Incentives, Productivity and Selection in Labor Markets: Evidence from Rural Malawi*

Raymond P. Guiteras
University of Maryland†

B. Kelsey Jack
Tufts University‡

October 2012

Abstract

An observed positive relationship between compensation and productivity cannot distinguish between two channels: (1) an incentive effect and (2) a selection effect. We use a simplified Becker-Degroot-Marschak mechanism (BDM), which provides random variation in piece rates conditional on revealed reservation rates, to separately identify the two channels in the context of casual labor markets in rural Malawi. We find that the incentive effect is more important than the selection channel. Women have lower reservation piece rates and higher productivity. Explicit monitoring incentives improve output quality at the expense of quantity, though workers appear to behave according to a gift exchange model in the absence of explicit quality monitoring. Results provide no evidence for income targeting but do suggest that productivity is influenced by the productivity of a worker's peers.

*We thank seminar and workshop participants at the University of Chicago, University College London / L.S.E., Duke University and RAND. James Berry, Jessica Goldberg, John Ham and John List provided useful comments. John Anderson and Stanley Mvula were invaluable in leading the field team. We gratefully acknowledge financial support from the University of Maryland Department of Economics.

†Email: guiteras@econ.umd.edu

‡Email: Kelsey.Jack@tufts.edu

1 Introduction

An observed positive relationship between compensation and productivity cannot distinguish between two channels: (1) an incentive effect, i.e. a higher piece rate motivates productivity; (2) a selection effect, i.e. higher-productivity workers have better outside options, so increasing compensation leads to a higher-productivity pool of workers. These channels have different implications for efficiency: the first effect represents a causal effect and increased output, whereas the second could be simply a zero-sum reallocation of labor. In the absence of information on reservation piece rates, even randomly assigned piece rates cannot solve this problem unless worker pools are held fixed.

We use a simplified Becker-Degroot-Marschak mechanism (Becker et al. 1964) to separately identify the two channels in the context of casual labor markets in rural Malawi. Workers choose the minimum piece rate at which they are willing to accept a contract to perform a simple task, sorting beans by type and quality. A piece rate offer is then generated randomly, determining whether the worker is given a contract and, if so, the piece rate. Random assignment to a quality monitoring treatment provides exogenous variation in the workers' incentives to trade off quantity of output for quality. The experiment is conducted over four consecutive days in each of twelve villages, spanning both the low and high labor demand seasons.

We use the resulting data set to examine a number of determinants of worker productivity in the casual labor market. First, we separate the effects of a higher piece rate on the type of worker contracted from the direct effect of a higher piece rate on productivity. We compare effort allocation toward quantity versus quality with and without explicit incentives for quality. Results show that workers are highly responsive to the piece rate in terms of the quantity of output produced, and that output quantity and quality are substitutes. Consequently, the introduction of explicit quality monitoring improves the average quality of production but at a quantity cost: workers are slower but more precise when errors are penalized. These intensive margin effects are larger and more precisely estimated than the effect of the piece rate or quality monitoring on worker selection into the task.

Second, higher piece rates create a clear incentive for a higher quantity of output. In the absence of explicit monitoring incentives, quality should therefore deteriorate as piece rates increase if quantity and quality are substitutes. We find worker behavior to be consistent with a gift exchange model in which workers share some of the surplus associated with higher piece rates through output quality. Specifically, quality does not deteriorate as piece rates increase, even in the absence of explicit quality monitoring.

Third, the design is stratified by gender, which is an important determinant of labor market outcomes in rural Malawi. Both extensive and intensive margin effects are sensitive to gender. In our setting, women both accept lower piece rates and produce more and higher quality output than do men. These gender differences are consistent with the value of the outside option available to women and its similarity to the observed task.

Finally, we exploit two sources of random variation to test other potential "behavioral" determinants of labor productivity: target earnings and peer effects. First, we use variation in the piece rate across days to test for reference-dependent income targeting and find little evidence that workers adjust their productivity based on cumulative earnings within the day. Second, we use variation in work group piece rates to test for peer effects. We find that the average exogenously determined piece rate for other workers in the work group has a positive effect on worker productivity, but the effect decreases as the size of the work group increases.

The study design offers a novel approach to characterizing labor market supply and productivity parameters in a data-scarce environment. While the specific parameters may not generalize beyond the study context, the findings provide several pieces of unique evidence. First, our design cleanly separates the selection margin from the incentive margin, and shows that higher piece rates are more important for generating high effort than for attracting high quality workers. Second, while previous studies have shown that men and women face different labor market opportunities, we characterize the margins on which these differences operate. Third, we offer clear evidence on a number of different behavioral determinants of labor market outcomes, including gift exchange, target earnings and peer effects. The study therefore relates to a number of different strands of literature in both labor and development economics.

Both observational and experimental studies have examined the relative importance of worker selection and worker effort in determining the total productivity effect of incentives (Lazear 2000; Dohmen and Falk 2011; Eriksson and Villeval 2008).[1] While these studies suggest that selection is an important determinant of worker ability, selection effects may be less important for determining overall output in settings where effort can be measured and tied directly to compensation. In these cases, the optimal piece rate depends on the elasticity of effort with respect to the piece rate (Stiglitz 1975).[2] The observability of effort will directly affect the firm's choice of incentive scheme, and the firm may value dimensions of effort that are difficult to measure, such as quality.[3]

The ability to shirk on difficult to observe dimensions may also affect both worker selection and worker effort. Existing evidence on changes in worker composition based on variation in supervision or monitoring is scarce.[4] While monitoring is typically used to increase effort levels (see Prendergast (1999)), some theoretical

[1]In the context of a U.S. factory producing windshield glass, Lazear (2000) concludes that approximately half of the productivity gains from a switch from wages to piece rates is due to changes in worker composition, i.e. selection. Dohmen and Falk (2011) document sorting on both productivity and other worker characteristics in a laboratory setting. Eriksson and Villeval (2008) also use a laboratory setting to generate exogenous variation in incentive schemes and observe both sorting and effort effects.

[2]For example, in a study of workers in a tree planting firm in British Columbia, Paarsch and Shearer (1999) estimate an elasticity of effort, as measured by the number of trees planted per day, with respect to the piece rate.

[3]A large literature examines the effects of incentive schemes on worker effort choice (e.g. Bandiera et al. (2005, 2010)), but cannot typically identify both worker effort and worker composition effects.

[4]Related work has examined the productivity impacts of monitoring. For example, Nagin et al. (2002) varied the information about monitoring provided to call center employees engaged in piece rate work. The cost of detection in their setting is a pay deduction, though employees presumably also adjust effort to account for the risk of dismissal. The authors find support for "rational cheating": workers shirk more when the probability of detection is lower. However, effort choice does not systematically vary with the value of the outside option. Higher powered managerial incentives may also generate more careful screening of high ability workers into employment, affecting the composition of worker abilities (Bandiera et al. 2007). Kaur et al. (2011) present evidence on worker demand for self-control. In their field study, workers select into incentive schemes that impose more effort commitment.

work and laboratory studies suggest that monitoring may crowd out workers' intrinsic motivation (Dickinson and Villeval 2008).[5] This prediction relies on a gift exchange model of worker effort, in which workers reciprocate higher wages by increasing effort on other dimension (Akerlof 1982). Empirical support for the gift exchange model has been mixed (Fehr and Schmidt 2004; Fehr et al. 2008; Gneezy and List 2006; Bellemare and Shearer 2009).[6] We are not aware of evidence on gift exchange in developing country labor markets, either from the laboratory or the field.

In addition to reciprocity, other non-standard preferences have been shown to affect worker effort choices. First, a substantial theoretical and empirical literature explores a model of reference dependent preferences, which leads to daily or weekly earnings targets that determine worker effort (Camerer et al. 1997; Farber 2005; Doran 2010; Fehr and Goette 2007).[7] Second, workers may be sensitive to the effort choices or incentives of their peers (Gaechter et al. 2010). Peer effects may be driven by the production technology (e.g., Mas and Moretti (2009)) or the incentive scheme (e.g., Bandiera et al. (2005)). Generally, studies of peer effects find a positive spillover from highly productive workers to less productive colleagues, particularly where incentives to free ride are small. To date, the empirical findings on both target earnings and peer effects are based on developed country studies, and may not

[5] Dickinson and Villeval (2008) find evidence in a laboratory setting that monitoring crowds out worker reciprocity but only when the relationship between the employer and the worker are long-term. Related to this, in a field experiment, Landry et al. (2011) show that conditional performance incentives lead to higher performance than do unconditional performance incentives. The authors interpret this as evidence against crowding out of reciprocity through explicit incentives.

[6] Outside of the laboratory, Gneezy and List (2006) document behavior consistent with the gift exchange model when the gift is in the form of a higher flat wage, but only for the first few hours of work following receipt of the gift, after which effects quickly dissipate. Bellemare and Shearer (2009), on the other hand, find evidence of higher effort on the day that workers receive a surprise gift bonus on top of their usual piece rate pay, though the reciprocity was not enough to make the gift profitable to the firm.

[7] See Section 2 for further discussion of the empirical literature. We are not aware of any studies to date that have directly tested the target earnings model in a developing country setting. Giné et al. (2010) use observational data on fishing decisions in Southern India and find that recent earnings affect participation decisions, but have less of an effect on labor supply than do expected earnings. Camerer et al. (1997) note that negative wage elasticity of labor supply may arise if labor and leisure are strong complements, which only occurs if daily consumption is determined by daily income. While this is less likely to explain empirical patterns in developed countries, it may be more relevant in developing country settings, where individuals are cash constrained.

extend to developing country settings in which workers are liquidity constrained, income generating opportunities are seasonal and social ties may play a stronger role in shaping behavior.

The paper proceeds as follows. Section (2) provides a simple theoretical model to motivate the experiment and frame the empirical analysis. Section (3) describes the experimental design and implementation. Section (4) presents the empirical results. Section (5) concludes.

2 Model

We describe a simple model of effort choice under a piece rate scheme that builds on Nagin et al. (2002). The model generates predictions about selection, effort, and the effects of monitoring. We extend the conceptual framework to discuss reciprocity, threshold earnings and peer effects.

2.1 Setup

A firm values output quantity, q, and loses revenue from output quality, Q, below a quality threshold $\bar{Q}$. It offers a piece rate r to workers for production of q and may also choose to monitor Q using a monitoring technology M to detect quality below a pre-determined quality threshold $\bar{Q}$. The monitoring technology is binary ($M \in \{0, 1\}$)), costs the firm θ when used, and is perfectly able to detect $Q < \bar{Q}$.

Workers are offered a piece rate r for each unit of output of acceptable quality ($Q > \bar{Q}$) and choose to allocate effort toward production of q and Q, which together determine the cost of effort $c(q, Q)$, which is increasing and convex in each argument. Workers are indexed by their productivity, $\gamma \geq 1$, which for simplicity we model as entering multiplicatively and symmetrically between quantity and quality:

$$c\left(q, Q; \gamma\right) = c\left(q, Q\right) / \gamma.$$

If the firm chooses to implement the monitoring technology, it affects worker payoffs only when $Q < \bar{Q}$. Workers maximize utility

$$U = r[q - M(\bar{Q} - Q)1\left\{Q < \bar{Q}\right\}] - \frac{c(q, Q)}{\gamma}.$$

The first order conditions determining the worker's utility-maximizing level and allocation of effort (q^*, Q^*) are:

$$\begin{aligned} r- &\quad \frac{\partial c\left(q^*, Q^*\right)}{\partial q} \quad \frac{1}{\gamma} \quad = 0 \\ rM- &\quad \frac{\partial c\left(q^*, Q^*\right)}{\partial Q} \quad \frac{1}{\gamma} \quad = 0 \quad \text{if } Q^* < \bar{Q} \\ &\quad \frac{\partial c\left(q^*, Q^*\right)}{\partial Q} \quad \frac{1}{\gamma} \quad = 0 \quad \text{if } Q^* = \bar{Q}. \end{aligned}$$

Note that the FOC on Q implies that $Q^* = 0$ when $M = 0$ and $Q^* < \bar{Q}$.[8] In the absence of monitoring that makes $\bar{Q}$ explicit, workers may still assume some quality threshold below which they will be penalized in some way. Average Q is likely to fall in the absence of monitoring as workers gain experience with $M = 0$.[9]

2.2 Selection

As the piece rate and monitoring technology are varied, workers will choose whether or not to accept a contract according to their utility under the contract and their outside option, $U_A(\gamma)$. Workers accept the contract if

$$U = r[q^* - M(\bar{Q} - Q^*)1\left\{Q^* < \bar{Q}\right\}] - \frac{c(q^*, Q^*)}{\gamma} \geq U_A(\gamma).$$

We index the outside option by the productivity parameter to emphasize that a worker's outside option will depend on her overall productivity, which may be re-

[8] Note that $Q^* > \bar{Q}$ cannot be optimal for the worker, since he is not paid for quality above the threshold.

[9] While the implementation of the experiment had no quality threshold when $M = 0$, we observe no worker choosing $Q^* = 0$, which suggests that they continue to believe that they will be penalized for sufficiently low levels of Q.

flected in γ, her productivity in this task. While we cannot sign this relationship unambiguously, $\partial U_A(\gamma)/\partial\gamma > 0$ if workers who are more productive in this task have better outside options. This is likely to be the case for workers with outside options that reward similar skills. The reservation rate, or the worker's minimum willingness to accept, is therefore given by

$$\underline{r} = \frac{U_A(\gamma) + c(q^*, Q^*)/\gamma}{[q^* - M(\bar{Q} - Q^*)1\left\{Q^* < \bar{Q}\right\}]}.$$

We are interested in comparative statics regarding monitoring (the relationship between $\underline{r}$ and M) and sorting (the relationship between $\underline{r}$ and γ). The first is relatively simple: $\underline{r}$ is clearly increasing in M. To understand sorting, we consider the partial derivative $\partial\underline{r}/\partial\gamma$.[10] The sign of this derivative depends on two competing effects from the two terms in the numerator: $\partial U_A(\gamma)/\partial\gamma > 0$ and $\partial\left(c(q^*, Q^*)/\gamma\right)/\partial\gamma < 0$. That is, the minimum piece rate required to attract a high quality worker depends on the comparative magnitudes of (a) the better outside option available to a higher quality worker and (b) the lower cost of effort in this task to a higher quality worker.

From the FOCs, q^* and Q^* are ambiguous in their response to the piece rate, and depend on quantity-quality tradeoffs ($\partial^2 c(q,Q)/\partial q\partial Q < 0$). The minimum piece rate increases with monitoring: $\partial\underline{r}/\partial M > 0$.

2.3 Effort

The worker's effort choice is given by the first order conditions above. In the absence of monitoring ($M = 0$), a higher piece rate unambiguously increases effort in the quantity dimension, $\partial q^*/\partial r > 0$.

The effect of monitoring on effort depends on whether quantity and quality are complements or substitutes in the worker's production function. For the task we study, they are likely to be substitutes ($\partial^2 c(q,Q)/\partial q\partial Q < 0$), in which case monitoring increases investment in output quality at a cost of some reduction in output

[10] By the envelope theorem, it is sufficient to consider the partial derivative, i.e. we do not need to consider changes in q^* and Q^* as γ changes since these are already optimized.

quantity. The worker balances the marginal costs and benefits of allocating effort between quality and quantity. From the first order conditions and the binary monitoring technology, the worker sets:

$$\frac{\partial c\left(q^{*},Q^{*}\right)}{\partial q}=\frac{1}{M}\frac{\partial c\left(q^{*},Q^{*}\right)}{\partial Q}.$$

Higher productivity workers are predicted to produce both higher quantity and higher quality output:

$$\begin{aligned}\tfrac{\partial q^{*}}{\partial\gamma} &= \tfrac{1}{\gamma^{2}}\tfrac{\partial c(q^{*},Q^{*})}{\partial q^{*}} &&> 0\\ \tfrac{\partial Q^{*}}{\partial\gamma} &= \tfrac{1}{\gamma^{2}}\tfrac{\partial c(q^{*},Q^{*})}{\partial Q^{*}} &&> 0,\quad Q^{*}<\bar{Q}.\end{aligned}$$

2.4 Gender

Worker gender can affect both sorting and effort if there is a clear relationship between gender and productivity, γ, or how productivity relates to the worker's outside option, U_A.[11] If females are more likely to face outside employment that values similar skills, then higher piece rates are more likely to select for high productivity female workers than male workers.[12]

Thus, at a given piece rate, the average productivity of men and women may differ, with implications for the production of both quantity and quality.[13]

[11] Studies in development economics on gender differences in labor supply date back several decades, and consistently document differences in supply elasticities by gender (Bardhan 1979; Rosenzweig 1978). In a recent field experiment, Goldberg (2011) randomly varies daily wages in rural Malawi and finds similar supply elasticities for men and women. In agricultural labor markets, gender differences may also vary with the labor season, if men's and women's labor inputs become valuable at different points in the agricultural cycle.

[12] From the participation constraint, the reservation wage is increasing in the outside option if $q^{*}>M(\bar{Q}-Q^{*})1\left\{Q^{*}<\bar{Q}\right\}$. On the other hand, Gneezy et al. (2003) provide compelling evidence that selection is not the full story: women perform less well under competitive incentives even when they perform equally well in non-competitive settings.

[13] Selection into different incentive schemes may be affected by factors other than the reservation wage. For example, Niederle and Vesterlund (2007) find that, in the laboratory, women choose competitive compensation schemes less often than men, even when their potential earnings are higher under competitive incentives. Gneezy et al. (2009) conduct a similar study across matrilineal

2.5 Reciprocity

In the absence of explicit incentives to produce quality ($M = 0$), higher piece rates should result in lower quality if quantity and quality are substitutes. However, a number of studies have generated evidence for gift exchange based efficiency wages, whereby workers reciprocate wages above the minimum rate by sharing surplus (Fehr et al. 2008).[14] For a given level of worker productivity, observing non-decreasing quality as piece rates increase, $\partial Q^*/\partial r \geq 0$ in the absence of explicit quality incentives suggests that the worker shares some of the surplus from higher piece rates with the firm.

2.6 Target earnings

The first order conditions indicate that workers will continue to produce output until their marginal cost of effort ($\partial c/\partial q$) equals the marginal benefit (r). If, instead, workers have a daily earnings target of $rq^T = T$, workers will slow their effort, or stop, once T is reached, if $q^T < q^*$.[15] One clear prediction is that quantity produced will be lower when the worker faces a higher piece rate. If a worker's primary source of disutility is time spent working, workers will stop working earlier on days with

and patriarchal societies and find that the lab results are reversed (women are more competitive) in the matrilineal society. Together these and numerous other studies suggest that men and women may sort into different incentive schemes based on underlying preferences. Women may also be differentially responsive to other aspects of the employment relationship, including the relationship with the employer (reciprocity) or relationships with colleagues (peer effects).

[14]Piece rates above reservation rates may improve output on other effort dimensions (output quality) for two reasons. First, higher rates make the threat of firing more costly, decreasing the firm's best-response audit probability, which could increase profits, even net of the greater payments to workers (Shapiro and Stiglitz 1984). Second, higher rates may induce reciprocity-based gift exchange, under which the worker shares some of the surplus by increasing other effort dimensions (Akerlof 1982).

[15]A number of studies in developed countries (most involving taxi drivers) have studied a target earnings model of labor supply, in which workers set an earnings target for the day or the week and stop working or reduce effort once they have met that target. Camerer et al. (1997) explain a negative labor supply elasticity for taxi drivers in New York with a model of reference dependent preferences. Others have returned to the puzzle with more data (Farber 2005) and exogenous variation in wages (Doran 2010), and found mixed results. Fehr and Goette (2007) randomly vary commission rates (share of revenue) for bike messengers and find labor supply decisions that are consistent with a target earnings model.

a higher piece rate. If disutility is mostly from intensity of effort, we would expect workers to work more slowly. Both would result in lower output when piece rates are high.

2.7 Peer effects

In a standard productivity model, workers respond to their own incentives to exert effort. However, worker effort costs may also be affected by non-standard factors such as the effort choices of their peers. Two possible channels are highlighted in the existing literature.[16] First, greater effort by peers may decrease a worker's own effort cost of effort. We model this as c_i' is decreasing in $q^*_{\sim i}$ or $Q^*_{\sim i}$, where $\sim i$ denotes worker i's peer. Second, workers may respond to relative incentives because of fairness norms, which affect the cost of effort. The predictions are ambiguous. On the one hand, workers who receive piece rates below those of their peers may be discouraged from exerting effort: $\partial c/\partial(r_i - r_{\sim i}) > 0$. Alternatively, workers who receive piece rates greater than those of their peers may suppress productivity out of fairness concerns: $\partial c/\partial(r_i - r_{\sim i}) < 0$. Peer effects are most likely to affect output quantity, which is more observable that quality.

3 Experimental design and implementation

To study productivity in the casual labor market, we create new demand for casual labor under controlled conditions that generate random variation in worker incentives. The context is informal day labor markets in rural Malawi, where such work is

[16] Bandiera et al. (2010) estimate the effects of social ties on those working near each other, in a piece rate setting where any externalities between workers are purely social. They find that workers adjust their effort upward or downward to more closely match the productivity of those with whom they have close social ties. In a more laboratory-like setting, Falk and Ichino (2006) find evidence of peer effects even in the absence of social ties. Though the incentives in their set up offer no rewards for cooperation, peer effects increase overall output by raising the productivity of the least productive workers. Relative wages may also matter. Gaechter et al. (2010) show that, in a laboratory experiment, peers' wages and effort choices affect one's own effort choice, but only when these are observable.

called *ganyu*. In Malawi, like in many rural agricultural settings in developing countries, labor markets are highly seasonal. Households both buy and sell labor, both for daily wages (Goldberg 2011) and in piece-rate-based jobs. In our study, workers are hired to sort harvested, dried beans into eight categories.[17] Sorted beans are sold for about 50 percent more than are mixed beans. This task is well-suited to our study for several reasons: it is a familiar, common task for *ganyu*, typically compensated by piece rates; output has clear quantity and quality dimensions; it is a task where output can respond strongly to effort (in this case, focus and concentration) but effort is not physically taxing.[18]

3.1 Experimental design

The subjects are first invited to a "day zero" training session at which the task is explained and they are shown examples of the categories of beans.[19] Then, on each of the next four days, we obtain each participant's reservation piece rate $\underline{PR_i}$ (truthful revelation is incentive-compatible in our design) and make a randomized piece rate offer PR_i, which determines whether the participant is hired ($PR_i \geq \underline{PR_i}$) and the piece rate, if hired, per unit (PR_i). Workers who are hired work for the remainder of the day, about six hours on average. We measure output q_i as the number of units ("scoops", approximately 800g) sorted in a six-hour day. We also record a quality measure Q_i, the number of errors in a random sample of beans from a category. A randomized monitoring treatment, described below, explores workers' multitasking problem (quantity vs. quality) and the impact of monitoring on this tradeoff.[20]

[17]Specifically: *nanyati* (light brown or red with stripes), *zoyara* (small white), *khaki* (beige), *zofira* (small red), *phalombe* (large red), *napilira* (red with white stripes), *zosakaniza* (mixed / other) and discards (e.g. rotten, soybeans, stones, etc.). The categories are derived from discussions with purveyors of sorted beans in the Lilongwe market.

[18]As discussed in Section 3.2.1, a physically demanding task could pose problems for the incentive compatibility of BDM by introducing dependence between the level of effort in one day and the disutility of effort in subsequent days.

[19]We also provide subjects with visual aids during the sorting process, including examples of each of the sorted bean categories.

[20]Throughout, we refer to those with whom we interact at any stage as *subjects*, those who participate in BDM as *participants*, and those hired to work as *workers*.

3.1.1 Randomization and the Becker-DeGroot-Marschak Mechanism

We use the Becker-Degroot-Marschak mechanism (BDM) to uncover reservation piece rates, determine who works and set the piece rate. In BDM, the participants first states her reservation piece rate, $\underline{PR}_i$.[21] A piece rate PR_i is then drawn at random from a jug. If the random draw is less than the reservation piece rate, i.e. $PR_i < \underline{PR}_i$, the participant is not hired. If the random draw is at least as high as the reservation piece rate, i.e. $PR_i \geq \underline{PR}_i$, then the participant is hired at a piece rate of PR_i. Using BDM provides two key advantages. First, by breaking the link between the stated reservation piece rate and the actual piece rate paid, it makes truthful revelation of minimum willingness to accept the dominant strategy for the participant. Second, it creates random variation in the actual piece rate paid to workers with identical reservation piece rates. That is, two participants with the same reservation piece rate, $\underline{PR}_i = \underline{PR}_j$, will face different actual piece rates, $PR_i \neq PR_j$, and this difference will be determined purely by chance. This random variation allows us to isolate the causal effect of the piece rate on productivity.[22]

We implement a simplified version of BDM, in which a surveyor presents an individual participant with a menu of 5 piece rates: $5, 10, 15, 20, 25$ MWK per unit sorted.[23] The participant indicates which of the rates she will accept, the lowest of which we record as her reservation piece rate.[24] She then draws the actual piece rate offer from a uniform distribution with the same support as the reservation piece rates. Her draw determines whether she will work, and if so at what rate.[25]

[21] In fact, we obtain a narrow interval rather than a point measure, as will be clear below.

[22] Berry, Fischer, and Guiteras (2011) emphasize a third benefit of BDM: the ability to estimate heterogeneous treatment effects. Chassang et al. (2012) provide theoretical foundations, placing BDM in the class of "selective trials."

[23] All figures are in Malawi Kwacha. At the time of the study, the official exchange rate was roughly 150 MWK per US dollar, although a parallel black market exchange rate of roughly 180 MWK existed. Given the price premium for sorted beans on the market, an employer would find it profitable to hire workers to sort beans at piece rates up to 40 MWK per unit sorted.

[24] To be precise, she reveals a range on her reservation piece rate. For example, if she indicates that 15 MWK is the lowest rate she will accept, the her true reservation piece rate is in the interval $(10, 15]$. We believe this loss in resolution is more than outweighed by the gain in simplicity.

[25] This description of the implementation of BDM is highly simplified. In reality, the surveyor also leads the subject through a series of checks designed to confirm that the subject is indeed

The table below shows the possible outcomes of the game, with reservation piece rates in rows and piece rate offers in columns. Note that we will only observe outcomes for participants who draw a piece rate at least as high as their reservation piece rate, which is why the matrix is upper triangular.

$\underline{PR}_i$	PR_i				
	5	10	15	20	25
5	$(5, 5)$	$(5, 10)$	$(5, 15)$	$(5, 20)$	$(5, 25)$
10		$(10, 10)$	$(10, 15)$	$(10, 20)$	$(10, 25)$
15			$(15, 15)$	$(15, 20)$	$(15, 25)$
20				$(20, 20)$	$(20, 25)$
25					$(25, 25)$
> 25					

Without knowledge of the reservation piece rate, differences in productivity across piece rates (columns) are confounded with differences in productivity across workers with different reservation piece rates (rows). The benefit of BDM is the ability to make comparisons of outcomes across rows and down columns. A comparison across a row shows the causal effect of the piece rate, holding the reservation piece rate constant. A comparison down a column shows the association between the reservation piece rate and output, holding the actual piece rate fixed. Since we can only observe individuals working at or above their reservation piece rates, the number of comparisons that can be made varies. For example, we will have a lot of information in the relationship between the piece rate and output for those with very low reservation piece rates (row 1), but none for those with very high reservation piece rates (row 5). Conversely, we will obtain no information on the association between the reservation piece rate and productivity when the actual piece rate is very low

willing to work at the rates she says she will accept, and indeed prefers not working to working at the rates she declines. Our complete script in English is provided in the Appendix. All subjects attend a training session prior to BDM implementation in which the surveyors perform a skit with several examples designed to communicate the incentive-compatibility of BDM. The script for this skit is available upon request. The BDM decisions are elicited in private, so only the participant and the interviewer know her piece rate, unless she chooses to reveal it. Of course, whether or not she works is observed by everyone.

(column 1), but we will have a lot of information on this association when the actual piece rate is very high (column 5).[26]

3.1.2 Output quality versus output quantity

A higher piece rate gives a worker a clear incentive to work faster. However, sheer quantity is not the only desired outcome: incorrect sorting of beans lowers the value of the final product. To investigate this tradeoff between quantity and quality, we randomize a monitoring treatment that increases workers' incentives to produce quality output.

Quality is measured by recording the overall error rate, as well as the error rate by category of bean, since some types of beans are more or less difficult to sort than others. In both the monitoring treatment and the control, two randomly assigned categories of beans are checked for errors each time a worker presents a sorted unit. Possible errors include mis-categorized beans, flawed beans (with holes or rotten areas), or other foreign materials. The number of errors for each of the checked categories is recorded for each unit sorted, and the category for evaluation changes with each unit.

In addition to measuring this quantity-quality relationship, we are interested in learning how this relationship changes when we make the workers' pay dependent on quality. We randomly assign half of the subjects each day, stratified by sex, to a monitoring treatment. Subjects in the monitoring group are told that each unit of sorted beans will be checked for quality. The procedure (both as implemented and as described to the subjects) is that two categories of beans will be randomly selected and then a quantity equal to the size of a small handful from each category will be checked for errors. Based on pilot results, a threshold of three errors was used for all categories, though the difficulty of this threshold varies across categories

[26] In principle, one could obtain information on the productivity of individuals working below their reservation piece rate either through coercion or by having subjects participate in a two-stage process wherein, in the first stage, they express their preferences in one BDM offering pure piece rates and another offering a high flat participation fee in addition to a piece rate, then in a second stage randomizing which of these will actually be carried out. The first is unethical and the second is too complex for our context.

because of variation in the number of beans in a scoop, the inherent difficulty of avoiding mistakes, and in the difficulty of detecting errors. Workers were not told and could not observe which category was being evaluated, and the category was randomly assigned for each scoop. If either sample fails, the workers are required to return to their workstation to correct errors. Upon resubmission, two categories were randomly selected again (with replacement of the original categories) and the procedure repeated. This acts as a time tax on carelessness, since they are not given a new unit of beans until the unit under consideration is approved. The monitoring and control groups are segregated to the extent possible during the day to keep the controls minimally aware of the monitoring treatment. We retain data on the number of errors each time the unit of beans is submitted. These quality checks are performed on both the monitoring and the non-monitoring treatments, though only the monitoring treatment is sent back for corrections if the errors exceed the threshold. To reduce Hawthorne effects, the checks for control group workers are performed after the worker has received her next unit of beans and returned to her workstation to continue sorting.

3.2 Implementation

The experiment was implemented in six villages, one in each of six districts in Central Malawi over a period of six weeks in the low labor demand season (July-August) and a second six week period during the high labor demand season (January-February). In each of the six Districts, a list of 12 or more suitable villages was obtained from a District Agriculture Extension Officer.[27] We then randomly selected 2 villages from each district, one for implementation during the low labor demand season and a second during the high labor demand season. The village was informed of the activities approximately one week in advance and an open invitation was issued to

[27] The villages were identified as locations where the collaborating NGO was not working. They were also selected on a number of characteristics, including distance from the district capital and distance from the road since these factors are likely to affect the functioning of labor markets in these villages.

attend the orientation session on a Monday afternoon.[28] Subjects who participated in the orientation session were registered and became eligible to participate in the subsequent days' activities. During the orientation session, the bean sorting task was explained and surveyors performed a skit to illustrate the BDM mechanism and show subjects that truthful revelation of their minimum willingness to accept was their best strategy. Subjects were informed that they would receive a participation fee of 50 MwK per day, plus their earnings from the day's work sorting. The participation fee was emphasized to minimize self-selection into the experiment on subsequent days: we wanted to draw as representative as possible a sample of the village population. Because of field capacity constraints, we limited the number of BDM participants on each day to 50. After the first three weeks of the first data collection period, the number was reduced to 40 to address implementation challenges caused by the high acceptance rates of even low piece rate offers. On a given work day, if more than 40 (50) of subjects arrived by the pre-specified start time, a lottery was conducted to select 40 (50) participants. Those who were not selected were compensated for their time with a bar of soap. This constraint was often binding: on average, 52.9 (s.d. 20.9) potential subjects arrived on time and were eligible to participate in the lottery if there was one (48.5 (s.d. 10.7) in the low season and 57.3 (s.d. 27.1) in the high season). A lottery was required on 15 of 24 days of the experiment in the low labor demand season, and on all 24 days in the high labor demand season.

For each subject who attends the initial afternoon training session, we observe attendance decisions for every subsequent work day, for a total of four attendance observations per individual. Conditional on attending in a given day and being selected to participate in BDM and the survey, we also observe her reservation piece rate.[29] Participants whose BDM draw was greater than or equal to their stated reservation piece rate received a contract. For contracted workers, we observe the

[28]To confirm that invitation procedures did not differ significantly across labor seasons, we returned to all villages to collect household rosters and confirm which households received invitations. Villages assigned to the peak labor demand period had a combined population of 2,645, while villages assigned to the low labor demand season had a combined population of 2,206.

[29]Individuals who participated in BDM in a previous session were given priority to maximize the balance within the panel of observations. This priority status did not depend on whether they received a contract.

number of bean units that a worker sorts, the quality for every unit sorted, and her seating location relative to other workers.

A short survey was administered to every participant to collect basic covariates, in particular those likely to be associated with the opportunity cost of time.[30] The participation fee was contingent on the participant completing the survey. Finally, at the end of the work day on the last day in the village, an incentive compatible measure of risk aversion is obtained by providing individuals with a choice among the real-money following lotteries: (a) 50 MWK with probability 1; (b) 70 MWK or 40 MWK, each with probability 0.5; (c) 100 MWK or 20 MWK, each with probability 0.5. Any individual who participated (regardless of the outcome of BDM or work output) during the week was invited to participate in the lottery on Friday afternoon.[31]

3.2.1 Repeated work decisions

The work activity is conducted four times in each village, giving subjects the opportunity to participate in the BDM exercise on multiple days. This could present a problem for the incentive-compatibility of BDM. In its traditional use to measure willingness to pay for products, the option to play BDM multiple times could lead the subject to bid below her true WTP in early rounds. However, BDM is still incentive-compatible if decisions are independent across days. This would not be the case if, for example, the work was very physically demanding and working one day affected one's capacity the next day. Another violation would occur if there were

[30] Survey data were collected in two parts. The first, more comprehensive part, covering basic demographics and other time-invariant variables, was conducted only once with each participant. That is, a subject who was selected to participate on a given day was not administered this part of the survey if she had participated (and therefore been surveyed) on a previous day. The second part was a very brief set of questions on the subject's potential alternative activities for that day. In both cases, the survey was conducted independent of the outcome of the BDM experiment. However, for logistical reasons, both were administered *after* the BDM experiment was conducted and the results were known, so it is possible that the responses were affected by the result of the experiment.

[31] There was a tradeoff involved in the timing of this lottery: if the lottery were held at the beginning of the week, the result of the lottery could affect the respondent's bids, but work outcomes could not have not affected the respondent's answers to the survey questions. We conducted the lottery at the end of the week because we thought it was more important for the bids to be unaffected by the lottery result than for the risk aversion measures be unaffected by the BDM and work results.

income effects, i.e. working one day increased NPV lifetime earnings appreciably and led to more consumption of leisure. We do not believe either of these are present in our current context: the work is by design not physically taxing, and these earnings are not large enough to plausibly affect willingness to work in a neoclassical model.

3.3 Descriptive statistics

3.3.1 Characteristics and participation

Characteristics of participants are described in Table 1, which breaks the sample into the low and high labor seasons (six weeks per season). In the low labor season, 355 subjects participate in the experiment; 334 participate in the high labor season. Individuals may work multiple days of the week, though attendance is not mandatory, which results in an unbalanced individual panel by day with 1005 observations in the low labor season and 870 observations in the high labor season. Individuals in the low labor season work an average of 2.8 days while individuals in the high labor season work an average of 2.6 days out of the possible 4 work days.

Selected individual survey measures gathered for participants in BDM bidding are shown in Table 1. (The full set of survey measures is reported in the Online Appendix.) Most measures were collected only once, though a question on what the worker would have done that day if she had not joined the study was collected each work day. A few characteristics of the sample described in Table 1 are worth noting. First, over 60 percent of the sample is female and between 20 and 30 percent are from female headed households. Around two-thirds of households grow beans (relevant since workers are hired to sort beans), and workers report performing about 1.4 days of casual labor (*ganyu*) the previous week. Around two-thirds also report that casual labor is a regular source of income for their family. A number of characteristics vary significantly with the season. Most notably, the daily wage reported for the most recent casual labor is significantly higher in the high labor demand season. Individuals who join during the high season report slightly fewer months of food shortage, indicating that they are better off than participants in the low season.[32] In

[32]Note that households are more likely to have run out of food in January than in July, which

the high season, workers are less likely to list housework as one of their alternative activities for the day and more likely to list working their own land.

Several factors may contribute to the observed differences across labor seasons. First, the underlying characteristics of the villages visited may differ across seasons. Although our villages were randomly assigned to season, given our small number of villages (12) we cannot appeal to the law of large numbers to argue that the villages are likely to be well-balanced. Second, different types of individuals may have selected into the study, explaining differences in average participant age or other income sources. Finally, seasonal variation in labor demand and productive activites may explain differences in reported casual labor wages and outside options on the day of data collection.[33]

Participation or take up results can be broken down into the number of individuals who showed up for work each day (attendees), those who participated in the BDM draws (participants) and those who received a work contract and completed the bean sorting task (contracts awarded to workers). Table 2 breaks these results down by day (columns 2-5) and by labor season (columns 6 and 7). On average, the number of attendees is increasing through the week, with more attendees during the peak labor season. The average share of registered subjects attending each day is lower for peak labor season, suggesting fewer repeat workers during this period. As shown in the bottom two rows of Table 2, subjects who received a contract the previous day are more likely to attend than those who did not. This may be due to the fact that the probability of receiving a contract is higher for those with lower willingness to accept, as revealed by the BDM.

We also collected household rosters to determine the share of invited households in a village that attended the study session. In both the high and low labor demand season, the probability of receiving an invitation, conditional on being eligible, was around 85 percent. Conditional on receiving an invitation, around 25 percent of individuals attended the study session in the low labor demand season, versus

suggests that this difference is not due to the salience of food shortages during food short months.

[33]These explanations are not mutually exclusive. For example, differences in income sources may be due both to self selection and underlying differences in the villages.

around 48 percent in the peak labor demand season ($p < 0.001$, after controlling for district fixed effects). The probability of attending, conditional on receiving an invitation, is about 5 percent higher for females than for males, though it does not differ significantly for males and females by labor season. Of course, the highly divergent attendance rates raise questions about underlying sample differences across the labor seasons.

3.3.2 Willingness to accept

The share of participants accepting each of the 5 piece rates offered under the BDM are shown in Table 3. The first row shows the mean minimum WTA revealed in BDM, while the following rows break responses down for each of the 5 piece rates offered. The columns show responses by day of the week (columns 2-5), labor season (columns 6 and 7) and monitoring treatment (columns 8 and 9). For the lowest piece rate of 5 MWK, acceptance rates were around 44 percent, while the acceptance rates for the highest piece rate of 25 MWK were close to 100 percent. Raw means of acceptance rates do not vary much by labor season or by monitoring treatment. A greater proportion of women are willing to accept at each piece rate, and the mean minimum WTA for women is approximately 2 MWK lower than for men.

The bottom two rows of Table 3 summarize "mistakes" in the BDM procedure. Very few participants (< 3 percent) refused a drawn price that they had accepted in their BDM decisions. A larger share (13 percent) respond that they would have been willing to accept a drawn rate that they had rejected in their BDM decisions. This number declines throughout the week, suggesting continued learning about the mechanism, and from the low-labor-demand season, which was implemented first, to the high-labor-demand season, as well as over weeks within season (although noisily, not reported) which suggests that the surveyors improved at convincing subjects that stating one's true minimum WTA was their best strategy. On the other hand, the question of whether a participant would have been willing to accept at a previously rejected rate is purely hypothetical and individuals may have wished to express a willingness to work in their responses to this non-binding question.

3.3.3 Quantity and quality of worker output

The primary measures of productivity, number of scoops sorted per day (q) and average number of errors per scoop (Q) are summarized in Table 4.[34] The mean number of units sorted per day across all days is 7.35 (s.d. 1.97), which is increasing throughout the week, and the mean number of errors per scoop is 1.88 (s.d. 1.01). The quantity of output is lower (0.59 scoops per day) and the quality of output is higher (0.66 fewer errors per scoop) in the monitoring treatment, suggested that workers sorted more carefully and therefore more slowly in the monitoring treatment. Females sort 0.76 more scoops per day than men, and commit slightly fewer errors per scoop (0.16). The monitoring treatment effects, trends over days and gender differences are further explored in the regression analysis below.

Figures 1 and 2 show the densities of the quantity of output by monitoring treatment and labor season, respectively. These nonparametric plots suggest that the differences in means discussed above generally hold across the full distribution of quantities. Figures 3 and 4 lead to the same general conclusion for quality of output, i.e. errors per scoop.

3.3.4 Correlations between characteristics and outcomes

To describe the association between individual characteristics and the three primary outcome measures, we estimate the pairwise correlation between each individual survey measure and the outcome, controlling for monitoring (M_{id}), village (V_j) and day-of-week (DoW_d) fixed effects.

$$y_{id} = \beta x_i + \lambda M_{id} + \delta V_j + \tau \mathrm{DoW}_d + \varepsilon_{id}. \tag{1}$$

Table 5 presents the pairwise correlations. Individual predictors of willingness to accept are shown in column 1. As noted in the descriptive tables, female participants have significantly lower reservation piece rates (WTA, Col. 1) than do male workers, and participants who report performing casual labor in the last week (*ganyu*) also

[34] Q is recorded the first time the workers bring a scoop of sorted beans to the enumerator, before they have been instructed to correct any errors above the threshold.

reveal lower willingness to accept, as do those who indicate they would have been working in another casual labor job if not participating in the study on the day of observation. Finally, subjects from households that grow beans or other (unspecified) agriculture report higher WTA, perhaps because of higher opportunity costs.

A number of characteristics are significantly related to quantity of scoops sorted (q_{id}, Col. 2) or to the number of errors per scoop (Q_{id}, Col. 3). Females produce higher quantity of output and also higher quality (fewer errors), as do more educated workers. Older workers sort fewer scoops and generate more errors. Workers who would have been doing other ganyu that day also sort fewer units of beans. Workers who expect to sort more scoops do, in fact, sort more scoops and they also make fewer errors.[35] For the most part, both output quality and quantity move together for these individual characteristics, suggesting that females and more educated workers are better at both dimensions of productivity, while older workers are worse at both.

4 Empirical Results

We present our empirical strategy and results together, by theme. Our three outcome measures are willingness to accept as measured by BDM, quantity of output measured by the number of units of beans sorted per day, and quality of output measured by the number of errors per unit. We first discuss selection, i.e. the relationship between minimum WTA and productivity. Second, we estimate incentive effects, i.e. the causal effect of piece rates on output. Third, we estimate differences between men and women in selection and incentive effects. Finally, we test for gift exchange (reciprocity), income targeting and peer effects.

4.1 Selection and productivity

In this subsection, we explore the relationship between reservation piece rates and productivity. First, we examine the determinants of minimum WTA. Second, we es-

[35]Productivity expectations are elicited after the BDM draw so expectation measures may partly reflect anticipation of the productivity effects of higher draws.

timate the relationship between minimum WTA and productivity, which we interpret as the selection channel of of the relationship between piece rates and productivity.

To study the determinants of reservation piece rates, we regress minimum WTA, as revealed in BDM, on characteristics of the market, specifically, the labor season (Peak = 1 for high labor season), the monitoring treatment, M_{id}, and day of the week, DoW_d:

$$\text{MinWTA}_{id} = \phi\text{Peak}_i + \lambda M_{id} + \zeta F_i + \tau\text{DoW}_d + \delta\text{Dist}_j + \varepsilon_{id}. \tag{2}$$

Table 6 shows estimation results in cross section (columns 1 and 2), with random effects (columns 3 and 4) and with individual fixed effects (column 5). All columns include district fixed effects (Dist_i). The peak labor season lowers minimum WTA slightly. The most interesting result is that the monitoring treatment does not have a significant effect on minimum WTA. Subjects do not appear to demand greater compensation for the more stringent standards imposed by monitoring. This absence of "sorting on monitoring" will prove important for interpretation of subsequent regressions. Minimum WTA falls over the course of the week, which cannot be explained solely by selection given the robustness to individual fixed effects (column 5). Minimum WTA is about 1 MWK higher on the first day than on later days in the week, relative to a mean of 10 MWK in the sample.

Next, we investigate selection: the association between output (quality and quantity) and minimum WTA. To isolate the selection channel, we estimate the relationship between minimum WTA as revealed through the BDM and our two outcome measures, quantity of output and quality of output, controlling for the actual piece rate received by the worker:

$$\begin{aligned} y_{id} \quad = \quad & \phi\text{Peak}_i + \lambda\text{M}_{id} + \alpha\text{PR}_{id} + \gamma\left(\text{PR}_{id} \times \text{M}_{id}\right) + \beta\text{minWTA}_{id} \\ & +\zeta F_i + \tau\text{DoW}_d + \delta\text{Dist}_j + \varepsilon_{id}. \end{aligned} \tag{3}$$

We interpret the coefficient β as the selection effect: the relationship between the reservation piece rate and productivity, holding the actual piece rate constant. Tables 7 and 8 show the effects of willingness to accept on the number of scoops sorted per day and the number of errors per scoop, respectively.[36] With respect to quantity, we observe *negative* selection: after controlling for the worker incentive provided by the piece rate (PR_{id}), the monitoring treatment and the day of the week, minimum willingness to accept is negatively related to quantity of output (Table 7), though the size of the coefficient is small (a 10 MWK increase in minimum WTA lowers the number of units sorted per day 0.20-0.30, relative to a mean of 7.4 units). The same specification with number of errors per scoops as the dependent variable shows no significant relationship between minimum willingness to accept and quality of output, and the coefficient on minimum willingness to accept is inconsistently signed (Table 8).

The preceding regressions impose a linear functional form. As an alternative, we estimate a more flexible model with indicators for each interval of minimum willingness to accept and piece rate draw. The resulting semiparametric relationship between minimum willingness to accept and output is plotted in Figures 5 and 6. Figure 5 shows a slightly negative relationship between minimum willingness to accept and quantity of output at low piece rates, which becomes insignificant at higher piece rates. The relationship between minimum willingness to accept and quality of output is insignificant at every piece rate category.

4.2 Incentives and productivity

By controlling for the willingness to accept revealed under BDM, we are able to isolate the direct effect of incentives on productivity. These effects are due solely to changes in the worker's effort choice in response to a change in the piece rate or monitoring of output quality, both of which are randomized.[37]

[36] While we present and discuss quantity and quality results separately, it is important to remember that they are jointly determined by the worker. That is, we should not think of determinants of quantity as operating with quality held fixed.

[37] Although monitoring is randomized, monitoring is not random conditional on minimum WTA, since BDM participants announced their minimum WTA knowing whether or not they were assigned to the monitoring group. As noted above, we do not observe that being assigned to monitoring

4.2.1 Piece rate

The straightforward test for the direct effect of incentives on productivity relies on (3) above but examines the effect of the piece rate on productivity, controlling for minimum willingness to accept. Thus, the coefficient of interest is α.

Table 7 shows the effect of the piece rate on quantity of output, controlling for the worker's reservation piece rate (columns 2 - 4) and individual characteristics (column 4). Increasing the piece rate by 10 MWK increases the number of scoops sorted per day by between 0.24 and 0.50 units, relative to a mean of 7.4. Going from the lowest draw (5) to the highest draw (25) increases output by between one-half to one unit per day. These effects are similar in magnitude to the selection effects described above, but with opposite sign. The quantity of output is also increasing with the day of the week, an effect that is robust to the inclusion of individual fixed effects (column 5). In the peak labor season, workers sort almost half a scoop more per day.

Table 8 shows the effect of a piece rate on quality of output, measured by the number of errors per scoop of sorted beans. The piece rate appears to have little direct effect on quality of output, though the coefficient is consistently positive indicating that errors may be increasing in the piece rate. The number of errors per scoop is decreasing in the day of the week, consistent with individuals gaining experience with the task. This effect is robust to the inclusion of individual fixed effects (column 5), suggesting that it is not driven by changes in the composition of workers over the course of the week. The labor demand season does not appear to affect quality of output.

4.2.2 Monitoring

The effect of stricter monitoring on quantity and quality of output is measured by γ in (3) above. If quantity and quality are substitutes, then workers must choose to allocate effort toward quantity or toward quality (reduce errors). If this is the case, γ will take on the same sign for the two output regressions (quantity and errors).

affects stated minimum WTA significantly.

The direct effect[38] of the monitoring treatment on quantity of output is shown in Table 7. The coefficient on monitoring is negative and significant, lowering output by between -0.56 and -0.78 scoops per day, or about a third of a standard deviation. The loss in quantity of output is accompanied by a reduction in the number of errors per scoop, as shown in Table 8. The coefficient on monitoring is between -0.63 and -0.76 scoops per day, or around three-quarters of a standard deviation. Monitoring does appear to divert effort toward output quality at a cost of some quantity.

4.3 Gender differences

To examine differential selection and effort choices by gender, we repeat the analyses and interact key regressors with a dummy variable indicating that the participant is female.. Differences in reservation piece rates are obtained by re-estimating (2) with interactions of the female variable with the labor season (Peak) and monitoring treatment (M). Table 9 shows the results for selection by gender. The labor season does appear to matter for female minimum WTA, which is significantly lower in the peak labor season. The monitoring treatment does not affect the participation decision for male or for female workers.

Differences in productivity, controlling for reservation piece rates, are obtained by re-estimating (3) allowing the effects of the labor season, monitoring treatment, reservation rates and the draw to vary by gender. Tables 10 and 11 show the effects on the quantity of output and quality of output respectively.

Monitoring reduces the quantity of output for both genders, but more so for females (Table 10). The pure incentive effect of the piece rate on quantity of output is similar across genders. However, the selection effect of minimum willingness to accept on quantity of output does appear to vary by gender. Men who exhibit a higher minimum WTA rate sort significantly fewer units of beans, while the effect for women is insignificant and inconsistently signed. This difference may be due in part to the fact that women's outside option may be more like bean sorting while men's outside option is more likely to be physical labor. Men produce more output

[38] By "direct," we mean holding selection constant by conditioning on minimum WTA.

in the high labor season, approximately half of a scoop, while women do not. This difference may also be related to the differences in men and women's outside options, and how they vary with the labor season, though it may also be due to differences in the types of men joining the study in the low and high labor demand seasons.

Both men and women are similarly responsive to monitoring in their allocation of effort toward output quality (Table 11). Among men, the error rate increases (quality decreases) with the piece rate, although this effect is small and not significant across specifications. The incentive effect of the piece rate has a small positive effect on the number of errors for men and no effect for women. Minimum WTA does not appear to affect the number of sorting errors for either men or for women. Men do, however, reduce the number of errors per unit sorted in the peak labor demand season, while error rates for women do not differ significantly across seasons. Again, this may have to do with the opportunity cost of time for men and women by labor season affecting their sensitivity to incentives or affecting selection into the study.

4.4 Worker reciprocity

Our experimental design provides an opportunity to test for the presence of gift exchange. In our setting, workers could choose to share surplus generated by a piece rate above their reservation piece rate by increasing the quality of output (reducing the number of errors). This type of reciprocity should be detectable primarily among workers in the no monitoring treatment, where there are no explicit incentives for reducing errors. Therefore, to test for empirical evidence of gift exchange, we examine the interaction of the monitoring treatment and the piece rate.

Without quality incentives, a higher piece rate generates explicit incentives to work faster at the expense of quality. In the absence of gift exchange, we would expect the number of errors increase more quickly with the piece rate in the no monitoring treatment than in the monitoring treatment.

We examine the interaction between the monitoring treatment and the piece rate in a regression framework with the number of errors as the outcome variable. Specifically, the coefficient of interest from (3) is γ and results are shown in columns

3-5 of Table 8. The interaction term is small and insignificant, i.e. the number of errors increases no faster in the no-monitoring treatment than in the monitoring treatment, which is consistent with gift exchange: unmonitored workers appear to be sharing some surplus.

We also estimate a more flexible, semiparametric model with interactions for each of the 5 piece rates offered under BDM:

$$\begin{aligned} y_{id} \quad = \quad & \sum_{r=\{5,10,\ldots,25\}} \alpha_r 1\left\{\mathrm{PR}_{id}=r\right\} + \sum_{r=\{5,10,\ldots,25\}} \beta_r 1\left\{\mathrm{minWTA}_{id}=r\right\} \\ & + \sum_{r=\{5,10,\ldots,25\}} \gamma_r \left(1\left\{\mathrm{PR}_{id}=r\right\} \times \mathrm{M}_{id}\right) \\ & + \phi \mathrm{Peak}_i + \delta \mathrm{Dist}_i + \tau \mathrm{DoW}_d + \varepsilon_{id}. \end{aligned} \tag{4}$$

The outcome variable is again the number of errors per scoop and the coefficients of interest are $\gamma_5, \gamma_{10}, \ldots, \gamma_{25}$. Results are presented in Table 12. In the no-monitoring treatment, the number of errors is not increasing in the drawn piece rate. All coefficients on piece rate draws are negatively signed and insignificant for workers in the no-monitoring treatment and positively signed and occasionally marginally significant in the monitoring treatment. Figure 7 plots these coefficients. As in the linear specification, these results are consistent with gift exchange.

Next, we look within day for evidence of declining reciprocity over time as found in Gneezy and List (2006). Over the course of the work day, the workers sort an average of seven units of beans. Declining reciprocity would be supported by a finding of increasing errors during the course of the day in the no-monitoring group, as they learn that low-quality work will not be punished. Figure 8 shows the number of errors by scoop over the course of the day.[39] The number of errors committed by workers in the monitoring treatment declines over the course of the day, consistent with learning about the cost of sorting errors. In the no monitoring treatment, the number of errors per scoop is constant throughout the day. This suggests that

[39] We limit to the first 10 scoops in a day since there are very few observations of more than 10 scoops in a day.

reciprocity through output quality, to the extent that it is present, does not erode over the course of the work day.

4.5 Target earnings

The canonical studies of income targeting examine cumulative earnings for taxi drivers, and analyze whether drivers tend to quit earlier on days when their hourly earnings are higher (Camerer et al. 1997; Farber 2005, 2008; Doran 2010). Cumulative earnings in our setting are likely less salient, because the workers in our study are paid at the end of each day rather than following each unit of production. Furthermore, because workers must be present at the end of the day to be paid, and there are costs to leaving and returning, the length of the work day is less flexible. In fact, it is very rare for workers to leave before the end of the day.

However, workers in our context can reduce intensity of effort once the earnings target has been reached by relaxing and working less hard. We therefore use the elapsed time per unit sorted as a proxy for intensity of effort, and test whether effort depends on cumulative earnings that day. Figure 9 shows the pattern of elapsed time per unit over the course of a day for male and female workers. Averaging across all days and piece rates, the average sorting time decreases, due partly to learning and partly to selection: only the faster workers are sorting the larger numbers of units.[40]

To test for target earnings, we estimate the following equation:

$$\begin{aligned} \text{Minutes}_{idj} \quad = \quad & \alpha_i + \gamma_{jp} 1\left\{\text{Unit}_{idj}\right\} \times 1\left\{\text{Draw}_{idj}\right\} \\ & + \phi \text{Peak}_i + \lambda \text{M}_{id} + \zeta F_i + \tau \text{DoW}_d + \delta \text{Dist}_j + \varepsilon_{idj} \qquad (5) \end{aligned}$$

where the dependent variable, Minutes_{idj}, is the time individual i takes to sort the j^{th} unit on day d, and the coefficient of interest γ_{jp} corresponds to the j^{th} unit sorted at a piece rate of p. We include individual random effects and estimate off variation in the piece rate by day. The prediction of a target income model is that

[40] The first unit is dropped from the analyses below because there were various startup activities that makes the time of sorting the first unit not comparable with the time for other units.

workers with a higher piece rate will slow down more later in the day, as they are more likely to have met an income target.

Figure 10 plots the predicted values given by this semiparametric model. Under a reference dependent labor supply model, we would expect to see an increase in the sorting time once workers hit their income target. While the time elapsed on the 10th unit does increase slightly relative to the 9th unit, the difference is insignificant and greatest for the lowest piece rate. Income targeting would predict that workers reduce effort sooner when piece rates are high. Finally, Figure 11 presents a simpler, parametric view, from a model as in Equation (5) but with linear and quadratic terms for Unit_{idj} and Draw_{idj} (and all interactions) instead of fully interacted dummies. The patterns are similar across piece rate draws, which is inconsistent with income targeting.

4.6 Empirical test for peer effects

In our design, workers self-select into work groups averaging 5-6 individuals. While these work groups are endogenously formed, workers in the same group experience different (random) piece rates.[41] We are able to use this variation in piece rates to identify the effect of peers' compensation on a worker's own effort. Output from individual i on day is regressed on the mean piece rate for other workers in the group, along with other controls:

$$y_{idg} = \alpha \text{PR}_{id} + \beta \text{minWTA}_{id} + \lambda \bar{\text{PR}}_{g,-i} + \psi \text{min}\bar{\text{W}}\text{TA}_{g,-i} + \gamma \text{M}_{id} + \delta \text{Dist}_{ig} + \tau \text{DoW}_d + \varepsilon_{idg}.$$

We are interested in the relationship between the mean piece rate in the group (for workers $-i$) and productivity. To address the potential effect of reservation rates on the mean piece rate, we also control for the mean reservation rate of other workers in the group. We also allow both worker i's draw and the size of the work group

[41] Workers are separated by monitoring treatment, and are not aware of others workers' wage draws when they choose where to sit.

to interact with the effect of other worker's piece rates.[42] Since the regressors of interest vary at the work group level, we compute standard errors robust to two-way clustering at the work group and individual (across days) level (Cameron et al. 2011).

Results are consistent with higher productivity peers increasing a workers own productivity, but at a decreasing rate as the size of the group increases. Table 13.A shows no direct effect of mean group rates on productivity (Columns 1 and 2). However, when interacted with the size of the group (Column 3), the mean piece rate among other workers has a positive and significant effect (0.065, s.e. 0.028). which is decreasing in group size (-0.019, s.e. 0.007). In other words. for every one-Kwacha increase in the average wages of a worker's peers, that worker increases output by 0.065 scoops. For every additional worker (beyond a minimum of one peer), in the work group, this effect decreases by 0.019 scoops. Interestingly, Table 13.B shows that there is no detectable effect on quality, which suggests that workers are motivated to keep pace with their peers but do not do so at the expense of quality.

5 Discussion

We implement a unique experimental design in the casual labor market in rural Malawi to separate selection and effort effects on productivity, to observe gender differences in these markets and to test for other behavioral determinants of productivity. In our setting, production of quantity and quality are substitutes and workers allocate effort between these two types of output. A monitoring treatment shifts effort toward production of quality at a quantity cost.

Both the selection and the effort margins affect productivity, though selection only for quantity of output when the output of male and female workers is examined separately. A higher piece rate actually selects for less productive men, while a higher piece rate selects for more productive women, in terms of quantity of output. Selection appears to have no direct effect on the quality of output. On the other hand, raising the piece rate significantly increases the quantity of output, controlling

[42] The regressions exclude a few workers who chose to work alone.

for workers' reservation rates. While explicit incentives for output quality do reduce the number of errors in production, errors do not increase with the piece rate even without punishment for low quality. This latter result is consistent with a gift exchange model of reciprocity in which workers share surplus with the firm in the form of output quality.[43]

To compare the two labor supply margins (participation versus effort), we estimate the change in the percentage of workers willing to accept a contract for every percent change in in the contract price. For a ten percent increase in the contract price, participation increases by 5.2 percent. The effort elasticity ignores effects on quality of output for the moment and estimates the percent change in quantity of output for every percent change in the contract price. For a ten percent increase in the contract price, output increases by 0.6 percent. Monitoring does not lower the elasticity of output with respect to the piece rate, but does lower the quantity of output overall. At the mean, introducing monitoring lowers output as much as a 30 MWK decrease in the piece rate. It takes workers in the monitoring treatment about 0.7 days longer to sort a 50 kilogram bag of beans than workers who are not being monitored. Well sorted beans sell for up to 4,000 MWK more per 50 KG bag, potentially justifying the cost of monitoring.

Our design allows us to test a number of behavioral models of worker behavior. We examine models of gift exchange, income targeting and peer effects. We find behavior consistent with a simple model of gift exchange, with workers sharing surplus by maintaining quality even without explicit incentives to do so. We do not find evidence of income targeting – effort over the course of the day evolves similarly for workers with high and low piece rates. We do find some support for peer effects – workers with peers who are receiving high piece rates appear to produce more without sacrificing quality.

[43]The cost effectiveness of monitoring in this setting requires further investigation of the costs to the firm of errors in output.

References

George A. Akerlof. Labor contracts as partial gift exchange. *The Quarterly Journal of Economics*, 97(4):pp. 543–569, 1982. ISSN 00335533. URL http://www.jstor.org/stable/1885099.

Oriana Bandiera, Iwan Barankay, and Imran Rasul. Social preferences and the response to incentives: Evidence from personnel data. *The Quarterly Journal of Economics*, 120(3):pp. 917–962, 2005. ISSN 00335533. URL http://www.jstor.org/stable/25098759.

Oriana Bandiera, Iwan Barankay, and Imran Rasul. Social incentives in the workplace. *Review of Economic Studies*, 77(2):417–458, 2010. ISSN 1467-937X. doi: 10.1111/j.1467-937X.2009.00574.x. URL http://dx.doi.org/10.1111/j.1467-937X.2009.00574.x.

Pranab K. Bardhan. Labor supply functions in a poor agrarian economy. *The American Economic Review*, 69(1):pp. 73–83, 1979. ISSN 00028282. URL http://www.jstor.org/stable/1802498.

G. Becker, M. DeGroot, and J. Marschak. Measuring utility by a single-response sequential method. *Behavioral Science*, 9:226–236, 1964.

Charles Bellemare and Bruce Shearer. Gift giving and worker productivity: Evidence from a firm-level experiment. *Games and Economic Behavior*, 67(1):233–244, 2009. URL http://www.sciencedirect.com/science/article/pii/S0899825608002194.

James Berry, Greg Fischer, and Raymond Guiteras. Eliciting and utilizing willingness to pay: Evidence from field trials in Northern Ghana. Mimeo, September 2011. URL http://www.econ.umd.edu/research/papers/584.

Colin Camerer, Linda Babcock, George Loewenstein, and Richard Thaler. Labor Supply of New York City Cabdrivers: One Day at a Time. *The Quarterly Journal of Economics*, 112(2):407–441, 1997. URL http://qje.oxfordjournals.org/content/112/2/407.short.

A. Colin Cameron, Jonah B. Gelbach, and Douglas L. Miller. Robust inference with multiway clustering. *Journal of Business and Economic Statistics*, 29(2):238–249, 2011. doi: 10.1198/jbes.2010.07136. URL http://pubs.amstat.org/doi/abs/10.1198/jbes.2010.07136.

Sylvain Chassang, Gerard Padro i Miquel, and Erik Snowberg. Selective trials: A principal-agent approach to randomized controlled experiments. *American Economic Review*, 2012. URL `http://www.nber.org/papers/w16343`. Forthcoming.

David Dickinson and Marie-Claire Villeval. Does monitoring decrease work effort?: The complementarity between agency and crowding-out theories. *Games and Economic Behavior*, 63(1):56 – 76, 2008. ISSN 0899-8256. doi: 10.1016/j.geb.2007.08.004. URL `http://www.sciencedirect.com/science/article/pii/S0899825607001364`.

Thomas Dohmen and Armin Falk. Performance pay and multidimensional sorting: Productivity, preferences, and gender. *The American Economic Review*, 101(2):556–590, 2011. doi: doi:10.1257/aer.101.2.556. URL `http://www.ingentaconnect.com/content/aea/aer/2011/00000101/00000002/art00005`.

Kirk Bennett Doran. Wages, daily income goals and daily labor supply. Mimeo, September 2010.

Tor Eriksson and Marie Claire Villeval. Performance-pay, sorting and social motivation. *Journal of Economic Behavior and Organization*, 68(2):412 – 421, 2008. ISSN 0167-2681. doi: 10.1016/j.jebo.2007.10.003. URL `http://www.sciencedirect.com/science/article/pii/S0167268108001273`.

Armin Falk and Andrea Ichino. Clean evidence on peer effects. *Journal of Labor Economics*, 24(1):39–57, 2006. URL `http://www.jstor.org/pss/10.1086/497818`.

Henry S. Farber. Is tomorrow another day? The labor supply of New York City cabdrivers. *Journal of Political Economy*, 113(1):46–82, 2005.

Henry S. Farber. Reference-Dependent Preferences and Labor Supply: The Case of New York City Taxi Drivers. *The American Economic Review*, 98(3):pp. 1069–1082, 2008. ISSN 00028282. URL `http://www.jstor.org/stable/29730106`.

Ernst Fehr and Lorenz Goette. Do workers work more if wages are high? evidence from a randomized field experiment. *The American Economic Review*, 97(1):298–317, 2007. URL `http://www.jstor.org/pss/30034396`.

Ernst Fehr and Klauss M. Schmidt. Fairness and incentives in a multi-task principal-agent model. *The Scandinavian Journal of Economics*, 106(3):453–474, 2004. URL `http://www.jstor.org/pss/3441119`.

Ernst Fehr, Karla Hoff, and Mayuresh Kshetramade. Spite and development. *The American Economic Review*, 98(2):494–499, May 2008. doi: doi:10.1257/aer.98.2.494. URL http://www.jstor.org/stable/29730070.

Simon Gaechter, Daniele Nosenzo, and Martin Sefton. The impact of social comparisons on reciprocity. Mimeo, May 2010. URL http://ideas.repec.org/p/cdx/dpaper/2010-10.html.

Xavier Giné, Mónica Martínez-Bravo, and Marian Vidal-Fernández. Intertemporal Substitution or Reference-Dependent Preferences? Evidence from Daily Labor Supply of South Indian Boat-owners. Mimeo, April 2010.

Uri Gneezy and John A. List. Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments. *Econometrica*, 74(5):pp. 1365–1384, 2006. ISSN 00129682. URL http://www.jstor.org/stable/3805928.

Uri Gneezy, Muriel Niederle, and Aldo Rustichini. Performance in competitive environments: Gender differences. *The Quarterly Journal of Economics*, 118(3): pp. 1049–1074, 2003. ISSN 00335533. URL http://www.jstor.org/stable/25053930.

Uri Gneezy, Kenneth L. Leonard, and John A. List. Gender differences in competition: Evidence from a matrilineal and a patriarchal society. *Econometrica*, 77 (5):pp. 1637–1664, 2009. ISSN 00129682. URL http://www.jstor.org/stable/25621372.

Jessica Goldberg. Kwacha Gonna Do? Experimental Evidence about Labor Supply in Rural Malawi. Mimeo, September 2011. URL http://www-personal.umich.edu/~jegoldbe/docs/goldberg_ganyu.pdf.

Supreet Kaur, Michael Kremer, and Sendhil Mullainathan. Self-control at work. Mimeo, 2011. URL http://www.people.fas.harvard.edu/%7Ekaur/papers/KaurKremerMullainathan_SelfControl.pdf.

Craig E. Landry, Andreas Lange, John A. List, Michael K. Price, and Nicholas G. Rupp. Is there a hidden cost of control in naturally-occurring markets? Evidence from a natural field experiment. Mimeo, September 2011. URL http://www.nber.org/papers/w17472.

Edward P. Lazear. Performance pay and productivity. *The American Economic Review*, 90(5):pp. 1346–1361, 2000. ISSN 00028282. URL http://www.jstor.org/stable/2677854.

Alexandre Mas and Enrico Moretti. Peers at work. *American Economic Review*, 99 (1):112–145, 2009. URL http://www.aeaweb.org/articles.php?doi=10.1257/aer.99.1.112.

Daniel S. Nagin, James B. Rebitzer, Seth Sanders, and Lowell J. Taylor. Monitoring, motivation, and management: The determinants of opportunistic behavior in a field experiment. *The American Economic Review*, 92(4):pp. 850–873, 2002. ISSN 00028282. URL http://www.jstor.org/stable/3083284.

Muriel Niederle and Lise Vesterlund. Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics*, 122(3):pp. 1067–1101, 2007. ISSN 00335533. URL http://www.jstor.org/stable/25098868.

Harry J. Paarsch and Bruce S. Shearer. The response of worker effort to piece rates: Evidence from the British Columbia tree-planting industry. *The Journal of Human Resources*, 34(4):pp. 643–667, 1999. ISSN 0022166X. URL http://www.jstor.org/stable/146411.

Canice Prendergast. The provision of incentives in firms. *Journal of Economic Literature*, 37(1):pp. 7–63, 1999. ISSN 00220515. URL http://www.jstor.org/stable/2564725.

Mark R. Rosenzweig. Rural wages, labor supply, and land reform: A theoretical and empirical analysis. *The American Economic Review*, 68(5):pp. 847–861, 1978. ISSN 00028282. URL http://www.jstor.org/stable/1811317.

Carl Shapiro and Joseph E. Stiglitz. Equilibrium unemployment as a worker discipline device. *The American Economic Review*, 74(3):pp. 433–444, 1984. ISSN 00028282. URL http://www.jstor.org/stable/1804018.

Joseph E. Stiglitz. Incentives, risk, and information: Notes towards a theory of hierarchy. *The Bell Journal of Economics*, 6(2):pp. 552–579, 1975. ISSN 0361915X. URL http://www.jstor.org/stable/3003243.

Table 1: Descriptive Statistics for Participants

Variable	Low Season	High Season	Diff.	
Number of participants	355	334		
Number of daily observations	1005	870		
Female	0.690	0.638	-0.052	
	(0.463)	(0.481)	[0.036]	
Age	34.647	35.184	0.537	
	(13.183)	(14.104)	[1.111]	
Number of adults in household	3.161	3.058	-0.103	
	(1.588)	(1.774)	[0.130]	
Years of education	3.913	4.565	0.652	***
	(3.353)	(3.169)	[0.252]	
Female headed household	0.201	0.303	0.102	***
	(0.401)	(0.460)	[0.033]	
Days of ganyu last week	1.402	1.425	0.023	
	(2.345)	(2.087)	[0.171]	
Daily wage from recent ganyu (MKW)	257.684	336.668	78.984	***
	(178.964)	(381.019)	[24.416]	
Ever participated in ganyu for international org.	0.023	0.131	0.108	***
	(0.150)	(0.338)	[0.020]	
Household produces beans	0.686	0.627	-0.059	
	(0.465)	(0.484)	[0.037]	
Household produces tobacco	0.478	0.279	-0.199	***
	(0.500)	(0.449)	[0.037]	
Household produces other agriculture	0.738	0.645	-0.093	***
	(0.440)	(0.479)	[0.035]	
Typical per year months without adequate food	3.562	3.119	-0.443	**
	(2.344)	(2.159)	[0.173]	
Alternative activity: housework	0.267	0.074	-0.193	***
	(0.443)	(0.262)	[0.034]	
Alternative activity: other ganyu	0.235	0.172	-0.063	*
	(0.425)	(0.378)	[0.038]	
Alternative activity: work own land	0.211	0.525	0.314	***
	(0.409)	(0.501)	[0.044]	
Alternative activity: work own business	0.085	0.118	0.033	
	(4.429)	(3.624)	[0.393]	

*, **, *** denote significance at 10%, 5% and 1%, respectively.

Notes: this table presents means of participants' characteristics during the low and high season, with standard deviations in parentheses, as well as differences in means, with the standard error of the estimated difference in brackets.

Table 2: Descriptive Statistics on Attendance and Participation

		Day of Week				Season	
	All Days	Day 1	Day 2	Day 3	Day 4	Low	High
Number of attendees	52.9	47.3	52.1	52.9	59.3	48.5	57.3
	(20.9)	(22)	(17.1)	(21.3)	(23.6)	(10.7)	(27.1)
Number of participants	39.1	37.7	39.8	38.6	40.3	41.9	36.3
	(7)	(8)	(6.1)	(8.1)	(6.1)	(5.9)	(6.9)
Number of contracts awarded	30.5	27.7	32.3	30.3	31.8	31.8	29.3
	(7.8)	(8.2)	(7.4)	(8.8)	(6.7)	(8.1)	(7.4)
Proportion of registered workers attending	.474	.425	.467	.475	.531	.727	.367
	[.007]	[.014]	[.014]	[.014]	[.014]	[.011]	[.008]
Attendance conditional on previous day:							
attending	.815		.746	.802	.888	.846	.788
	[.009]		[.018]	[.016]	[.013]	[.012]	[.013]
participating in BDM	.851		.772	.83	.948	.861	.838
	[.01]		[.02]	[.017]	[.01]	[.013]	[.015]
participating in BDM and being awarded a contract	.879		.813	.858	.961	.888	.869
	[.01]		[.021]	[.018]	[.01]	[.013]	[.015]
participating in BDM but not being awarded a contract	.755		.664	.716	.899	.775	.727
	[.025]		[.043]	[.048]	[.03]	[.031]	[.04]

Notes: An attendee is defined as any subject who registers on the orientation day and is present at the beginning of a work day. A maximum of 40 attendees participate in BDM each day (50 in the first three weeks, see discussion in text). If participation is oversubscribed, 40 (50) of the attendees are selected by lottery for participation. Standard deviations in parentheses. Standard error of estimated proportion in brackets.

Table 3: Descriptive Statistics on BDM

		Day of Week				Season		Monitoring		Gender	
	All Days	Day 1	Day 2	Day 3	Day 4	Low	High	No	Yes	Male	Female
Mean minimum WTA (MWK)	10.32	11.06	9.99	10.52	9.76	10.55	10.06	10.52	10.12	11.74	9.67
	(5.90)	(6.56)	(5.89)	(5.60)	(5.44)	(6.06)	(5.70)	(5.99)	(5.81)	(6.63)	(5.41)
Share agreeing to piece rate of:											
5	0.44	0.43	0.47	0.39	0.45	0.42	0.45	0.43	0.45	0.38	0.46
10	0.68	0.63	0.71	0.67	0.71	0.68	0.69	0.66	0.70	0.58	0.73
15	0.84	0.77	0.85	0.85	0.87	0.82	0.85	0.83	0.84	0.73	0.88
20	0.96	0.94	0.96	0.97	0.96	0.95	0.97	0.95	0.96	0.92	0.97
25	0.99	1.00	1.00	1.00	0.99	1.00	0.99	0.99	0.99	0.99	1.00
Ex post refused contract	0.026	0.026	0.025	0.027	0.026	0.034	0.017	0.027	0.025	0.034	0.023
Ex post would have accepted	0.126	0.233	0.112	0.061	0.072	0.176	0.054	0.128	0.124	0.109	0.138

Notes: Sample is all participants in BDM. Minimum WTA is the participant's bid in BDM. Ex post refused contract if the participant ultimately rejects a piece rate she had agreed to prior to the draw. Ex post would have accepted if a participant who did not receive a contract, i.e. drew higher than her minimum WTA, states in the exit survey that she would have accepted the piece rate drawn had she been given the opportunity. Standard deviations in parentheses.

Table 4: Descriptive Statistics on Output (Quantity and Quality)

	All	Day of Week				Season		Monitoring		Gender	
		Day 1	Day 2	Day 3	Day 4	Low	High	No	Yes	Male	Female
Quantity: scoops of beans sorted	7.35	6.02	7.21	7.90	8.12	7.15	7.57	7.65	7.06	6.81	7.57
	(1.97)	(1.61)	(1.74)	(1.97)	(1.87)	(1.81)	(2.12)	(2.06)	(1.85)	(1.97)	(1.93)
Quality: errors per scoop	1.88	2.20	1.92	1.76	1.69	1.88	1.88	2.22	1.56	2.00	1.84
	(1.01)	(1.19)	(1.04)	(0.91)	(0.83)	(1.05)	(0.97)	(1.01)	(0.90)	(1.17)	(0.93)

Notes: This table provides simple descriptive statistics on the output of participants, in terms of quantity of beans sorted and quality of the sorting. The sample is all participants in BDM who received contracts. Standard deviations in parentheses.

Table 6: Determinants of Willingness to Accept

	Cross Section				Random Effects				Fixed Effects	
	(1)		(2)		(3)		(4)		(5)	
Peak Season	-0.780	**	-0.753		-0.901	**	-0.869	*		
	(0.377)		(0.469)		(0.386)		(0.477)			
Monitoring Treatment	-0.431		-0.470		-0.297		-0.418		-0.189	
	(0.269)		(0.321)		(0.221)		(0.266)		(0.235)	
Female	-2.321	***	-2.789	***	-2.635	***	-2.965	***		
	(0.449)		(0.529)		(0.453)		(0.544)			
Second day	-0.904	***	-0.795	*	-0.998	***	-0.854	**	-1.040	***
	(0.343)		(0.410)		(0.319)		(0.379)		(0.330)	
Third day	-0.326		0.088		-0.177		0.237		-0.108	
	(0.353)		(0.413)		(0.336)		(0.396)		(0.352)	
Fourth day	-1.086	***	-0.935	**	-1.072	***	-0.879	**	-1.077	***
	(0.351)		(0.416)		(0.339)		(0.401)		(0.353)	
Indiv. Controls	No		Yes		No		Yes			
District FEs	Yes		Yes		Yes		Yes			
Indiv. FEs	No		No						Yes	
Mean Dep. Var.	10.32		10.60		10.32		10.60		10.32	
SD Dep. Var.	5.899		5.993		5.899		5.993		5.899	
Adj. R-squared	0.0555		0.123						0.0179	
Observations	1857		1331		1857		1331		1857	

*, **, *** denote significance at 10%, 5% and 1%, respectively.

Notes: this table presents regressions of minimum willingness to accept (WTA) on season (peak labor demand), monitoring, whether the participant was female, and day-of-week fixed effects, with the first day as the ommitted category. Columns (2) and (4) control for individual covariates. All regressions include district fixed effects, although in column (5) these are absorbed by the individual fixed effects. Standard errors robust to clustering at the participant level are in parentheses.

Table 7: Determinants of Quantity (Number of Scoops Sorted per Day)

	Random Effects				Fixed Effects
	(1)	(2)	(3)	(4)	(5)
Monitoring treatment	-0.564 ***	-0.558 ***	-0.782 ***	-0.749 ***	-0.769 ***
	(0.066)	(0.064)	(0.195)	(0.196)	(0.195)
Draw		0.031 ***	0.024 ***	0.022 ***	0.025 ***
		(0.005)	(0.007)	(0.008)	(0.007)
Monitoring X Draw			0.013	0.008	0.012
			(0.010)	(0.011)	(0.010)
Minimum WTA	-0.006	-0.020 *	-0.020 *	-0.032 ***	-0.026 **
	(0.010)	(0.011)	(0.011)	(0.010)	(0.011)
Peak season	0.412 ***	0.410 ***	0.409 ***	0.490 ***	
	(0.130)	(0.131)	(0.131)	(0.152)	
Female	0.731 ***	0.739 ***	0.741 ***	0.776 ***	
	(0.139)	(0.139)	(0.139)	(0.154)	
Day 2	1.169 ***	1.175 ***	1.178 ***	1.303 ***	1.189 ***
	(0.086)	(0.084)	(0.084)	(0.094)	(0.084)
Day 3	1.847 ***	1.847 ***	1.848 ***	2.025 ***	1.871 ***
	(0.091)	(0.091)	(0.091)	(0.106)	(0.091)
Day 4	2.064 ***	2.072 ***	2.072 ***	2.121 ***	2.094 ***
	(0.083)	(0.081)	(0.081)	(0.097)	(0.081)
Indiv. Controls	No	No	No	Yes	
District FEs	Yes	Yes	Yes	Yes	
Indiv. Effects	Rand.	Rand.	Rand.	Rand.	Fixed
Mean Dep. Var.	7.350	7.350	7.350	7.473	7.350
SD Dep. Var.	1.975	1.975	1.975	1.953	1.975
Observations	1461	1461	1461	1028	1461

*, **, *** denote significance at 10%, 5% and 1%, respectively.

Notes: this table presents regressions of quantity sorted (number of scoops) on whether the participant was assigned to the monitoring treatment, the actual piece rate the participant received (Draw), the interaction of monitoring and the actual piece rate (Monitoring X Draw), the minimum piece rate the participant was willing to accept (Minimum WTA), and day-of-week fixed effects, with the first day as the ommitted category. All regressions include season (peak labor) and district fixed effects, although in column (5) these are absorbed by the individual fixed effects. Standard errors robust to clustering at the participant level are in parentheses.

Table 8: Determinants of Quality (Number of Errors per Scoop)

	Random Effects				Fixed Effects
	(1)	(2)	(3)	(4)	(5)
Monitoring treatment	-0.628 ***	-0.627 ***	-0.670 ***	-0.765 ***	-0.696 ***
	(0.042)	(0.042)	(0.136)	(0.162)	(0.136)
Draw		0.008 **	0.007	0.005	0.005
		(0.004)	(0.006)	(0.007)	(0.006)
Monitoring X Draw			0.002	0.008	0.004
			(0.008)	(0.009)	(0.008)
Minimum WTA	0.001	-0.003	-0.003	-0.005	0.000
	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)
Peak season	-0.064	-0.064	-0.064	-0.073	
	(0.057)	(0.057)	(0.057)	(0.073)	
Female	-0.171 **	-0.168 **	-0.168 **	-0.281 ***	
	(0.071)	(0.071)	(0.071)	(0.081)	
Day 2	-0.309 ***	-0.308 ***	-0.307 ***	-0.285 ***	-0.310 ***
	(0.076)	(0.076)	(0.076)	(0.088)	(0.076)
Day 3	-0.462 ***	-0.462 ***	-0.462 ***	-0.324 ***	-0.476 ***
	(0.073)	(0.073)	(0.073)	(0.092)	(0.072)
Day 4	-0.529 ***	-0.528 ***	-0.528 ***	-0.369 ***	-0.541 ***
	(0.073)	(0.073)	(0.073)	(0.089)	(0.074)
Indiv. Controls	No	No	No	Yes	
District FEs	Yes	Yes	Yes	Yes	
Indiv. Effects	Rand.	Rand.	Rand.	Rand.	Fixed
Mean Dep. Var.	1.883	1.883	1.883	1.862	1.883
SD Dep. Var.	1.013	1.013	1.013	0.984	1.013
Observations	1461	1461	1461	1028	1461

*, **, *** denote significance at 10%, 5% and 1%, respectively.

Notes: this table presents regressions of quality (number of errors per scoop) on whether the participant was assigned to the monitoring treatment, the actual piece rate the participant received (Draw), the interaction of monitoring and the actual piece rate (Monitoring X Draw), the minimum piece rate the participant was willing to accept (Minimum WTA), and day-of-week fixed effects, with the first day as the ommitted category. All regressions include season (peak labor) and district fixed effects, although in column (5) these are absorbed by the individual fixed effects. Standard errors robust to clustering at the participant level are in parentheses.

Table 9: Determinants of Willingness to Accept
Gender Differences

	Cross Section				Random Effects			
	(1)		(2)		(3)		(4)	
Peak Season								
Among Men	-0.234		-0.577		-0.423		-0.565	
	(0.790)		(0.802)		(0.793)		(0.823)	
Among Women	-1.029	**	-0.891		-1.126	***	-1.083	**
	(0.415)		(0.547)		(0.427)		(0.550)	
Monitoring								
Among Men	-0.887		-0.960		-0.622		-0.893	*
	(0.574)		(0.606)		(0.439)		(0.481)	
Among Women	-0.223		-0.168		-0.155		-0.135	
	(0.288)		(0.365)		(0.252)		(0.312)	
Indiv. Controls	No		Yes		No		Yes	
District FEs	Yes		Yes		Yes		Yes	
Mean Dep. Var.	10.32		10.60		10.32		10.60	
SD Dep. Var.	5.899		5.993		5.899		5.993	
Observations	1857		1331		1857		1331	

*, **, *** denote significance at 10%, 5% and 1%, respectively.

Notes: this table presents differential effects by gender of season (peak labor demand) and monitoring. Additional regressors not reported are as in Table 3.A: whether the participant was female (i.e. the level effect, as opposed to the interactions reported in this table), day-of-week fixed effects, district fixed effects, and, in columns (2) and (4), individual covariates. Standard errors robust to clustering at the participant level are in parentheses.

Table 10: Determinants of Quantity (Number of Scoops Sorted per Day)
Differential Effects by Gender

	Random Effects			
	(1)	(2)	(3)	(4)
Monitoring				
Among Men	-0.325 ***	-0.339 ***	-0.452	-0.078
	(0.125)	(0.121)	(0.364)	(0.297)
Among Women	-0.659 ***	-0.645 ***	-0.914 ***	-1.079 ***
	(0.076)	(0.074)	(0.229)	(0.245)
Draw				
Among Men		0.032 ***	0.028 **	0.025 *
		(0.010)	(0.013)	(0.013)
Among Women		0.030 ***	0.022 ***	0.021 **
		(0.006)	(0.008)	(0.009)
Monitoring X Draw				
Among Men			0.006	-0.013
			(0.018)	(0.016)
Among Women			0.016	0.019
			(0.012)	(0.014)
Minimum WTA				
Among Men	-0.034 **	-0.049 ***	-0.049 ***	-0.037 **
	(0.017)	(0.018)	(0.018)	(0.015)
Among Women	0.008	-0.006	-0.006	-0.026 *
	(0.012)	(0.013)	(0.013)	(0.013)
Peak Labor				
Among Men	1.013 ***	1.020 ***	1.024 ***	0.995 ***
	(0.225)	(0.226)	(0.226)	(0.217)
Among Women	0.180	0.173	0.167	0.242
	(0.156)	(0.157)	(0.157)	(0.185)
Indiv. Controls	No	No	No	Yes
District FEs	Yes	Yes	Yes	Yes
Mean Dep. Var.	7.350	7.350	7.350	7.473
SD Dep. Var.	1.975	1.975	1.975	1.953
Observations	1461	1461	1461	1028

*, **, *** denote significance at 10%, 5% and 1%, respectively.

Notes: this table presents differential effects by gender of whether the participant was assigned to the monitoring treatment (Monitoring), the actual piece rate the participant received (Draw), the interaction of monitoring and the actual piece rate (Monitoring X Draw), the minimum piece rate the participant was willing to accept (Minimum WTA), and season (peak labor demand) on the quantity of beans sorted (number of scoops per day). Additional regressors not reported include whether the participant was female (i.e. the level effect, as opposed to the interactions reported in this table), day-of-week fixed effects, district fixed effects, and, in column (4), individual covariates. Standard errors robust to clustering at the participant level are in parentheses.

Table 11: Determinants of Quality (Number of Errors per Scoop)
Differential Effects by Gender

	Random Effects			
	(1)	(2)	(3)	(4)
Monitoring				
Among Men	-0.626 ***	-0.636 ***	-0.712 **	-0.880 ***
	(0.089)	(0.090)	(0.298)	(0.334)
Among Women	-0.629 ***	-0.628 ***	-0.618 ***	-0.660 ***
	(0.047)	(0.047)	(0.148)	(0.169)
Draw				
Among Men		0.019 **	0.017	0.013
		(0.009)	(0.014)	(0.015)
Among Women		0.004	0.004	0.003
		(0.004)	(0.007)	(0.007)
Monitoring X Draw				
Among Men			0.004	0.018
			(0.017)	(0.018)
Among Women			-0.001	-0.000
			(0.008)	(0.009)
Minimum WTA				
Among Men	0.004	-0.004	-0.004	-0.014
	(0.010)	(0.011)	(0.011)	(0.012)
Among Women	-0.001	-0.002	-0.002	0.002
	(0.006)	(0.007)	(0.007)	(0.008)
Peak Labor				
Among Men	-0.237 *	-0.232 *	-0.230 *	-0.222
	(0.127)	(0.127)	(0.127)	(0.138)
Among Women	0.005	0.005	0.005	0.031
	(0.062)	(0.062)	(0.063)	(0.076)
Indiv. Controls	No	No	No	Yes
District FEs	Yes	Yes	Yes	Yes
Mean Dep. Var.	1.883	1.883	1.883	1.862
SD Dep. Var.	1.013	1.013	1.013	0.984
Observations	1461	1461	1461	1028

*, **, *** denote significance at 10%, 5% and 1%, respectively.

Notes: this table presents differential effects by gender of whether the participant was assigned to the monitoring treatment (Monitoring), the actual piece rate the participant received (Draw), the interaction of monitoring and the actual piece rate (Monitoring X Draw), the minimum piece rate the participant was willing to accept (Minimum WTA), and season (peak labor demand) on the quality of output (number of errors per scoop). Additional regressors not reported include whether the participant was female (i.e. the level effect, as opposed to the interactions reported in this table), day-of-week fixed effects, district fixed effects, and, in column (4), individual covariates. Standard errors robust to clustering at the participant level are in parentheses.

Table 12: Effect of Compensation and Monitoring on Quality of Output

	(1)		(2)		(3)	
Draw = 5 (base)	0		0		0	
	.		.		.	
Draw = 10	-0.152		-0.140		-0.199	
	(0.170)		(0.170)		(0.174)	
Draw = 15	-0.183		-0.174		-0.092	
	(0.152)		(0.156)		(0.161)	
Draw = 20	-0.209		-0.178		-0.149	
	(0.153)		(0.155)		(0.162)	
Draw = 25	0.043		0.054		-0.020	
	(0.154)		(0.157)		(0.160)	
Monitoring	-0.844	***	-0.843	***	-0.836	***
	(0.173)		(0.173)		(0.175)	
Monitoring X (Draw = 10)	0.160		0.157		0.146	
	(0.220)		(0.220)		(0.221)	
Monitoring X (Draw = 15)	0.325		0.323		0.316	
	(0.202)		(0.202)		(0.203)	
Monitoring X (Draw = 20)	0.336		0.329		0.319	
	(0.192)		(0.192)		(0.194)	
Monitoring X (Draw = 25)	0.124		0.121		0.114	
	(0.198)		(0.198)		(0.200)	
Female	-0.174	*	-0.176	*	-0.174	*
	(0.071)		(0.071)		(0.071)	
Min. WTA categories	No		Yes		Yes	
Min. WTA X Draw	No		No		Yes	
Mean Dep. Var.	1.883		1.883		1.883	
SD Dep. Var.	1.012		1.013		1.013	
Overall R-squared	0.175		0.181		0.187	
Observations	1462		1461		1461	

*, **, *** denote significance at 10%, 5% and 1%, respectively.

Notes: this table presents regressions of quality of output (number of errors detected per scoop) on the actual piece rate the participant faced (Draw), whether the participant was assigned to the monitoring treatment (Monitoring Treatment), the interaction between the actual piece rate and monitoring, and an indicator for whether the participant was female. Other regressors not reported are the minimum piece rate the participant was willing to accept (Minimum WTA), in levels and interacted with monitoring, and season, district and day-of-week fixed effects. Standard errors robust to clustering at the participant level are in parentheses.

Table 13.A: Peer Effects: Quantity (Number of Scoops per Day)

	(1)		(2)		(3)	
Mean draw in work group (excluding own)	-0.007		-0.009		0.065	**
	(0.009)		(0.024)		(0.027)	
Number of individuals in work group	-0.027		-0.027		0.299	**
	(0.042)		(0.042)		(0.121)	
Mean Draw X Own Draw			0.000			
			(0.001)			
Mean Draw X Group Size					-0.019	***
					(0.007)	
Mean of the minimum WTA in work group (excluding own)	-0.003		-0.003		-0.004	
	(0.015)		(0.015)		(0.015)	
Minimum WTA = 10	0.108		0.108		0.102	
	(0.092)		(0.092)		(0.091)	
Minimum WTA = 15	-0.185		-0.186		-0.197	
	(0.163)		(0.163)		(0.163)	
Minimum WTA = 20	-0.325		-0.325		-0.344	
	(0.227)		(0.227)		(0.228)	
Minimum WTA = 25	-0.748		-0.746		-0.770	*
	(0.470)		(0.467)		(0.456)	
Draw = 10	0.247	**	0.236	*	0.252	**
	(0.120)		(0.136)		(0.118)	
Draw = 15	0.329	**	0.307		0.338	**
	(0.135)		(0.209)		(0.132)	
Draw = 20	0.647	***	0.613	**	0.653	***
	(0.130)		(0.299)		(0.128)	
Draw = 25	0.626	***	0.582		0.642	***
	(0.139)		(0.412)		(0.137)	
Monitoring	-0.563	***	-0.563	***	-0.577	***
	(0.087)		(0.087)		(0.088)	
Female	0.770	***	0.769	***	0.780	***
	(0.156)		(0.157)		(0.156)	
Peak Labor Season	0.414	**	0.414	**	0.410	**
	(0.176)		(0.176)		(0.176)	
District FEs	Yes		Yes		Yes	
Indiv. Effects	Random		Random		Random	
Mean Dep. Var.	7.338		7.338		7.338	
SD Dep. Var.	1.973		1.973		1.973	
Num. work groups	383		383		383	
Num. Indiv.	608		608		608	
Observations	1440		1440		1440	

*, **, *** denote significance at 10%, 5% and 1%, respectively.

Notes: this table presents regressions of quantity sorted (number of scoops) on the mean draw in the participant's work group (excluding the participant's own draw), the group size, and their interaction. Other controls include: mean minimum WTA in the group (excluding the participant's own minimum WTA), indicators for the participant's minimum WTA, indicators for the participant's draw, and fixed effects for peak labor season, whether the participant was assigned to the monitoring treatment, whether the participant is female. All regressions include district fixed effects and day-of-week fixed effects. Observations are at the individual-by-day level. Standard errors are twoway-clustered at the individual and work-group level.

Table 13.B: Peer Effects: Quality (Number of Errors per Scoop)

	(1)	(2)	(3)
Mean draw in work group (excluding own)	0.006	0.004	0.002
	(0.007)	(0.018)	(0.024)
Number of individuals in work group	0.012	0.012	-0.008
	(0.037)	(0.037)	(0.103)
Mean Draw X Own Draw		0.000	
		(0.001)	
Mean Draw X Group Size			0.001
			(0.006)
Mean of the minimum WTA in work group (excluding own)	0.013	0.013	0.013
	(0.010)	(0.010)	(0.010)
Minimum WTA = 10	-0.047	-0.047	-0.047
	(0.064)	(0.063)	(0.064)
Minimum WTA = 15	-0.009	-0.009	-0.008
	(0.086)	(0.086)	(0.086)
Minimum WTA = 20	-0.189 **	-0.189 **	-0.188 **
	(0.089)	(0.089)	(0.089)
Minimum WTA = 25	0.567	0.569	0.567
	(0.388)	(0.389)	(0.388)
Draw = 10	-0.056	-0.064	-0.056
	(0.113)	(0.161)	(0.113)
Draw = 15	0.016	-0.000	0.016
	(0.110)	(0.219)	(0.110)
Draw = 20	-0.002	-0.027	-0.003
	(0.111)	(0.299)	(0.112)
Draw = 25	0.122	0.089	0.121
	(0.106)	(0.373)	(0.107)
Monitoring	-0.630 ***	-0.630 ***	-0.630 ***
	(0.050)	(0.050)	(0.050)
Female	-0.161 *	-0.161 *	-0.162 *
	(0.087)	(0.087)	(0.087)
Peak Labor Season	-0.067	-0.067	-0.067
	(0.067)	(0.067)	(0.068)
District FEs	Yes	Yes	Yes
Indiv. Effects	Random	Random	Random
Mean Dep. Var.	1.887	1.887	1.887
SD Dep. Var.	1.015	1.015	1.015
Num. work groups	383	383	383
Num. Indiv.	608	608	608
Observations	1440	1440	1440

*, **, *** denote significance at 10%, 5% and 1%, respectively.

Notes: this table presents regressions of quality of output (number of errors per scoop) on the mean draw in the participant's work group (excluding the participant's own draw), the group size, and their interaction. Other controls include: mean minimum WTA in the group (excluding the participant's own minimum WTA), indicators for the participant's minimum WTA, indicators for the participant's draw, and fixed effects for peak labor season, whether the participant was assigned to the monitoring treatment, whether the participant is female. All regressions include district fixed effects and day-of-week fixed effects. Observations are at the individual-by-day level. Standard errors are twoway-clustered at the individual and work-group level.

Figure 1: Number of scoops sorted, by monitoring treatment

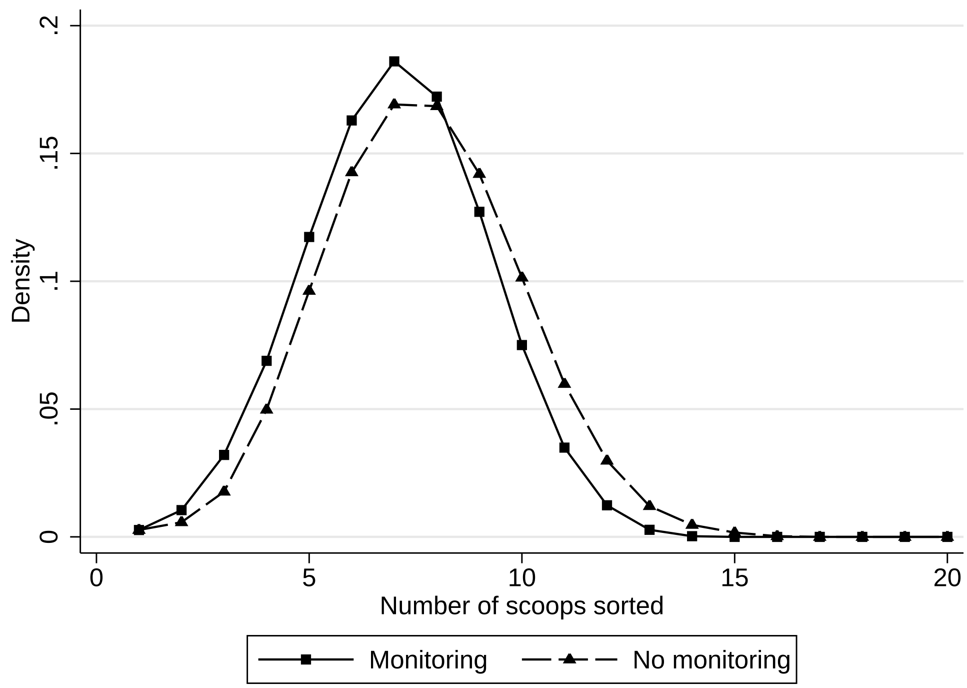


Figure 2: Number of scoops sorted, by season

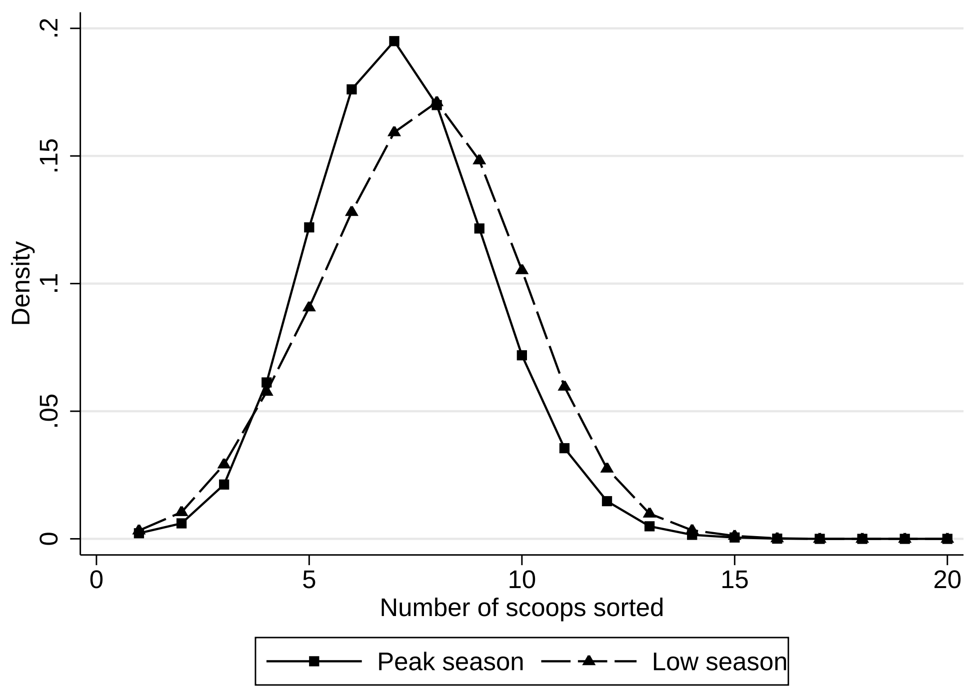


Figure 3: Errors per scoop, by monitoring treatment

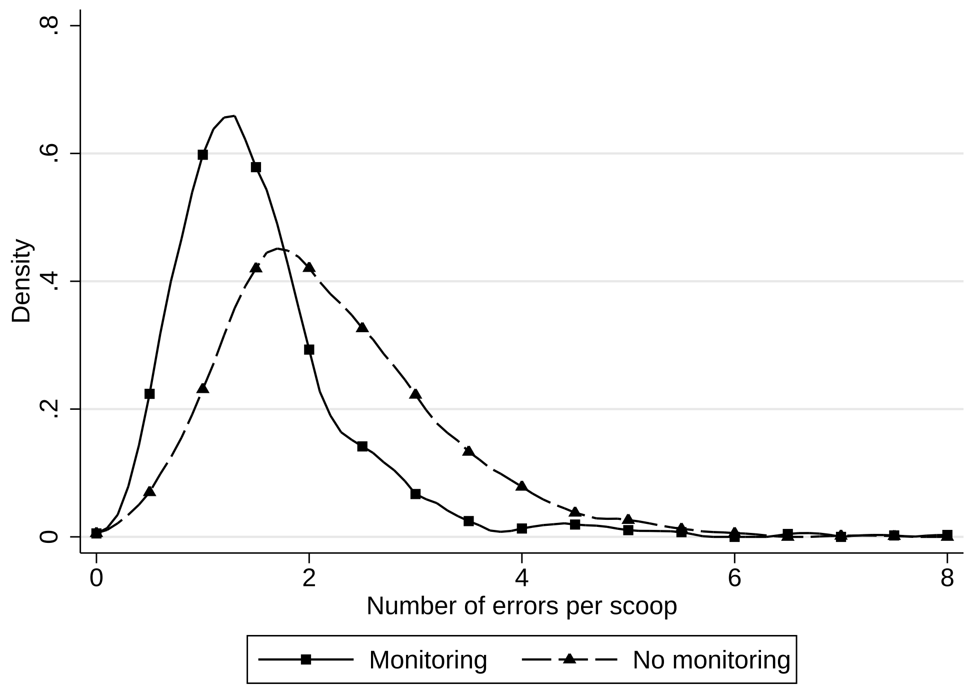


Figure 4: Errors per scoop, by season

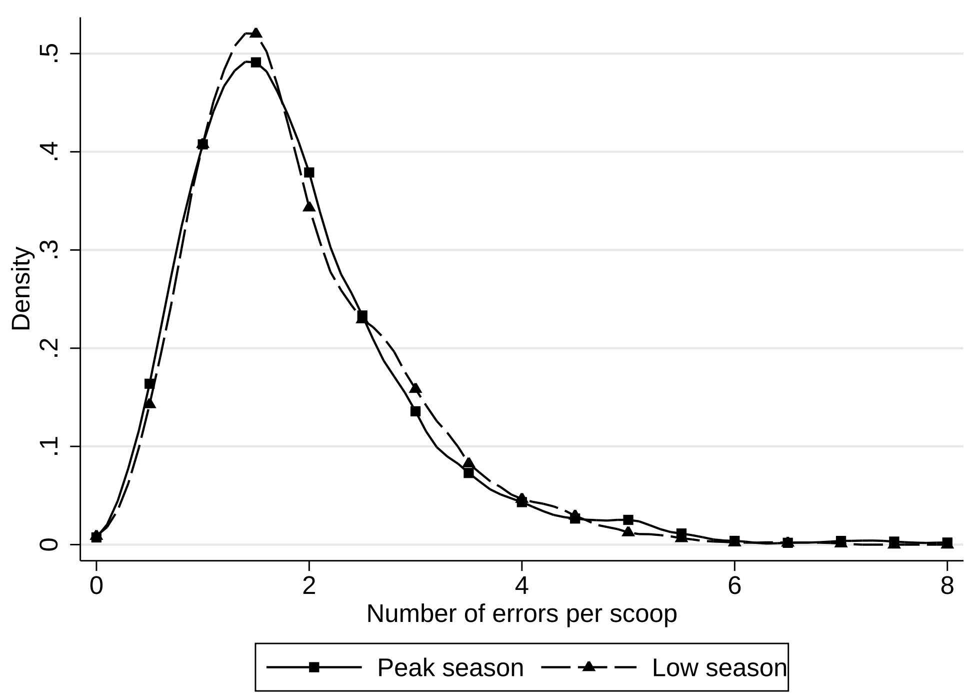


Figure 5: Marginal effect of minimum WTA on number of scoops,
(by piece rate interval)

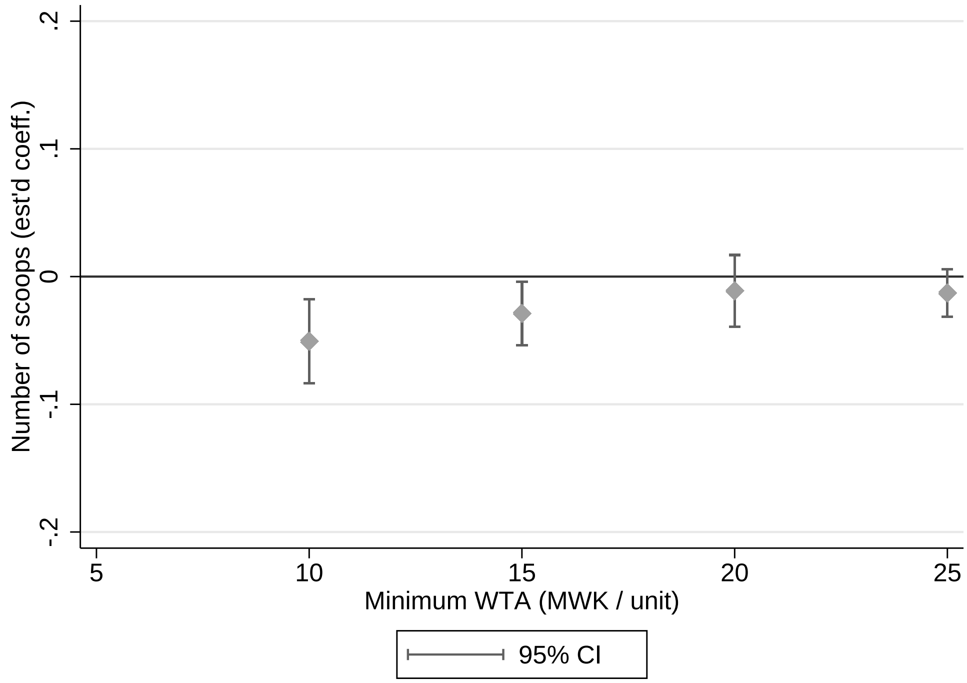


Figure 6: Marginal effect of minimum WTA on errors per scoop,
(by piece rate interval)

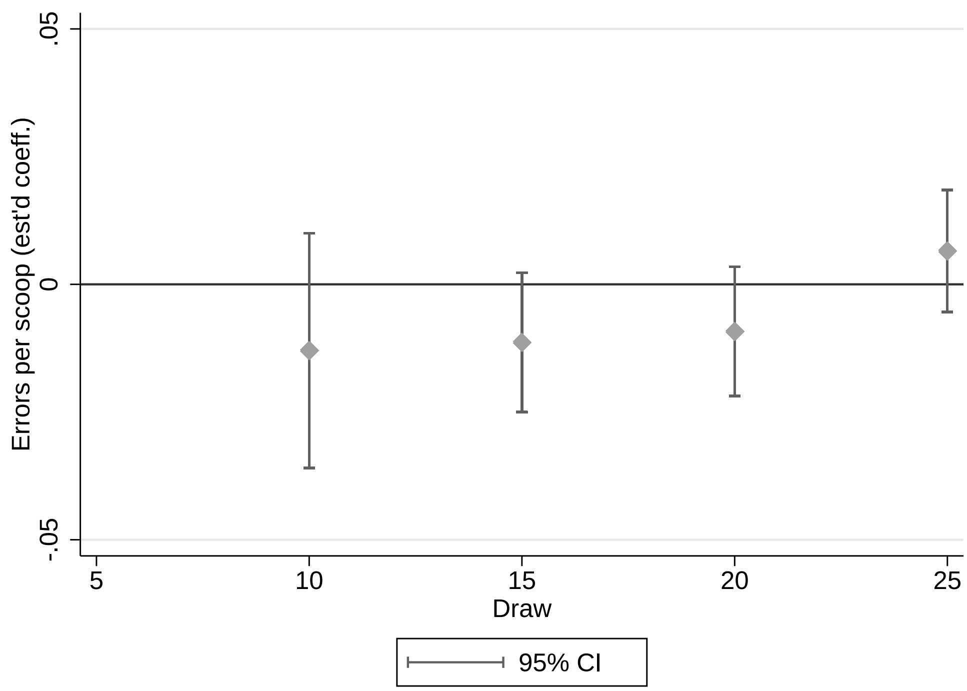


Figure 7: Errors per scoop, by piece rate draw and monitoring treatment (estimated coefficients)

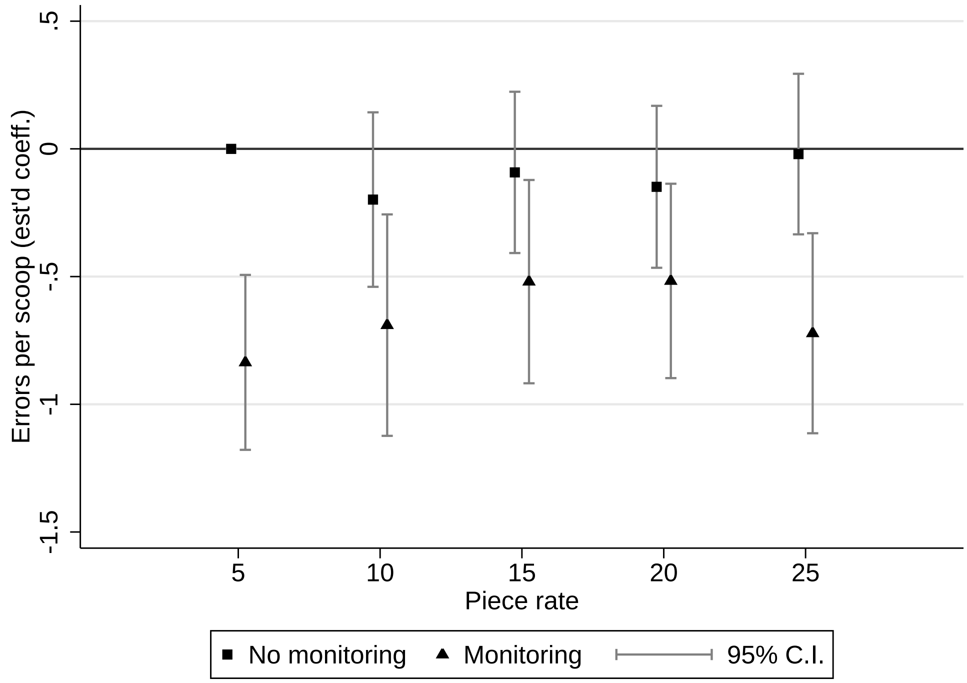


Figure 8: Errors per scoop, by scoop and by monitoring treatment

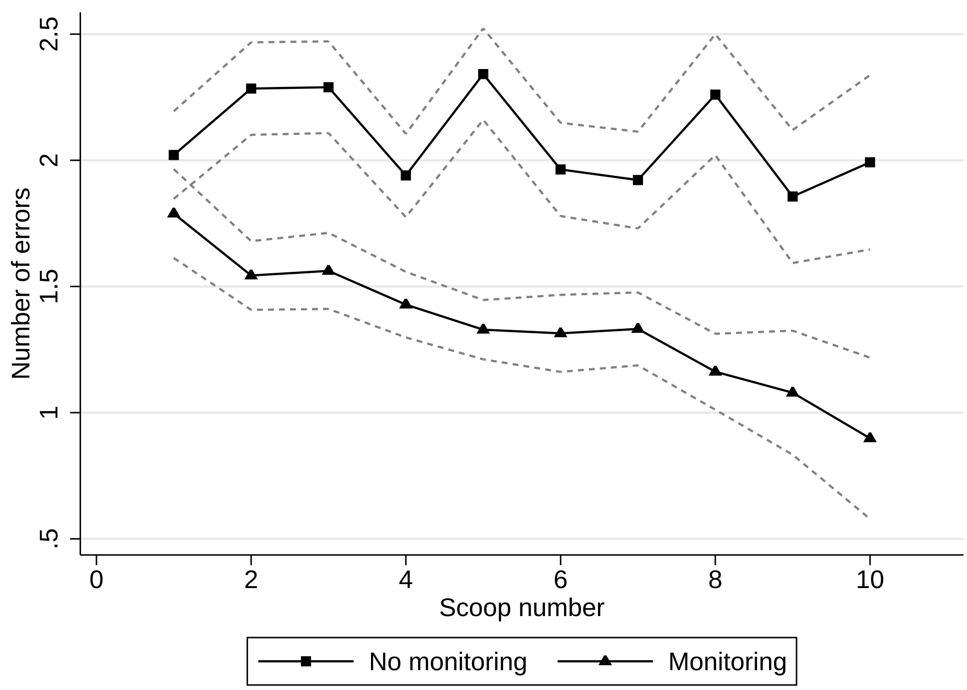


Figure 9: Evolution of sorting time by unit over the day

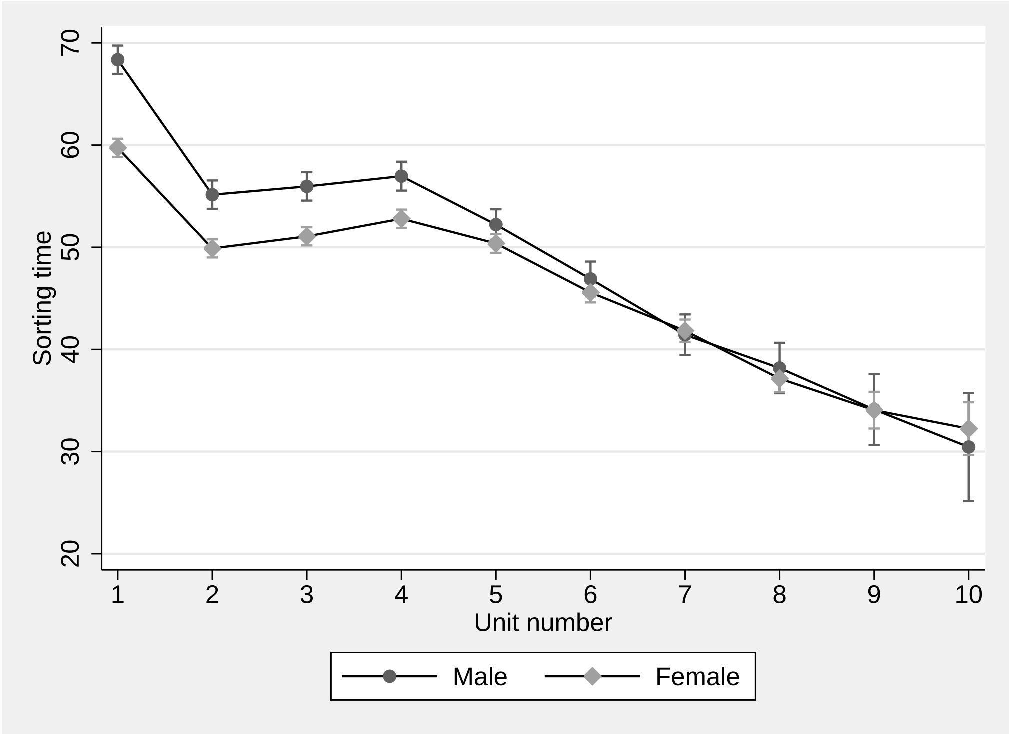


Figure 10: Relationship between sorting time and piece rate (semiparametric)

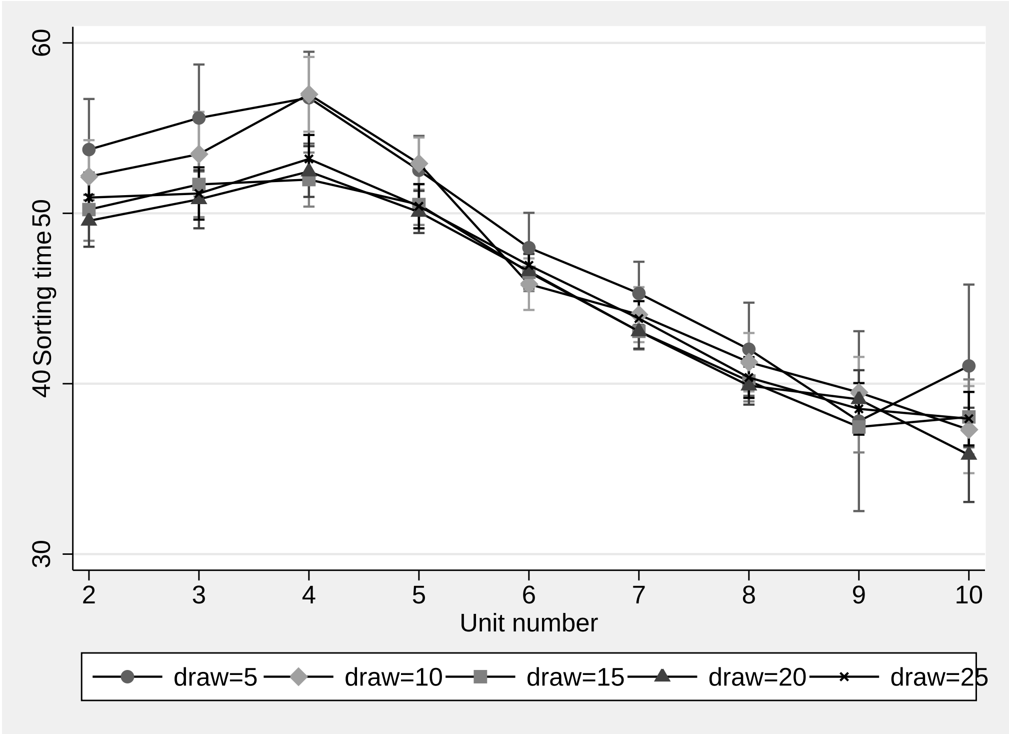
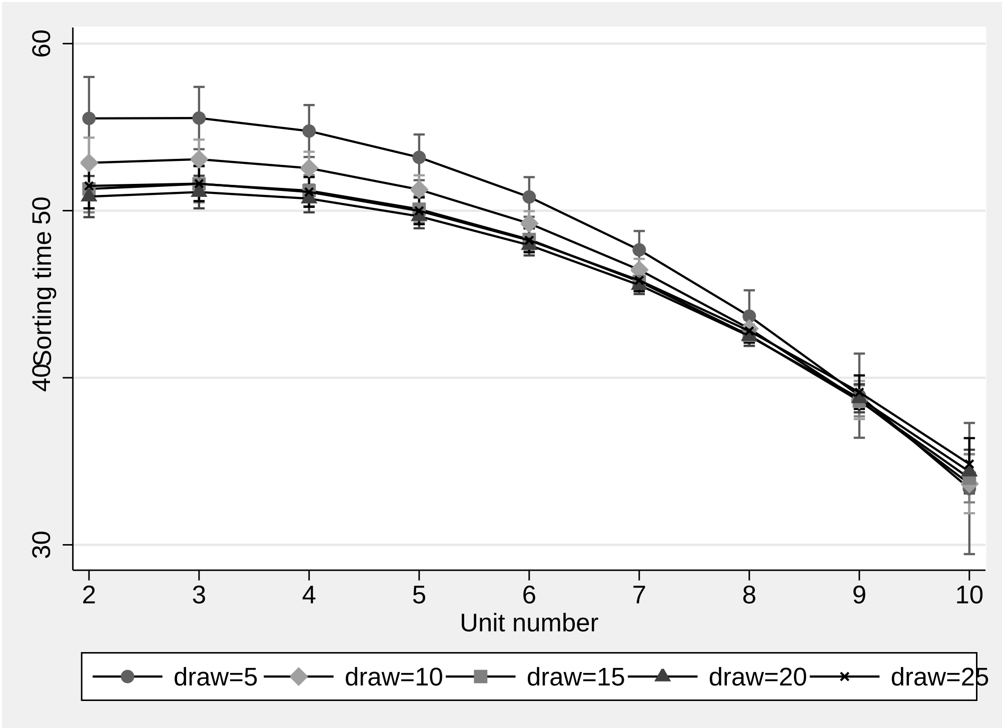


Figure 11: Relationship between sorting time and piece rate (quadratics with interactions)



Appendix: BDM script

Labor allocation survey 2010 BDM response form NO MONITORING (NM)

Surveyor initials: ________________________ GVH: ________________________

Respondent ID: ________________________ Time: __________________ (24:00)

Respondent name: ____________________________ Date: ________________________

Before we proceed with your wage choices, do you have any questions?

If you agree to proceed, I will ask you whether you will be willing to work for several different wages. As we showed in the example, you will draw a token that determines which of these wages you receive. If you say no to that wage and you draw it, you will not be given a contract and you will not have a chance to change your mind. If you say yes to that wage and you draw it, you will be expected to work and be paid that rate for each scoop of beans that you sort. If you are part way through a scoop at 4pm, you will receive a portion of the pay for that scoop.

Do you wish to proceed? <<*circle one*>> Yes No

We will look at your beans to make sure you are sorting according to instructions. If you are not sorting according to instructions, we will ask you to go back and sort the beans again according to the instructions. Is this clear?

BDM Question Responses ***Respondent answers***

<<*For each of the following wages, confirm 3 times and describe implications of yes/no*>>

	Question:	*Code:*	*1st*	*2nd*	*3rd*	*FINAL*
1	If you pick MK 5, would you accept that as your wage?	Yes: 1 No: 0				
	Most people can sort at least 4 scoops in a day. If you sort 4 scoops and your rate is MK 5 per scoop, you will take home MK20 plus 50, for a total of MK70. You may earn more or less depending on how hard you work.					
2	If you pick MK 10, would you accept that as your wage?	Yes: 1 No: 0				
	Most people can sort at least 4 scoops in a day. If you sort 4 scoops and your rate is MK 10 per scoop, you will take home MK40 plus 50, for a total of MK90. You may earn more or less depending on how hard you work.					
3	If you pick MK 15, would you accept that as your wage?	Yes: 1 No: 0				
	Most people can sort at least 4 scoops in a day. If you sort 4 scoops and your rate is MK 15 per scoop, you will take home MK60 plus 50, for a total of MK110. You may earn more or less depending on how hard you work.					
4	If you pick MK 20, would you accept that as your wage?	Yes: 1 No: 0				
	Most people can sort at least 4 scoops in a day. If you sort 4 scoops and your rate is MK 20 per scoop, you will take home MK80 plus 50, for a total of MK130. You may earn more or less depending on how hard you work.					
5	If you pick MK 25, would you accept that as your wage?	Yes: 1 No: 0				
	Most people can sort at least 4 scoops in a day. If you sort 4 scoops and your rate is MK 25 per scoop, you will take home MK100 plus 50, for a total of MK150. You may earn more or less depending on how hard you work.					

BDM Draw Responses

	Question:	*Code:*	*Answer:*	*Instructions:*
6	What price did you draw from the basket?	MK		
7	*Did the respondent say he/she would accept a contract at that wage?*	Yes: 1 No: 0		1→8 0→9
8	*If yes, read:* You said you would accept a wage of <<*say wage*>>, so we will offer you a labor contract at that wage.			→ contract

Contract

I will work sorting beans until up to 4pm at the latest, and will receive __________ <<*Fill in wage*>> per scoop that I sort during that time.

ID Number: ________________________ Signature: ______________________

Comments *(refused the contract offer, was called away during the day, etc)*: