Anatomy of a Contract Change[1]

Rajshri Jayaraman
ESMT Berlin

Debraj Ray
New York University and University of Warwick

Francis de Véricourt
ESMT Berlin

July 2014, revised July 2015
Forthcoming, *American Economic Review*

ABSTRACT

We study a contract change for tea pluckers on an Indian plantation, with a higher government-stipulated baseline wage. Incentive piece rates were lowered or kept unchanged. Yet, in the following month, output *increased* by 20–80%. This response contradicts the standard model and several variants, is only partly explicable by greater supervision, and appears to be "behavioral." But in subsequent months, the increase is comprehensively reversed. Though not an unequivocal indictment of "behavioral" models, these findings suggest that non-standard responses may be ephemeral, and should ideally be tracked over an extended period of time.

1. INTRODUCTION

We study the productivity impact of a contract change. The setting is a tea plantation in India. The activity in question is tea-plucking, the output of 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 as part of a customary practice of renegotiation once every three years. This plantation represented only a fraction of the negotiating parties, which involved representatives from some 20 unions and large plantations operating in the tea-growing region, with no dominant players on either side of the table. Negotiations commenced a few months prior to the expiration of the going contract, and went through several rounds.

The new contract increased the baseline wage by approximately 30%. Almost all of this increase was mandated by state government legislation. In June 2008, the government announced a preliminary notification under the Minimum Wages Act, 1948, setting the minimum daily wage payable to plantation workers in the state at Rs.102 (2.25 USD). The baseline in force in our plantation was, at the time, Rs.78 (1.72 USD); actually, even lower if a penalty clause is netted out. The new wage

[1]An earlier version of this paper was titled "Productivity Response to a Contract Change," and circulated as NBER Working Paper No.19849, January 2014. Jayaraman: European School of Management and Technology (ESMT), Schlossplatz 1, 10178 Berlin, Germany; rajshri.jayaraman@esmt.org. Ray: Department of Economics, New York University, New York, NY 10012; debraj.ray@nyu.edu. de Véricourt: European School of Management and Technology (ESMT), Schlossplatz 1, 10178 Berlin, Germany; francis.devericourt@esmt.org. We are grateful to the editor and two anonymous referees, as well as 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, Hunter College, University of Warwick and the University of Heidelberg for useful comments and suggestions. Ray acknowledges funding from the National Science Foundation under Grant SES-1261560.

contract, instituted at the beginning of September 2008, set a baseline wage of Rs.104, and covered over 10,000 pluckers from different plantations in the same 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.

The new contract also altered marginal incentives. It eliminated an existing penalty for shortfalls below a minimum "standard." It left the remaining piece rates unchanged, but they now came into play at somewhat higher output thresholds. 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.

Standard principal-agent theory yields an unambiguous prediction for worker productivity: it must weakly decline. But output per worker *increased* by over 80%, from 30 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 existence is undeniable, lying somewhere between 20% and 40% — 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, static model of incentives, which would imply reduced or at best unchanged effort.

Four main possibilities suggest themselves. 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. Third, there are possibilities of intertemporal substitution in plucking. Finally, notions of gratitude, reciprocity and gift-giving — "behavioral responses," in short — might play a role, as opposed to a monetary payoff-maximizing response to a changed incentive structure.

The first explanation appears to carry no weight at all. Briefly — but in more detail below — temporary workers, who are hired seasonally, are *less* responsive to the contract change than permanent workers, who cannot be fired. The second channel — heightened supervisory effort — appears to account for some of the increase: around a quarter of increased productivity. Third, we rule out any significant possibilities for intertemporal substitution in plucking. We also argue against the existence of learning or nutrition-based channels. Our conclusion is that 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. Starting around the second month after the contract change, an output decline sets in, until 4 months after the contract change, *output returns to pre-change levels*. The reversal is evident in the reduced form, but to assess its extent, we estimate a simple structural model from the pre-change data. We apply those estimates to the data *after* the contract change. Not surprisingly, our predictions are far removed from the output distribution observed in the month after the contract change. This confirms what we've already noted: that output in the immediate aftermath of the new contract departs 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 output increase (relative to the "standard prediction") essentially vanishes by the end of month

4. In fact, by week 17 following the contract change, 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[2] has significant "behavioral effects" in the short run. Yet, these effects dissipate just as significantly as time goes by, "ultimately" giving way to outcomes reasonably consistent with the predictions of classical incentive theory. These findings warn against the claim of dominant behavioral responses based on outcomes just following a change. Once the euphoria dies down, such effects may vanish. In particular, we caution against the exclusive use of regression discontinuity methods that emphasize only what happens at the point of the event. Both short- and long-term effects are important, and ideally, we would like to identify both. The way to do that is to track responses over time.[3]

Our paper connects with a research program outlined by Gneezy, Meier, and Rey-Biel (2011), who examine "some general aspects of how extrinsic incentives may come into conflict with other motivations." They emphasize how the introduction of "extrinsic incentives" (money) might erode "intrinsic motivations" (reciprocity, gratitude or fair play). That is, in situations in which the *baseline* interaction is non-monetary — say blood donations or social work — financial incentives can cut into and displace pro-social motivations, perhaps with a negative net effect; see, e.g., Gneezy and Rustichini (2000), and Mellström and Johannesson (2008), and well as Gneezy, Meier, and Rey-Biel (2011), Frey and Jegen (2001) and Charness and Kuhn (2011), who review the empirical evidence.[4]

In contrast, our baseline interaction is monetary, and a social component is introduced via the perceived generosity of the contract change. The immediate response of workers does appear to be a form of pro-social reciprocity. Yet over time, monetary incentives — the original foundation of the interaction — ultimately hold sway. That is, not only might monetary considerations displace social motivations in some settings (as in the earlier literature), but they may also be hard to dislodge when they are *status quo*, which is certainly true of most employment settings.

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 introduces the long run response; Section 6 studies it structurally. Section 7 concludes.

2. Contract Change in a Tea Plantation

2.1. **Setting.** We study a tea plantation in India, owned by a large producer, in a state where tea is an important source of employment. This plantation is one of several in the region, collectively forming

[2]Our paper belongs to a growing empirical literature on the implications of the contract form in a natural employment setting; see, e,g., Shaban (1987), Lazear (2000), Ackerberg and Botticini (2002), Shearer (2004), Bellemare and Shearer (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.

[3]These results from the field echo experiments by Gneezy and List (2006), who find that positive productivity responses to gift exchange are eroded in a matter of hours, and Bellemare and Shearer (2009), who over a 2-week period find that productivity increases only on the day of the gift. However, some behavioral changes may well be longer-lived. For instance, the pure framing of incentive bonuses as penalties rather than rewards appears to have a persistent effect on productivity; see Hossain and List (2012).

[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) and Lacetera, Macis, and Slonim (2012). Theoretical research that bears on these issues includes Bénabou and Tirole (2003) and Besley and Ghatak (2005) and is reviewed in Kőszegi (2013)

the dominant source of employment in the locality. Our study plantation has over two hundred fields on which clonal tea bushes grow in rows. These are pruned to form a flat surface, resembling a trimmed hedge, constituting a "plucking table" approximately one meter tall. This is roughly waist-height for the workers on the plantation, making plucking less physically arduous and facilitating supervision.

Around 70% of the pluckers in our plantation are female. They typically work full days or not at all. They are permanently assigned to small "gangs," composed of about 20–40 members, who tend to a fixed set of geographically proximate fields. Each workday, gangs are assigned to plucking or non-plucking duties; the latter include pruning, weeding, and the spraying of pesticides. A gang's daily job duties as well as field assignments are decided centrally, by managers. So is the decision regarding 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 a 4-year field life cycle, with minor adjustments for weather conditions. Each gang has a supervisor, who is paid a fixed wage. (Gang and supervisor assignment were unchanged after the new contract.)

Fresh, unprocessed tea leaf is manually plucked from the bushes, either by hand or with metal shears. Leaves are collected in bags carried on the backs of the pluckers; these bags are weighed and the amount of green leaf plucked is recorded by the supervisor for each plucker on each day.[5] There are no complementarities in production across different workers. Increased output is therefore a simple consequence of increased individual speed in plucking, permitting more bushes to be plucked.[6]

Approximately 65% of the pluckers are permanently employed and have a median tenure of approximately 21 years. They are entitled to work on all workdays and 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 35% are temporary workers and are hired on a season-by-season basis. They are not entitled to the same non-pecuniary benefits. However, *all* workers 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.

All told, tea plucking is a pretty routine task. 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. The company has a considerable history in this industry. It is therefore reasonable to suppose that asymmetric information regarding worker ability or effort is not a central consideration here. So the usual complications to standard settings — such as multitasking (Holstrom and Milgrom, 1991), ratchet effects (Gibbons, 1987) or career concerns (Gibbons and Murphy, 1992) — are unlikely to be focal in our environment. Finally, permanent workers cannot be legally dismissed, which makes termination an ineffectual instrument in inducing effort (Dutta, Ray, and Sengupta, 1989). On the other hand, temporary workers can be fired; this distinction will inform part of the analysis below.

[5]Indeed, the supervisor closely monitors this small group, recording each member's daily activity. That includes presence or absence, 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.

[6]An individual plant can also be plucked more intensively, though over plucking can damage later yields. As we will discuss later, this is not a concern since gangs are made responsible for particular rows in a field, thereby internalizing this dynamic externality, and also because the majority of workers have at least a decade of experience.

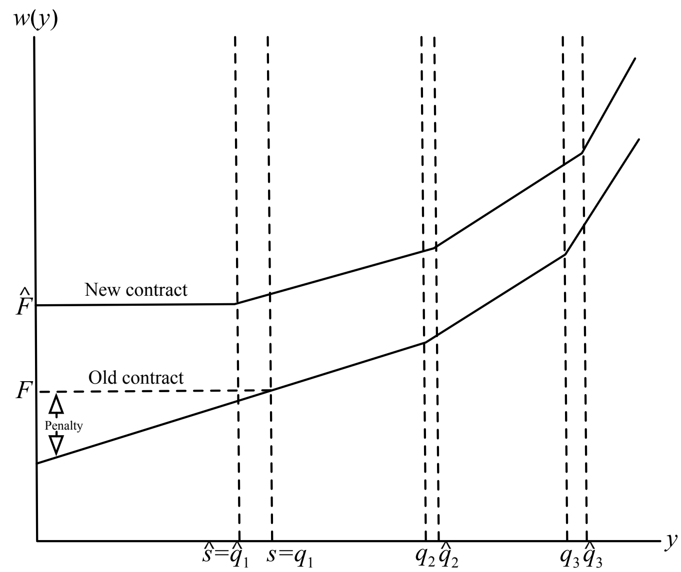


Figure 1. Wage Contract. *Notes.* This figure depicts the daily wage as a function of output y for a typical contract before and after the contract change. F and $\hat{F}$ are the fixed daily wages, paid conditional on participation; q_i and $\hat{q}_i$ (for various i) denote output thresholds corresponding 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.

2.2. **Contract.** Workers are paid a baseline wage on the days they work. Otherwise they receive no payment. In addition, they receive marginal incentives: piece rates per kilogram of plucked leaf. These rates vary across different slabs of output, and the thresholds at which the they kick in (though, as it happens, not the rates themselves) vary by estate-climate type (we observe two such types in our data) and plucking method. 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 lower curve in Figure 1 depicts a typical pre-change contract. The minimum standard is shown by s and the fixed daily wage by F. There are three piece rates that apply to different output slabs. Oddly, the contract has no kink at s: the minimum standard is located "within" the first incentive slab, and the piece rate to the "left" of s is framed as a penalty. In total, we observe 4 such contracts in our data, with cutoffs as well as the minimum standards set higher for shears, or for high-yield estates.

2.3. **Contract Change.** 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. Negotiations began 3–4 months prior to contract expiration and went through several rounds. They involved representatives from approximately 20 unions and employers, with no dominant players on either side of the table. In the end, 10,000 pluckers from different plantations were covered by the new agreement. Our plantation had around 2,000 pluckers.

From the point of view of establishing the exogeneity of this contract change to the individual worker, it is imperative to note that state government legislation effectively drove the increase in the baseline wage. In June 2008, the state confirmed a preliminary notification under the Minimum Wages Act, 1948, setting the minimum wage for plantation workers at Rs.101.52 (approximately 2.25 USD at 2008 exchange rates). The daily wage being paid at the study plantation was, at the time, Rs. 77.55 (approximately 1.72 USD in 2008). Writ petitions and on-going press reports clearly indicated that owners objected to the new baseline wages. But 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 of Rs. 103.76, an increase of around 30%. In fact, because the penalty provisions for dropping below the standard were also abolished, the effective increase in the daily fixed wage component was 44%.[7]

This much is indubitably exogenous, not only to the individual plucker, but — because of state legislation — to the plantation as a whole. The plantations did collectively react to the change by a flattening of the piece rate structure. The piece rate for the first slab up to the standard was now effectively zero, because of the absence of a pecuniary penalty. This change was the biggest driver of the flatter structure. The piece rates for the remaining slabs were unchanged at Rs. 0.45, 0.55 and 0.85, respectively. However, the *thresholds* at which these rates applied were (slightly) shifted to the right. 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 increased baseline wage. Online Appendix O.1 provides a simple model that explains such a response. A typical new contract is shown as the upper curve in Figure 1.

In summary, then, the core driver of the contract change was exogenous to all members of our plantation, and this is the change in the fixed wage. The new piece rate structure was arguably an endogenous reaction to this core change. But all said and done, both the timing and structure of the new contract were exogenous to the individual *worker*. Of course, that does not preclude alterations in worker behavior before the (anticipated) change. The data do indicate abnormally low output in the first three weeks of August 2008, just before the change. We will deal with this by considering different counterfactuals.

2.4. **Output Data.** Our data are obtained from supervisor entries in personnel records for all employed workers. 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.

Table 1 describes the basic data structure. In 2008, we have daily observations for roughly 2,000 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 and Week 0 refer to the month and the week before the (date of the) contract change and Months 1–4 and Weeks 1–17 refer to the corresponding months or weeks following the contract change. Note that the number of worker-day

[7]If a worker came in and hypothetically plucked no output, the penalty would amount to Rs. 16–28, depending on yield class and plucking method. Presumably, there were 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	24	26	25	25	24
No. Unique Workers	1992	1983	1956	2049	2137
No. Obs.	33552	37339	33752	35426	34197
	Study plantation: 2007[b]				
No. Days	25	25	26	25	24
No. Unique Workers	2319	2116	2110	2276	2164
No. Obs.	39822	36067	37730	38705	33231
	Control plantation: 2008[c]				
No. Days	25	26	24	25	24
No. Unique Workers	639	666	730	661	642
No. Obs.	10256	11615	10694	11309	10368

Table 1. Data Structure. *Notes.* This table describes the number of observations in geographic locations and time periods for which we have data. The contract change was instituted at the start 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.

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.[8]

We exploit two 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.[9] The second counterfactual is a plantation located in a hill station at a 100 km. great circle distance north of the study plantation. This "control" plantation is observed contemporaneously in 2008, but is on a different 3-year contract cycle, and did not experience a contract change over the observation period. Two features make it a good counterfactual. First, it is owned and operated by the same company, and shares the same management, technologies, cropping and labor practices as the treatment plantation. Second, the natural environment is comparable: the plantations are located at roughly the same altitude, have similar geological characteristics (slopes, soil quality, etc.), are subject to the same monsoon cycles, and have almost identical rainfall patterns over the study period.

As already discussed, workers are permanently assigned to gangs that pluck fields based on a predetermined schedule and plucking method set by management. 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. We correct for rainfall with an appropriate lag structure discussed

[8]We restrict attention to days on which workers participated *and* were assigned to plucking duty. The reason for this focus is twofold. First, 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.

[9]Hence, participation does not change dramatically over this period. In fact, we will show shortly that it cannot account for the productivity changes we observe. Moreover, our results are robust to the inclusion of worker fixed effects.

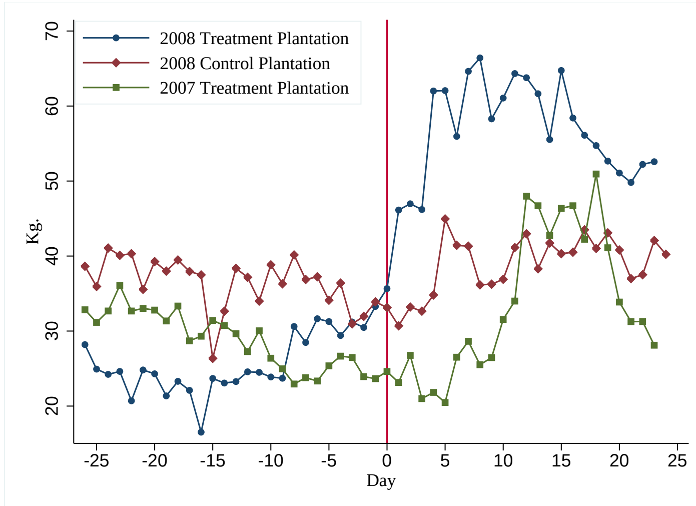


Figure 2. Time Series: Average Daily Output. *Notes.* The time series depicted in this graph describe average daily output in the treatment plantation in 2008 and 2007 and in the control plantation in 2008 for one month on either side of Day 0, the date of the contract change.

later, constructed from daily Tropical Rainfall Measuring Mission (TRMM) grid cells whose centroid is within 20 km. of the village closest to the plantation.[10]

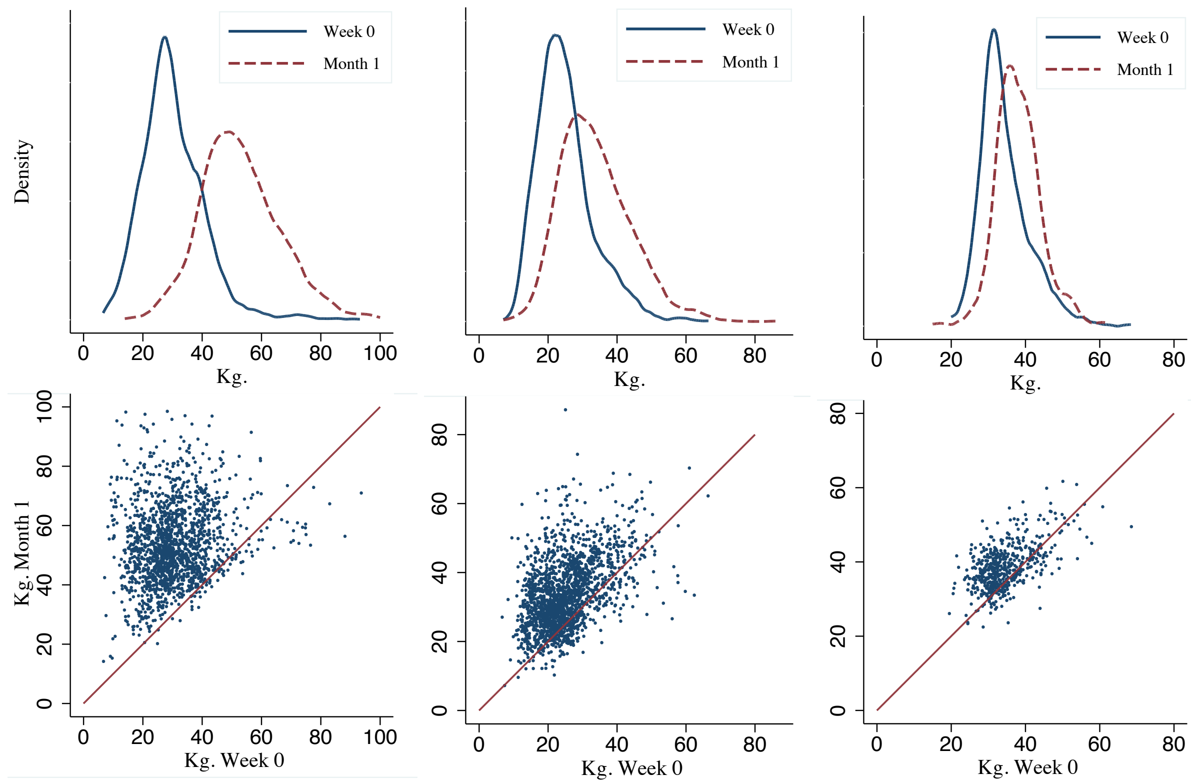
3. Productivity Response in the Short Term

The basic facts are evident from a cursory examination of average daily output over Month 0 (August) and Month 1 (September). Figure 2 shows that output increased sharply in the treatment plantation in the month after the contract change. No comparable increase is visible in the control plantation over the same period or in the treatment plantation for the corresponding time period in the previous year.

The rest of this section establishes the fact of this large increase, by taking care of several potential confounds. First, as already mentioned, output in the first 3 weeks of Month 0 in 2008 is low, both relative to the control plantation in 2008, and relative to Month 0 in the treatment plantation in 2007. This low output persists even when we correct for other (time-varying) variables, such as rainfall. In order to avoid the possibility that our estimates simply reflect an unusually low pre-change output, we restrict our pre-change sample to Week 0 — the week directly preceding the date of the contract change. Output during this period is comparable to corresponding outputs in the control plantation and in the treatment plantation in the previous year. Analogous results using all of Month 0 are even more pronounced, and are provided in the Online Appendix; see Section O.3 there.

Yet even the Week 0 baseline leaves little doubt regarding the change. Figure 3 depicts marginal densities of daily individual output (top panel), as well as worker-specific scatters (bottom panel), for Week 0 and Month 1 average output. The first panel is the situation of interest: output in the study plantation in the last week of August and in all of September 2008. The middle panel shows the study plantation again, this time over the same periods in 2007. The right panel depicts output in the

[10]We are grateful to Eric Stroebl for rainfall data.



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

Figure 3. Kernel Density and Scatter Plots: Average Daily Output. *Notes.* The top three panels of this figure depict kernel density estimates for the average daily output of workers in Week 0 (solid line) and Month 1 (dashed line). The bottom three scatter plot depict 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^o. 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.

control plantation in the corresponding period in 2008 (with no change in contract). There is a mild increase in output 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 jump in the treatment plantation in 2008 is dramatic. Average output increased by 83%, from 30 to 55 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.[11] This is a remarkable increase, and we will discuss various factors that might bear on it. In this section, we address some of the preliminary considerations.

[11]Figure O.3 in the Online Appendix shows that both the relative and absolute increases in this and subsequent graphs 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 55 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.

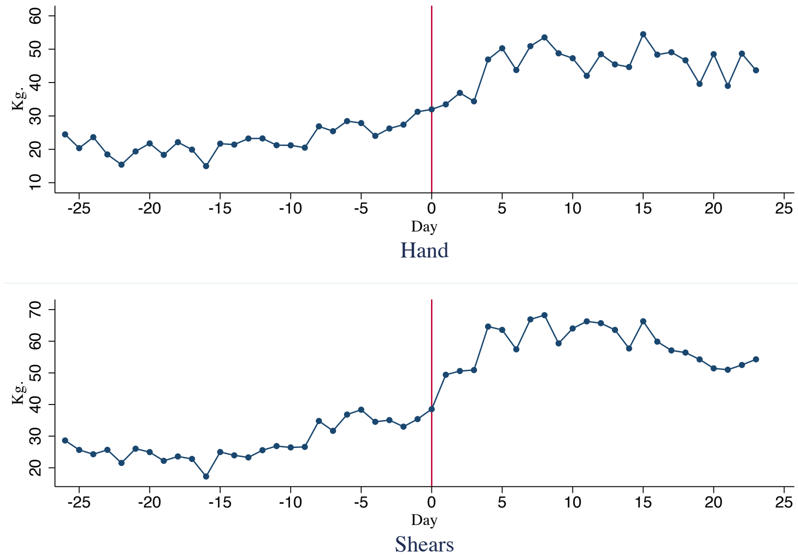
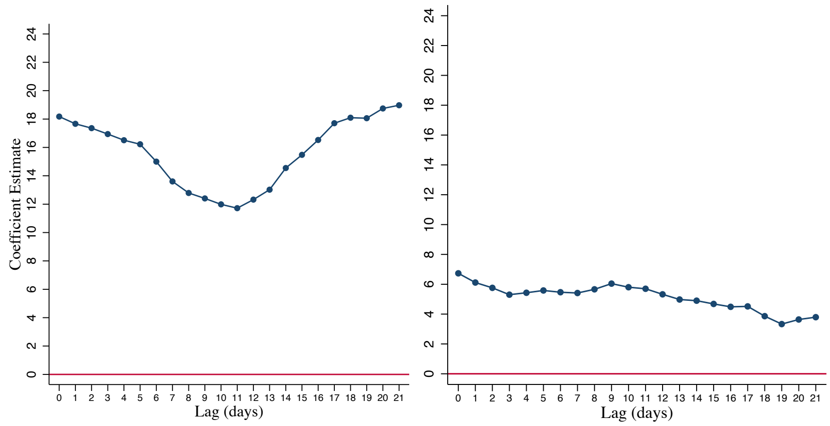


Figure 4. Time Series: Average Daily Output Disaggregated by Hand and Shears. *Notes.* This figure depicts average daily output in the treatment plantation in 2008. The top panel averages over plucker-days assigned to hand plucking and the bottom panel to plucker-days assigned to shears plucking.

Start of the Plucking Season. August is the first month of the plucking season, and there is a natural tendency for output to grow as the season picks up. Different estates on the plantation are more likely to fall under a lower yield class in August, and a higher yield class 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 these classes by looking at the change in the minimum standard: 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 it is nowhere close to what happened on the treatment plantation in 2008.

Participation Rates. The scatter plots presented in Figure 3 demonstrate the within-worker increase in output, conditional on workforce participation. 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.

Hands Versus Shears. The higher output could reflect an increase in the use of shears in the month(s) following the contract change, since output is higher with shears than with hand plucking. However, Figure 4 shows that the change we observe does not simply represent a technological shift, as output in the treatment plantation increased for *both* hands and shears. Additional graphical support — kernel densities for hands and shears, and accompanying scatter plots at the worker level — is available in Figure O.4 in the Online Appendix. At the same time, we do 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 in the regression analysis below.



(a) 2008 First Difference

(b) Double Difference

Figure 5. 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 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.

Rainfall. The South-West monsoon is relevant for the period under consideration. In our tea-growing region, it generally lasts 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 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 the increased use of shears, the frequency of past recent plucking, 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 depicts the first difference and 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 first difference estimate τ_1 from the following regression:

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

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 a plucking method dummy variable and field fixed effects. "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 5.

The right panel presents the OLS double difference estimate τ_2, recording how productivity in the treatment plantation in 2008 changed relative to its earlier change in 2007:

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

where Treat_{it} 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. Both panels indicate that regardless of the choice of lag

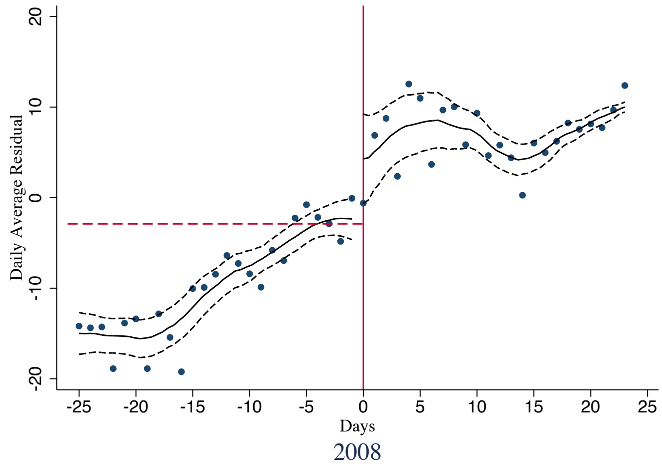


Figure 6. Time Series: Average Daily Residual. *Notes.* This figure depicts average daily residual in the treatment plantation in 2008. Each dot in the figure represents a different day. The figure is overlaid with a local polynomial smoothed using an Epanechnikov kernel (solid curve). Dashed curves denote the 95% confidence interval. The horizontal line denotes the 2007 Month 0 average residual in the treatment plantation.

structure, the change in output remains firmly positive following the new contract. 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 40% (see Table 2 for a more precise estimate). Figure 5 also shows that an 11-day lag in rainfall yields the most conservative estimate of the productivity increase. In all future estimates for the treatment plantation, we calculate the output residuals of an OLS regression with field and plucking method fixed effects, quadratic terms for the number of days a field has been plucked using hand and 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 lack field-level data.

Figure 6 plots daily average residuals for Months 0 and 1 in 2008, fitted with a local polynomial using an Epanechikov kernel.[12] The series displays a sharp discontinuity on the date of the new contract. Regression discontinuity estimates indicate a 4–5 kilogram increase in output, in the vicinity of 15%, in a matter of 2-5 days; see Table A1 in the Appendix. The Appendix also plots residuals for the treatment plantation in 2007 and the control plantation in 2008 (Figure A2).The treatment plantation in 2007 also displays an uptick of productivity in the region of the discontinuity (see Panel (a) of Figure A2), but the increase is far more muted: regression discontinuity estimates (not reported) are only half the size of those in the treatment plantation. In the control plantation, such a discontinuity is entirely absent (see Panel (b) of Figure A2): regression discontinuity estimates (not reported) are negative and statistically insignificant.

Finally, Table 2 presents first difference estimates (columns 1-3) and double difference estimates (columns 4-5) of equations (1) and (2) respectively, using Week 0 as the baseline, and including all

[12]As before, there is a depression in the first 3 weeks of Month 0, but by 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). Figures O.5 and O.6 in the Online Appendix shows kernel densities and scatter plots analogous to those in Figure 3, but for residuals thus calculated, rather than output.

	Dependent Variable: Output (Daily Kg. Tea) Month 1 over Week 0				
	First Difference			Double Difference	
	OLS	OLS	FE	OLS	OLS
	(1)	(2)	(3)	(4)	(5)
Counterfactual				2007 Treatment	2008 Control
After × Treat (τ_2)				6.13**	13.34***
				(2.388)	(1.828)
Treat (τ_0)				9.29***	-0.60
				(1.213)	(0.982)
After (τ_1)	12.03***	12.41***	12.76***	9.14***	4.23***
	(1.622)	(1.632)	(0.429)	(2.356)	(1.170)
Rainfall	0.48***	0.50***	0.46***	0.36***	0.54***
	(0.054)	(0.055)	(0.015)	(0.077)	(0.100)
Shears days	1.98	2.08	4.88***	-0.10	
	(5.554)	(5.856)	(1.353)	(4.181)	
Shears days^2	6.79	7.27	2.54	4.78	
	(7.037)	(7.358)	(1.843)	(5.121)	
Hand days	-29.06***	-30.07***	-27.28***	-18.37***	
	(7.030)	(7.087)	(1.963)	(5.216)	
Hand days^2	32.47***	33.99***	31.33***	21.83***	
	(9.997)	(10.067)	(2.717)	(6.793)	
Shears dummy	12.27***	12.00***	12.36***	12.45***	
	(1.467)	(1.514)	(0.481)	(1.245)	
Field FE	yes	yes	yes	yes	yes
Worker FE	no	no	yes	no	no
Plucked at least 4 days before and after	no	yes	no	no	no
No. Observations	44,910	38,801	44,910	89,147	59,837
Adjusted R-squared	0.523	0.533	0.558	0.546	0.278
No. Unique Workers			1,947		
Week 0 mean output in 2008 Treatment	30.64	30.67	30.64	30.64	30.64
	(0.434)	(0.438)	(0.434)	(0.434)	(0.434)

Table 2. Short Run Regression Estimates. *Notes.* This table presents first difference estimates for equation (1) in Columns 1-3 and double difference regression estimates for equation (2), with the 2007 Treatment Plantation and the 2008 Control Plantation as counterfactuals, in Columns 4 and 5, respectively. In each case, Week 0 is compared to Month 1. Column 1 contains the full 2008 treatment plantation sample. Column 2 imposes the further restriction that workers worked at least 4 days before and after the contract change, in order to deal with worker selection. OLS estimates are presented in columns 1, 2, 4 and 5, with standard errors clustered by day. Column 3 imposes worker fixed effects (FE), with robust standard errors. Standard errors in parentheses. $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$.

the controls described earlier. Column 1 includes all the 2008 treatment plantation observations in Week 0 and Month 1. 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. (or 40%) is undeniable across these three first difference specifications. This estimate is qualitatively similar — at 13.3 kg. — according to the double difference estimate in Column 5, which uses the 2008 control plantation as a counterfactual. The double difference in Column 4, which uses the 2007 treatment plantation counterfactual, yields our lowest linear point estimate, but even this amounts to a 20% increase. (Each estimate almost doubles when Month 0 is used as a baseline; see Online Appendix, Table O.1.) The results are qualitatively similar when we control for output outliers. When we drop either the top 1% or the top 5% of observations in the output distribution, the first difference results are virtually unchanged. The double difference estimates increase for the 2007 counterfactual and fall for the 2008 counterfactual. All estimates, however, remain positive and statistically significant.

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

The discussion above establishes that following the contract change, there was an immediate output increase of a sizable magnitude. We now consider different factors that might bear on this increase, among them (1) static incentives (2) dynamic incentives (3) heightened supervision (4) intertemporal substitution of plucking effort (5) learning, and (6) "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. Under fairly general conditions, classical incentive theory tells us that effort and output should not go up. 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. Assume that the wage functions are continuous and increasing, with bounded slope. (Typically they will be comprised of a baseline wage and various piece rates, as in our plantation.) Suppose further that (C1) the wage function shifts up, so $\hat{w}(y) \geq w(y)$ for all y, (C2) it flattens, so $\hat{w}(y) - w(y)$ is non-increasing in y, and (C3) the minimum standards don't go up, so $\hat{s} \leq s$. All three are features of our observed contract change.

The worker's payoff function is given by

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

where u is increasing and concave in w, effort cost c is increasing and convex in y (with unbounded slope as $y \to \infty$), and L is a nondecreasing, convex "penalty 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.[13]

The worker chooses effort (or equivalently y) to maximize payoffs, subject to the link between w and y 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, although 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.

[13] We control for all stochastic shocks that are realized prior to the supply of effort, and hence 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.

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

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

The proposition simply formalizes what we would suspect right away: a contract change that increases baseline wages, flattens the incentive structure, and lowers standards without changing supervisory effort, should induce workers to decrease effort. Recall that in the specific case studied here, the main "flattening" occurred because of the elimination of monetary per-unit penalties below the minimum standard. If supervision could be stepped up to compensate for this removal (and we argue below that it was, to some degree), the change could keep effort choices unaltered, though no better. This is precisely why the increase in output following the contract change is of interest.

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.

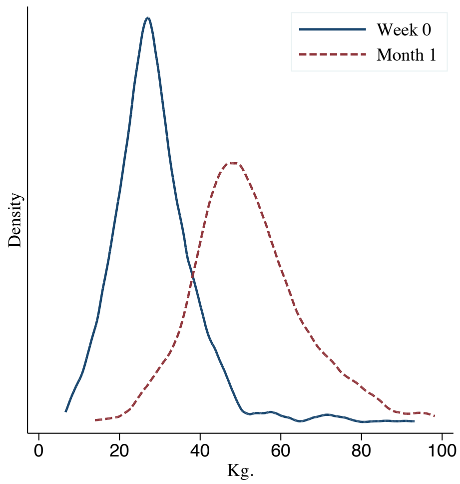
4.2. **Dynamic Incentives.** It is possible that the new contract might have been accompanied by sharper dynamic incentives. For instance, the very fact of higher wages might serve as an inducement to provide higher effort, using the threat of firing or contract non-renewal, as in Shapiro and Stiglitz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

However, permanent workers cannot be legally dismissed. 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 the other hand, this fact 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.

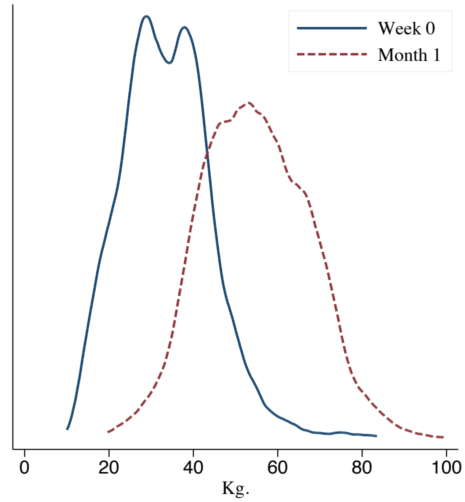
Figure 7 depicts kernel densities analogous to those in the top left panel of Figure 3, disaggregated by permanent and temporary workers. (For corresponding comparisons using Month 0, see Figure O.7 in the Online Appendix.) Output for both sets of workers shift to the right. This does not square with the dynamic incentives argument. Whatever the cause of the output increase, it applies to both permanent and temporary workers.

Table 3 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, both permanent and temporary workers increase output in response to the contract change. Indeed, contrary to what one would expect, the response of permanent workers to the contract change is almost twice as large as that of temporary workers.

We therefore find it 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 exhorted to perform well. This motivates our next consideration.



(a) Permanent Workers



(b) Temporary Workers

Figure 7. 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 permanent workers in Panel (a) and temporary workers in Panel (b).

4.3. **Supervision.** Recall that the government-mandated minimum wage increase was predictably resisted by plantation owners. Writ petitions seeking a stay on the minimum wage notification were submitted by planters, but were dismissed by the state's high court. Plantation owners were all too aware of the incentive effects of piece rates, but these could not be raised in tandem with the higher fixed wage: the costs of doing so were perceived to be too high. It is therefore important to see if other non-pecuniary means of maintaining effort — supervision, in a word — 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 given their long association with the plantation, are intensely loyal. For permanent workers, supervision is limited to a mix of exhortation to work, coupled with worse treatment of the worker and her family (within limits, of course, as the workers are unionized), or later assignment to more arduous tasks.

We impose a natural restriction on supervisory authority. Presumably, supervisors only have power in the output range below the explicitly announced minimum standard. Any drop below this minimum is tantamount to a breach of contract, even though the punishments may be limited. Now consider outputs that exceed the standard, and for which an additional piece rate is paid. 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," and the use of the piece rate undercuts the possibility of any legal or contractual breach when output varies in this range. This is perfectly illustrated by the pre-contract use of the term "penalty" below the standard. If we neglect supervisory authority, there is no difference between a per-unit penalty and a piece rate, but the very use of the term "penalty" suggests that extra-pecuniary supervisory authority can be brought to bear on workers in this sub-standard range, and not above.

	Dependent Variable: Output (Daily Kg. Tea)					
		Permanent			Temporary	
	OLS	OLS	FE	OLS	OLS	FE
	(1)	(2)	(3)	(4)	(5)	(6)
After (τ_1)	13.88***	13.99***	14.71***	7.34***	7.60***	7.94***
	(1.649)	(1.659)	(0.498)	(2.315)	(2.378)	(0.889)
Rainfall	0.47***	0.49***	0.46***	0.50***	0.54***	0.50***
	(0.056)	(0.058)	(0.018)	(0.071)	(0.065)	(0.031)
Shears days	5.09	4.99	7.75***	0.43	1.61	2.28
	(6.843)	(6.947)	(1.606)	(6.328)	(5.411)	(2.878)
Shears days^2	1.65	2.27	-2.32	10.76	10.96	8.04**
	(8.774)	(8.944)	(2.245)	(8.978)	(7.189)	(3.688)
Hand days	-27.12***	-27.95***	-25.36***	-23.21**	-24.73***	-24.70***
	(8.003)	(8.147)	(2.264)	(8.720)	(7.691)	(4.226)
Hand days^2	32.23***	34.11***	31.97***	16.75	17.09	20.62***
	(11.248)	(11.691)	(3.200)	(12.718)	(11.015)	(6.267)
Shears dummy	12.22***	12.20***	11.72***	12.12***	11.68***	12.44***
	(1.621)	(1.610)	(0.542)	(2.057)	(2.264)	(1.066)
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	32,474	29,539	32,474	12,436	9,024	12,436
Adjusted R-squared	0.537	0.546	0.577	0.532	0.550	0.548
No. Unique Workers			1,308			650
Week 0 mean output in 2008 Treatment	29.06	29.17	29.06	35.10	35.09	35.10
	(0.492)	(0.496)	(0.492)	(0.878)	(0.911)	(0.878)

Table 3. Short Run Regression Estimates: Permanent and Temporary Workers. *Notes.* This table presents regression estimates for equation (1) (First difference) in 2008, disaggregated by permanent (columns 1–3) and temporary workers (columns 4–6). The estimates compare Week 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 errors clustered by day. Column 5 estimates individual fixed effects (FE), with robust standard errors. Standard errors in parentheses. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$.

In short, apart from the monetary inducement of the piece rate itself, it is difficult, if not impossible, to appeal to a legal contract to coerce a worker to consistently perform above minimum standards. We therefore adopt the position that supervisory authority applies, if at all, to the underperformers, and use this conceptual restriction to back out an estimate for the contribution made by supervision.

Following up on this point, we separate all workers into two categories. Using data from the period before the contract change, we can see if the monthly average of each worker falls below ("underperforms") or above ("overperforms") the applicable pre-change standard.[14] Table 4 records the subsequent increase in output coming from underperformers versus overperformers.

[14]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.

	Underperformers		Overperformers		S (%)
	Levels	Logs	Levels	Logs	
	(1)	(2)	(3)	(4)	(5)
Month 1	16.78***	0.45***	9.43***	0.22***	25.7
over Week 0	(2.032)	(0.051)	(1.904)	(0.038)	
No. Observations	23,064	23,064	19,666	19,666	
Month 1	28.00***	0.81***	19.88***	0.48***	23.2
over Month 0	(1.873)	(0.052)	(1.963)	(0.047)	
No. Observations	36,769	36,769	29,295	29,295	

Table 4. Supervisory Effort. *Notes.* This table presents OLS estimates for τ_1 analogous to those in Table 2 (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. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$.

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 4. 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 45%, if Week 0 is taken to be the baseline.

Yet at the same time, the overperformers *also* increased their output. The corresponding number is 22% for the Week 0 baseline. (The estimates, as usual, double for both under- and overperformers if Month 0 is taken as the baseline.) As already argued, it is extremely unlikely that this increase could come from supervision. Under that assumption we can estimate the fraction of the 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. Then 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

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

where q is the output share of underperformers in the week or month prior to the contract change. This is reported as "supervision share" in the final column of Table 4; by this metric, better supervision accounts for at most a quarter of the output increase following the new contract.

Admittedly, this number is open to reinterpretation, but to increase it any further, one must assume *either* that individuals above the standard 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.

4.4. **Unplucked Leaves and Intertemporal Substitution.** Output in the treatment plantation in 2008 was unusually low in the first 3 weeks of August, which raises the concern that the increased output in September may, in part, reflect the accumulation of unpicked mature shoots from the earlier period. Some degree of intertemporal substitution is certainly possible from one day to the next. For instance, Sunday is a holiday and there is a significant and positive day-of-the-week effect on Mondays and Tuesdays, relative to the weekly average. However, a basic primer on the morphology of the tea plant in this region will easily serve to rule out effects lasting for more than 5–7 days.

The top of the plucking table has a base of maintenance leaves, above which mature shoots are harvested by workers. Mature shoots are made up of 4 leaves (above the "maintenance leaf") and a bud; see Appendix Figure A1. The plantation practices coarse plucking, which means that workers pluck shoots —"flushes"— comprising 3 leaves and a bud.[15] The fourth "janam" leaf, located just above the maintenance leaf, must be left intact if the shoot is to grow again.

It is imperative to note that mature shoots must be plucked within 5–7 days, because after this period, new leaves unfold from the bud. Overgrown shoots (with more than 4 leaves) must be plucked and discarded: their leaves are no longer tender and they eventually develop floral buds that hinder commercial use. So, both over-plucking and *under*-plucking can damage plant yield, both in the long-term and also within several days, the latter because of overgrown shoots. It is therefore no surprise that Hall (2000, Ch. 2, p. 48) writes: "Pluckers barely have enough time to harvest all the leaves from the fast and furious growth of each bush before the whole cycle has to be repeated."

This little primer makes it clear that underplucking, far from creating an "renewable resource" of leaves for future plucking, can actually lower subsequent output, as overgrown shoots must be plucked and discarded, or the bush pruned. Therefore, low output in August cannot account for the increased output even in the days following the contract change, let alone during the entire month.

4.5. **Learning.** Perhaps workers were unaware that a contract change had occurred, or failed to understand the new terms. In the former (unlikely) case, productivity would be unresponsive to the change. It wouldn't have increased. As it so happens, the new contract was clearly communicated and explained to workers by both the unions and the employer. The most salient feature was the base wage. Of course, workers may not have learned the full magnitude of its increase until month-end when they saw their paychecks, but the direction and substance of the increase were obvious in view of public discussion around the wage legislation. Since this did not affect marginal incentives, however (or at best would have lowered those incentives under a concave utility function for income), learning the precise magnitude of the base wage increase should not have raised productivity.

Alternatively, workers may have had an imperfect understanding of the new piece rate structure. In order to *increase* productivity, however, they would have had to believe that piece rates were now

[15] Coarse plucking is standard practice in these plantations, with output targeted to the wholesale tea auction market.

higher. This is implausible. Unions viewed a key achievement of the new contract as setting the piece rate for the first slab to zero; in other words, doing away with the "penalty" below the standard. This was the most notable change in the piece rate structure and was likely to have been widely advertised.

4.6. **Nutrition.** A large body of empirical evidence, following Leibenstein (1957), suggests that at low levels of income, there is a close connection between wages, nutrition, and manual labor productivity, with low wages perpetuating poor nutrition, which in turn hampers productivity; see Strauss and Thomas (1999) for a comprehensive survey of the literature. The dramatic productivity increase we saw following the contract change was accompanied by an equally dramatic increase in earnings: average monthly wages increased by 73% between Month 0 and Month 1. This could, in principle, have improved nutritional intake, feeding directly into higher productivity.

There are three main arguments that run against this nutrition-wage hypothesis. First, while it is true that daily wages increased, it is equally true that wages — though calculated daily — are only paid at month end, and local credit markets are highly imperfect. It is extremely unlikely, therefore, that workers could have borrowed against month-end wage payments, which themselves are endogenous to worker effort and therefore subject to moral hazard.

Second, on average, earnings prior to the change lie well above the poverty line, which includes minimum caloric needs. We estimate workers' average earnings in the month prior to the contract in 2008 change to be Rs. 2,011; in 2007, it was Rs. 2,132. The rural poverty line in the state in 2008 was about Rs. 600.A rudimentary census of plantation housing residents suggests an average household size of 4. Assuming that there are two workers with similar monthly earnings in each household, this would place all household members at least at 1.6 times the poverty line. At this level of per capita income, it seems unlikely that workers were undernourished.

Third, workers may have modest earnings, but they are entitled by law to non-pecuniary benefits that extend to basic food security. Section 11 of the Plantation Labour Act (1951) stipulates that "[t]he State Government may make rules requiring that in every plantation wherein one hundred and fifty workers are ordinarily employed, one or more canteens shall be provided and maintained by the employer for the use of the workers," and that "[such] rules may provide for . . . the foodstuffs which may be served therein and the charges which may be made therefor." Our study plantation, with almost 2000 workers, clearly falls under the purview of this law. Indeed, as far as the plantations are concerned, the state government stipulated rules in 1955 requiring that food be provided in canteens on a non-profit basis. While the company has discretion regarding the range of food on offer, the non-profit requirement puts an upper bound on prices. Quite apart from this legal requirement, the company has a strategic interest in ensuring that workers are not under-nourished, given their long-term employment relationships.

4.7. **Behavioral Responses.** Our tentative conclusion is that the remaining gap must be chalked up to "behavioral responses." Our data do not permit us to identify a specific mechanism within this category, but some explanations seem more likely than others. Income targeting, for example, seems unlikely. In the month following the contract change, average earnings were Rs. 2,907. This represents an 73% increase over average monthly earnings prior to the contract change. Short of clinging to the dubious assertion that earnings targets increased by over 70% following the contract change, income targeting is an implausible explanation.

The sharp increase in monthly earnings is mainly due to the 44% increase in the fixed component of daily wages. increase in the baseline wage. That wage increased by 30% and the zero piece rate for the first slab added an additional 14 percentage points to this, amounting to a 44% increase in the fixed component of daily wages. This was a large increase, and it may have served to render the baseline wage more salient; see Englmaier, Roider, and Sunde (2014) for evidence on productivity responses to the salience of particular contractual features. Yet, increased salience of the fixed wage component cannot serve alone to increase effort: if utility is concave in income, it can only dampen effort, or at best leave it unchanged.

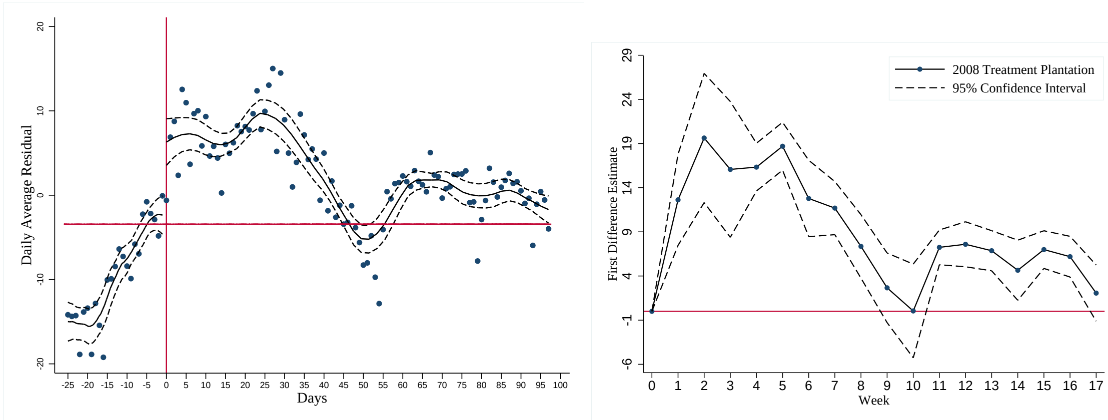
On the other hand, workers may well have responded by increasing their effort in the spirit of gratitude or reciprocity. As a counterargument, it could be pointed out that the 30% base wage increase was effectively mandated by minimum wage legislation, exceeded 3-year inflation (since the last contract) by only about 5 percentage points, and could therefore hardly be interpreted as an act of generosity by the owners. However, doing away with the penalty was not required by law and it increased the effective fixed wage by an additional 14%, so workers may well have perceived this as a gift. It also seems plausible that these factors influenced underperformers more than overperformers. They may have been more likely to feel underpaid and therefore more likely to respond positively to a perceived gift; see Cohn, Fehr, and Goette (forthcoming). Alternatively, they experienced a more generous proportional wage increase, and so may have been "more than proportionately" grateful.

The "perceived generosity" argument also squares with the observation that permanent workers responded more strongly than temporary workers (recall Table 3). Any argument based on dynamic incentives would have gone the other way. For permanent workers, the contract change implies a larger change in permanent income and therefore a more generous act, compared to temporary workers who received the same contract, but with no guarantees after the expiry of their term. Permanent workers, who have a long-term relationship with the employer, are both well-aware of the increased costs the new contract imposes on the owners and may be more likely to perceive the substantial wage increase as being pro-social. Both these factors may also explain their differential response relative to temporary workers; see Hennig-Schmidt, Sadrieh, and Rockenbach (2010) and Hossain and Li (2013), respectively.

Other behavioral explanations are possible, and are consistent with a larger relative response for underperformers. For instance, workers might derive utility from a positive bonus, over and above their utility for money. That would have induced larger responses from low-end workers. Or perhaps those workers responded positively to the removal of the penalty. Finally, it is possible that the short-term increase was a Hawthorne effect of some kind (Mayo, 1933, Ch. 3); specifically, a "novelty effect" (Clark and Sugrue, 1991, p.333). These are references to generally positive but often temporary gains in productivity when a novel change occurs in the workplace. In short, we can't pin down the exact form of the response, and indeed, with the plethora of effects that go under the rubric "behavioral," it would be difficult to do so except under highly controlled conditions. Yet our elimination of the usual "non-behavioral" suspects leaves us comfortable with the conclusion that "behavioral effects" — black-boxed though they may be — played a role in the output increase.

5. Productivity Response in the Longer Term

If our data were limited to this immediate post-contract period, the overall impression would be one of a rather dramatic and counterintuitive increase in output. The increase is robust to the inclusion of



(a) Average Daily Residual

(b) First Difference

Figure 8. The Longer Run: Average Daily Residual in 2008 Treatment Plantation. *Notes.* Panel (a) displays daily residuals until 4 months after the contract change, for the treatment plantation in 2008, accounting for rainfall and other controls. The horizontal line denotes average output in Week 0 in the treatment plantation in 2008. Panel (b) presents estimates of weekly first differences relative to week 0, for each week following the contract change, until Month 4. Dashed lines show the 95% confidence interval.

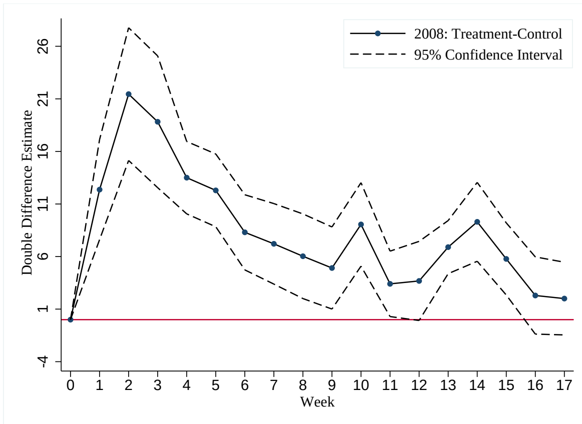
various controls, it flies in the face of the dampening of marginal static incentives, it is not explicable by the use of dynamic incentives such as firing threats, and increased supervisory effort explains at best a fraction of the overall change. We are left, then, with the intriguing hypothesis that a "behavioral" response took place following the contract change, one that reciprocally rewarded a better deal with higher effort. We have discussed these matters in some detail in the previous section.

We now 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.

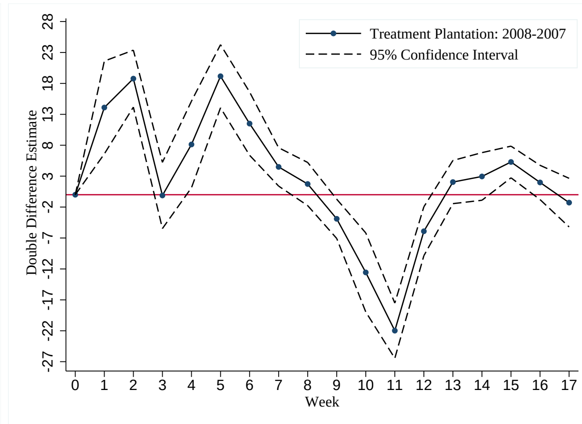
Figure 8 plots residuals; it extends Figure 6 to four months after the contract change. The horizontal line in the first panel denotes average Week 0 residual in 2008, which is almost identical to the average residual in the treatment plantation in Month 0 of 2007 and in the control plantation in Month 0 of 2008. This panel plots the daily residuals to Month 4. It shows that the Month 1 productivity response 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. In contrast, as already mentioned, Figure A2 in the Appendix shows no jump with subsequent reversion for the daily residuals in the treatment plantation in 2007 and the control plantation in 2008. Moreover, by Month 4, these residuals from the two time series are at similar Month 4 levels compared to the treatment plantation in 2008. The phenomenon we uncover is clearly specific to the treatment plantation in 2008: there is a Month 1 jump, which reverses by the end of Month 4, both in itself and relative to our two counterfactuals.

The second panel of Figure 8 provides estimates for the vector ω_1 in the following equation:

(4) $$\text{Output}_{it} = \alpha + \omega_1 \text{Week}_t + \psi \text{ Controls}_{i(t)} + \rho \text{ Rainfall}_t + \varepsilon_{it},$$



(a) Double Difference Relative to Control Plantation



(b) Double Difference Relative to Treatment Plantation in 2007

Figure 9. Average Daily Residual: Double Differences by Week for Two Counterfactuals. *Notes.* This figure presents estimates of double differences by week—the vector ω_2 in equation (5)—for two counterfactual scenarios. Panel (a) uses the control plantation on the same dates in 2008, and Panel (b) uses the treatment plantation on the corresponding dates in 2007. Dashed lines show the 95% confidence interval.

where Week_t is a vector of 17 dummies, one for each of the calendar weeks following the contract change, the exclusion being Week 0. The coefficients on these dummies record how output in each subsequent week changed relative to Week 0; see Column 1 of Table A2 in the Appendix for the estimates. There is a clear downward trend in the coefficients after the initial spike in weeks 2–5, and in the last week of observation, the increase is no longer statistically significant.

We follow up this line of reasoning with two more pooled OLS regressions that consider all observation days from Week 0 onward for 2008 and two explicit counterfactuals: the control plantation on the same dates in 2008, and the treatment plantation on the corresponding dates in 2007. Figure 9 presents the point estimates for the vector ω_2 in the following regression:

$$\text{(5)}\quad \text{Output}_{it} = \alpha + \omega_0\,\text{Treat}_{it} + \omega_1\text{Week}_t + \omega_2(\text{Week}_t \times \text{Treat}_{it}) + \beta\,\text{Controls}_{it} + \rho\,\text{Rainfall}_t + \varepsilon_{it}$$

where Treat_{it} is a dummy variable equal to 1 in the treatment plantation in 2008. Once again the exclusion is Week 0. The double difference estimates of ω_2 in Figure 9 are entirely consistent with the first difference estimates in Figure 8. The increase — relative to either counterfactual — is large and significant in the initial weeks following the contract change, but it decreases in later periods. In the last two weeks before the end of the season, the increase is no longer statistically significant. See Columns 2 and 4 in Table A2 in the Appendix for the double difference coefficient estimates.

For a separate take on the attenuation of the output increase, Figure A3 in the Appendix depicts kernel *densities* of daily residuals — not just averages — in weeks 4, 8, 12 and 16 after the contract change. These are compared to the corresponding distribution of residuals in Week 0. Notice that the two densities are very different in the first panel, but the difference then dissipates in subsequent weeks. The decline is true of both permanent and temporary workers: there is no evidence that dynamic incentives induce temporary workers to exert more effort than permanent workers in either the short or the long run; see Table A3 in the Appendix. In fact, by Week 8, the response for temporary workers turns negative and remains more or less so until the end of the observation period.

We are going to argue that this reversal is fundamentally due to an erosion of the short-term behavioral response. We make this case in three steps. First, we recall from Section 4 that the initial output increase could perhaps be partly explained by heightened supervision. A decline in supervision could therefore explain a part of this reversal, but only a part. In fact, over the subsequent three-month output decline, the relative contribution of supervision appears to *increase*, which suggests that the "behavioral" response dies faster than supervision does. Table A4 in the Appendix continues the estimates in Table 4 for the treatment plantation in later months. We retain the same classification into "underperformers" and "overperformers," but extend Table 4 to the full observation period by regressing output on a full set of controls, along with weekly dummies starting from Week 1, the exclusion being Week 0, separately for overperformers and underperformers. Column 5 in Table A4 presents week-by-week supervisory effort estimates, calculated from the weekly dummy coefficients for these sub-samples. The table shows that the relative contribution of supervisory effort tends to increase over the observation period. Our estimates of supervisory contribution climb from around a quarter in the month following the contract change, to a half in the last month of observation.

Second, it might be argued that the decline in output from Month 2 onwards is the consequence of exuberant plucking in Month 1, which lowered subsequent plant yields in subsequent months. The general point is related to our discussion on intertemporal substitution in Section 4.4, where we worried that underplucking in an earlier period may provide a more abundant yield in subsequent periods. We ruled that possibility by invoking the morphology of the tea plant: this *could not* happen because the plant yields would be lower and not higher in September if they were under plucked in the first few weeks of August. Here by contrast workers *could*, in principle, overpluck the plant today thereby lowering its yield in the future.

But work organization allows us to safely rule out this eventuality. Recall from Section 2.1 that within a given field workers are always assigned to the same row(s) of bushes. This is explicitly designed to preserve dynamic incentives: if, indeed, a worker overplucked a bush in Month 1, *she* would suffer directly from lower yields in subsequent periods. This disincentive for overplucking would be diluted if gangs rotated through completely different fields over the course of time, rarely, if ever, returning to the same field, compelling workers in different gangs to pluck the same rows. This is not the case. As described in Online Appendix O.2, the data are consistent with each gang being assigned to an exclusive set of fields.

In the third and final step, we argue for the erosion of behavioral responses by estimating an entirely standard model off the pre-change data, and using those estimates to predict post-change output. As we shall see, by Month 4 the standard model does well in explaining the observed distribution of output. This exercise is conducted in the next Section.

6. A Structural Approach to Understanding Post-Contract Output

We've seen so far that the large output increase in Months 1 and 2 ebbs away in Months 3 and 4, and by the end of Month 4 we are down to practically no increase at all. At this point the data cease. Whether the ebb continues in later periods is not something we can directly observe.

Where might theory stand on this matter, using a simple, static model of incentives? Recall that the main flattening of the incentive structure came from the elimination of monetary penalties below the minimum standard, which were just piece rates in disguise. The remaining changes pertained to relatively minor shifts in output thresholds for the payment of (unchanged) piece rates. So the basic model would predict roughly unchanged effort above the minimum standard, with a decline below the standard *if* supervisory effort were unchanged. With an increase in supervison to counteract the removal of the penalty, it would not be out of line to predict that output should remain unchanged.[16] Based on this discussion, our best guess is that by the end of Month 4, the simple model of incentives in Section 4.1 could apply quite well. We now back up that assertion by estimating the parameters of a simple structural model off the pre-contract data, and applying those parameters to predict the "out-of-sample" post-contract outcomes.

6.1. **Approach.** We place more structure on the model of Section 4.1. First, we suppose u is linear; specifically, that $u(w) = w$. That is, we ignore the income effects from the change in the baseline wage, and concentrate entirely on marginal incentives. Second, we write the cost function as

$$c(y, \mu) = \frac{\mu}{\theta} \left[\exp(\theta y) - 1 \right], \tag{6}$$

where μ is a shock observed before the effort (or output) decision is made, and θ 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

$$L(x) = \max\{\beta x, 0\}, \tag{7}$$

where x is the shortfall $s-y$ in output from the contractually stipulated standard, and β is the pecuniary equivalent of the unit cost for falling below the standard. With linear u, 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 5. 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 $w(y)$.

We estimate four parameters: β (the loss below minimum standard), θ (the curvature of effort disutility), and the two parameters defining the gamma distribution for the 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.[17] To this end, consider the individual's effort choice problem: for each realization μ, choose y to maximize

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

[16] This discussion neglects possible income effects on effort that could arise because of the hike in the baseline wage. We ignore them, as we are dealing with low-income workers, and indeed, in the model we estimate, we presume that utilities are linear in consumption.

[17] Individual heterogeneity is unavoidable in confronting the data that we have.

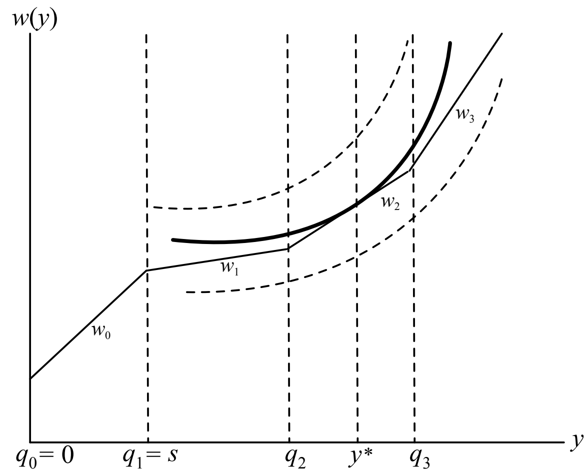


Figure 10. The Individual Optimization Problem. *Notes.* This figure depicts the optimal choice of y in the individual problem. It shows that apart from a necessary first-order condition, a global condition is required to prevent movement to a different incentive slab. This additional condition provides restrictions on the curvature of the disutility of effort which assists in the identification of the parameter θ.

	Threshold q_0		**Standard** $q_1 = s$		**Threshold** q_2		**Threshold** 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 5. Incentives 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. (Classes 1 and 4 were not observed.) 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.

The data has positive output realizations for all workers. Therefore, a necessary (but not sufficient) first order condition must hold for every output not exactly at s or at any of the other thresholds:

$$w_i = \mu \exp(\theta y), \tag{8}$$

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 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 (8).

Two restrictions help us avoid this identification problem. The first is an additional optimality condition: there must be no "global deviations" to other incentive slabs even when the first-order condition (8) on a particular slab applies. Figure 10 explains. The first-order condition is met at y^*, but for global optimality the indifference curve that is tangent to the wage function at y^* must lie everywhere above the wage function elsewhere: this imposes a *separate* restriction on curvature θ which cannot be "exchanged" with μ. Second, 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 — but it still prevents μ and θ from freely "substituting" for each other in the cost function is as described in (6), thereby facilitating identification of each parameter separately. Once θ and the distribution of μ are pinned down, β can be backed out easily from the distribution of observed outputs below the minimum standard s.

6.2. **Estimation.** Here are the details of the estimation strategy, which uses *only* 2007 data from roughly 2,000 workers in the treatment plantation:

1. *Estimate* μ. Fix θ and β. For each contract, characterized by a yield-class/plucking-method combination $k = 1, \ldots, 4$, and for each worker j, use equation (8) to estimate the mean and variance of $\mu(j,k)$, and so the scale and shape parameters for the gamma distribution that determines $\mu(j,k)$.

2. *Simulate 2007 output.* For each worker j and each class-method configuration k, draw $\mu(j,k)$ from the gamma distribution in Step 1. For each output slab i of the form $[q_i, q_{i+1})$, we know the marginal piece rate (because β is fixed from Step 1) and so can calculate the level of output y_{ij} that solves equation (8). Check to see if the y^*_{ij} lies in the slab i. If so, keep the value, otherwise return a missing value. It is easy to prove that for some slab(s), there *must* be non-missing values. For each slab i with non-missing value y^*_{ij}, calculate worker payoff. If all payoffs are negative, set output $y^*_j = 0$. Otherwise, the chosen 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 $(\theta,\beta) = \{(0.1, 0.1), ...(2, 1.0)\}$. Choose (θ,β) so that the empirical distribution of the simulated data most closely resembles the 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 and the simulated cumulative output distribution G. Figure A4 in the Appendix shows that the D-statistic is minimized and the mean difference in output is closest to 0 at $(\theta,\beta) = (0.9, 1)$.[18] In addition, t-tests for the null hypothesis of equal means in the simulated and actual distributions for different values of (θ,β) indicate that only for $\theta = 0.9$ is the p-value different from zero. Moreover, the p-value reaches its maximum at

[18]Tables furnishing the D-statistics and mean differences from which this graph was constructed are available in the Online Appendix Tables O.2 and O.3, respectively.

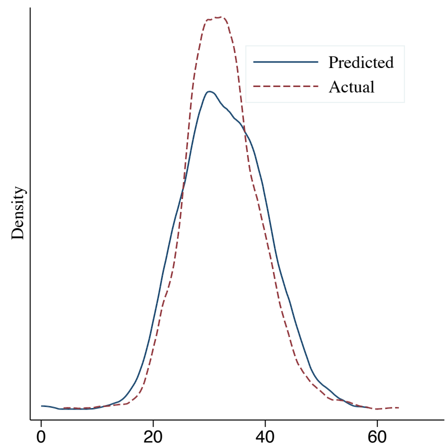


Figure 11. 2007 Kernel Density: Predicted and Actual. This figure compares actual and predicted distribution of output (in kg.) in the treatment plantation in 2007, with $(\theta, \beta) = (0.9, 1)$.

$(\theta, \beta) = (0.9, 1.0)$; see Online Appendix Table O.4. (The p-value falls for larger values of β with $\theta = 0.9$; not reported.) We therefore set $(\theta, \beta) = (0.9, 1)$.

Figure 11 compares the output predicted by the structural model under the old contract for $(\theta, \beta) = (0.9, 1)$ with the actual 2007 data. Table 6 presents summary statistics comparing the actual data (row 1) to the simulated data (row 2). It shows that the simulated and actual distributions are very similar by most of these measures.

Data	Mean	Median	SD	Skewness	Kurtosis	Interquartile Range
Actual	32.39	32.00	6.61	0.35	3.89	8.43
Simulated	32.65	32.36	7.54	0.03	3.27	10.39

Table 6. Comparison of summary statistics from Actual and Simulated 2007 data. *Notes.* Rows 2 denotes sample statistics for the simulated data with $(\theta, \beta) = (0.9, 1.0)$.

6.3. **Predicting Output Post-Change.** The contract change generates a new effective wage function. Table 7 summarizes it. We now use our estimated model to predict the post-change distribution of output. Specifically, we set $(\theta, \beta) = (0.9, 1.0)$ and take 100 random draws of $\mu(j, k)$ from the corresponding gamma distribution for each (j, k). For each draw, we calculate the optimal output y_j^* as 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 to 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.

Figure 12 compares the predicted distribution under the new contract to the actual distribution in 4-week intervals—for Weeks 4, 8, 12 and 16—following the contract change. By week 16, we see almost complete convergence to the predicted distribution.

	Threshold $\hat{q}_0$		**Standard $\hat{q}_1 = \hat{s}$**		**Threshold $\hat{q}_2$**		**Threshold $\hat{q}_3$**	
	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.)	$\beta \rightarrow$		$0.40 \rightarrow$		$0.55 \rightarrow$		$0.85 \rightarrow$	

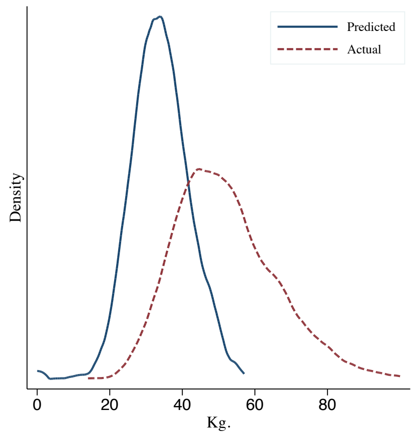
Table 7. Incentive Slabs for Yield Classes 2 and 3, Hands and Shears Under the New Contract. *Notes.* This table describes the piece rates for yield classes 2 and 3 under the new contract, and thresholds at which they became active. (Yield classes 1 and 4 were not observed on any estates.) Each threshold $\hat{q}_i$ 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.

Figure 13 shows deviations of actual output from the mean predicted output of the structural model. Panel (a) compares weekly averages. In the first 4 weeks following the contract change, output is approximately 20 kg. higher than predicted by the structural model. This difference drops by a half to about 10 kg. between weeks 5-8, and again by a half to roughly 5 kg. in the subsequent 7 weeks, until by week 17 — the last week of observation — the prediction of the structural model is statistically indistinguishable from the actual output at the 5% level. Panel (b) shows the deviation of actual output from the mean predicted output of the structural model, in terms of 2-day averages (i.e. for a bin size of 2). 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 5 kg. deviation in Month 3, but drops again in Month 4. In the last week of observation, the predictions of the structural model coincide with actual output.

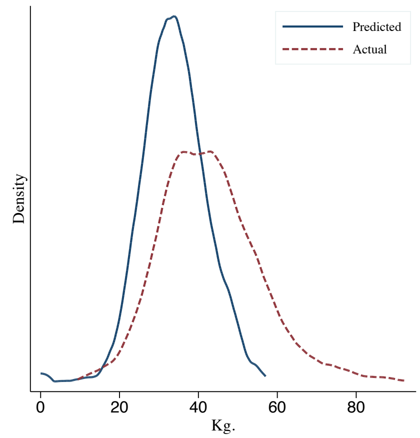
Our structural exercise underlines and supports the earlier observation that an output response, although initially quite dramatic, fades away in the succeeding months. What this exercise adds to the previous observations is that a standard and parsimonious static model of incentives does a remarkably 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: the behavioral response is eroded 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.

7. Conclusion

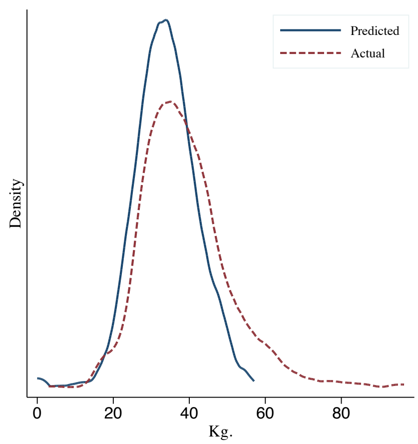
This paper studies the productivity impact of a contract change for tea pluckers in an Indian plantation. The contract raised the baseline wage by over 30%, but lowered marginal incentives by eliminating a linear penalty for underperformance, and (less significantly) by applying the existing piece rates at higher output thresholds. 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



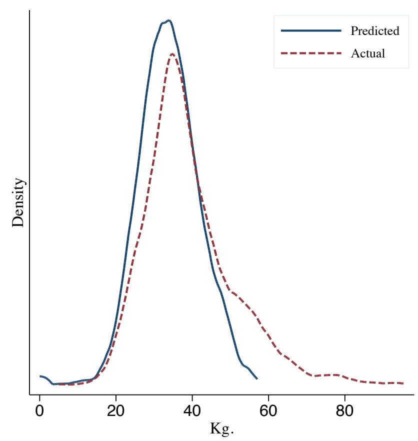
(a) Week 4



(b) Week 8



(c) Week 12

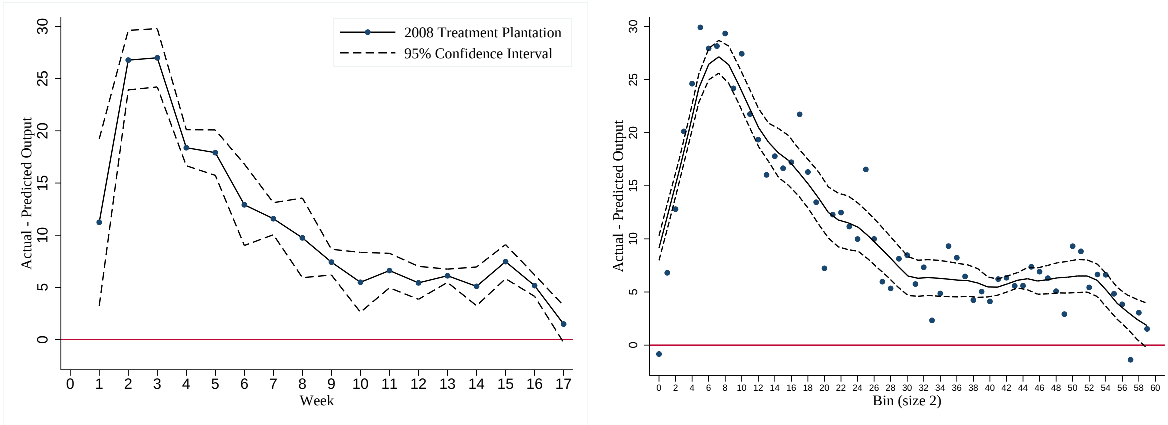


(d) Week 16

Figure 12. 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 Weeks 4, 8, 12 and 16 after the contract change.

and field type. This is 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. There are several possibilities that we attempt to account for. One possible explanation 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 square with this hypothesis for two reasons. First, most workers are permanent and *cannot* be fired by law. Second, the positive productivity response of temporary workers is considerably *lower* than than of permanent workers.



(a) Week by Week

(b) Two-Day Moving Averages

Figure 13. Week-by-Week and Two-Day Differences in Actual and Predicted Daily Outputs. *Notes.* Panel (a) presents the difference between actual average weekly output and average output as predicted by the structural model under the new contract, using parameter values $(\theta, \beta) = (0.9, 1)$. Panel (b) does the same for two-day averages; each dot denotes a 2-day average of actual output (bin size 2). The figure is overlaid with a local polynomial smoothed using an Epanechnikov kernel (solid curve). Dashed curves in both panels denote 95% confidence intervals.

A second hypothesis is that managers stepped up their supervisory effort, 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 estimate that supervision accounts for about a quarter of the output increase, but a large fraction of the increase is still unaccounted for.

We discuss and eliminate other possibilities, such as intertemporal substitution of unplucked leaves, learning effects, or nutritional effects on productivity. That leaves us with 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 raised morale (Solow (1979)). Our data do not permit us to identify a specific mechanism within the behavioral category, but some explanations seem more likely than others. (Income targeting, for example, seems unlikely.) It also seems plausible that these factors influenced underperformers more than overperformers. They experienced a more generous effective proportional wage increase than overperformers. As a result, they may have been "more than proportionately" grateful, and indeed, they do respond more than proportionately.

But 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 this point, our data end. It is entirely possible that the output reversal could have continued. The basic theory predicts either a decline of effort, or no change if supervisory effort was stepped up, which appeared to have been the case. At any rate, the increase is attenuated 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.[19]

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) and Gneezy, Meier, and Rey-Biel (2011). This literature emphasizes how the introduction of the financial incentives might erode more social incentives (reciprocity, gratitude or fair play).

In this paper the baseline relationship is an employment contract. The transaction is monetary to begin with, and gratitude, reciprocity and pro-social behavior are secondary considerations. Do pro-social motivations ultimately hold sway? It would appear not: they matter in the short run, but do not persist. Ultimately, in this labor market setting, monetary incentives come to dominate 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 nurtured. But, particularly in markets where the fundamental relationship is delineated along economic lines, we need to be alert to to the possibility that long- and short-term effects differ, and consequently 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.

[19]In the 5 month period that we study here, there is no significant change in worker participation. As such, this model does not account for selection effects. It is possible that in the longer-run the new contract, with its higher fixed wage component, will attract less-productive workers, and this selection effect will erode productivity even further, as in Lazear (2000).

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.

Charness, Gary and Peter Kuhn. 2011. "Lab labor: What can labor economists learn from the lab?" *Handbook of labor economics* 4:229–330.

Clark, R.E. and B. M. Sugrue. 1991. "Research on Instructional Media, 1978–1988." In *Instructional Technology: Past, Present, and Future*, edited by G.J. Anglin. Englewood, CO: Libraries Unlimited.

Cohn, Alain, Ernst Fehr, and Lorenz Goette. forthcoming. "Fair Wages and Effort Provision: Combining Evidence from a Choice Experiment and a Field Experiment." *Management Science* .

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.

Englmaier, Florian, Andreas Roider, and Uwe Sunde. 2014. "The Role of Communication of Performance Schemes: Evidence from a Field Experiment." Unpublished Manuscript.

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.

Hall, Nick. 2000. *The Tea Industry*. Elsevier, New Holland.

Hennig-Schmidt, Heike, Abdolkarim Sadrieh, and Bettina Rockenbach. 2010. "In search of workers' real effort reciprocitya field and a laboratory experiment." *Journal of the European Economic Association* 8 (4):817–837.

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.

Hossain, Tanjim and King King Li. 2013. "Crowding out in the labor market: a prosocial setting is necessary." *Management Science* 60 (5):1148–1160.

Hossain, Tanjim and John List. 2012. "The Behavioralist Visits the Factory: Increasing Productivity Using Simple Framing Manipulations." *Management Science* 58:2151–2167.

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.

Leibenstein, Harvey. 1957. *Economic backwardness and economic growth*. New York: Wiley.

Mayo, E. 1933. *The Human Problems of an Industrial Civilization*. New York, NY: Macmillan.

Mellström, Carl and Magnus Johannesson. 2008. "Crowding out in blood donation: was Titmuss right?" *Journal of the European Economic Association* 6 (4):845–863.

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.

Strauss, John and Duncan Thomas. 1999. "Health, Nutrition, and Economic Development." *Journal of Economic Literature* 36 (2):766–817.

Appendix

Proof of Proposition 1. We prove Proposition 1 for a more general family of utility functions to which the separable payoff function of Section 4.1 belongs. Denote by $U(w, y, s)$ the worker's utility, which is non-decreasing concave in wage w, non-increasing and concave in effort (or produced output) y and non-increasing in the standard s. Utility U includes the payoff consequence of not meeting the standard: as the standard increases, the pressure on the worker may increase. We further assume that U is submodular in (w, y, s), i.e. the marginal utility of wages is non-increasing in effort and standard. In particular, the payoff function $U(w, y, s) = u(w) - c(y) - L(s - y)$ of Section 4.1 satisfies all these conditions.

Given a wage function w and standard $s \geq 0$, the worker chooses y to maximize her utility. Let $\mathcal{Y}$ and $\hat{\mathcal{Y}}$ be the optimal sets before and after the contract change, respectively, which are assumed to be non-empty.[20] We may restate Proposition 1 as:

If $\hat{y} \in \hat{\mathcal{Y}}$ and $y \in \mathcal{Y}$, then either $\hat{y}, y \in \hat{\mathcal{Y}} \cap \mathcal{Y}$ or $\hat{y} \leq y$.

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

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

The first inequality holds since $U(\cdot, \cdot, \hat{s})$ is submodular with $\hat{w}(y) \geq w(\hat{y})$ from (C1). The second one holds since $U(\cdot, y)$ is concave with $\delta \geq 0$ from (C1). The third one holds since $U(\cdot, y, \hat{s})$ is non-decreasing, where $\hat{w}(y) - \delta \geq w(y)$ 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 $y \in \mathcal{Y}$, and is strict if and only if $\hat{y} \notin \mathcal{Y}$. Thus, either $\hat{y} \leq y$ or if $\hat{y} > y$ then $0 \leq U(\hat{w}(\hat{y}), \hat{y}, \hat{s}) - U(\hat{w}(y), y, \hat{s}) \leq U(w(\hat{y}), \hat{y}, s) - U(w(y), y, s) \leq 0$, which implies that $y, \hat{y} \in \hat{\mathcal{Y}} \cap \mathcal{Y}$. □

[20] This holds for instance when $\lim_{y \to \infty} U(w(y), y, s) = -\infty$, a condition satisfied under the assumptions of Section 4.1.

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

	All	Shears[a]	Hand[b]
Optimal Bandwidth[c]	1.69	4.58	3.32
Bandwidth = 2			
RD Estimate	4.752***	4.966***	3.891***
	(0.429)	(0.577)	(0.603)
2 X Optimal Bandwidth			
RD Estimate	3.999***	5.388***	4.891***
	(0.547)	(0.457)	(0.576)
3 X Optimal Bandwidth			
RD Estimate	5.306***	5.085***	4.330***
	(0.478)	(0.374)	(0.468)
Observations	110,355	87,363	22,992

Table A1. Regression Discontinuity Estimates: Months 0 and 1. *Notes.* Dependent variable is the OLS residual for output levels per person per day. *a*. Observations with shears. *b*. Observations with hand. *c*. Imbens and Kalyanaraman (2012) optimal bandwidth for sharp design. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$.

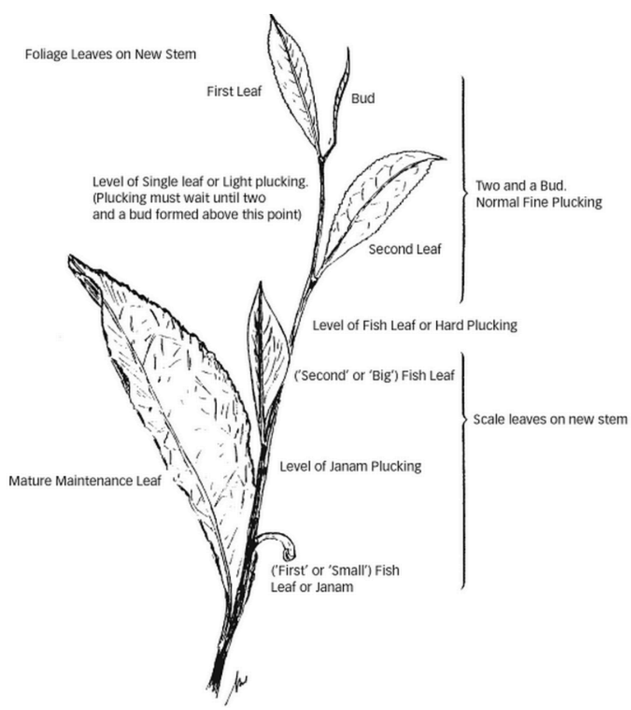
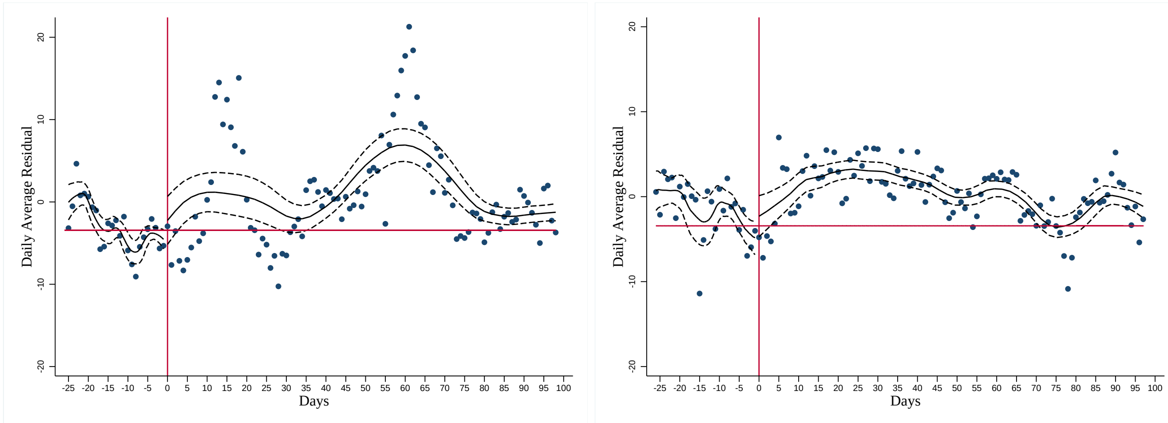


Figure A1. Tea Shoot. *Notes.* Source: Hall (2000, Ch. 2, p. 51). Coarse plucking corresponds to the "Level of Janam Plucking" in this diagram.



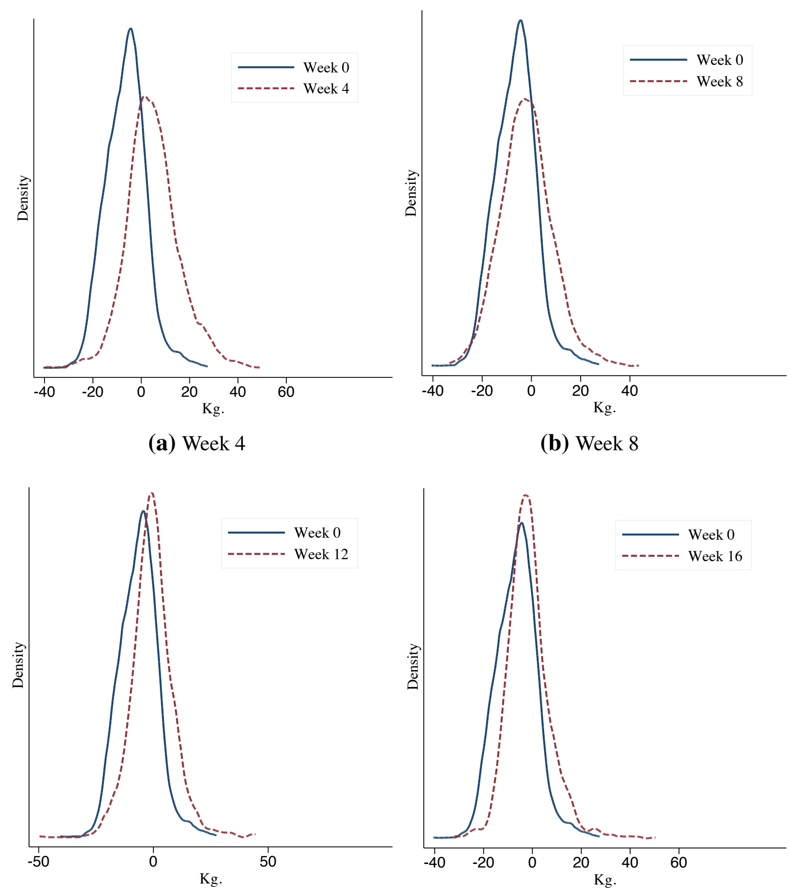
(a) 2007 Treatment Plantation

(b) 2008 Control Plantation

Figure A2. Average Daily Residuals for Counterfactuals. *Notes.* This figure echoes Figure 8 for the treatment plantation, 2007 (Panel a), and the control plantation, 2008 (Panel b). Panel (b) residuals only account for rainfall. The horizontal line is Week-0 average output in the treatment plantation, 2008.

Counterfactual: for DD	Dependent Variable: Output (Daily Kg. Tea) Treatment Plantation in 2007 [a]		Control Plantation in 2008[b]	
	FD (1)	DD (2)	FD (3)	DD (4)
Week 1	12.63***	14.10***	13.18***	12.37***
	(2.601)	(3.807)	(3.079)	(2.401)
Week 2	19.63***	18.77***	23.11***	21.45***
	(3.689)	(2.336)	(3.754)	(3.184)
Week 3	16.07***	-0.12	22.02***	18.82***
	(3.879)	(2.716)	(4.282)	(3.162)
Week 4	16.33***	8.13**	20.88***	13.50***
	(1.364)	(3.551)	(0.920)	(1.737)
Week 5	18.70***	19.16***	23.05***	12.30***
	(1.368)	(2.590)	(1.416)	(1.745)
Week 6	12.78***	11.50***	18.02***	8.31***
	(2.182)	(2.596)	(2.164)	(1.803)
Week 7	11.69***	4.50***	16.07***	7.21***
	(1.516)	(1.572)	(1.253)	(1.931)
Week 8	7.35***	1.73	12.20***	6.03***
	(1.813)	(1.770)	(2.091)	(2.030)
Week 9	2.66	-3.90**	7.51***	4.91**
	(1.987)	(1.609)	(1.986)	(1.969)
Week 10	0.04	-12.56***	5.12*	9.06***
	(2.676)	(3.240)	(2.773)	(2.002)
Week 11	7.24***	-21.98***	11.61***	3.40**
	(0.993)	(2.265)	(0.860)	(1.570)
Week 12	7.59***	-5.91***	10.72***	3.68*
	(1.286)	(2.006)	(1.202)	(1.898)
Week 13	6.86***	2.05	11.44***	6.90***
	(1.146)	(1.772)	(0.966)	(1.274)
Week 14	4.65***	2.95	9.14***	9.30***
	(1.717)	(1.954)	(1.310)	(1.892)
Week 15	6.99***	5.29***	11.77***	5.78***
	(1.076)	(1.298)	(0.756)	(1.729)
Week 16	6.19***	1.98	10.52***	2.30
	(1.162)	(1.406)	(1.027)	(1.852)
Week 17	2.06	-1.28	6.59***	2.01
	(1.604)	(1.994)	(1.161)	(1.744)
No. Obs.	142,612	283,108	149,153	195,583
Adj. R-Squared	0.405	0.477	0.192	0.197

Table A2. Change in Weekly Average Output Relative to Week 0. *Notes.* Each column records the coefficient of a different OLS regression. Columns 1 and 3 present first difference (FD) estimates for ω_1 from equation (4) for the treatment plantation in 2008, and columns 2 and 4 the double difference (DD) estimate for ω_2 from equation (5). *a*. DD comparison plantation is the treatment plantation in 2007, controlling for rainfall and other time-varying characteristics. *b*. DD comparison plantation is the control plantation in 2008, controlling for rainfall only. The absence of additional controls in column 3 accounts for the difference in the FD estimates in columns 1 and 3. Standard errors are clustered by day. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$.



(a) Week 4

(b) Week 8

(c) Week 12

(d) Week 16

Figure A3. 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 weeks 4, 8, 12 and 16 after the contract change.

	Dependent Variable: Output (Daily Kg. Tea)							
	Permanent Workers				Temporary Workers			
	Levels		Logs		Levels		Logs	
	FD	DD	FD	DD	FD	DD	FD	DD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Week 1	13.93***	15.22***	0.34***	0.37***	9.41***	8.01**	0.21***	0.18**
	(2.719)	(4.029)	(0.056)	(0.085)	(2.414)	(3.528)	(0.046)	(0.081)
Week 2	19.70***	17.68***	0.47***	0.32***	18.60***	20.15***	0.39***	0.41***
	(3.997)	(2.428)	(0.079)	(0.061)	(3.561)	(2.719)	(0.065)	(0.073)
Week 3	17.08***	0.15	0.43***	-0.10	12.13***	-3.58	0.28***	-0.19***
	(4.115)	(2.962)	(0.081)	(0.068)	(3.618)	(2.808)	(0.067)	(0.059)
Week 4	17.47***	7.83**	0.45***	0.12	11.87***	7.01**	0.29***	0.14*
	(1.421)	(3.593)	(0.031)	(0.085)	(1.710)	(3.484)	(0.032)	(0.083)
Week 5	19.67***	19.64***	0.50***	0.46***	16.16***	16.36***	0.37***	0.35***
	(1.388)	(2.655)	(0.029)	(0.070)	(1.772)	(3.111)	(0.033)	(0.081)
Week 6	13.60***	11.80***	0.37***	0.25***	10.21***	10.10***	0.24***	0.21***
	(2.314)	(2.736)	(0.049)	(0.062)	(2.492)	(3.152)	(0.049)	(0.069)
Week 7	13.36***	5.85***	0.38***	0.13***	6.85***	-1.15	0.19***	-0.03
	(1.569)	(1.637)	(0.034)	(0.042)	(1.560)	(1.898)	(0.032)	(0.051)
Week 8	8.65***	2.74	0.27***	0.01	2.45	-3.20	0.09**	-0.12**
	(1.977)	(1.879)	(0.044)	(0.047)	(2.177)	(2.460)	(0.046)	(0.061)
Week 9	3.67*	-3.29*	0.16***	-0.15***	-0.94	-7.46***	0.03	-0.20***
	(2.096)	(1.825)	(0.043)	(0.050)	(2.140)	(2.091)	(0.042)	(0.052)
Week 10	1.14	-11.65***	0.11*	-0.34***	-2.87	-16.92***	-0.00	-0.43***
	(2.793)	(3.568)	(0.061)	(0.082)	(2.450)	(2.730)	(0.048)	(0.064)
Week 11	7.69***	-22.53***	0.25***	-0.55***	5.30***	-21.94***	0.16***	-0.51***
	(1.093)	(2.247)	(0.026)	(0.049)	(1.329)	(2.999)	(0.027)	(0.063)
Week 12	9.16***	-7.09***	0.26***	-0.25***	2.01	-6.43**	0.07**	-0.20***
	(1.348)	(2.096)	(0.030)	(0.053)	(1.432)	(2.589)	(0.030)	(0.063)
Week 13	7.56***	2.57	0.24***	0.02	4.03***	-1.09	0.12***	-0.01
	(1.282)	(1.985)	(0.031)	(0.058)	(1.221)	(1.958)	(0.025)	(0.052)
Week 14	5.73***	4.14**	0.20***	0.10**	-0.43	-2.32	0.04	-0.04
	(1.788)	(1.962)	(0.040)	(0.052)	(2.026)	(2.538)	(0.042)	(0.066)
Week 15	8.76***	6.73***	0.27***	0.16***	1.55	-0.04	0.07**	0.02
	(1.164)	(1.361)	(0.027)	(0.038)	(1.229)	(1.896)	(0.026)	(0.051)
Week 16	7.87***	3.54**	0.24***	0.05	0.23	-5.58***	0.03	-0.16***
	(1.278)	(1.547)	(0.030)	(0.045)	(1.274)	(1.663)	(0.025)	(0.046)
Week 17	3.42**	0.64	0.13***	0.00	-2.53	-8.08***	-0.03	-0.20***
	(1.537)	(1.960)	(0.040)	(0.055)	(1.951)	(2.518)	(0.046)	(0.066)
No. Obs.	106,195	211,096	106,195	211,096	36,417	72,012	36,417	72,012
Adj. R-Squared	0.406	0.476	0.444	0.529	0.459	0.538	0.468	0.560

Table A3. Change in Weekly Average Output Relative to Week 0 by Permanent and Temporary Workers. *Notes.* This table provides the analog of Table A2, 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. $^{*}p < 0.10$,$^{**}p < 0.05$,$^{***}p < 0.01$.

	Underperformers		Overperformers		S(%)
	Levels	Logs	Levels	Logs	
	(1)	(2)	(3)	(4)	(5)
Week 1	18.02***	0.46***	7.46***	0.16***	38
	(2.808)	(0.062)	(2.391)	(0.046)	
Week 2	24.33***	0.61***	14.81***	0.31***	24
	(3.639)	(0.076)	(4.359)	(0.078)	
Week 3	22.11***	0.59***	10.60**	0.23***	34
	(3.906)	(0.081)	(4.316)	(0.078)	
Week 4	20.83***	0.57***	13.28***	0.29***	24
	(1.525)	(0.037)	(1.623)	(0.031)	
Week 5	22.70***	0.61***	16.34***	0.37***	18
	(1.444)	(0.033)	(1.639)	(0.030)	
Week 6	16.00***	0.47***	10.74***	0.24***	24
	(2.078)	(0.048)	(2.779)	(0.052)	
Week 7	15.08***	0.46***	9.69***	0.23***	43
	(1.701)	(0.040)	(1.751)	(0.035)	
Week 8	10.81***	0.36***	5.26**	0.15***	32
	(1.902)	(0.045)	(2.210)	(0.044)	
Week 9	6.68***	0.26***	-0.70	0.03	72
	(2.221)	(0.049)	(2.184)	(0.041)	
Week 10	4.87*	0.23***	-4.04	-0.05	100
	(2.666)	(0.061)	(2.899)	(0.056)	
Week 11	10.53***	0.36***	5.12***	0.15***	32
	(1.154)	(0.031)	(1.457)	(0.029)	
Week 12	12.05***	0.37***	4.51***	0.12***	41
	(1.460)	(0.035)	(1.498)	(0.030)	
Week 13	10.25***	0.35***	4.55***	0.12***	39
	(1.248)	(0.033)	(1.479)	(0.030)	
Week 14	8.76***	0.31***	1.58	0.07*	53
	(1.661)	(0.041)	(1.958)	(0.039)	
Week 15	12.02***	0.38***	3.24**	0.09**	52
	(1.272)	(0.032)	(1.614)	(0.036)	
Week 16	10.08***	0.33***	3.59**	0.09***	47
	(1.259)	(0.034)	(1.522)	(0.031)	
Week 17	6.11***	0.22***	-1.58	-0.01	100
	(1.632)	(0.044)	(1.992)	(0.045)	
No. Obs.	76,360	76,360	55,247	55,247	
Adj. R-Squared	0.413	0.451	0.460	0.505	

Table A4. Supervisory Effort over Weeks 1-17. *Notes.* This table presents OLS coefficient estimates for weekly dummy variables—the exclusion being Week 0—in levels and logs for the 2008 Treatment Plantation. Each column pertains to a different regression, each of which include the full set of controls. 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"). Column 6 displays the estimated share of output increase S due to supervision, as given by equation (3). Standard errors are clustered by day. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$.

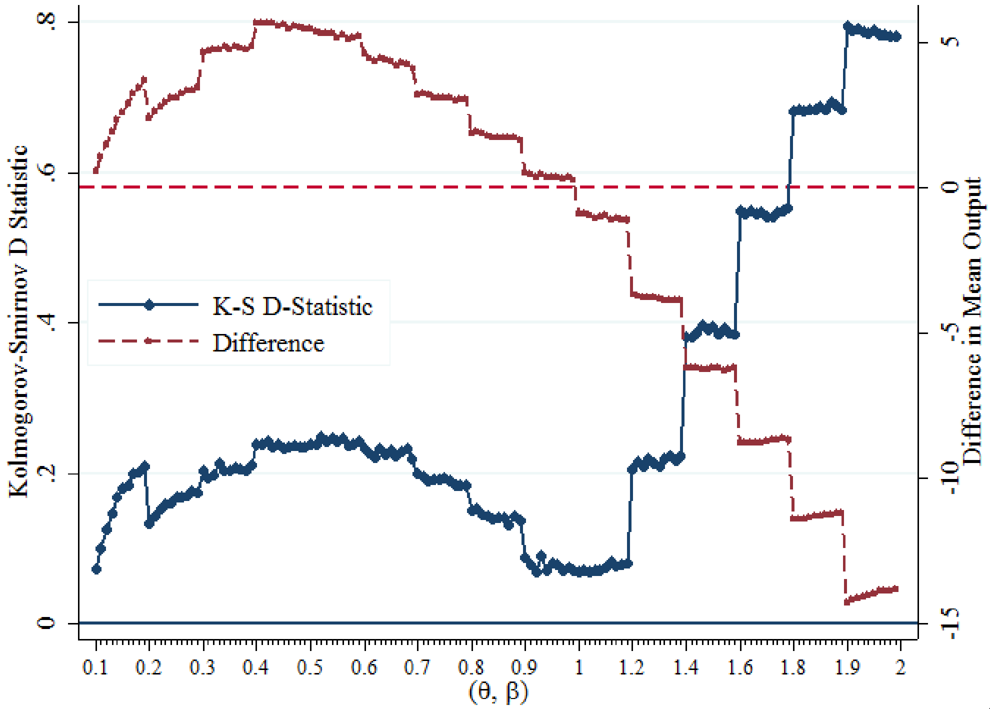


Figure A4. Kolmogorov-Smirnoff D Statistic and Difference between Predicted and Actual Mean Output by (θ, β) Combinations. *Notes.* This figure depicts two statistics testing the difference in predicted and actual output for 2007. The left y-axis denotes the Kolmogorov-Smirnoff distance statistic $D = \sup_y |F(y) - G(y)|$, 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 shows the difference in mean output; the absolute value of this difference is 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.