]
Optimal Bandwidth[c]	3.41	5.85	1.16
RD Estimate	5.309***	6.624***	6.793***
	(0.583)	(0.647)	(0.624)
2 X Optimal Bandwidth RD Estimate	6.727***	5.202***	4.755***
	(0.429)	(0.431)	(0.430)
3 X Optimal Bandwidth RD Estimate	5.513***	8.242***	4.684***
	(0.345)	(0.355)	(0.373)
Observations	99,942	80,339	19,603

Table 2. Regression Discontinuity Estimates: Month 0 and Month 1. *Notes.* Dependent variable is the OLS residual for output levels. *a*. Person-day observations with shears. *b*. Person-day observations with hand. *c*. Imbens and Kalyanaraman (2012) optimal bandwidth for sharp design. $^{*}p < 0.10,^{**}p < 0.05,^{***}p < 0.01$.

Table 3, which presents OLS first difference estimates of τ_1 in equation (1), extends the analysis to the entire month following the contract change. As promised, it takes care of compositional effects for hands and shears, by controlling for the number of days (and days squared) that had been previously spent on each plucking mode, including a "shears dummy" if the a worker had been assigned to pluck a field using shears. It also controls for field fixed effects and rainfall.[7]

The first three columns of this table use Week 0 as the baseline. Column 2 accounts for worker participation effects by restricting attention to those workers who plucked tea for at least 4 days before and after the contract change, and Column 3 includes worker fixed effects. The expansion of output by approximately 12.5 kg. is undeniable. It corresponds to an increase of 40%. This increase doubles when all of Month 0 is taken as a baseline (columns 4-6).

4. Possible Drivers of the Short-Term Increase in Output

The discussion so far establishes that following the contract change, there was an immediate output increase of a sizable magnitude. There are different factors that might be responsible for this large increase. In the following sections, we examine the different possibilities, itemized here:

[7]These are controls used throughout to estimate residuals and we will not mention them again.

	Month 1 over Week 0			Month 1 over Month 0		
	OLS	OLS	FE	OLS	OLS	FE
After	12.64***	12.58***	12.96***	24.06***	23.99***	23.99***
	(1.814)	(1.769)	(0.403)	(1.914)	(1.941)	(0.326)
Rainfall	0.51***	0.54***	0.50***	0.42***	0.43***	0.42***
	(0.066)	(0.066)	(0.015)	(0.070)	(0.071)	(0.014)
Shears days	-0.28	-0.05	-0.05	0.33	0.38	0.52***
	(0.702)	(0.710)	(0.173)	(0.562)	(0.538)	(0.138)
Shears days^2	0.17	0.14	0.11***	0.05	0.04	-0.01
	(0.149)	(0.148)	(0.034)	(0.124)	(0.117)	(0.032)
Hand days	-3.52**	-3.75**	-3.05***	-1.36	-1.34	-1.29***
	(1.460)	(1.510)	(0.293)	(0.871)	(0.881)	(0.217)
Hand days^2	0.73**	0.78**	0.66***	0.24	0.24	0.23***
	(0.273)	(0.285)	(0.063)	(0.180)	(0.182)	(0.051)
Shears dummy	13.87***	13.65***	14.64***	6.82***	6.81***	7.04***
	(2.333)	(2.331)	(0.585)	(1.678)	(1.654)	(0.338)
Field FE	YES	YES	YES	YES	YES	YES
Worker FE	NO	NO	YES	NO	NO	YES
Plucked at least 4 days before and after	NO	YES	NO	NO	YES	NO
No. Observations	41,241	34,491	41,241	63,666	61,258	63,666
Adjusted R-squared	0.508	0.525	0.536	0.582	0.584	0.604
No. Unique Workers			1,985			2,126

Table 3. Short Run Regression Estimates. *Notes.* This table presents regression estimates for τ_1 in equation (1) (First difference) in 2008. Columns 1–3 compare Week 0 to Month 1. Columns 4–6 compare Month 0 to Month 1. Columns 2 and 5 impose the further restriction that workers worked at least 4 days before and after the contract change, in order to deal with worker selection. Columns 1–2 and 4–5 estimate equation (1) using OLS, with standard errors clustered by day. Column 5 estimates fixed effects (FE), with robust standard errors. Standard errors in parentheses. $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$.

1: Static incentives generated by the new contract
2: Dynamic incentives
3: Heightened supervision
4: "Behavioral" responses

4.1. **Static Incentives.** The new contract raised base wages and, if anything, lowered the piece rates at the margin. In addition, a penalty for falling below the minimum standard was removed. Classical incentive theory tells us that this cannot increase worker effort (after controlling for field, plucking method, and all shocks such as rainfall or family illness).

More formally, let w denote the original wage function and $\hat{w}$ the new wage function, and let s and $\hat{s}$ be the minimum production standards in each case. The new contract changes the wage function and the production standards so that,

C1: $\hat{w}(y) \geq w(y)$ for all y,
C2: $\hat{w}(y) - w(y)$ is non-increasing in y,
C3: $\hat{s} \leq s$.

Condition C1 captures the upward shift of the wage function, and C2 the flattening of the function; C3 records the decrease in minimum production standards. All three are features of our observed contract change.

We make no particular assumption about the wage functions except that they are continuous and increasing, with bounded slope. Typically they will be comprised of a baseline wage and various piece rates, as in our plantation.

The worker supplies effort e to produce output y. We take the two to be measured in the same units, so that $y = e$.[8]

The worker's payoff function is given by

$$u(w) - c(e) - L(s - e),$$

where u is increasing and concave in w, c is increasing and convex in e (with unbounded slope as $e \to \infty$) and L is a nondecreasing, convex function with $L(x) = 0$ for all $x \leq 0$. Note that the consequence of not meeting the standard is, to some extent, a choice variable that can be influenced by the employer via supervisory effort.

The worker chooses effort to maximize payoff, subject to the link between w and e created by the wage function. Notice that the optimization problem is not concave even if payoffs have the "correct" curvature, as the wage functions might be non-concave, and in our setting, they certainly are. So, while an optimum always exists under our assumptions, multiple optima are possible. Yet, the following result states, roughly speaking, that no optimal effort under the new contract can exceed any effort that was optimal before the change.

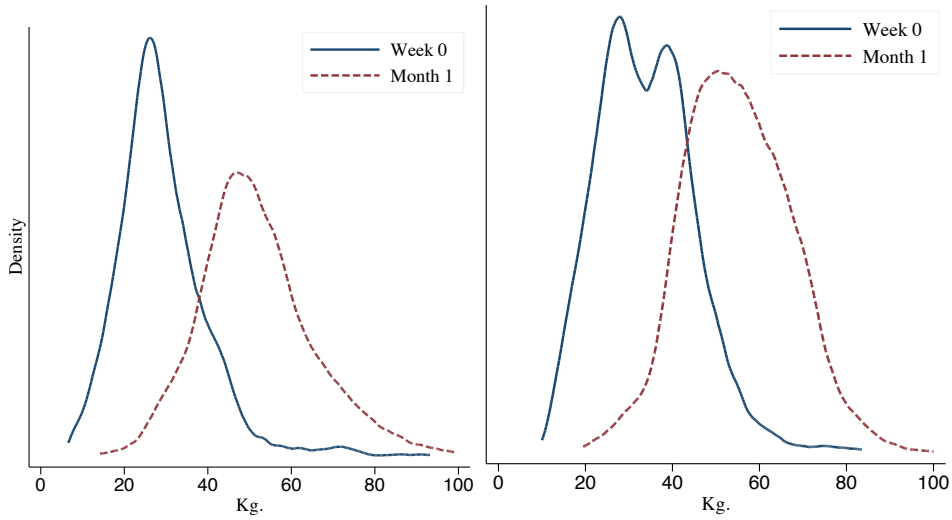
PROPOSITION **1.** *Let e be an optimal effort choice for w and $\hat{e}$ an optimal choice for $\hat{w}$. Then either both choices are optimal for both problems, or $e \geq \hat{e}$.*

This result holds in fact for more general and non-separable utility functions, provided we impose submodularity restrictions on the interaction between w, e and s. See the Appendix for a proof.

The proposition simply formalizes what we would suspect right away: that a contract change that increases and flattens wages (and lowers standards without changing supervisory effort) should induce workers to decrease effort, or at best keep the set of optimal effort choices unchanged. This is precisely why the increase in output following the contract change is of interest. That said, we will return to this model below in order to estimate some parameters off the data preceding the contract change. That will allow us to examine just how much of the post-contract variation can be captured by these parameters, especially in the longer-term.

We now turn to some possible explanations for the immediate increase in output.

[8]We control for all stochastic shocks that are realized prior to the supply of effort, and so don't carry these in the notation. In addition, there is little loss of generality in removing all stochastic shocks after effort is supplied, because the condition and quantity of the leaves on the bushes are observed at the time of applying effort. In any case, the approach can directly be extended to more general output functions of the form $y(e, \epsilon)$ where y is non-decreasing in both effort e and (random) perturbation ϵ.



(a) Permanent Workers **(b)** Temporary Workers

Figure 9. Kernel Density: Average Daily Output by Permanent and Temporary Workers. *Notes.* This figure depicts kernel density estimates for the average daily output of workers in Week 0 (solid line) and Month 1 (dashed line), as in Figure 3, but disaggregated by temporary and permanent workers.

4.2. **Dynamic Incentives.** It is possible that the new contract might have been accompanied by sharper dynamic incentives. For instance, the very fact of higher wages might serve as an inducement to provide higher effort, using the threat of firing, as in Shapiro and Stiglitz (1984) and Dutta, Ray, and Sengupta (1989).

However, permanent workers cannot be fired, by law. And this is not just *de jure*, it is *de facto*: the median tenure of permanent workers on our plantation is an impressive 21 years. On other other hand, the very fact that permanent workers cannot be fired yields a potential test for dynamic incentives, which involves the comparison of output responses across permanent and temporary workers. The latter are only hired seasonally. The significant improvement in contractual terms can therefore reasonably be expected to act as an "efficiency wage." If dynamic incentives are indeed at the heart of the story, temporary workers should be contributing the bulk of the increase in plantation output following the contract change.

Figure 9 depicts kernel densities analogous to those in the left panel of Figure 3, disaggregated by permanent and temporary workers. Both sets of workers display similar shifts in output. This does not square with the dynamic incentives argument. Whatever the cause of the output increase, it applied equally to permanent and temporary workers.

Table 4 drives this point home by computing first differences along the lines of equation (1), carrying out the exercise separately for permanent and for temporary workers. We consider the specification in which the difference is computed using Week 0 as baseline. (Using a Month 0 baseline yields estimates for both permanent and temporary worker which are 10 kg. higher than the Week 0 baseline estimates.) Across all specifications, the reactions of permanent and temporary workers are very similar. If anything, the response of permanent workers to the contract change is slightly larger.

	Permanent			Temporary		
	OLS	OLS	FE	OLS	OLS	FE
After	13.45***	13.28***	13.86***	11.21***	11.18***	11.33***
	(1.831)	(1.747)	(0.487)	(2.098)	(2.200)	(0.765)
Rainfall	0.52***	0.55***	0.51***	0.49***	0.53***	0.48***
	(0.070)	(0.070)	(0.019)	(0.071)	(0.072)	(0.028)
Shears days	0.06	0.29	0.18	-0.85	-0.72	-0.39
	(0.787)	(0.857)	(0.206)	(0.978)	(0.885)	(0.332)
Shears days^2	0.10	0.08	0.08**	0.28	0.24	0.14**
	(0.167)	(0.182)	(0.039)	(0.235)	(0.209)	(0.066)
Hand days	-2.42	-2.59*	-1.89***	-5.83***	-6.67***	-5.28***
	(1.436)	(1.458)	(0.292)	(1.621)	(1.740)	(0.746)
Hand days^2	0.58**	0.64**	0.50***	1.00***	1.08***	0.91***
	(0.283)	(0.294)	(0.064)	(0.317)	(0.344)	(0.180)
Shears dummy	15.88***	15.81***	16.18***	8.62***	8.00***	9.75***
	(2.325)	(2.288)	(0.675)	(2.763)	(2.699)	(1.201)
Field FE	YES	YES	YES	YES	YES	YES
Worker FE	NO	NO	YES	NO	NO	YES
Plucked at least 4 days before and after	NO	YES	NO	NO	YES	NO
No. Observations	28,820	25,310	28,820	12,421	8,951	12,421
Adjusted R-squared	0.529	0.546	0.562	0.518	0.544	0.518
No. Unique Workers			1,326			670

Table 4. Short Run Regression Estimates: Permanent and Temporary Workers. *Notes.* This table presents regression estimates for τ_1 in equation (1) (First difference) in 2008, disaggregated by permanent (columns 1–3) and temporary workers (columns 4–6). The estimates compare 1 week before the contract change to one month after. Columns 2 and 5 impose the further restriction that workers worked at least 4 days before and after the contract change, in order to deal with worker selection. Columns 1–2 and 4–5 estimate equation (1) using OLS, with errors clustered by day. Column 5 estimates individual fixed effects (FE), with robust standard errors. Standard errors in parentheses. $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$.

Given these observations, we find it very difficult to attribute the output jump to any form of dynamic incentives connected with non-renewal of the contract. However, it is entirely possible that under-performing workers were pressured in other ways, or positively exhorted to perform well. This takes us to the next item: supervision.

4.3. **Supervision.** Recall that the government-mandated minimum wage increase, which had to be paid as the fixed, baseline component of the wage, was predictably resisted by the plantation owners. As already noted, writ petitions seeking a stay on the minimum wage notification were submitted by planters, and were dismissed by the state's high court just before the new contract was instituted. Plantation owners were all too aware of the incentive effects of piece rates, but now these could not be raised in tandem with the higher fixed wage: the costs of doing so were perceived to be too high.

	Underperformers		Overperformers		S (%)
	Levels	Logs	Levels	Logs	
Month 1	17.49***	0.47***	9.80***	0.23***	25
over Week 0	(2.128)	(0.052)	(2.134)	(0.043)	
Month 1	27.51***	0.79***	20.25***	0.48***	22
over Month 0	(1.808)	(0.051)	(2.030)	(0.046)	

Table 5. Supervisory Effort. *Notes.* This table presents OLS estimates analogous to those in Tables 3 and 4 (with the full set of controls, not shown here) in levels and logs for the 2008 Treatment Plantation, but disaggregated by person-method observations which were below the standard in Month 0 of 2008 ("underperformance"), and those which were above the standard in that month ("overperformance"). Coefficients in the top half of the table compare outputs to that in Week 0, and those in the bottom half, outputs to that in Month 0. The final column denotes the estimated share of output increase S due to supervision, as given by equation (3). Standard errors are clustered by day. $^{*}p < 0.10,^{**}p < 0.05,^{***}p < 0.01$.

It is therefore important to see if other non-pecuniary means of maintaining effort — or in a word, supervision — could have been intensified.

Supervisors are paid a fixed wage, between Rs. 2600–3400 per month, depending on seniority. They are drawn from the ranks of pluckers, are overwhelmingly female, and because of their long association to the plantation, are intensely loyal to its interests. At least for permanent workers, who cannot be fired, supervision is limited to a mix of exhortation to work, coupled with punishments such as bad treatment of the worker and her family (within limits, of course, as the workers are unionized), or later assignment to less palatable or more arduous tasks.

There are two ranges in which supervisory effort presumably has different impacts. The first is the output range that falls below the explicitly announced and commonly understood minimum standard, given the yield class and plucking method. In this range, the supervisor presumably has real authority, because any drop below specified minimum standards is tantamount to a breach of contract, even though the punishments are limited. The second range contains all outputs that exceed the standard, and for which an additional piece rate is paid. It is hard to imagine that the supervisor would have any significant authority in this range. The very fact that a piece rate is paid, to begin with, suggests that outputs in this zone are to be treated, to some extent, as a "bonus." Apart from the monetary inducement of the piece rate itself, it is very difficult, if not impossible to appeal to a legal contract to exhort or coerce a worker to consistently perform above the minimum standard.

Following up on this point, we can separate all workers functioning under some given plucking method into two categories. Using month-long data from the period just before the contract change, which is August 2008 or Month 0, we can see if the monthly average of each worker falls below ("underperforms") or above ("overperforms") the applicable standard before the contract change.[9] Table 5 records the subsequent increase in output coming from underperformers versus overperformers.

[9]August is the first month of the plucking season. The yield class, and therefore the standard, is generally lower, and this is taken into account when defining under- and overperformance.

A theory based entirely on supervisory effort below the minimum standard would generate no output increase at all from the overperformers (simply apply an obvious variant of Proposition 1). All the output increase would come from underperformers. This isn't quite what we see in Table 5. It is true that the output increase from underperformers is significantly larger *in absolute terms* (and therefore *a fortiori* so in relative terms) than the corresponding increase from the overperformers. It is possible to argue that some of this effect is surely due to regression to the mean, as the underperformers return to some steady state of performance. But in light of the fact that we average outputs over an entire month to define underperformers, we consider this effect to be relatively unimportant. The underperformers raised their output by approximately 80%, if Month 0 is taken to be the baseline, and by approximately 50%, if Week 0 is the baseline.

Yet at the same time, the overperformers *also* increased their output. The corresponding numbers are are around 50% and 25% for the Month 0 and Week 0 baselines respectively. As we've already argued, it is extremely unlikely that this increase could come from supervision. Tentatively, then, we can estimate the fraction of the output increase in Month 1 due to supervision. Say the underperformers increase by percentage a, while the overperformers increase by a corresponding percentage b, where $a > b$. First, attribute *none* of b, which is all above the standard, to supervision. Remove that same increase from the underperformers, which reflects the assumption that whatever "behavioral response" is driving the overperformers is also driving the underperformers. Attribute the rest, which is $a - b$, to better supervision. Then the contribution share from supervision, S, is given by the formula

$$S = \frac{q(a-b)}{qa + (1-q)b} \quad , \tag{3}$$

where q is the output share of underperformers in the week or month (depending on the baseline) prior to the contract change. This is what we report as "supervision share" that we report in the final column of Table 5. It indicates that heightened supervision accounts for at most a quarter of the output increase following the contract change.

Admittedly, this number is open to reinterpretation and debate, but to increase it any further, one must assume either that individuals who were above the standard to begin with were nevertheless coaxed or coerced into doing better, or that individuals who were below the standard did not exhibit any of the same behavioral responses as their compatriots above the standard, and were made to perform better by supervision alone. For instance, if we presume that *all of* the increase a must be credited to supervision, then the supervisory contribution nearly doubles, but we are still left with a large and unexplained gap.

Our tentative conclusion is that the remaining gap (and perhaps more) must be chalked up to "behavioral responses." A contract was generously improved — or at least the improvement was perceived to be generous — and workers reciprocated. But which of several potential responses — gratitude, reciprocity, obedience or fear — might it be? The data that we have does not permit us to go further down this road, but it seems plausible that underperformers were more likely to be subject to most of these influences than overperformers. They experienced a more generous effective proportional wage increase than overperformers. As a result, they may have been "more that proportionately" grateful or reciprocal. Other behavioral explanations are also consistent with a larger relative response for underperformers. For instance, workers might like to earn a positive

bonus, over and above their utility for money. With the piece rate in the lowest slab set to zero under the new contract, this would have induced larger responses from lower-productivity workers. Or perhaps low-end workers responded positively to the removal of the penalty. These or other explanations would imply that our estimate of S may well be an upper bound for supervisory effort.

In sum, although we cannot be quite sure about the exact share attributable to heightened supervisory effort, it appears very likely that "behavioral effects" played a significant role in the Month 1 output increase.

5. Productivity Response in the Longer Term

If our data were limited to this immediate post-contract period we would be left with the impression of a rather dramatic and counterintuitive increase in output. It is counterintuitive because the new contract dampened incentives at the margin. The increase is robust to the inclusion of various controls, it is not explicable by the use of dynamic incentives such as firing threats, and increased supervisory effort explains at best a fraction of the overall change. We are left, then, with the intriguing hypothesis that a "behavioral" response took place following the contract change, one that reciprocally rewarded a better deal with higher effort.

Our task is not to examine the nature of the behavioral response: we do not have the data to do that. Rather, we ask a different question: for how long did the response last? Fortunately, we have data on worker productivity for up to 4 months into the new contract; that is, to the end of the annual plucking season. Our findings can be summarized in a single sentence: the output differential reversed itself in the months that followed, with most of the increase eroded by Month 4.

We begin by examining the daily time series of residuals by extending Figure 8 to 4 months after the contract change. Figure 10 does this. The horizontal line denotes average Week 0 residuals in 2008, which is almost identical to the average residual in the treatment plantation in Month 0 of 2007 and in the control plantation in Month 0 of 2008. The figure shows that the productivity response observed in Month 1 persists through the second half of Month 2, but then tapers off until, by the end of Month 4, output retreats to the levels in the week before the contract change.

Why is Week 0 a good baseline for this comparison? The reason is that Month 4 takes us to the very end of the plucking season, and output there is similar to that at the beginning of the season. This is immediately confirmed by an examination of the corresponding residuals for the treatment plantation in 2007 and the control plantation in 2008 (Figures A8 and A9 respectively, in the Appendix). As already noted, neither series exhibits a discontinuity around September 2, and by Month 4, the outputs from the two time series are at Week 0 levels — and at similar Month 4 levels compared to the treatment plantation in 2008. Thus no matter how we look at it, the Month 1 jump on the treatment plantation in 2008 appears to be reversed by the end of Month 4, both in itself and relative to our two counterfactuals.

A different perspective on the long-run change is provided by Figure 11. In each panel, the solid line depicts the kernel density of average daily residuals in the treatment plantation in Week 0. The dashed line in each panel is the kernel density of average daily residuals in the treatment plantation for each of the four months after the contract change. Notice that the two densities are very different in the first panel, but then the difference begins to reverse, until by Months 3 and 4 it has virtually

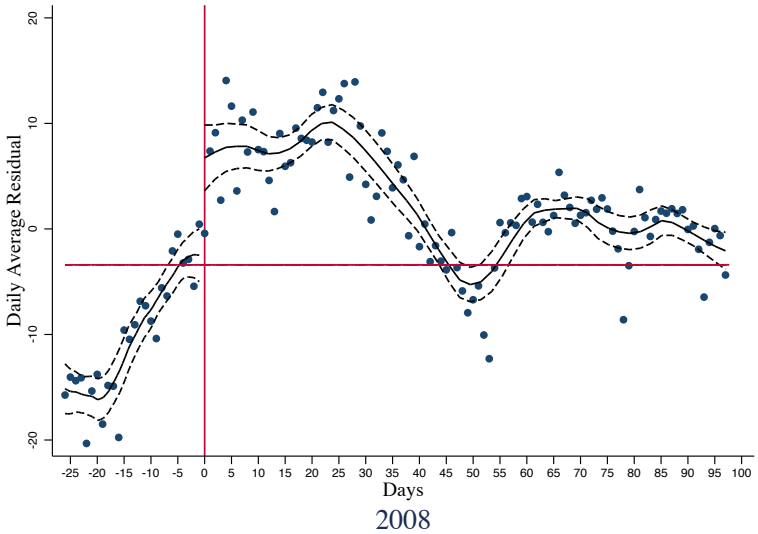


Figure 10. Time Series: Average Daily Residual. *Notes.* This figure extends the time series for residuals in Figure 8 to 4 months after the contract change, for the treatment plantation in 2008, accounting for rainfall and other time-varying controls. The horizontal line denotes average output in Week 0 in the treatment plantation in 2008.

vanished. A similar shift down can be observed in month-by-month scatter plots of average residuals before and after the contract change; see Appendix Figure A10.

It should be noted that the comparison is with the week just before the contract change, when average daily output was comparable to that in the treatment plantation in Month 0, 2007, and to that in the control plantation in Month 0, 2008. If the comparison is made, instead, to the density of average outputs over the entire Month 0, 2008, then the difference is initially still more pronounced, and the reversal is not complete by Month 4 (output is still higher). As we believe that the output at the beginning of Month 0 was unusually low, we prefer the comparison in Figure 11, but for completeness, we provide the corresponding diagram using monthly averages as benchmark in Figure A11, in the Appendix.

Table 6 puts numbers on the figures, restricting the "before" period to Week 0, and explicitly resurrects our two counterfactuals. The table provides OLS first difference estimates of τ_1 from equation (1), and double difference of τ_2 from (2) for our two counterfactuals: the treatment plantation in 2007 and the control plantation in 2008. Note, as before, that the first counterfactual corrects for all potentially time-varying controls whereas the second corrects only for rainfall; accordingly, we report the first differences separately for the treatment plantation in both these specifications.

There are large increases in output in the treatment plantation in Months 1 and 2, but in Months 3 and 4 they fade. The entire increase effectively vanishes in the last two weeks of Month 4 (see the bottom rows) — this, despite the fact that there is no change in participation during these last weeks. We note again that if we use the entire Month 0 average as baseline, we do get increases for first and

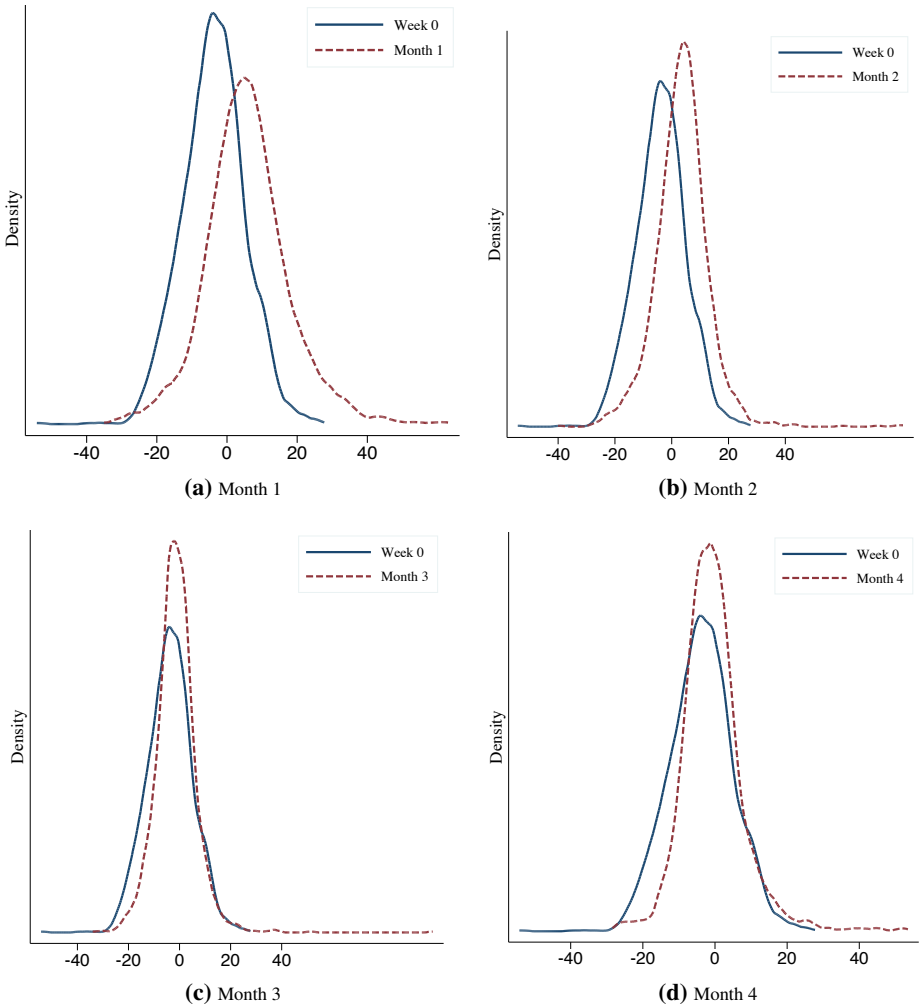


Figure 11. Kernel Densities: Average Daily Residuals. *Notes.* The solid line in each panel is the kernel density of average daily residuals in the treatment plantation one week before the contract change. The dashed line in each panel is the kernel density of average daily residuals in the treatment plantation in each of the four months after the contract change.

double differences all the way through, as Figures 10 and A11 indicate we would.[10] (It is still true, of course, that the differences decline.)

The relative decline characterizes both permanent and temporary workers: there is no evidence that dynamic incentives matter, not in Month 1 as we've already seen, and equally not in Months 2–4; see Table A1 in the Appendix.

[10]Our concern continues to be that the first few weeks of Month 0 is an aberration, perhaps even a work slowdown prior to the contract change, which is why we emphasize the last week of that month both for computing the increase and the subsequent reversal of output.

	Levels		Logs		Levels		Logs	
	FD	DD	FD	DD	FD	DD	FD	DD
	Counterfactual: Treatment Plantation in 2007[a]				Counterfactual: Control Plantation in 2008[b]			
Month 1	12.64***	6.74***	0.33***	0.37***	17.32***	13.66***	0.44***	0.34***
	(1.814)	(2.523)	(0.039)	(0.034)	(2.293)	(1.885)	(0.054)	(0.045)
Month 2	14.02***	6.64***	0.37***	0.24***	14.84***	12.98***	0.41***	0.35***
	(1.596)	(1.410)	(0.034)	(0.025)	(0.850)	(2.003)	(0.021)	(0.050)
Month 3	7.95***	-10.82***	0.24***	0.23***	10.15***	7.01***	0.31***	0.22***
	(0.949)	(2.424)	(0.025)	(0.032)	(0.646)	(1.330)	(0.019)	(0.037)
Month 4	8.64***	4.94***	0.23***	0.10**	9.84***	6.32**	0.29***	0.23***
	(1.458)	(1.243)	(0.032)	(0.041)	(1.220)	(2.451)	(0.030)	(0.063)
Week 49	9.17***	9.00***	0.24***	0.20***	8.96***	16.09***	0.27***	0.45***
	(2.214)	(1.454)	(0.048)	(0.051)	(1.045)	(2.747)	(0.028)	(0.077)
Week 50	11.86***	6.91***	0.29***	0.15***	13.03***	3.19*	0.37***	0.16***
	(1.489)	(1.421)	(0.029)	(0.043)	(0.861)	(1.568)	(0.019)	(0.042)
Week 51	7.15**	3.11**	0.17*	0.04	5.28**	2.63	0.23***	0.10
	(2.707)	(1.336)	(0.084)	(0.042)	(1.778)	(2.366)	(0.059)	(0.068)
Week 52	0.51	1.71	0.07	0.02	2.15	1.42	0.14	0.09
	(3.600)	(1.857)	(0.100)	(0.064)	(2.783)	(2.289)	(0.090)	(0.065)

Table 6. Change in Monthly Average Output Relative to Week 0. *Notes.* Each cell records the coefficient of a different OLS regression with controls, in which the dependent variable is output (columns 1-2 & 5-6) and log(output) (columns 3-4 & 7-8). Columns 1-4 include the full set of controls and columns 5-8 only control for rainfall (hence the difference in estimates in columns 1 and 5.) Columns 1 and 3 present first difference (FD) estimates for τ_1 from equation (1) for the treatment plantation, and columns 2 and 4 the double difference (DD) estimate for τ_2 from equation (2). *a*. Comparison plantation is the treatment plantation in 2007, controlling for rainfall and other time-varying characteristics. *b*. Comparison plantation is the control plantation in 2008, controlling for rainfall only. Weeks 49-52 refer to week-by-week differences in Month 4. Standard errors are clustered by day. $^{*}p < 0.10,^{**}p < 0.05,^{***}p < 0.01$.

Recall from Section 4 that the initial output increase could perhaps best be explained by a purely behavioral response such as reciprocity, with some help from heightened supervision. Other factors such as dynamic incentives are markedly absent. But over the three-month decline in output that occurs later, the relative contribution of supervision appears to increase, which suggests that the residual "behavioral" response dies away faster than the overall reversion of output to pre-contract levels. Table 7 continues the estimates in Table 5 for the treatment plantation in later months, as always with a full set of controls. Recall that we partitioned the sample into person-method observations which were below the standard in Month 0 of 2008 ("underperformers"), and those that were above ("overperformers"). We retain exactly this classification here, but add to Table 5 by examining what happens in Months 2–4.

The table shows that the *relative* contribution of supervisory effort tends to increase over Months 1–4. Our estimates of supervisory contribution climb from around a quarter to a third regardless of whether we measure this by the output difference over Month 0, or the output difference over Week 0.

	Month 1, 2, 3, 4 over Week 0					Month 1, 2, 3, 4 over Month 0				
	Underperformers		Overperformers		S	Underperformers		Overperformers		S
	Levels	Logs	Levels	Logs	(%)	Levels	Logs	Levels	Logs	(%)
1	17.49***	0.47***	9.80***	0.23***	25	27.51***	0.79***	20.25***	0.48***	22
	(2.128)	(0.052)	(2.134)	(0.043)		(1.808)	(0.051)	(2.030)	(0.046)	
2	15.79***	0.45***	12.37***	0.28***	16	20.24***	0.66***	17.98***	0.46***	16
	(1.239)	(0.026)	(2.129)	(0.042)		(0.979)	(0.036)	(1.485)	(0.042)	
3	10.40***	0.34***	5.65***	0.15***	29	17.56***	0.60***	13.15***	0.37***	21
	(0.826)	(0.024)	(1.440)	(0.032)		(0.840)	(0.033)	(1.217)	(0.038)	
4	10.79***	0.33***	6.51***	0.14***	31	16.03***	0.56***	10.46***	0.28***	30
	(1.296)	(0.030)	(1.810)	(0.042)		(1.084)	(0.041)	(1.511)	(0.047)	

Table 7. Supervisory Effort over Months 1–4. *Notes.* This table continues Table 7 in levels and logs for the 2008 Treatment Plantation, for Months 2–4 following the contract change. The table compares person-method observations which were below the standard in Month 0 of 2008 ("underperformers"), with those which were above the standard in that month ("overperformers"). Columns 2–5 compare outputs to that in Week 0, and columns 7–10 compare outputs to that in Month 0. Columns 6 and 11 display the estimated share of output increase S due to supervision, as given by equation (3). Standard errors are clustered by day. $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$.

This is about as far as we can go with our reduced-form analysis of the data. We can obtain some more insight by estimating the standard model off the pre-change data, and then using those estimates to predict the post-change distribution. As we shall see, by Month 4 the standard model does a good job in explaining the observed distribution of output.

6. A Structural Approach to Understanding Post-Contract Output

6.1. **Approach.** Recall the model presented in Section 2. We will now place further structure on that model, and estimate its parameters from the pre-change data. First, we suppose that u is linear; specifically, $u(w) = w$. This specification implies that we ignore the income effects from the change in the baseline wage, and concentrate entirely on the marginal incentives provided by the piece rates. Second, we write the cost function as

$$c(e, \mu) = \frac{\mu}{\theta}\left[\exp(\theta e) - 1\right], \tag{4}$$

where μ is a random shock observed by the worker before the effort decision is made, and θ is a parameter that controls the curvature of effort disutility. We presume that μ is drawn from a gamma distribution, the shape and scale parameters of which will form part of the estimation exercise. Next, we suppose that

$$L(x) = \max\{\beta x, 0\}, \tag{5}$$

where x is to be interpreted as the shortfall $s - e$ in effort from the contractually stipulated standard. Given the linearity of the utility function, we can fold β into the piece rate structure. Thus, under the old contract, workers face an effective piece rate of $w_0 \equiv \beta + 0.40$ up to the minimum standard $q_1 = s$, where Rs. 0.40 is the per-kilo monetary penalty for dropping below s. The next piece rate

is $w_1 = 0.40$, defined on $[s, q_2)$, followed by $w_3 = 0.55$ on $[q_2, q_3)$ and $w_4 = 0.85$ on $[q_3, \infty)$. The combined effective piece rates and thresholds are summarized for the relevant yield classes (which are 2 and 3, yield classes 1 and 4 not being observed) in Table 8. To complete the description of the wage function, we must recall the baseline wage, which is given by Rs. 77.55. This fixed wage is now effectively $77.55 - (0.40 + \beta)s$. Call the full wage function described here $w(y)$.

	Start at q_0		**Standard**, $q_1 = s$		**Threshold 2**, q_2		**Threshold 3**, q_3	
	Hands	Shears	Hands	Shears	Hands	Shears	Hands	Shears
Yield Class 2	0	0	23	28	34	39	50	55
Yield Class 3	0	0	28	33	44	49	59	64
Piece Rate/Loss (Rs.)	0.40+β →		0.40 →		0.55 →		0.85 →	

Table 8. Incentive Slabs for Yield Classes 2 and 3, Hands and Shears Under the Old Contract. *Notes.* This table describes the piece rates for yield classes 2 and 3 under the old contract, and thresholds at which they became active. (Yield classes 1 and 4 were not observed on any estates.) Each threshold q_i is four numbers, one for each yield class and plucking method. Effective piece rates for each of the incentive slabs, with rounding to the nearest kilogram, are defined on $[q_k, q_{k+1} - 1]$ for $k = 0, 1, 2$ and $[q_3, \infty)$. The slab $[q_0, q_1 - 1] = [0, s - 1]$ has an effective piece rate of $0.40 + \beta$, where β is the coefficient on the loss function below the standard.

Our goal, then, is to estimate four parameters: β (the loss below minimum standard), θ (the curvature of the disutility of effort), and the two parameters that characterize the gamma distribution for the random shock μ. These are the minimum degrees of freedom that are needed to fit the data: in particular, the random shock μ captures both a description of individual types as well as any additional idiosyncratic shocks that an individual might experience before she chooses her effort. Individual heterogeneity is unavoidable in confronting the data that we have.

With the redefinition of the wage function in place, we may rewrite the individual's effort choice problem: for each realization μ, choose $y = e$ to maximize

$$w(y) - \frac{\mu}{\theta}\left[\exp(\theta y) - 1\right]$$

The data only has positive output realizations for all participants. Therefore, if we accept the above description of the maximization problem, a necessary (but not sufficient) first order condition must hold for every output not exactly at the minimum standard or at any of the other thresholds:

$$w_i = \mu \exp(\theta y), \tag{6}$$

whenever y lies in the slab (q_i, q_{i+1}), where we set $q_4 = \infty$, and where the w_i's are defined as above. This is the first step that informs the estimation: data on individual output y_{it} permit us to back out which effective piece rate w_i applies to each individual, except for the value of β. But there is a potential identification problem which we must address at this stage.

The problem has to do with the interchangeability of θ and μ. Given any observed distribution of outputs, we can "explain" that distribution by choosing a relatively narrow distribution for μ and a correspondingly small value of θ which imparts enough elasticity in effort responses to traverse the space of observations. But we can, if we wish, choose a more diffused distribution for μ, provided that

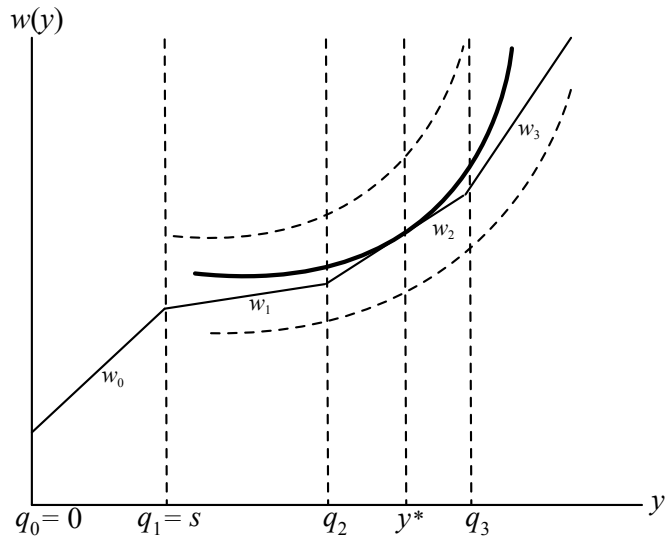


Figure 12. The Individual Optimization Problem. *Notes.* This figure depicts the optimal choice of y in the individual optimization problem. It shows that apart from a necessary first-order condition, a global condition is required to prevent movement to a different incentive slab. This additional condition provides restrictions on the curvature of the disutility of effort which assists in the identification of the parameter θ.

we increase the disutility curvature by raising θ. This is the sense in which the first-order condition alone may fail to pin down both θ and the distribution of μ; note how μ and $\exp\theta$ are multiplicatively related in (6).

Two restrictions, one conceptual and the other parametric, help us avoid this identification problem. The conceptual restriction is generated by an additional condition for optimality, for this is a non-concave problem. The payoff function must have enough curvature to avoid "global deviations" to other incentive slabs even when the first-order condition (6) on a particular slab applies. Figure 12 explains this. The first-order condition is met at y^*, but there must still be enough curvature so that the indifference curve that is tangent to the wage function at y^* lies everywhere above the wage function elsewhere: this imposes a *separate* restriction on curvature θ which cannot be "exchanged" with μ. In addition, we impose the parametric restriction that μ comes from a gamma distribution. This distribution is both extremely flexible – much more so than the more commonly used (log) normal family – and has the convenient property that it prevents μ and θ from freely "substituting" for each other in the cost function is as described in (4), thereby facilitating identification of each parameter separately.[11] Once θ and the distribution of μ are pinned down, β can be backed out easily by examining the distribution of observed outputs conditional on being below the minimum standard s.

6.2. **Details of the Estimation Exercise.** Here are the details of the numerical exercise that obtains estimates for the parameters of interest.

[11]The beta distribution also has this convenient property, but its range is limited to the unit interval. In the supplemental appendix, we replicate the structural estimation exercise using the beta distribution and show that the basic results are qualitatively identical.

1. *Estimate* μ. Fix θ and β. For each possible contract, characterized by a yield class and plucking method combination $k = 1, \ldots, 4$, use equation (6) to estimate the mean and variance of $\mu(j,k)$, where j indexes a person. Use these estimates to calculate the scale and shape parameters for up to 4 different gamma distributions — one each for the two yield classes and plucking methods — separately for each of the roughly 2000 workers in the sample.

2. *Simulate 2007 output.* For each individual j and each of the four class-method configurations k, randomly draw a $\mu(j,k)$ from a gamma distribution with parameters calculated from Step 1. With β fixed from Step 1, we know each of the effective piece rates w_i, $i = 1, \ldots, 4$. For each of these piece rates w_i, calculate the level of output y_{ij} that solves equation (6). Check to see if the y^*_{ij} is feasible, in the sense that it lies in the output slab $[q_i, q_{i+1})$ corresponding to i. If so, keep the value, otherwise return a missing value. For non-missing values, y^*_{ij} calculate the value of utility $U_{ij}(y^*_{ij})$. If all U_{ij}'s are negative, set optimal output $y^*_j = 0$. Otherwise, the optimum level of output y^*_j is that value of y_{ij} corresponding to the highest (positive) utility.

3. *Iterate and average.* Repeat Step 2 for 50 random draws of $\mu(j,k)$, and calculate the average optimal output over these 50 draws for each individual j under each contract k. Calculate the proportion of days in 2007 that j spent under each contract k. Apply these weights to the average individual output under each contract to calculate a weighted average output.

4. *Choose best fit.* Repeat Steps 1–3 for 200 possible values of θ and β corresponding to the 0.1 point grid $(\theta, \beta) = \{(0.1, 0.1), ...(2, 1.0)\}$. Chose (θ, β) such that the empirical distribution of the simulated data most closely resembles the actual 2007 data according to two criteria: a t-test of the difference in means and the Kolmogorov-Smirnoff distance statistic $D = \sup_y |F(y) - G(y)|$, which is the supremum of the absolute distances between the actual cumulative output distribution $F(\cdot)$ and the simulated cumulative output distribution $G(\cdot)$. Table A2 (see Appendix) reports the difference between the simulated and actual 2007 output corresponding to the value of θ and β in the row and column headings. This difference is minimized at $(\theta, \beta) = (0.9, 1.0)$. Table A3 (see Appendix) reports the corresponding p-values for a t-test whose null hypothesis is that the means are equal. Only for $\theta = 0.9$ are these p-values different from zero, and it reaches its maximum at $(\theta, \beta) = (0.9, 1.0)$. (The p-value falls for larger values of β with $\theta = 0.9$; not reported.) Table A4 (see Appendix) shows the Kolmogorov-Smirnoff distance statistic. It reaches its lowest values at $\theta = 0.9$ and $\theta = 1.0$.

In our structural prediction, we use $(\theta, \beta) = (0.9, 1)$, since satisfies both criteria. This can be seen in Figure A12 in the Appendix, which presents a graphical description of the D-statistic and the mean differences in output – both of which are closest to 0 at $(\theta, \beta) = (0.9, 1)$.

Figure 13 compares the output our structural model predicts under the old contract for $(\theta, \beta) = (0.9, 1)$ with the actual 2007 data. Table 9 presents summary statistics comparing the actual (row 1) data to the simulated data for the three values of (θ, β) for which we cannot reject the null hypothesis of equal means at the 10% level. In each case, the simulated and actual distributions are very similar.

6.3. **Predicting Output Post-Change.** The contract change generates a new effective wage function. Table 10 summarizes it.

Using the estimate for (θ, β) and the distributions for μ in the previous exercise, we predict the distribution of output after the contract change, according to the structural model. Specifically, we set $(\theta, \beta) = (0.9, 1.0)$ and, using the estimates of the gamma distribution for each individual j under each

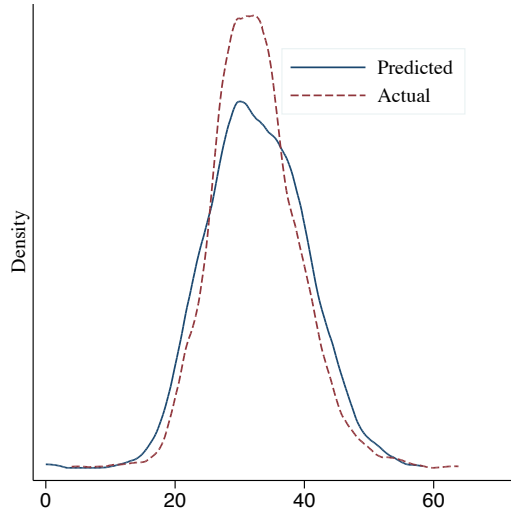


Figure 13. 2007 Kernel Density: Predicted and Actual. This figure plots the actual distribution of output (in kg.) in the treatment plantation in 2007, and compares it to the predicted distribution with $(\theta, \beta) = (0.9, 1)$.

	Mean	Median	SD	Skewness	Kurtosis	Interquartile Range
Actual	32.39	32.00	6.61	0.35	3.89	8.43
$(0.9, 0.8)$	32.69	32.49	7.56	0.03	3.28	10.35
$(0.9, 0.9)$	32.72	32.48	7.50	-0.02	3.17	10.52
$(0.9, 1.0)$	32.65	32.36	7.54	0.03	3.27	10.39

Table 9. Comparison of summary statistics from Simulated and Actual 2007 data. *Notes.* Rows 2-4 denote sample statistics for the simulated data with corresponding values of (θ, β).

contract k calculated earlier, take 100 random draws of $\mu(j, k)$. For each random draw, we calculate optimal output y_j^* as described in Step 2 of the previous section, but under the parameters of the new contract. We then calculate the average individual output over the 100 draws under each possible contract and construct a weighted average across contracts, where the assigned weights are equal the proportion of days the individual spent under each contract following the contract change. This weighted average is our prediction of individual output under the new contract, based on our earlier structural estimates.

Figure 14 compares this predicted distribution under the new contract to the month-by-month actual distribution under the new contract. By Month 4, we see almost complete convergence to the predicted distribution.

Table 11 indicates that one month after the contract change, output is 21 kg. higher than predicted by the structural model. This difference drops to 12 kg. in Month 2 and 7 kg. in Month 3. In the final week, the structural model underpredicts the actual output average by just 3 kg.

	Start at q'_0		**Standard**, $q'_1 = s'$		**Threshold 2**, q'_2		**Threshold 3**, q'_3	
	Hands	Shears	Hands	Shears	Hands	Shears	Hands	Shears
Yield Class 2	0	0	22	28	36	43	52	59
Yield Class 3	0	0	27	33	46	53	61	68
Piece Rate/Loss (Rs.)	$\beta \rightarrow$		$0.40 \rightarrow$		$0.55 \rightarrow$		$0.85 \rightarrow$	

Table 10. Incentive Slabs for Yield Classes 2 and 3, Hands and Shears Under the New Contract. *Notes.* This table describes the piece rates for yield classes 2 and 3 under the new contract, and thresholds at which they became active. (Yield classes 1 and 4 were not observed on any estates.) Each threshold $\hat{q}_i$ is four numbers, one for each yield class and plucking method. Effective piece rates for each of the incentive slabs, with rounding to the nearest kilogram, are defined on $[\hat{q}_k, \hat{q}_{k+1} - 1]$ for $k = 0, 1, 2$ and $[\hat{q}_3, \infty)$. The slab $[\hat{q}_0, \hat{q}_1 - 1] = [0, \hat{s} - 1]$ has an effective piece rate of β, where β is the estimated coefficient on the loss function below the new standard $\hat{s}$. There are no longer any monetary penalties in this range.

	Month 1	Month 2	Month 3	Month 4	Week 49	Week 50	Week 51	Week 52
				$(\theta, \beta) = (0.9, 1)$				
Difference	21.448***	12.394***	7.034***	6.168***	6.207***	8.766***	6.428***	2.788***
	(0.220)	(0.204)	(0.199)	(0.200)	(0.256)	(0.256)	(0.249)	(0.239)

Table 11. Difference in Actual and Predicted Output After Contract Change. *Notes.* This table depicts the difference between mean actual output and mean predicted output from the structural model with parameters $(\theta, \beta) = (0.9, 1.0)$. Robust standard errors in parentheses. $^{***}p < 0.01$.

In addition, we can look at the full time series using a regression discontinuity design with the (rounded) optimal bin sizes calculated earlier. The results are qualitatively similar to those in Table 11. Figure 15 shows the deviation of actual output from the mean predicted output of the structural model, in terms of 3-day averages (i.e. for a bin size of 3). There is a sharp initial increase in Month 1 following the contract change, and then a gradual tapering down, which seems to converge to around 7 kg. deviation in Month 3, but then dips down again in the last 2 weeks of observations. The results are qualitatively similar when we use bins of size 5; see Figure A13 in the Appendix.

Our structural exercise underlines and supports the earlier observation that an output response, while initially quite dramatic, fades away in the succeeding months. What this exercise adds to the previous observations is that a standard and parsimonious static model of incentives does a remarkably good job in predicting output following the contract change, once a few months have passed. It does underperform observed output just after the contract, though, so this is not at all to say that the initial, sizable jump in output is uninteresting. It is in fact particularly interesting from a behavioral perspective, as we have argued. Nevertheless, our longer-term analysis places that initial jump in context: behavioral responses fade away as the standard economics of moral hazard appears to reassert its dominant position. It is in this sense that our study warns against placing excessive emphasis on "behavioral responses," a response that may well be significant in the immediate aftermath of a contract or policy change, but not in the longer term.

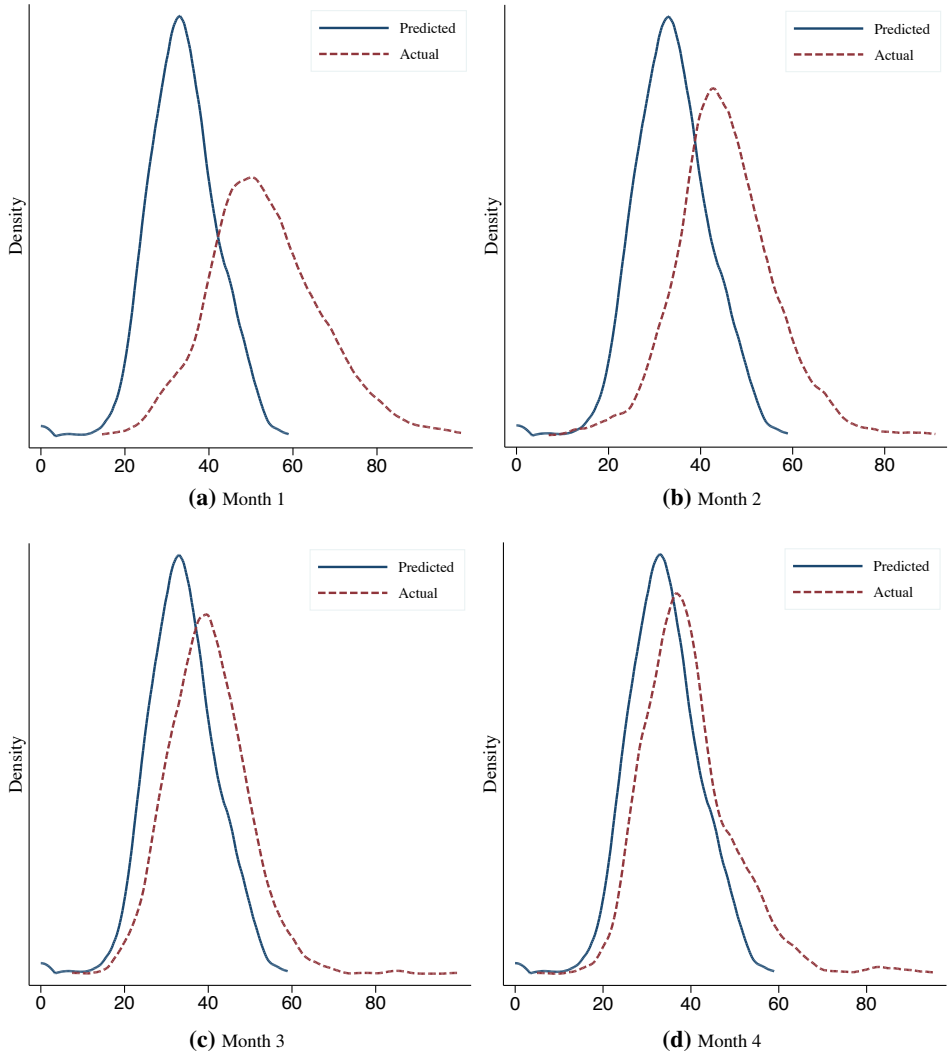


Figure 14. Kernel Density: Actual and Predicted Daily Outputs. *Notes.* The solid line in each panel is the kernel density of average daily outputs (in kg.) in the treatment plantation in 2008, as predicted by the structural model under the new contract with parameters $(\theta, \beta) = (0.9, 1.0)$. The dashed line in each panel is the kernel density of actual average daily outputs (in kg.) in the treatment plantation in 2008 for each of the four months after the contract change.

7. Conclusion

This paper studies the productivity impact of a contract change for tea pluckers in an Indian plantation. The contract raised the baseline wage by over 30%, but lowered marginal incentives by shifting the existing piece rates to higher minimum thresholds, and by eliminating a linear penalty for underperformance. In the one month following the contract change, there is a dramatic increase in productivity. It is robust to all sorts of controls, including plucking method and field type. This is

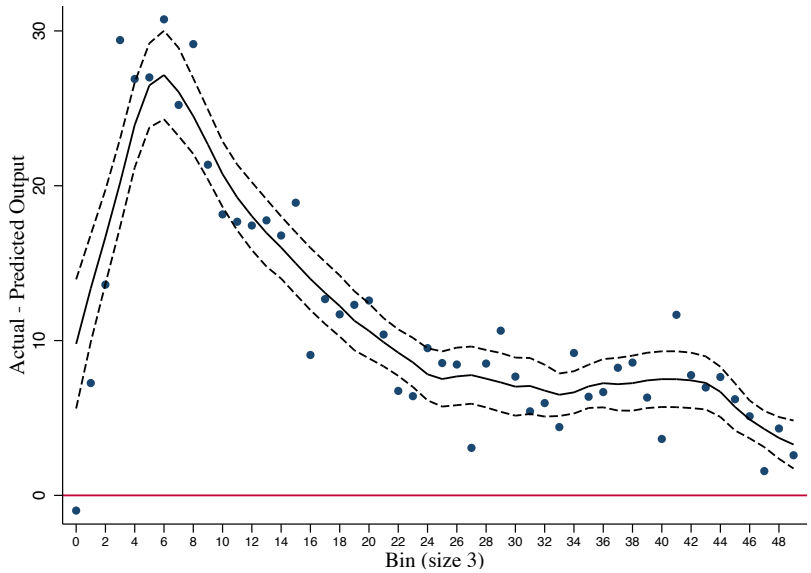


Figure 15. Deviation of Actual Output from Predicted mean: Bin size 3. *Notes.* This figure depicts the difference between actual average output and predicted average output under the new contract. Each dot denotes a 3-day average of actual output (bin size of 3). The figure is overlaid with a local polynomial smoothed using an Epanechnikov kernel (solid curve). Dashed curves denote the 95% confidence interval. Predicted average output is calculated as described in the text, with parameter values $(\theta, \beta) = (0.9, 1)$

a surprising result, and appears to directly contradict the predictions from a standard principal-agent model based on agent moral hazard.

The first part of the paper documents this output increase, and attempts to understand it. We describe a standard model, and argue that — if anything — there should be a *decline* in output. But there are three other possibilities that we attempt to account for. The first is that the contract change represented a large payoff increase to the workers, and therefore could have served as a dynamic efficiency wage, with contract termination acting as a now stronger threat. The data do not agree with this hypothesis for two reasons. First, the workers are, in the main, permanent and *cannot* be fired by law. Second, while there are indeed temporary workers, there is no difference at all in their output response as compared to permanent workers. In fact, the latter exhibit a somewhat more pronounced response to the contract change.

The second hypothesis is that managers stepped up their supervisory input, essentially replacing monetary incentives by exhortations, encouragement, coercion and threat (though as we've argued, these last two items are limited in scope and they would have had different impacts on temporary versus permanent workers). We can address this question by studying the subset of workers who fell under the minimum standard before the contract change, and see whether these "underperformers" increase their output more sharply than the "overperformers" – those who were performing above the standard. Here we do perceive a difference. The percentage increase in the output of underperformers does exceed that of the overperformers, so it seems likely that heightened supervision has some role to play. We do estimate that supervision accounts for between a quarter and a third of the output increase, but a large fraction of the increase is still unaccounted for.

That leaves the third hypothesis, which — as far as this paper is concerned — is essentially a black box: that there was a "behavioral response" of some kind. Perhaps workers felt the new contract was a form of "gift exchange" (Akerlof (1982), Fehr, Goette, and Zehnder (2009)) or was otherwise "fair" (Akerlof and Yellen (1990)), or perhaps raised morale (Solow (1979)). While we do not test these hypotheses directly, it is clear from a process of elimination that there was a significant "behavioral response."

However, the second part of the paper documents a reversal. In subsequent months the initial increase in output is comprehensively eroded. Four months after the contract change, the increase is muted and in the last two weeks, it appears to vanish altogether. At any rate, the increase is eroded enough so that an entirely standard model with no behavioral or dynamic features that we estimate off the *pre-change* data, fits the observations four months after the contract change remarkably well.

These findings speak to the importance of examining responses to a policy change, not just immediately after the change but for a substantive period of time afterward. Our study suggests that classical monetary incentives ultimately dominate, despite a possibly "behavioral" response in the shorter term. More generally, our findings speak to a literature in behavioral economics that emphasize both the interaction between "intrinsic" and "extrinsic" motivations, as well as the dynamic evolution of those motivations following a policy change; see Gneezy and Rustichini (2000), Ashraf, Bandiera, and Jack (2012) and Gneezy, Meier, and Rey-Biel (2011). This literature emphasizes how the introduction of the "extrinsic" (financial incentives) might erode the "intrinsic" (reciprocity, gratitude or fair play).

In this paper, "extrinsic" and "intrinsic" are switched: the baseline relationship is an employment contract, so that the intrinsic incentives are monetary to begin with, and gratitude, reciprocity and pro-social behavior are extrinsic. Does the extrinsic erode the intrinsic when their roles are reversed? It would appear not: extrinsic motivations do appear to hold sway in the short run, but in this case, the invasion of one motive by the other is far from permanent. The more depressing conclusion — depressing, at least, to these authors — appears to be that it is not a case of extrinsic vs intrinsic so much as the ultimate dominance of monetary incentives over their non-pecuniary counterparts. This is not to argue that agents are never driven by notions of the social good, or that loyalty to an employer cannot be built up. But we need to be alert to longer-term effects, and in particular to the hurried classification of many important economic phenomena as fundamentally "behavioral."

REFERENCES

Ackerberg, Daniel and Maristella Botticini. 2002. "Endogenous matching and the empirical determinants of contract form." *Journal of Political Economy* 110 (3):564–591.

Akerlof, G.A. 1982. "Labor contracts as a partial gift exchange." *The Quarterly Journal of Economics* 97 (4):543–569.

Akerlof, George and Janet Yellen. 1990. "The fair wage-effort hypothesis and unemployment." *The Quarterly Journal of Economics* 105 (2):255–283.

Ashraf, Nava, Oriana Bandiera, and Kelsey Jack. 2012. "No margin, no mission? A field experiment on incentives for pro-social tasks." CEPR Discussion Paper No. DP8834.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2005. "Social preferences and the response to incentives: Evidence from personnel data." *Quarterly Journal of Economics* 120 (3):917–962.

———. 2007. "Incentives for managers and inequality among workers: Evidence from a firm-Level experiment." *Quarterly Journal of Economics* 122 (2):729–773.

———. 2009. "Social connections and incentives in the workplace: Evidence from personnel data." *Econometrica* 77 (4):1047–1094.

———. 2010. "Social incentives in the workplace." *The Review of Economic Studies* 77 (2):417–458.

Bellemare, Charles and Bruce Shearer. 2009. "Gift giving and worker productivity: Evidence from a firm-level experiment." *Games and Economic Behavior* 67 (1):233–244.

Bénabou, Roland and Jean Tirole. 2003. "Intrinsic and extrinsic motivation." *Review of Economic Studies* 70:489–520.

Besley, Timothy and Maitreesh Ghatak. 2005. "Competition and incentives with motivated agents." *American Economic Review* 95 (3):616–636.

Dutta, Bhaskar, Debraj Ray, and Kunal Sengupta. 1989. "Repeated principal-agent games with eviction." In *The Economic Theory of Agrarian Institutions*, edited by P. Bardhan. Clarendon Press, Oxford.

Fehr, Ernst and Lorenz Goette. 2007. "Do workers work more when wages are High? Evidence from a randomized field experiment." *American Economic Review* 97 (1).

Fehr, Ernst, Lorenz Goette, and Christian Zehnder. 2009. "A behavioral account of the labor market: The role of fairness concerns." *Annual Review of Economics* 1:355–384.

Frey, Bruno S and Reto Jegen. 2001. "Motivation crowding theory." *Journal of economic surveys* 15 (5):589–611.

Gibbons, Robert. 1987. "Piece-rate incentive schemes." *Journal of Labor Economics* 5 (4):413–429.

Gibbons, Robert and Kevin Murphy. 1992. "Optimal incentive contracts in the presence of career concerns: theory and evidence." *Journal of Political Economy* 100:468–505.

Gneezy, Uri and John List. 2006. "Putting Behavioral economics to work: Testing for Gift exchange in labor markets using field experiments." *Econometrica* 74 (5):1365–1384.

Gneezy, Uri, Stephan Meier, and Pedro Rey-Biel. 2011. "When and why incentives (don't) work to modify behavior." *The Journal of Economic Perspectives* 25 (4):191–209.

Gneezy, Uri and Aldo Rustichini. 2000. "Pay enough or don't pay at all." *Quarterly Journal of Economics* 115 (3):791–810.

Holstrom, Bengt and Paul Milgrom. 1991. "Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design." *Journal of Law, Economics, and Organization* 7:24–52.

Imbens, Guido and Karthik Kalyanaraman. 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *The Review of Economic Studies* 79 (3):933–959.

Kőszegi, Botond. 2013. "Behavioral Contract Theory." *Journal of Economic Literature*. Forthcoming.

Lacetera, Nicola, Mario Macis, and Robert Slonim. 2012. "Will There Be Blood? Incentives and Displacement Effects in Pro-social Behavior." *American Economic Journal: Economic Policy* 4 (1):186–223.

Lazear, Edward. 2000. "Performance pay and productivity." *American Economic Review* 90 (5):1346–1361.

Shaban, Radwan Ali. 1987. "Testing between competing models of sharecropping." *The Journal of Political Economy* 95 (5):893–920.

Shapiro, Carl and Joseph Stiglitz. 1984. "Equilibrium unemployment as a worker discipline device." *The American Economic Review* 74 (3):433–444.

Shearer, Bruce. 2004. "Piece pates, fixed wages and incentives: Evidence from a field experiment." *Review of Economic Studies* 71 (2):514–534.

Solow, R.M. 1979. "Another possible source of wage stickiness." *Journal of Macroeconomics* 1 (1):79–82.

Topkis, Donald. 1998. *Supermodularity and complementarity*. Princeton University Press, Princeton, N.J.

APPENDIX

Proof of Proposition 1. In this appendix, we prove Proposition 1 for a very general family of utility functions to which the separable payoff function of Section 4.1 belongs. We denote by $U(w, e, s)$ the worker's utility, which is non-decreasing concave in wage w, concave in effort e and decreasing in the standard s. Utility U captures the consequence of not meeting the standard: as the standard increases, the pressure on the worker may increase and hence his utility decrease. We further assume that U is submodular in (w, e, s), i.e. the marginal utility of wages is non-increasing in effort and standard. In particular, payoff function $U(w, e, s) = u(w) - c(e) - L(s - e)$ of Section 4.1 satisfies all these conditions. In general, however, utility U does not need to be differentiable or separable.

Given a wage function $w(\cdot)$ and standard $s \geq 0$, the worker exercises an effort level that maximizes her utility. Multiple optimums are however possible and we need to consider $\mathcal{E}$ and $\hat{\mathcal{E}}$, the sets of optimal effort levels before and after the contract change, respectively, which are assumed to be non-empty.[12]

Proposition 1 states that either a worker decreases her effort with the contract change ($\hat{e} \leq e$) or she exerts effort levels that are optimal under both contracts ($\hat{e}, e \in \hat{\mathcal{E}} \cap \mathcal{E}$).[13] More formally,

Proposition 1. *If $\hat{e} \in \hat{\mathcal{E}}$ and $e \in \mathcal{E}$, then either $\hat{e}, e \in \hat{\mathcal{E}} \cap \mathcal{E}$ or $\hat{e} \leq e$.*

Proof: Take, $\hat{e} \in \hat{\mathcal{E}}$ and $e \in \mathcal{E}$ and assume $\hat{e}$ is such that $\hat{e} > e$ and define δ such that $\delta \equiv \hat{w}(\hat{e}) - w(\hat{e})$. It follows that,

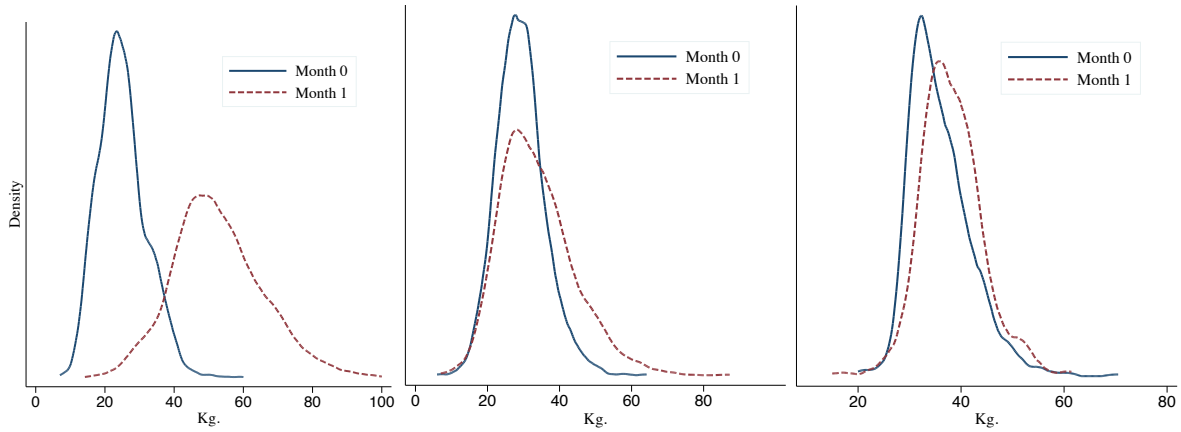
$$\begin{aligned}
U(\hat{w}(\hat{e}), \hat{e}, \hat{s}) &- U(\hat{w}(e), e, \hat{s}) \\
&= U(\hat{w}(\hat{e}), \hat{e}, \hat{s}) - U(\hat{w}(\hat{e}), e, \hat{s}) + U(\hat{w}(\hat{e}), e, \hat{s}) - U(\hat{w}(e), e, \hat{s}) \\
&\leq U(w(\hat{e}), \hat{e}, \hat{s}) - U(w(\hat{e}), e, \hat{s}) + U(\hat{w}(\hat{e}), e, \hat{s}) - U(\hat{w}(e), e, \hat{s}) \\
&\leq U(w(\hat{e}), \hat{e}, \hat{s}) - U(w(\hat{e}), e, \hat{s}) + U(w(\hat{e}), e, \hat{s}) - U(\hat{w}(e) - \delta, e, \hat{s}) \\
&= U(w(\hat{e}), \hat{e}, \hat{s}) - U(\hat{w}(e) - \delta, e, \hat{s}) \\
&\leq U(w(\hat{e}), \hat{e}, \hat{s}) - U(w(e), e, \hat{s}) \\
&\leq U(w(\hat{e}), \hat{e}, s) - U(w(e), e, s) \\
&\leq 0.
\end{aligned}$$

The first inequality holds since $U(\cdot, \cdot, \hat{s})$ is submodular with $\hat{w}(e) \geq w(\hat{e})$ from (C1). The second one holds since $U(\cdot, e)$ is concave with $\delta \geq 0$ from (C1). The third one holds since $U(\cdot, e, \hat{s})$ is non-decreasing, where $\hat{w}(e) - \delta \geq w(e)$ from (C2). The fourth one comes from the submodularity of $U(\cdot, \cdot, \cdot)$ in the three variables and (C3). The last inequality holds since $e \in \mathcal{E}$, and is strict if and only if $\hat{e} \notin \mathcal{E}$. Thus, either $\hat{e} \leq e$ or if $\hat{e} > e$ then $0 \leq U(\hat{w}(\hat{e}), \hat{e}, \hat{s}) - U(\hat{w}(e), e, \hat{s}) \leq U(w(\hat{e}), \hat{e}, s) - U(w(e), e, s) \leq 0$, which implies that $e, \hat{e} \in \hat{\mathcal{E}} \cap \mathcal{E}$. □

[12] This holds for instance when $\lim_{e \to \infty} U(w(e), e, s) = -\infty$.

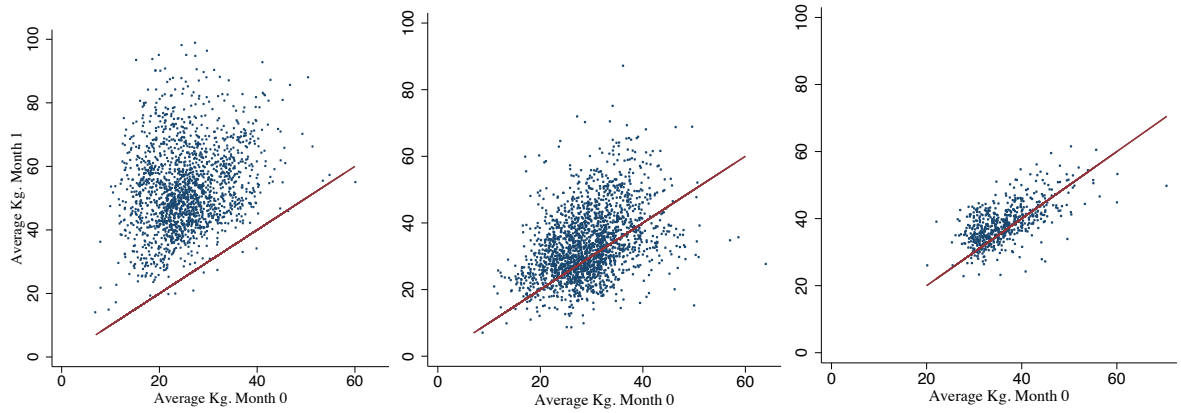
[13] More formally, the result states that $\hat{\mathcal{E}}$ is lower than $\mathcal{E}$ according to the induced set ordering; see Topkis (1998).

Additional Diagrams and Tables. We collect here some additional diagrams and tables referred to in the main text.



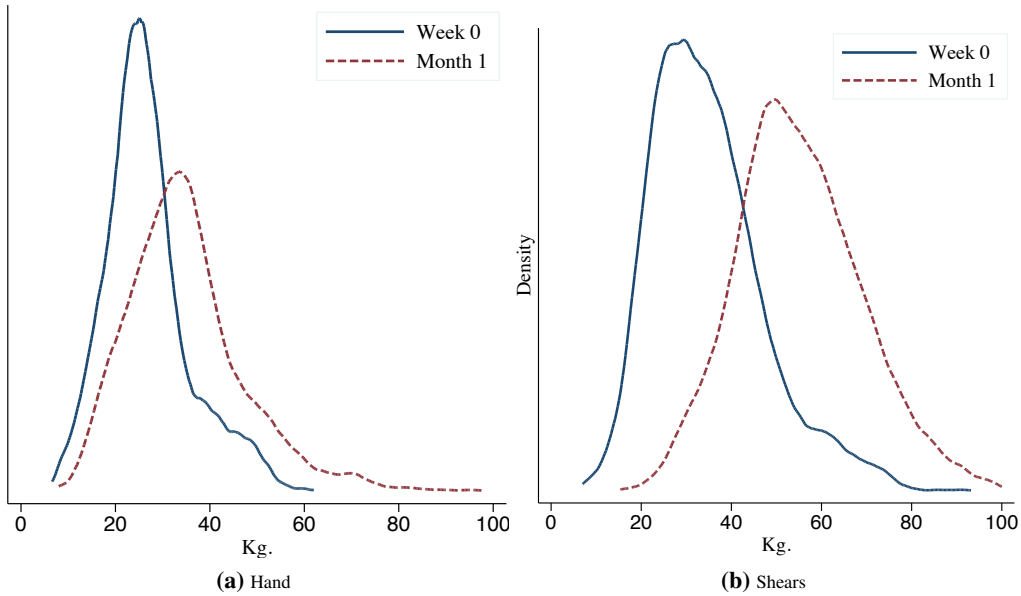
(a) 2008 Treatment Plantation **(b)** 2007 Treatment Plantation **(c)** 2008 Control Plantation

Figure A1. Kernel Density: Average Daily Output. *Notes.* This figure depicts kernel density estimates of workers' average daily output and is identical to Figure 3, except that it uses Month 0 instead of Week 0 as the baseline.



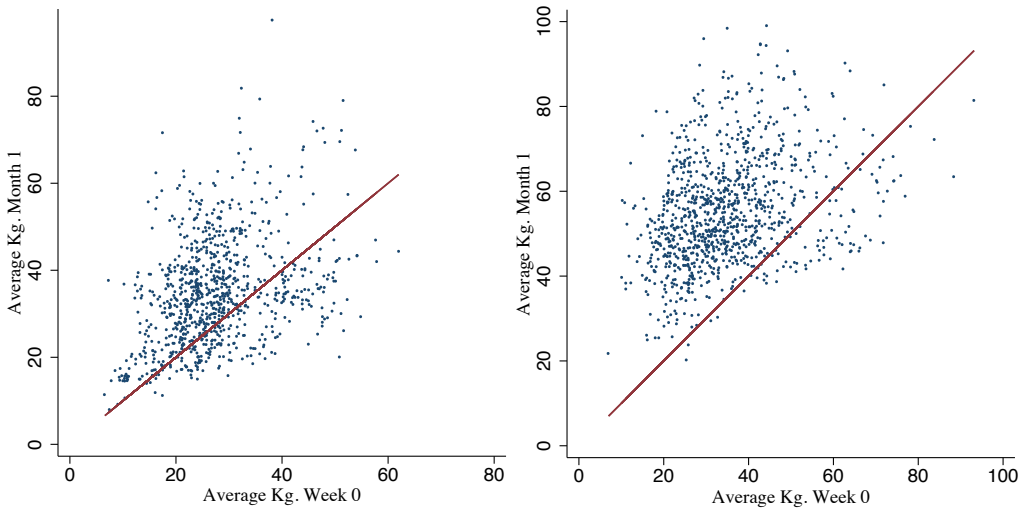
(a) 2008 Treatment Plantation **(b)** 2007 Treatment Plantation **(c)** 2008 Control Plantation

Figure A2. Scatter Plot: Average Daily Output. *Notes.* This scatter plot depicts workers' average daily output and is identical to Figure 4, except that it uses Month 0 instead of Week 0 as the baseline.



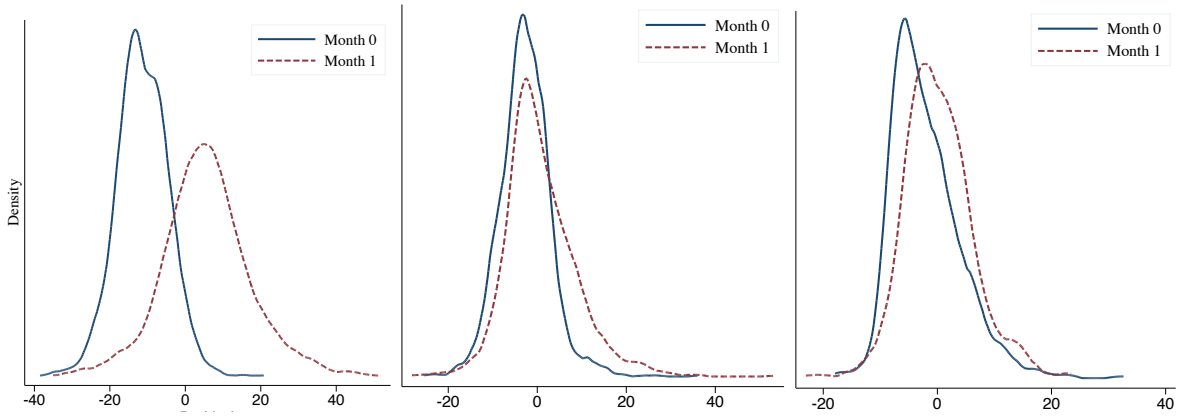
(a) Hand **(b)** Shears

Figure A3. Kernel Density: Average Daily Output Disaggregated by Hand and Shears. *Notes.* This figure depicts kernel density estimates of workers' average daily output analogous to Panel (a) in Figure 3, but disaggregated by hand (Panel a) and shears (Panel b).



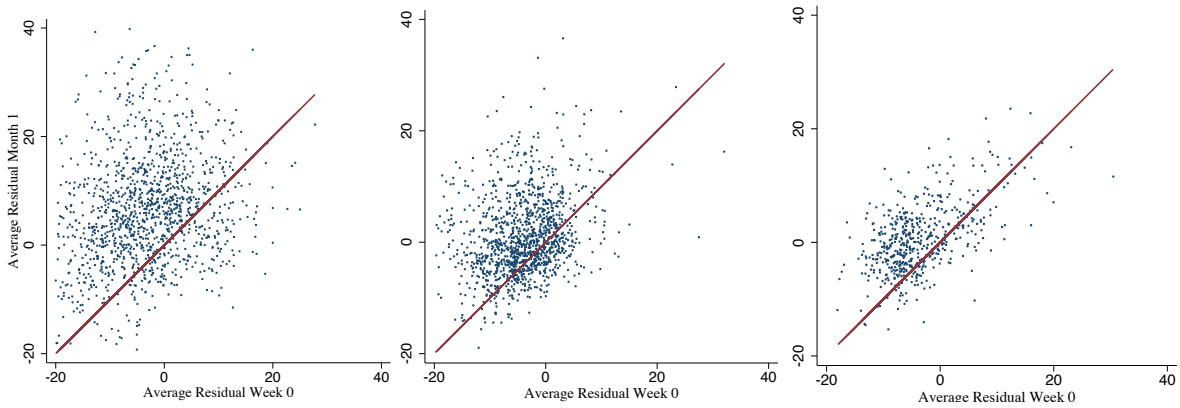
(a) Hand **(b)** Shears

Figure A4. Scatter Plot: Average Daily Output Disaggregated by Hand and Shears. *Notes.* This scatter plot depicts average daily output per worker analogous to Panel (a) in Figure 4, but disaggregated by hand (Panel a) and shears (Panel b).



(a) 2008 Treatment Plantation **(b)** 2007 Treatment Plantation **(c)** 2008 Control Plantation

Figure A5. Kernel Density: Average Daily Residual. *Notes.* This figure depicts kernel density estimates analogous to Figure 7, except that it uses Month 0 instead of Week 0 as the baseline.



(a) 2008 Treatment Plantation **(b)** 2007 Treatment Plantation **(c)** 2008 Control Plantation

Figure A6. Scatter Plot: Average Daily Residuals. *Notes.* This scatter plot is analogous to Figure 4, except that it plots average daily residuals rather than average daily output.

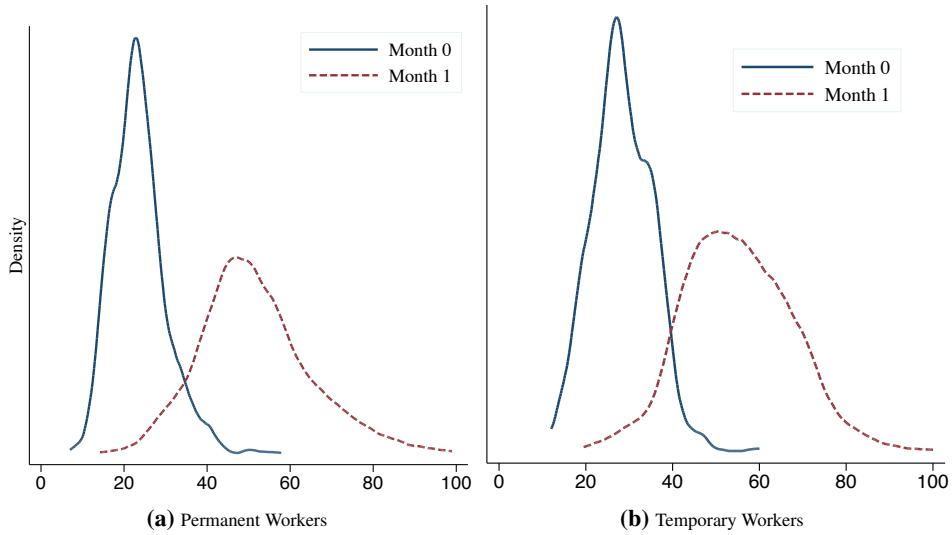


Figure A7. Kernel Density: Average Daily Output by Permanent and Temporary Workers. *Notes.* This figure is analogous to Figure 9, except that it uses Month 0 instead of Week 0 as the baseline.

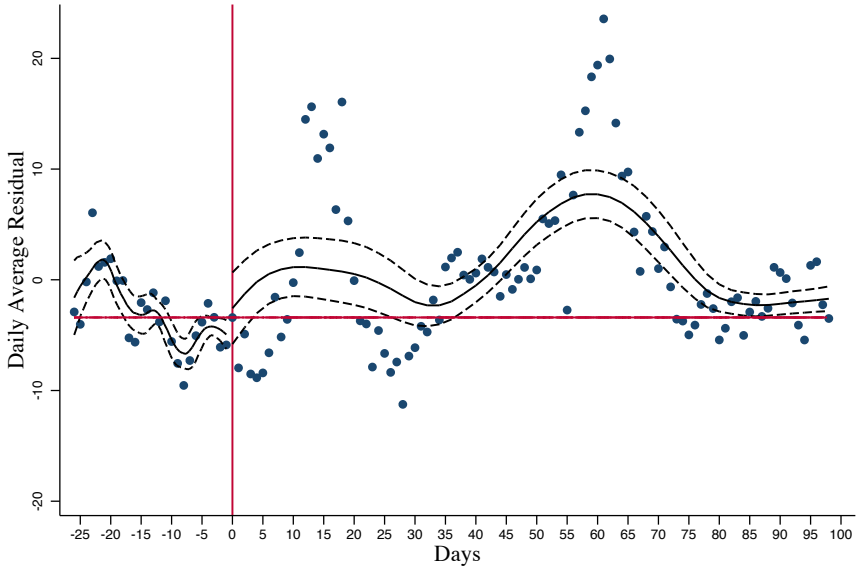


Figure A8. Time series: Average Daily Residual 2007. *Notes.* This figure is analogous to Figure 10, but for treatment plantation in 2007. The horizontal line denotes average output in Week 0 in the treatment plantation in 2008.

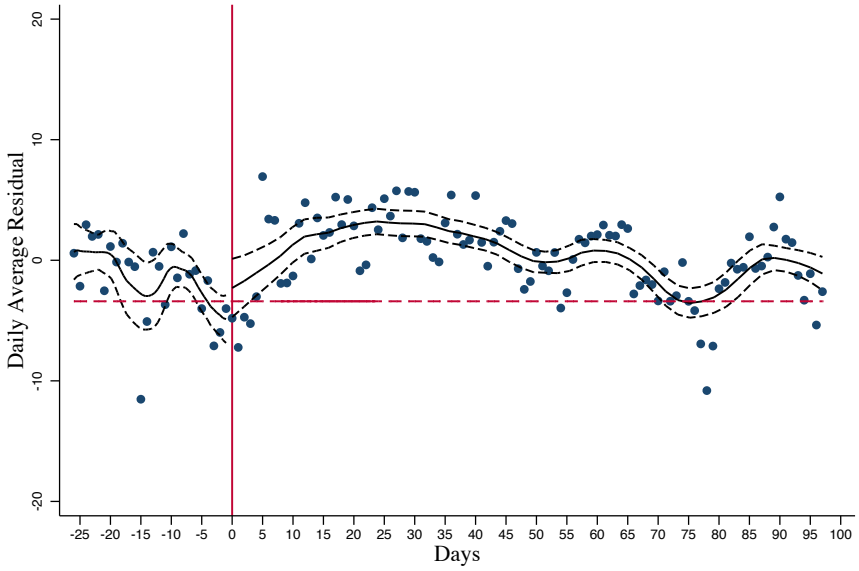
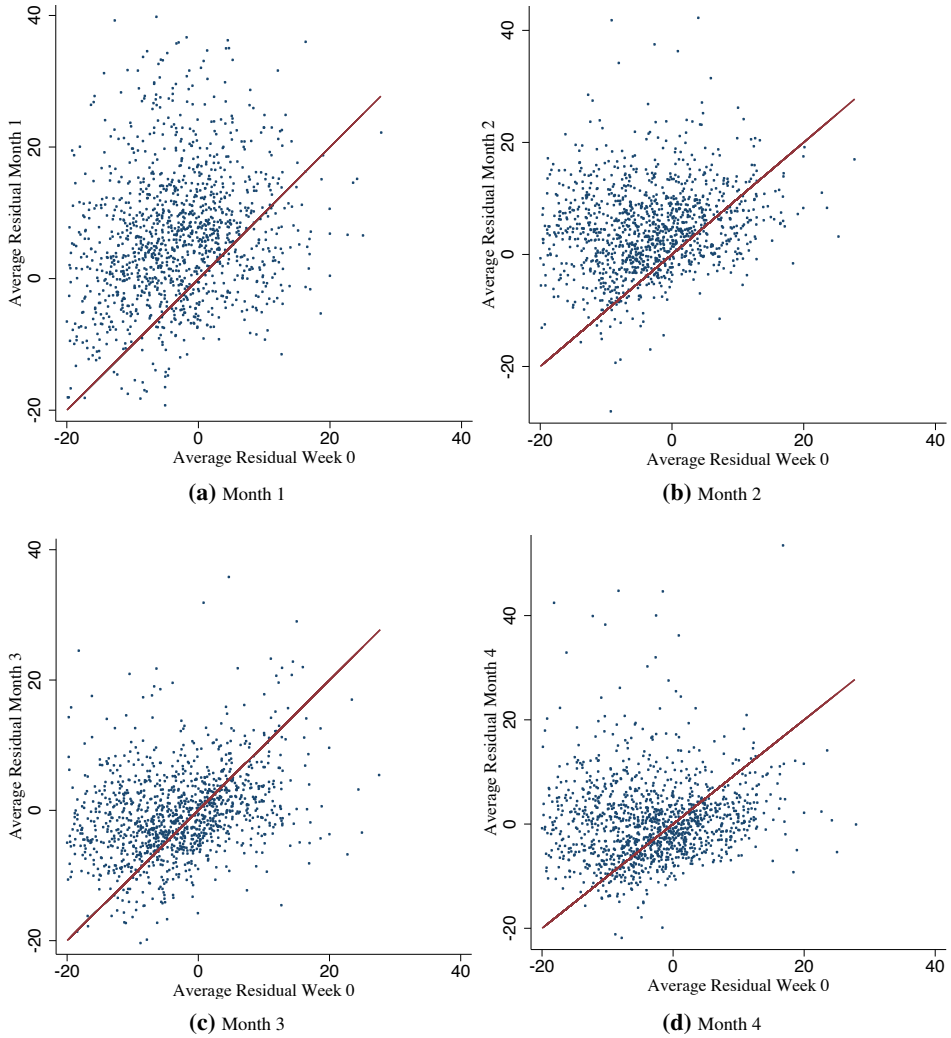


Figure A9. Time series: Average Daily Residual 2008 Control Plantation. *Notes.* This figure is analogous to Figure 10, but for the control plantation in 2008. Hence, the residuals here only account for rainfall. The horizontal line denotes average output in Week 0 in the treatment plantation in 2008.

	Levels		Logs		Levels		Logs	
	FD	DD	FD	DD	FD	DD	FD	DD
	Permanent Workers				Temporary Workers			
Month 1	13.45***	6.12**	0.34***	0.09	11.21***	6.73**	0.30***	0.16**
	(1.831)	(2.684)	(0.041)	(0.066)	(2.098)	(2.808)	(0.045)	(0.065)
Month 2	13.22***	7.03***	0.35***	0.13***	16.43***	6.95***	0.43***	0.16***
	(1.663)	(1.433)	(0.038)	(0.044)	(1.913)	(1.847)	(0.038)	(0.050)
Month 3	7.25***	-12.17***	0.22***	-0.32***	9.60***	-8.57***	0.28***	-0.19***
	(1.098)	(2.626)	(0.031)	(0.066)	(1.270)	(2.592)	(0.034)	(0.060)
Month 4	8.08***	5.26***	0.21***	0.10***	9.98***	3.11*	0.27***	0.05
	(1.390)	(1.105)	(0.031)	(0.038)	(1.891)	(1.714)	(0.041)	(0.047)
Week 49	6.99***	7.93***	0.16***	0.18***	16.16***	12.66***	0.44***	0.29***
	(2.129)	(1.544)	(0.038)	(0.051)	(3.038)	(2.103)	(0.074)	(0.066)
Week 50	9.30***	6.58***	0.23***	0.14***	15.87***	7.76***	0.37***	0.18**
	(1.246)	(1.222)	(0.026)	(0.039)	(1.884)	(2.403)	(0.043)	(0.067)
Week 51	7.27**	3.07**	0.14	0.02	6.47	4.11*	0.21	0.11*
	(2.606)	(1.459)	(0.103)	(0.047)	(6.222)	(2.339)	(0.158)	(0.066)
Week 52	1.33	2.42	0.04	0.04	-4.05	-0.20	0.01	0.00
	(3.580)	(1.876)	(0.101)	(0.068)	(5.970)	(2.441)	(0.157)	(0.074)

Table A1. Change in Monthly Average Output Relative to Week 0 by Permanent and Temporary Workers. *Notes.* This table provides the analog of Table 6, disaggregated by permanent and temporary workers. The counterfactual for the double difference estimates is the treatment plantation in 2007. Standard errors are clustered by day. $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$.



(a) Month 1 **(b)** Month 2

(c) Month 3 **(d)** Month 4

Figure A10. Scatter Plots: Average Daily Residuals. *Notes.* This figure plots average daily residuals in the treatment plantation in 2008, correcting for rainfall and other time-varying controls, in Week 0 (x-axis) Months 1-4 after (y-axis) the contract change.

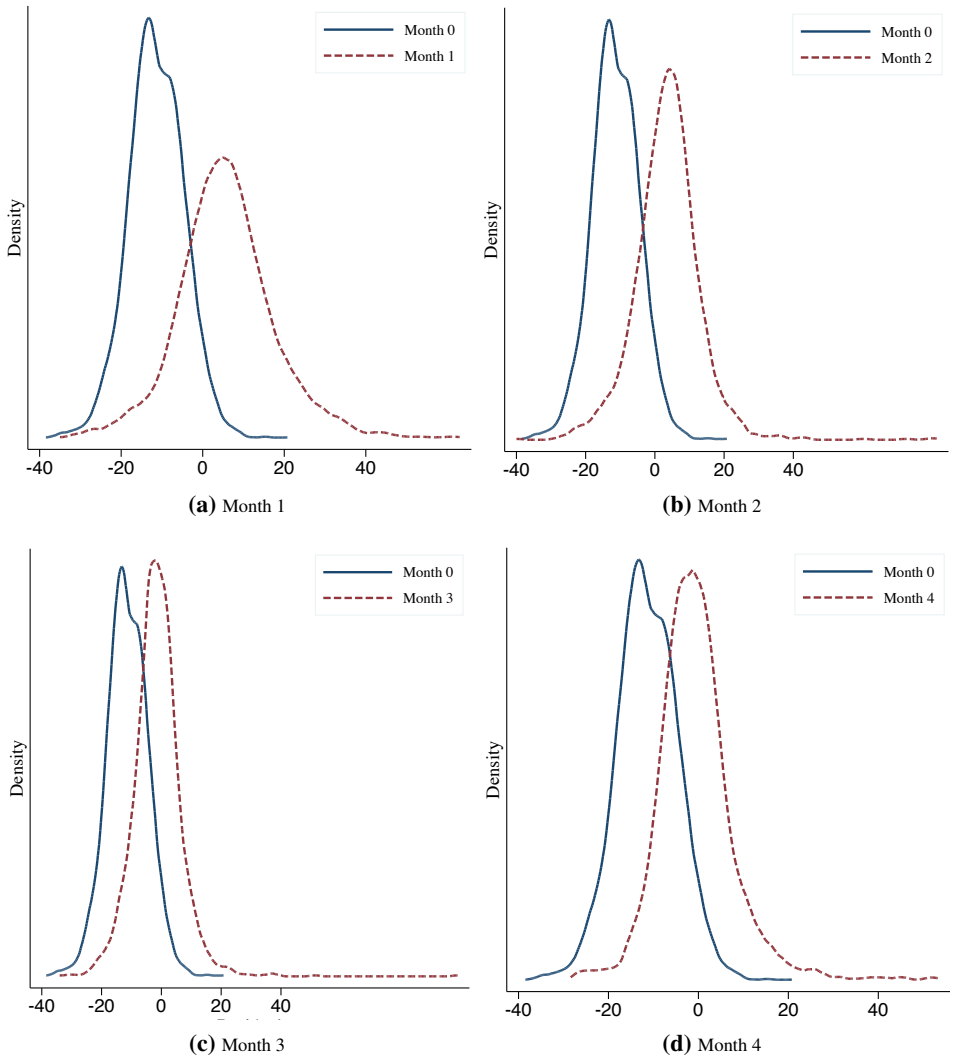


Figure A11. Kernel Density: Average Daily Residuals. *Notes.* This figure depicts kernel density estimates analogous to Figure 11, except that it uses Month 0 instead of Week 0 as the baseline.

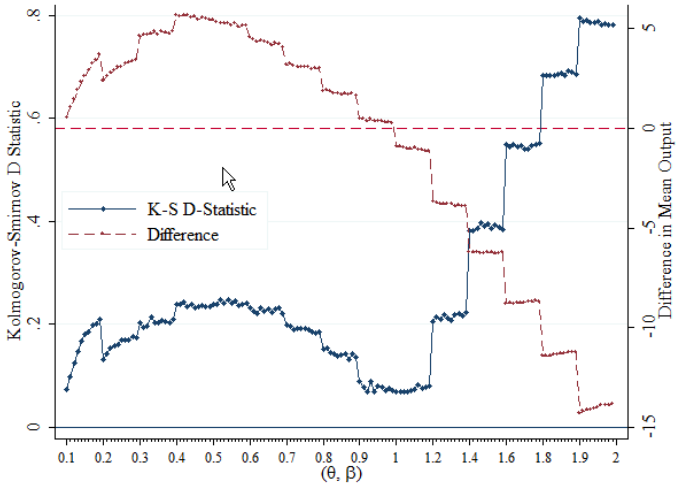


Figure A12. Kolmogorov-Smirnoff D Statistic and Difference between Predicted and Actual Mean Output by (θ, β) Combinations. *Notes.* This figure depicts two statistics testing the difference in predicted and actual output for 2007. The left y-axis denotes the Kolmogorov-Smirnoff distance statistic $D = \sup_y |F(y) - G(y)|$, which is the supremum of the absolute distances between the actual and simulated output distributions. A smaller D indicates more similar distributions. The right y-axis the difference in mean output, obviously minimized at zero (the dashed line). Each x-axis tick denotes a (θ, β) combination, where θ values are labeled on the big ticks and the associated β values of 0.1, 0.2, ...1.0 are marked by the 10 corresponding small ticks. For example, the first tick on the x-axis corresponds to $(\theta, \beta) = (0.1, 0.1)$, the second to $(\theta, \beta) = (0.1, 0.2)$, etc.

	β									
θ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.1	0.552	1.030	1.469	1.894	2.299	2.591	2.880	3.239	3.406	3.689
0.2	2.363	2.619	2.775	2.930	3.059	3.077	3.223	3.323	3.345	3.420
0.3	4.629	4.684	4.721	4.740	4.854	4.721	4.855	4.789	4.725	4.847
0.4	5.666	5.641	5.668	5.664	5.552	5.601	5.456	5.530	5.476	5.427
0.5	5.436	5.330	5.296	5.279	5.284	5.141	5.253	5.079	5.156	5.180
0.6	4.582	4.459	4.322	4.422	4.367	4.312	4.176	4.275	4.219	4.058
0.7	3.199	3.245	3.157	3.063	3.090	3.092	3.059	2.990	3.006	3.013
0.8	1.877	1.908	1.849	1.767	1.722	1.699	1.726	1.712	1.724	1.629
0.9	0.501	0.464	0.355	0.471	0.366	0.355	0.339	0.293	0.325	0.257
1.0	-0.923	-0.928	-0.960	-1.045	-1.020	-0.978	-1.103	-1.081	-1.136	-1.134
1.2	-3.668	-3.756	-3.782	-3.803	-3.793	-3.816	-3.900	-3.871	-3.909	-3.894
1.4	-6.219	-6.208	-6.200	-6.248	-6.248	-6.237	-6.232	-6.292	-6.274	-6.234
1.6	-8.794	-8.780	-8.791	-8.769	-8.766	-8.734	-8.696	-8.702	-8.641	-8.690
1.8	-11.420	-11.400	-11.412	-11.370	-11.332	-11.295	-11.274	-11.263	-11.223	-11.212
2.0	-14.289	-14.190	-14.131	-14.101	-14.051	-13.992	-13.874	-13.906	-13.895	-13.839

Table A2. Difference in Mean Output: Predicted and Actual 2007 Distribution. *Notes.* This table contains the difference between the mean of actual output and the mean of the predicted output in 2007. Each cell presents this difference for a different (θ, β) combination for which the simulation was run. The mean of the predicted output pertains to the weighted average of output for each individual for 50 replications for each of the 4 possible contracts (hand, shears and yield classes 2 and 3).

	β									
θ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.1	0.009	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.7	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.9	0.012	0.021	0.074	0.019	0.069	0.076	0.091	0.144	0.103	0.200
1.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1.2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1.4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1.6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1.8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table A3. P-value of t-test for Difference in Mean Output: Predicted and Actual 2007 Distribution. *Notes.* This table denotes p-values corresponding to the differences in Table A2. The null hypothesis of these t-tests is that the mean of the actual and predicted distributions are equal, against the two-sided alternative. A p-value$\geq$ 0.10 means that we cannot reject the null of equal means at the 10% level of significance.

	β									
θ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.1	0.073	0.099	0.124	0.147	0.167	0.181	0.185	0.200	0.201	0.210
0.2	0.132	0.143	0.153	0.157	0.160	0.169	0.169	0.170	0.177	0.175
0.3	0.202	0.194	0.197	0.213	0.203	0.202	0.208	0.206	0.204	0.211
0.4	0.238	0.238	0.242	0.234	0.239	0.233	0.234	0.236	0.234	0.234
0.5	0.239	0.239	0.248	0.241	0.247	0.242	0.246	0.237	0.239	0.242
0.6	0.233	0.226	0.222	0.232	0.225	0.230	0.223	0.229	0.232	0.220
0.7	0.199	0.196	0.189	0.191	0.192	0.193	0.190	0.184	0.183	0.184
0.8	0.151	0.153	0.145	0.144	0.139	0.141	0.142	0.131	0.142	0.137
0.9	0.088	0.078	0.069	0.089	0.070	0.081	0.079	0.071	0.075	0.071
1.0	0.068	0.070	0.069	0.070	0.071	0.073	0.083	0.077	0.078	0.080
1.2	0.205	0.215	0.210	0.218	0.213	0.208	0.219	0.222	0.217	0.222
1.4	0.382	0.381	0.386	0.396	0.390	0.395	0.386	0.392	0.388	0.384
1.6	0.549	0.544	0.549	0.545	0.548	0.541	0.541	0.546	0.549	0.552
1.8	0.683	0.684	0.683	0.683	0.685	0.687	0.684	0.693	0.689	0.685
2.0	0.795	0.789	0.790	0.787	0.786	0.788	0.783	0.784	0.781	0.781

Table A4. Kolmogorov-Smirnoff D Statistic: Predicted and Actual 2007 Distribution. *Notes.* This table presents Kolmogorov-Smirnoff distance statistics $D = \sup_y |F(y) - G(y)|$, which is the supremum of the absolute distances between the actual and simulated output distributions. A smaller D indicates more similar distributions. Each cell presents D for a different (θ, β) combination for which the simulation was run.

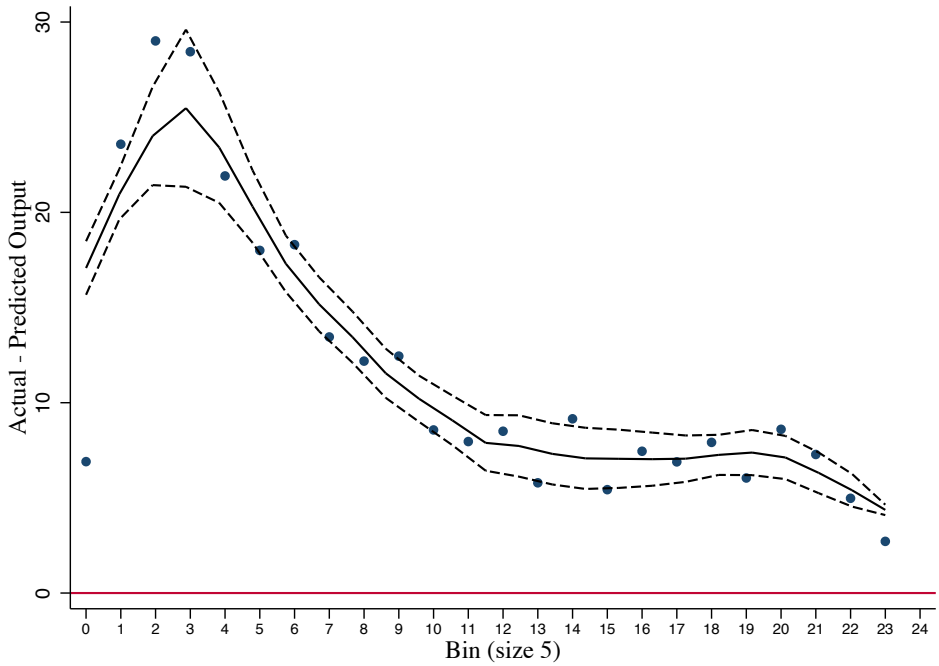


Figure A13. Deviation of Actual Output from Predicted Mean: Bin size 5. *Notes.* This figure depicts the analog of Figure 15, but with a bin size of 5

SUPPLEMENTAL APPENDIX: STRUCTURAL ESTIMATES WITH BETA DISTRIBUTION

Here we replicate the structural estimation exercise under the assumption that μ follows a beta distribution. This distribution is more restrictive than the gamma distribution, since it requires that μ lie in the unit interval. As a result, the structural model does not predict the underlying 2007 data as well as it does under the gamma distribution: the Kolmogorov-Smirnov distance statistic is always higher (see Table A6) and even if the absolute difference in means is lower for some parameter values (see Table A5), we are unable to reject the null hypothesis of equal means for the predicted and actual outputs for each possible (θ, β) combination. (For this reason, we don't even bother reporting a p-value table analogous to Table A3.) Still, the distance between the predicted and actual 2007 distributions according to the Kolmogorov-Smirnoff D and mean difference criteria are minimized at $(\theta, \beta) = (0.1, 0.1)$; see Figures A14 and A15 and Tables A5 and A6. At these parameter values, our results are qualitatively identical to those in our main results using the gamma distribution; see Table A7 and Figures A16-A18.

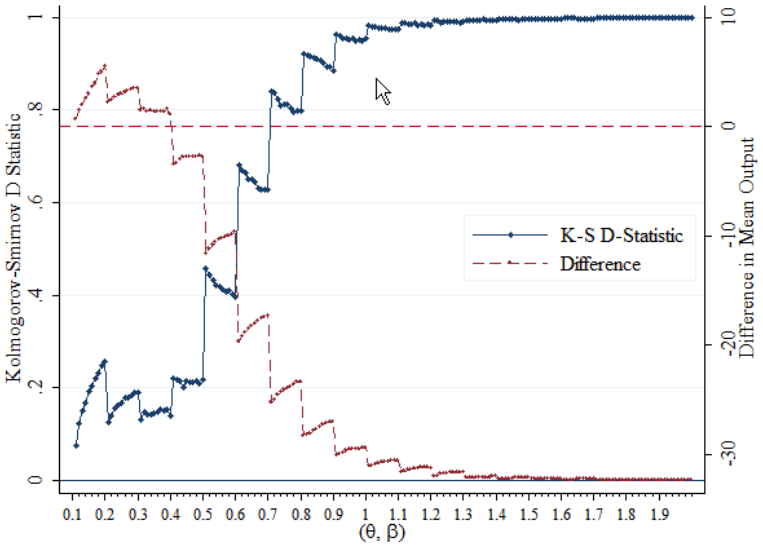


Figure A14. Kolmogorov-Smirnoff D Statistic and Difference between Predicted and Actual Mean Output by (θ, β) Combinations. *Notes.* This figure depicts the analog of Figure A12 under the assumption that μ follows a beta distribution rather than a gamma distribution.

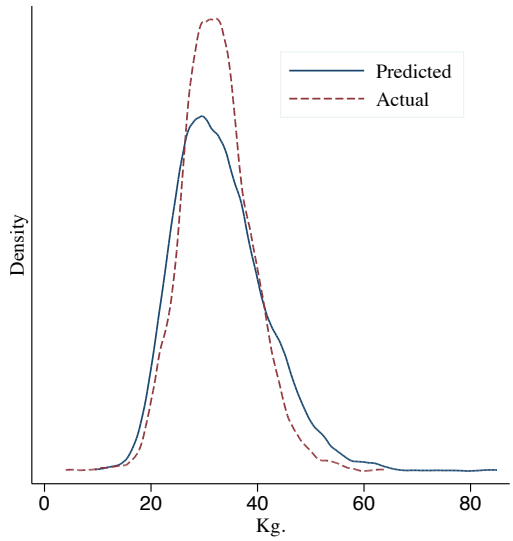


Figure A15. 2007 Kernel Density: Predicted and Actual. *Notes.*This figure plots the analog of Figure 13, but for the beta distribution with $(\theta, \beta) = (0.1, 0.1)$.

	β									
θ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.1	0.7	1.5	2.0	2.6	3.0	3.6	4.0	4.8	5.0	5.5
0.2	2.2	2.5	2.7	2.9	3.0	3.1	3.3	3.4	3.5	3.5
0.3	1.5	1.6	1.4	1.5	1.4	1.3	1.4	1.3	1.6	1.1
0.4	-3.4	-3.3	-3.0	-2.8	-2.7	-2.7	-2.7	-2.8	-2.7	-2.8
0.5	-11.6	-11.2	-10.8	-10.5	-10.4	-10.2	-10.1	-10.0	-9.9	-9.7
0.6	-19.7	-19.2	-18.8	-18.5	-18.2	-18.0	-17.8	-17.6	-17.4	-17.3
0.7	-25.3	-25.0	-24.6	-24.3	-24.1	-24.0	-23.8	-23.6	-23.4	-23.3
0.8	-28.4	-28.2	-28.1	-27.8	-27.7	-27.5	-27.3	-27.1	-27.0	-27.0
0.9	-30.1	-30.0	-29.9	-29.7	-29.6	-29.6	-29.5	-29.6	-29.4	-29.4
1	-31.1	-31.1	-30.9	-30.8	-30.7	-30.7	-30.6	-30.6	-30.6	-30.6
1.1	-31.6	-31.5	-31.5	-31.4	-31.3	-31.2	-31.2	-31.2	-31.2	-31.2
1.2	-32.0	-31.9	-31.8	-31.7	-31.7	-31.7	-31.7	-31.7	-31.7	-31.7
1.3	-32.1	-32.1	-32.1	-32.1	-32.1	-32.1	-32.1	-32.1	-32.0	-32.0
1.4	-32.3	-32.3	-32.3	-32.2	-32.2	-32.2	-32.1	-32.1	-32.1	-32.1
1.5	-32.3	-32.3	-32.3	-32.3	-32.3	-32.3	-32.3	-32.3	-32.3	-32.3
1.6	-32.3	-32.3	-32.3	-32.3	-32.3	-32.3	-32.3	-32.3	-32.3	-32.3
1.7	-32.3	-32.3	-32.3	-32.3	-32.3	-32.3	-32.3	-32.3	-32.3	-32.3
1.8	-32.4	-32.3	-32.3	-32.3	-32.3	-32.3	-32.3	-32.3	-32.3	-32.3
1.9	-32.4	-32.4	-32.4	-32.4	-32.4	-32.4	-32.4	-32.3	-32.3	-32.3

Table A5. Difference in Mean Output: Predicted and Actual 2007 Distribution. *Notes.* This table contains the difference between the mean of actual output and the mean of the predicted output in 2007. It is the analog of Table A2 under the assumption that μ follows a beta distribution.

	β									
θ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.1	0.076	0.122	0.150	0.168	0.193	0.204	0.221	0.232	0.247	0.257
0.2	0.127	0.140	0.155	0.163	0.167	0.178	0.179	0.185	0.188	0.188
0.3	0.131	0.147	0.143	0.143	0.144	0.148	0.152	0.151	0.153	0.138
0.4	0.221	0.218	0.216	0.201	0.214	0.211	0.211	0.215	0.209	0.217
0.5	0.458	0.443	0.432	0.420	0.419	0.413	0.406	0.409	0.403	0.396
0.6	0.682	0.671	0.663	0.651	0.650	0.644	0.631	0.628	0.627	0.627
0.7	0.840	0.836	0.822	0.810	0.813	0.812	0.804	0.796	0.798	0.797
0.8	0.922	0.918	0.916	0.911	0.909	0.906	0.902	0.894	0.892	0.885
0.9	0.963	0.960	0.956	0.954	0.951	0.955	0.950	0.952	0.949	0.954
1	0.982	0.980	0.979	0.976	0.977	0.976	0.975	0.973	0.973	0.974
1.1	0.989	0.987	0.987	0.986	0.987	0.983	0.986	0.983	0.985	0.984
1.2	0.992	0.993	0.990	0.990	0.990	0.991	0.990	0.991	0.989	0.991
1.3	0.995	0.994	0.994	0.995	0.994	0.995	0.994	0.995	0.994	0.993
1.4	0.997	0.997	0.997	0.996	0.997	0.996	0.995	0.996	0.996	0.995
1.5	0.997	0.997	0.997	0.997	0.997	0.997	0.996	0.997	0.997	0.997
1.6	0.998	0.998	0.998	0.998	0.997	0.997	0.997	0.997	0.998	0.998
1.7	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998
1.8	0.999	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998
1.9	0.999	0.999	0.999	0.999	0.999	0.998	0.998	0.998	0.998	0.998

Table A6. Kolmogorov-Smirnoff D Statistic: Predicted and Actual 2007 Distribution. *Notes.* This table is the analog of Table A4 under the assumption that μ follows the beta distribution.

	Month 1	Month 2	Month 3	Month 4	Week 49	Week 50	Week 51	Week 52
				$(\theta, \beta) = (0.1, 0.1)$				
Difference	19.920***	10.865***	5.505***	4.679***	4.678***	7.237***	4.900***	1.260***
	(0.393)	(0.385)	(0.382)	(0.383)	(0.415)	(0.414)	(0.410)	(0.404)

Table A7. Difference in Actual and Predicted Output After Contract Change. *Notes.* This table depicts the analog of Table 11, but for the beta distribution with parameters $(\theta, \beta) = (0.1, 0.1)$.

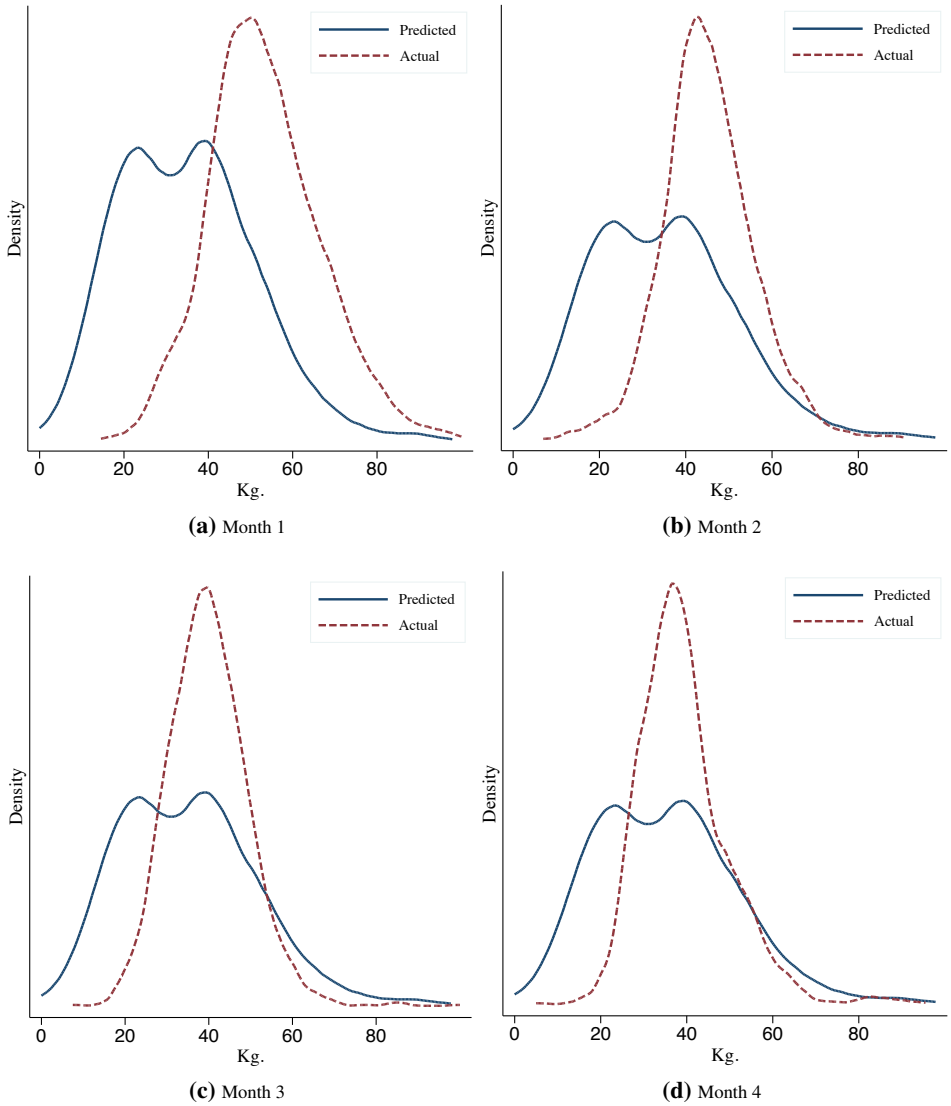


Figure A16. Kernel Density: Actual and Predicted Daily Outputs. *Notes.* This figure plots the analog of Figure 14, but for the beta distribution with parameters $(\theta, \beta) = (0.1, 0.1)$.

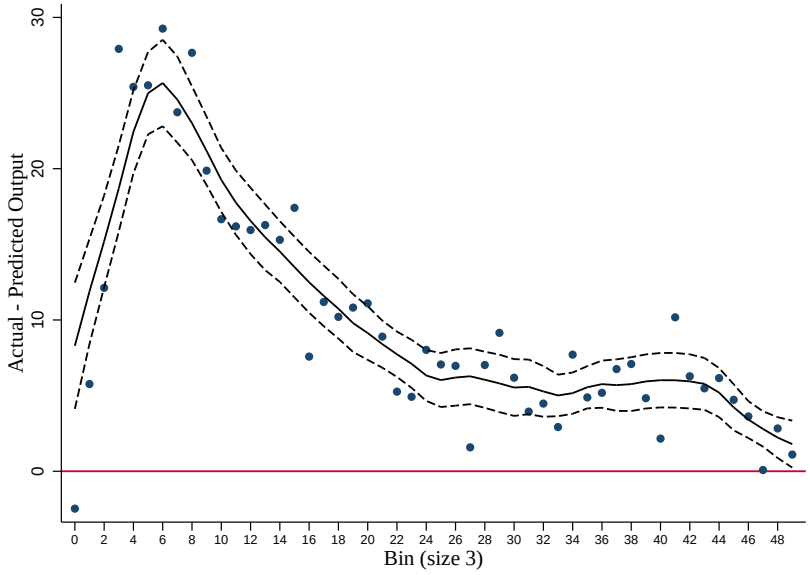


Figure A17. Deviation of Actual Output from Predicted Mean: Bin size b. *Notes.* This figure depicts the analog of Figure 15, but for predicted output using the beta distribution with parameter values $(\theta, \beta) = (0.1, 0.1)$.

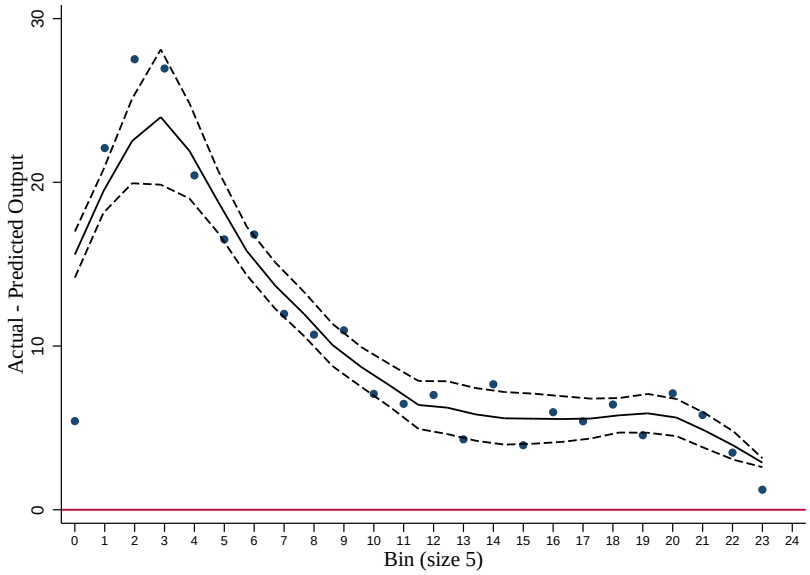


Figure A18. Deviation of Actual Output from Predicted Mean: Bin size b. *Notes.* This figure depicts the analog of Figure A13, but for predicted output using the beta distribution with parameter values $(\theta, \beta) = (0.1, 0.1)$.