Wage Differentials and Workers' Effort:

Experimental Evidence from Uganda.

John Sseruyange[1], Erwin Bulte[2]

[1]School of Economics, Makerere University. Plot 51, Pool Road. P.O BOX 7062, Kampala Uganda. E-mail:johnsseruyange@gmail.com
[2]Development Economics Group, Wageningen University, P.O. Box 8130, 6700 EW Wageningen, Netherlands.

Abstract: We organize a real-effort field experiment with varying piece rates to assess the impact of wages and social comparisons on productivity. In addition to analyzing how piece rates affect productivity during the "paid stage" of the experiment, we also consider how social comparisons affect effort supply during a voluntary and unpaid follow-up task. Our main results are that effort supply is relatively unresponsive to variation in own earnings, but responds strongly to pay inequality. While we only obtain weak support for the hypothesis that positive social comparisons invite extra effort during paid stages of the experiment, the effect of social comparisons is important during voluntary tasks. Specifically, positive social comparisons positively affect productivity during unpaid tasks, and negative comparisons have the opposite impact.

Keywords: Social comparisons, real effort task, fair wage, efficiency wages

1. Introduction

While conventional economic models assume that people only care about their own income and consumption level, a rich literature suggests most people also care both about how their income and consumption levels compare to those of others (e.g., Leibenstein 1950, Duessenberry 1952, Frank 1985). The implications of concerns about one's relative position have been studied for a range of issues, including happiness and well-being. A related literature considers the behavioral effects of social comparisons, including studies of job choice and effort provision. Social comparisons among workers may occur within firms as well as in the wider labor market, affecting job satisfaction (Card et al. 2012) so that relative pay can act as a compensating differential. Concerns about relative wages may also help explain wage compression (Charness and Kuhn 2007), secrecy rules regarding earnings, or the sorting of heterogeneous workers across firms. Since within-firm workers make for a more salient reference group than outside workers (e.g. Clark and Senik 2010), relative pay concerns also affect whether specific tasks are contracted out or organized within firm boundaries (Nickerson and Zenger 2008).[1]

In this paper we consider the interface of social comparisons and efficiency wages, or gift-exchange between employer and worker. We focus on effort supply in a context where workers interact and are able to compare their earnings. Social comparisons may shape reference values, and help individuals to decide whether they have been treated "fairly." The literature on the fair wage-effort hypothesis suggests effort supply is governed by a desire for reciprocity – workers respond to higher wages by working harder. The presence of such reciprocal workers affects how labor markets work, and could result in non-competitive wages and involuntary unemployment (e.g., Akerlof 1982, Akerlof and Yellen 1990). Reciprocity in

[1] Wage differences may of course create dynamic incentives, as workers may seek to acquire experience and invest in their human capital to increase their productivity and qualify for high wages in the future.

the workplace may be rationalized by reference-dependent preferences, but alternative rationalizations also exist. Social comparisons are an important mechanism helping workers to form reference values.

If social comparisons shift reference values they may also affect effort supply – especially in a context of imperfect contractibility. A small literature, summarized below, analyzes this issue. We contribute to this literature by examining whether absolute and relative earning levels affect effort supply in a real-effort field experiment in Africa. Our study is distinct for three reasons. First, in addition to focusing on effort supplied during the (paid) experimental task, we also elicit willingness to supply effort to a voluntary unpaid task *after the experiment*. This allows us to probe reciprocity beyond the quid-pro-quo relationship defined by the contract. Second, we use an experimental design based on exogenous variation in *piece rates*, rather than in the hourly wage as in most other studies. Bracha et al. (2015) also analyze the implications of varying piece rates, but focus on labor supply (time spent on a task), rather than productivity, or output per unit of time. Third, we use a non-conventional sample of African respondents – mainly smallholders and casual workers – that is quite distinct from the ones used in existing work. This allows us to verify the robustness of insights obtained earlier, in a different cultural context.

An important reason for studying piece rates in our experiment is the simple fact that piece rates represent the main payment regime in the context of informal or semi-informal labor markets in developing countries. It also avoids confusion about whether or not higher earnings are justified by greater productivity. Of course a drawback of using piece rate treatments is that the identification of reciprocity is not straightforward, as extrinsic motives for effort supply vary with the level of the piece rate. As mentioned by Bellemare and Shearer (2009), gift exchange occurs in a piece rate setting as long as the employer's valuation of the produced output is sufficiently high (exceeding the marginal cost due to the piece rate). To cleanly identify the effect on reciprocity we therefore analyze effort supply during a voluntary (unpaid) task. We also include two control groups (where social comparisons are eliminated by design) with high and low piece rates to study productivity.

We find that effort choices are not very sensitive to own earnings: piece rate levels do not explain variation in effort in the absence of social comparisons. However, relative wages have an impact on effort supply, but not always. Unlike earlier work we do *not* find that, relative to a control group, low wages reduce effort during the paid stage of the experiment. We document some evidence that positive comparisons make people more productive. Importantly, social comparisons significantly affect productivity during unpaid stages of the experiment. Low-piece rate workers supply less effort during the unpaid task when shirking is “cheap” and workers do not have to sacrifice their own income if they choose to be unproductive. This finding supports earlier studies based on fixed wages, where withholding effort is also privately cheap. Reciprocity carries over to settings beyond the contract, which is arguably an important result as many organizations depend on voluntary or badly monitored contributions of their co-workers for (economic) success.

The paper is organized as follows. In the next section we briefly summarize the existing literature on social comparisons and effort supply, focusing on recent experimental studies. In section 3 we introduce our experiment, summarize our data, and outline our identification strategy. Section 4 contains our empirical results, focusing both on effort supply for the paid and voluntary tasks. The conclusions and discussion ensue.

2. Social comparisons and effort supply

Starting with seminal work of Fehr et al. (1993), many papers have analyzed the fair wage-effort hypothesis, postulating that fairness considerations affect the supply of effort by workers if effort is not perfectly contractible. To decide about what constitutes a “fair” level of earnings, a worker can refer back to earlier payments (Cohn et al. 2010, Clark et al. 2010, Bracha et al. 2015) or consider how the surplus is (vertically) distributed between herself and the employer (Hennig-Schmidt et al. 2010). Importantly, the evaluation of "what is fair" may also be based on horizontal comparisons among peers engaged in the same activity. Social comparisons help workers to form reference values, and own earnings are likely to be evaluated as "fair" when exceeding the reference level. Conversely, wages below the reference value may be seen as "unfair" and invite feelings of disappointment or even anger.

The literature contains several explanations for why fairness evaluations may affect effort supply in the context of fixed wages. Clark et al. (2010) framed their results in terms of a concern for status, but observe that inequality aversion would yield similar results. Workers with high earnings supply more effort and incur a greater effort cost to reduce the gap between their "net earnings" and those of others. Relatedly, Hennig-Schmidt et al. (2010) introduce tensions due to cognitive dissonance. To alleviate tensions between an offered and reference wage, workers may provide extra effort (for wages exceeding the reference value) or shirk (in case the reverse is true). An alternative explanation would be a model based on a reciprocity norm prescribing how much effort to provide in response to a certain wage level – where higher wages imply workers should work harder. The salience of such a norm could vary with perceived fairness of the wage level, so workers feel free to ignore the norm when feeling badly treated (see, for example, the theory of "motivated reasoning" discussed by Benabou and Tirole 2016).

The empirical evidence for social comparisons as a determinant of effort is strong, but not overwhelming.[2] Several studies document that workers' effort does not always respond to co-workers' wages in a simple fashion (e.g. Charness and Kuhn 2007, Hennig-Schmidt et al.

[2] The persistence of reciprocity and gift exchange between employer and worker over time has also been debated (e.g., Gneezy and List 2006; Kube et al. 2012, Bellemare and Shearer 2009).

2010, Bartling and von Siemens 2011). This may be due to the fact that assessing the appropriate reference value is not always straightforward. For example, fixed wage differentials may be justified when workers are not equally productive. When workers are paid a fixed wage, then paying higher wages to more productive workers appears reasonable. In the context of heterogeneous workers, wages paid to high-productivity individuals should not map automatically on reference values of low-productivity workers.[3]

Nevertheless, the majority of the empirical work supports the hypothesis that social comparisons affect effort supply. This has been established in studies based on observational data (Karnes 2009, Clark et al. 2010), and in lab games (Clark et al. 2010, Gachter and Thöni 2010, Bracha et al. 2015).[4] Breza et al. (2015) organize a field experiment in an Indian manufacturing firm, offering different daily wages with or without a clear justification for wage differences (based on baseline productivity differences). While absolute wages of groups of workers are similar, reference values are manipulated by exogenous variation in wages of the relevant co-workers and by variation in the observability of productivity across tasks. The empirical evidence suggests negative comparisons – being paid less for the same task than your peer, without a clear rationale or justification – reduce effort supply. In contrast, positive comparisons do not increase effort on average.

This is consistent with experimental evidence by Cohn et al. (2014), who study effort supply by pairs of workers in response to wage cuts. Cutting both wages decreases performance (the fair wage-effort hypothesis), but unilaterally cutting only one wage decreases performance of the affected worker more than twice as much (and leaves effort of the other worker

[3] Gachter and Thöni (2010) also demonstrate that "intentions matter more than consequences," or that workers more readily accept wage differentials that appear random than the result of discrimination by the employer. For earlier work on intention-based reciprocity, refer to Dufwenberg and Kirchsteiner (2004).

[4] The design of Bracha et al. (2015) is a bit different from the other studies. They do not study how much effort is allocated to a task (per unit of time), but ask how many minutes students want to work on a specific task when receiving a high or low piece rate.

unaffected). Cohn et al. (2015) also report that the “removal of perceived unfairness,” rather than positive reciprocity, determines performance. Workers who perceive to be underpaid at the base wage increase their performance after their wage is increased, but those who feel fairly paid do not change their performance after a wage increase.

3. **Experimental Design and data**

We designed and implemented a real-effort field experiment in Kamuli district, Uganda, in November and December of 2015, and collected data in 9 villages. We arrived in the villages one day before the experiment and asked the chief to provide us with a census of the households. After randomly selecting households to participate in the experiment, we invited selected households to send one (adult) representative. More than 95% of the households complied, and declining households were replaced by another one randomly drawn from the same population. We clearly announced the experiments were for (university) research purposes, and not part of any electoral campaign starting up in preparation of the 2016 elections in Uganda. Participants were informed there would be an opportunity to earn money. To avoid within-village information spill-overs we ran several sessions simultaneously, overseen by multiple enumerators, and not consecutively. To avoid between-village spill-overs the selected villages were geographically spread out and we never announced in which villages we would collect data next.

In total, we recruited 600 respondents from a peri-urban environment to engage in the experiment: 400 subjects were randomly assigned to our two treatment arms, A and B (or 200 subjects per arm), and 200 enrolled in two control arms (100 per control arm). Within experimental arms we allocated respondents to groups of 10 villagers, participating in the experiment together. Subjects were aware that assignment to experimental arms was random. All participants received a show-up fee of UG shillings 5,000 (1 USD ≈ 3,300 UGS), or about half of the average daily wage in the townships where we collected our data (UGS 8,000). This context is a setting that is intermediate between the conventional "lab" (without interaction between subjects, other than via the game) and the "field" (where subjects are unaware they participate in an experiment). Many subjects in our study spent time together before (and presumably after) the experiment, which enhances the saliency of social comparisons.

The experiment consisted of 2 stages (stages 1 and 2), or 4 sub-stages (1a, 1b, 2a, and 2b). In stages 1a and 2a we asked the subjects to sort beans for 30 minutes, in exchange for payment.[5] They were handed 12 kilograms of mixed dried beans in one large container: 3 kilograms from each of four different types, known by their local names as: *nambale short*, *nambale long*, *NAADs*, and *Masavu*. During stages 1a and 2a we paid participants a fixed piece rate per gram sorted (based on type). Subjects from treatment arm A were paid a relatively low piece rate of UGS 5 per gram sorted, and subjects from treatment arm B received a high piece rate of UGS 9 per gram. Subjects from the two treatment arms worked in nearby but different rooms, and during the sorting there was no contact between them. Subjects from the control groups also received a high (100 respondents) or low piece rate (100 respondents). At the end of stage 1a sorted output was measured, and individual earnings were computed and announced to individual subjects.

Immediately after stage 1a we implemented stage 1b, in which we invited subjects to engage in a voluntary task for which they would earn no money. We emphasized that people were free to leave for the break, but that it would be appreciated if they wanted to help by sorting some more beans. Those willing to engage in additional, unpaid bean sorting were

[5] Hence, our subjects were not free to choose the time spent sorting beans. For example, Fehr and Goette (2007) study whether agents work harder when wages go up, allowing workers to choose the number of hours spent working as well as effort per hour. In their study, the number of hours increased, and effort per hour decreased. Our subjects have fewer degrees of freedom, so our predictions with respect to effort per unit of time are unambiguous. Fehr and Goette (2007) do not study social comparisons, but do document evidence in favor of reference-dependent utility.

invited to sort another type of beans (from another container) for an additional five-minute interval. Subjects picked different types of beans from the mix for the paid and unpaid stages of the experiment. Specifically, they picked the *nambale long* type during the paid stage, and the *masavu* type during the unpaid task. Since sorting the latter type is much easier (unlike *nambale long*, it is quite distinct from the other three types of beans in the mix), productivity in the unpaid task is higher, all else equal, so we cannot compare productivity levels across paid and unpaid stages of the experiment. After stage 1b, and out of sight of the subjects, we mixed all sorted beans for the next experimental session. The beans included in the experiment were local varieties, well-known to the great majority of the respondents.

The difference between stages 1 and 2 of the experiment is in the information that participants have about their relative earnings. During stages 1a and 1b participants knew only their own piece rate (or the piece rate of others in their group). Social comparisons entered in stages 2a and 2b, after players (presumably) learned about the piece rates earned by their co-workers from other groups. This was accomplished as follows. During sessions, two groups of subjects were always sorting simultaneously in two adjacent rooms. After stage 1b we gave participants a 20 minutes break and provided them with snacks and refreshments. During this break, participants from the two groups were brought to a common room where they mingled and discussed. We assume information about earnings spread naturally across the participants during this break (which was invariably the case, according to the exit interviews).

Participants from the two treatment arms were always matched with individuals from the other treatment arm – bringing together high and low piece rate workers. In contrast, workers from control groups were matched with other control group workers earning the same rate. Hence, social comparisons are only salient for subjects from treatment arms A and B. Specifically, after discussing their piece rates and earnings with subjects in arm B, subjects in arm A learn they have been treated relatively badly even if absolute earnings were considerable—see below. In contrast, subjects from arm B presumably feel privileged. We did not provide any explanation for the gap in piece rates, and explored whether social comparisons affect productivity in the follow-up tasks.

After the break, participants returned to their working environment, and engaged in a second round of 30 minutes of bean sorting (stage 2a) and afterwards were again invited to contribute to an unpaid task (2b). During stage 2a all subjects worked for the same piece rate as before, and sorted the same beans as before. After this second stage, we paid subjects their earnings. Subjects participated in only one session. Mean earnings in the low (high) piece rate treatment equalled about UGS 1,600 (3,000) during the first stage, and UGS 2,000 (4,100) during the second stage. Mean earnings for the full experiment, including the show-up fee, range from UGS 8,400 for the respondents receiving the low piece rate to UGS 12,100 for subjects receiving the high rate. The former amount is roughly a full day of wages for unskilled labor in our study region, so even subjects in our low piece rate regime were paid rather well given that the experiment took less than half a day.

The analysis rests on a comparison of productivity across groups with and without social comparisons, both for the paid and unpaid task. To identify the effect of own wage and co-worker's wages on effort we do not rely on before-after comparisons but include control groups in a difference-in-differences analysis. The reason is that (i) workers may learn by doing, so their productivity could improve across rounds, and (ii) workers might have difficulty grasping the implications of a specific piece rate for potential earnings (that is: they may find it difficult to predict how many grams they can sort in half an hour, or how much they can potentially earn). The dynamics of effort across rounds for the control groups capture experience and updated information about productivity, enabling proper identification of the additional impact due to social comparisons.

We also analyze the data in a regression framework. First, to analyze whether own wage affects effort, we use the first stage data (N=600) and estimate the following model:

$$y_i = \beta_0 + \beta_1 High_i + \beta_2 X_i + \epsilon_i \tag{1}$$

where y_i captures the quantity of beans sorted by respondent i (in grams), $High_i$ is a dummy variable taking the value of "1" if the subject received a high piece rate, X_i is a vector of controls, and ϵ_i is an error term. We always cluster standard errors at the experimental group level. If higher piece rates invite additional effort, then we find $\beta_1>0$.

We use data from the second stage to estimate the following model:

$$y_i = \beta_0 + \beta_1 High_i + \beta_2 SC + \beta_3 SC \times High_i + \beta_4 X_i + \epsilon_i \tag{2}$$

where SC_i is a dummy variable indicating whether the respondent was in one of the experimental arms (i.e. spent the break with subjects earning a different piece rate). This variable take the value of "1" if subjects were matched with subjects earning a different piece rate, and takes a value of "0" for subjects from the control arms. The estimate coefficient β_1 again picks up the effect of own wage on effort, β_2 picks up the effect of unequal payment, and β_3 picks up any additional effect of unequal payment for high piece rate earners. Average quantities sorted for the various experimental groups during the 2nd stage of the experiment are therefore as follows:

Low piece rate, control group: β_0

High piece rate, control group: $\beta_0+\beta_1$ (3)

Low piece rate, treatment arm A: $\beta_0+\beta_2$

High piece rate, treatment arm B: $\beta_0+\beta_1+\beta_2+\beta_3$

Table 1 summarizes the observables of our participants, across the three experimental arms. The great majority of our subjects are (married) Christian women of the Musoga tribe, with little education. The average subject has approximately 4 children and has access to 1.7 acres for farming. Observe that participants are statistically identical for most observables, but that Christians and members of the Musoga tribe are slightly underrepresented in the control group. Religion and ethnic affiliation are not correlated with productivity in bean sorting. Overall, random assignment to arms resulted in rather balanced experimental groups, but we will control for observables in our regression models to increase the precision of our estimates.

<< Insert Table 1 here >>

4. Empirical results

We first report the outcomes of a series of pairwise comparisons starting with the paid stages of the experiment. Consider stage 1a of the experiment, when all workers were uninformed about how their piece rate compared to that of others. Sorted quantities for the various experimental arms are provided in the first row of Table 2. Not surprisingly, subjects in treatment arms A and B sort the same quantity of beans as their counterparts in the control groups. That is, in the absence of social comparisons, subjects earning a low piece rate in treatment arm A sort the same quantity of beans as subjects in the control group receiving the same piece rate (columns 1-2; *p*-values of t-test of equality of sample means is reported in column 3). Similarly, subjects in the high piece rate treatment B sort as much as subjects in their control group earning a high piece rate (columns 4-5; *p*-values from t-test reported in column 6).

<< Insert Table 2 here >>

More surprisingly, perhaps, is the finding that there does not appear to be a significant difference between subjects earning a low or high piece rate. While subjects in the latter group sort 338.6 grams, or some 14 grams more on average than their counterparts from the low piece rate group, this difference is not statistically significant at conventional levels (p=0.19).[6] Effort does not respond strongly to "own wages." We obtain the same result when we analyze the data in a regression framework; columns 1 and 2 of Table 3. While pooling data from treatment and control arms increases statistical power (as does controlling for observables), clustering standard errors at the group level reduces statistical power, and the net effect is that we find no significant difference in productivity between treatment arms. This could simply reflect that the low piece rate was also relatively high, compared to potential earnings outside the experiment.

<< Insert Table 3 here >>

Next, we consider productivity in stage 2a, or after subjects in treatment arms A and B have been able to compare their earnings to those of others. Quantities sorted for payment are provided in the second main row of Table 2. The first thing to observe is that *all* subjects have increased their productivity across rounds. This applies to workers in both experimental arms as well as in the control groups. Workers have learned to sort more efficiently, or have obtained a better understanding of the monetary stakes during the experiment (after learning about their stage 1a productivity and earnings).

The simple comparison of group means suggests that productivity levels vary across groups after the break. First, observe that productivity of low piece rate workers does not suffer from negative comparisons. Specifically, output of subjects in arm A increases by 27%, which matches the increment in productivity of their control group (p=0.41). However, outcomes

[6] The difference in output for high and low piece rate workers in the control group is also not statistically significantly different from zero (Δ=4,5 gram, p=0.19).

appear different for subjects in treatment arm B, who on average increased their output level by no less than 36%. This increase in output exceeds the increase in output of high piece rate workers in their control group (p=0.05), as well as that of workers in treatment group A (p=0.00). In other words, while negative social comparisons do not significantly reduce productivity in our piece rate experiment, positive comparisons seem to invite an increase in productivity. However, these results are only marginally significant in the regression analysis, reported in columns 3 and 4 of Table 3. While coefficients β_2 and β_3 have the expected signs (i.e., $\beta_2<0$ or a negative productivity effect of negative social comparisons, and $\beta_3>0$ or positive productivity effect of positive comparisons), they are imprecisely estimated. Specifically, β_3 is only significant at p=0.11 in both models.

The finding that negative social comparisons leave effort of the two experimental groups unaffected contrasts with findings of, for example, Gachter and Thöni (2010), Cohn et al. (2014, 2015) and Breza et al. (2015). These studies document an asymmetric impact of social comparisons on effort, finding that low-wage workers provide less effort and that productivity of high-wage workers is unaffected. For example, Breza et al. (2015) find that, on average, output declines by 22% for a given wage when a worker earns less than his co-workers,[7] and document little support for the hypothesis that performance increases when people earn more than their peers. See also Cohn et al. (2015) on differential effort responses to wage increases for underpaid and adequately paid workers, and Card et al. (2012) on the asymmetric effect of wage inequality on job dissatisfaction.

We believe a major reason for the divergence in findings is due to the fact that our subjects are paid on a piece rate basis – a payment regime where shirking is privately costly,

[7] Workers with low relative pay also more frequently miss work days, and give up nearly 10% of their income to avoid a workplace where they earn less than their colleagues.

especially given the rather high piece rates that we offered (even in the low piece rate treatment). Extrinsic motives appear sufficiently strong to dominate other concerns.

Next, to further probe whether social comparisons are neutralized by extrinsic incentives we analyse how social comparisons affect voluntary work. Recall that in stages 1b and 2b of the experiment we asked subjects to sort additional bags of beans without payment. First consider productivity prior to social comparisons, in the pre-information stage. The first row of Table 4 documents there is no variation in effort across the experimental arms during stage 1b – statistically these four measures of output are identical. Higher piece rates do not improve productivity for an unpaid voluntary follow-up task. We find the same result in the regression analysis (columns 1 and 2 of Table 5). Consistent with productivity during the paid stage of the experiment, this finding could reflect that both low- and high-piece rate workers believe they have been well-paid during the first stage.

<< Insert Tables 4 and 5 about here >>

How does productivity during the voluntary task evolve after information about relative earnings had been shared? Observe there are two different dimensions to reduced reciprocity. First, subjects may refuse to participate in the task altogether. However such rejections were extremely rare and refusal rates hover around 1-2% across all experimental arms. We therefore refrain from an econometric analysis of this type of reduced reciprocity.

Second, subjects may participate in the task but supply little effort. This is analyzed in the second row of Table 4. For this analysis we have included the non-compliers and gave them an output level of zero grams, so the results are akin to an intention to treat analysis. Similar results are obtained when focusing on the subsample of compliers (details available on request). Productivity of low piece rate workers is now negatively affected by the social comparison. This result is significant (p=0.067). For voluntary tasks we find that, compared to the low piece rate members of their control group, workers in treatment arm A sort almost 11% less beans. This is akin to findings in the literature based on variation in fixed wages: subjects supply less effort hen shirking is cheap. This finding seems to emerge across cultural contexts. The same result emerges in the regression analysis summarized in columns 3 and 4 of Table 5. Negative social comparisons lower voluntary output by 12-13 grams.

By comparison, positive comparisons in the high piece rate treatment B extend beyond the "work for payment stages" of the experiment, because they sort some 20% more beans than high piece rate workers from their control group (p=0.00). This immediately implies that, for the unpaid task, productivity in treatment arm B is much higher than in arm A (p=0.00). This result both emerges in the series of t-tests (Table 4) and the regression analysis (Table 5). Considering everything, it appears as if extrinsic motives dominate effort supply decisions during the paid stages of the experiment, and social comparisons have only a small role to play. However, subjects are aware of whether or not they have been treated "fairly." Both positive and negative comparisons affect effort supply during other stages of the work relation – beyond the contracted task – when shirking does not diminish own income.

5. Discussion and conclusions

Efficiency wage theory proposes that employers are willing to pay wages above the market-clearing wage, because in return workers are willing to supply extra effort. When effort is unobservable and non-contractible, such gift-exchange may be sustained as an equilibrium outcome if workers are altruistic (kind), inequality-averse, or wish to respect a reciprocal norm – motives that presumably vary across cultural contexts. In this paper we analyze productivity in the context of a task where African workers are paid a piece rate, and where we vary piece rates across subjects to invite positive and negative social comparisons. An innovation of the paper is that we allow information about piece rate differentials to "spread naturally" among our sample of workers, rather than artificially priming its salience during the instruction stage.

While behavioral responses to variation in own piece rates are modest in our experiment, we conclude that social comparisons are an important determinant of productivity during voluntary follow-up tasks. Our results support the hypothesis of gift exchange between workers and employers, and point to social comparisons as an important mediating factor. In the absence of social comparisons, productivity differences between high and low piece rate workers are rather small. Introducing social comparisons does not fundamentally alter this finding, and outcomes appear to be dominated by extrinsic motives. We obtain weak support for the hypothesis that positive social comparisons invite extra effort during paid stages of the experiment, but this finding is only marginally significant.

Social comparisons enter much more prominently when extrinsic motives are eliminated, during the unpaid stages of the experiment. Our subjects supply less effort in a voluntary task after discovering their earlier piece rate was relatively low. In contrast, high piece rate workers become more productive following a positive comparison. This suggests the consequences of social comparisons extend beyond "contractible tasks" for which subjects receive payment, and that both positive and negative comparisons may matter for productivity.

Some of our findings deviate from earlier work. While perhaps part of this difference may be attributed to cultural differences across sample populations, we believe the nature of the payment regime helps to explain the divergence. Gift-exchange has been tested mainly in experiments based on hourly wages (and imperfect contracts), where working harder is readily interpreted as an act of altruism or reciprocal behavior. However, since differences in hourly wages may be justified by underlying productivity differences, fixed wages may send an ambiguous signal about whether payments are fair or not. Instead, we base our study on exogenous variation in piece rates, where under-supplying effort is privately costly. Our findings are consistent with a scenario where extrinsic motives for effort dominate intrinsic ones. Writing about their results based on variation in fixed wages, Cohn et al. (2014, p.897) write "… we cannot translate our findings to a piece rate environment. With such a schedule, workers would still have an incentive to produce output." This is exactly what we find: any attenuating impact of negative comparisons appears to be dominated by extrinsic motives. But we also find that social comparisons may re-enter elsewhere in the labor relationship.

The effect of social comparisons on productivity can be formalized in various ways, or there are multiple candidate mechanisms linking comparisons to the supply of effort. Candidate mechanisms discussed in the literature include inequality aversion, status concerns, kindness, a desire for reciprocity or fairness, or context-specific respect for behavioral norms. Further research is necessary to distinguish between alternative theories.

It is important to mention a few caveats to our experimental design. First, while subjects mingled during the break, they "worked" in separate rooms. So subjects may believe that the difficulty of the task varied with the piece rate, which would attenuate incentives for social comparisons. However, the types of beans sorted during the experiment were well-known to the respondents (they are an important part of their daily diet), so during the break they could establish that the tasks were actually identical by discussing the details of their sorting experience. Second, the show-up fee was generous and so were the piece rates we offered even in the low piece rate treatment. While this does not eliminate the scope for social comparisons, we believe even low piece rate workers may have been rather satisfied with their earnings, possibly diminishing adverse impacts of negative comparisons. This could be the case, for example, if respondents not only compare their earnings to those of subjects in the other treatment arm, but also to those outside the experiment. Third, the experiments took less than half a day, which is of course a limited time span. It is an open question to what extent experimental findings such as ours extend beyond their specific setting.[8]

Notwithstanding these issues, we hope our experimental findings speak to the design of earning structures within organizations. Earlier work suggests pay inequality may reduce overall satisfaction and willingness to cooperate (see Pfeffer and Langton 1993, but also Bartling and von Siemens 2011 for evidence to the contrary). Our results imply that "low earners" will undersupply effort for voluntary tasks. Insofar as accurate and timely execution of such voluntary tasks by *all* workers is important to economic success, lowering the morale of part of the workforce implies an organizational risk. If favorable social comparisons invite reciprocal behavior, it seems better to create a reference group *outside* the organization. If workers collectively compare themselves to workers from other organizations, and feel privileged or well-treated, this should arouse reciprocal behavior. Interestingly, this implies a cross-organizational externality: one organization's earning structure will affect the morale and productivity of workers in other organizations. Understanding the strategic considerations implied by such outcomes is left for future research.

References

Akerlof, G. (1982). Gift exchange and efficiency wage theory: 4 views. *American Economic Review* 74, 470-492

Akerlof, G. and J. Yellen (1990). The fair wage efort hypothesis and unemployment. *Quarterly Journal of Economics* 105, 79-83

Bartling, B., Von Siemens, F. (2011). Wage inequality and team production: An experimental analysis. *Journal of Economic Psychology* 32, 1-16

[8] Another potential concern is that good news (positive comparisons) may spread more easily among the treatment group than bad news to avoid bad feelings among peers. In this case, treatment A is less intensively treated than treatment B. This potential concern follows directly from our choice to let knowledge about relative earnings diffuse "naturally" rather than via announcements of the experimenter.

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

Bellemare, C., S. Kroeger and A. van Soest (2008). Measuring inequity aversion in a heterogeneous population using experimental decisions and subjective probabilities. *Econometrica* 76, 815-839

Benabou, R. and Tirole, J, (2016). Mindful economics: The production, consumption and value of beliefs. *Journal of Economic Perspectives* 30(3), 141-164

Bracha, A., Gneezy, U. and G. Loewenstein (2015). Relative pay and labor supply. *Journal of Labor Economics* 33(2), 297-315

Breza, E., Kaur, S. and Shamdasani, Y. (2015). The morale effects of pay inequality. Columbia Business School. Discussion Paper

Card, D., Mas, A., Moretti, E., and Saez, E. (2012). Inequality at work: The effect of peer salaries on job satisfaction, *American Economic Review* 102(6), 2981-3003

Charness, G. and Kuhn, P. (2007). Does pay inequality affect worker effort? Experimental evidence. *Journal of Labor Economics* 25(4), 693-723

Clark, A., Masclet, D., and Villeval, N.C. (2010). Effort and comparison income: Experimental and survey evidence. *Industrial and Labor Relations Review* 63(3), 407-426

Clark, A. and C. Senik (2010). Who compares to whom? The anatomy of income comparisons in Europe. *Economic Journal* 120, 573-594

Cohn, A., Fehr, E., B. Herrmann, and Schneider, F. (2014). Social comparison and effort provision: Evidence from a field experiment. *Journal of the European Economic Association* 12(4), 877-898

Cohn, A., E. Fehr and L. Goette (2015). Fair wages and effort provision: Combining evidence from a choice experiment and a field experiment. *Management Science* 61(8), 1777-1794

Duessenberry, J. (1952). *Income, saving and the theory of consumer behavior*. Cambridge MA: Harvard University Press

Dufwenberg, M. and Kirchsteiner, G. (2004). A theory of sequential reciprocity. *Games and Economic Behavior* 47, 268-298

Fehr, E., Kirchsteiner, G. and Rield, A. (1993). Does fairness prevent market clearing? An experimental investigation. *The Quarterly Journal of Economics* 108: 437-459

Fehr, E. and L. Goette (2007). Do workers work more if wages are high? Experimental evidence from a randomized field experiment. *American Economic Review* 97(1), 298-317

Fehr, E., and Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly journal of Economics*, 817-868.

Fehr, E., L. Goette and C. Zehnder (2009). A behavioral account of the labor market: The role of fairness concerns. *Annual Review of Economics* 1: 355-84

Frank, R. (1985). *Choosing the right pond: Human behavior and the quest for status.* New York: Oxford University Press

Gachter, S. and Thoni, C. (2010). Social comparison and performance: Experimental evidence on the fair wage-effort hypothesis. *Journal of Economic Behavior and Organization* 76, 531-543

Gneezy, U., and List, J.A. (2006). Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments. *Econometrica*, *74*(5), 1365-1384

Hennig-Schmidt, H., B. Rockenbach and A. Sadrieh (2010). In search of workers' real effort reciprocity -- A field and a laboratory experiment. *Journal of the European Economic Association* 8(4), 817-837

Kube, S., Maréchal, M. A., and Puppe, C. (2012). The currency of reciprocity: Gift exchange in the workplace. *The American Economic Review*, *102*(4), 1644-1662.

Karnes, R. E. (2009). A change in business ethics: The impact on employer–employee relations. *Journal of Business Ethics, 87*(2), 189-197.

Leibenstein, H. (1950). Bandwagon, snob, and Veblen effects in the theory of consumer demand. *Quarterly Journal of Economics* 64(2): 183-207

Pfeffer, J. and Langton, N. (1993). The effect of wage dispersion on satisfaction, productivity and working collaboratively: Evidence from college and university faculties. *Administrative Science Quarterly* 38(3), 382-407

Appendix 1: Detailed Variables Definition

Male: dummy variable taking value "1" if participant is male.

Married/Engaged: dummy variable taking value "1" is participant is married or engaged.

Number of children: integer indicating number of children in the household.

Education: highest grade completed by the respondent.

Religion (Christian): dummy variable taking value "1" if participant religious affiliation is Christianity.

Tribe (Musoga): dummy variable taking value "1" if participant belongs to Basoga tribe.

Land size: land owned by the participant in acres.

High: dummy that takes a value of "1" for participants with high piece rate.

SC: dummy that takes a value of "1" if participants were matched with other participants earning a different piece rate.

Table 1: Descriptive Statistics and Balance Test

	Treatment arm (A) N = 200	Treatment arm (B) N=200	Control (C) N =200	*p*-values A = B	*p*-values A = C	*p*-values B = C
Male (dummy)	0.270	0.295	0.285	0.580	0.640	0.755
Married/Engaged (dummy)	0.630	0.645	0.640	0.756	0.769	0.883
Number of children	4.075	3.840	3.605	0.495	0.041	0.304
Education (years)	2.300	2.350	2.390	0.649	0.246	0.605
Religion (Christians, dummy)	0.805	0.820	0.720	0.702	0.008	0.002
Tribe (Musoga, dummy)	0.930	0.900	0.880	0.283	0.031	0.386
Land size (acres)	1.760	1.720	1.580	0.898	0.111	0.214

Table 2: Piece Rates, Social Comparisons and Contractual Effort I

	Low piece rate			**High piece rates**		
	Treatment A	Control	*p*-value (t-test)	Treatment B	Control	*p*-value (t-test)
Output (stage 1A)	324,5 (6,067) N=200	320, 0 (10,676) N=100	0.71	338,6 (8,371) N=200	340,1 (11,087) N=100	0.92
	Break with information update	*Break without information update*		*Break with information update*	*Break without information update*	
Output (stage 2A)	412,9 (8.116) N=200	425,2 (13.245) N=100	0.41	461,5 (8.616) N=200	433,4 (11.245) N=100	0.05
Stage 2 – Stage 1 (t-test)	88.3 *p*=0.00	105.2 *p*=0.00		122.9 *p*=0.00	93.3 *p*=0.00	

Notes: Grams of beans sorted for different experimental groups, standard errors of the means reported in parentheses. *p*-values refer to outcomes of a simple t-test.

Table 3: Piece Rates, Social Comparisons and Contractual Effort II

	Beans sorted for payment: first stage		Beans sorted for payment: second stage	
High	16.100	15.569	8.170	10.051
	(12.81)	(12.91)	(18.49)	(18.25)
SC			-12.310	-6.959
			(16.33)	(16.33)
*SC * HIGH*			40.430	36.719
			(25.40)+	(25.37)+
Controls	No	Yes	No	Yes
Constant	323.026	299.170	425.200	423.496
	(7.08)***	(16.86)***	(1.248)***	(16.509)***
R-squared	0.006	0.041	0.028	0.117
Observation	600	600	600	600

Notes: Robust standard errors, clustered at the group level, reported in parentheses. Included controls are the variables included in Table 1. ***Coefficient significant at 1%, **Coefficient significant at 5%, *Coefficient significant at 10%, + coefficient significant at 12%.

Table 4: Piece Rates, Social Comparisons and Voluntary Effort I

	Low piece rate			**High piece rate**		
	Treatment A	Control	*p*-value (t-test)	Treatment B	Control	*p*-value (t-test)
Output (stage 1B)	114.0 (3.257) N=200	116.1 (3.857) N=100	0.692	118.6 (3.442) N=200	119.4 (4.385) N=100	0.885
	Break with information update	*Break without information update*		*Break with information update*	*Break without information update*	
Output (stage 2B)	123.4 (4.282) N=200	136. 0 (4.501) N=100	0.068	153.3 (3.193) N=200	127.0 (4.366) N=100	0.000
Stage 2 – Stage 1 (t-test)	9.4 *p*=0.03	19.9 *p*=0.00		34.7 *p*=0.00	7.6 *p*=0.13	

Notes: Grams of beans sorted for different experimental groups, standard errors of the means reported in parentheses. *p*-values refer to outcomes of a simple t-test.

Table 5: Piece Rates, Social Comparisons and Voluntary Effort II

	Beans sorted voluntarily: first stage		Beans sorted voluntarily: second stage	
High	4.190	4.271	-8.970	-10.195
	(6.39)	(6.24)	(7.21)	(6.91)
SC			-12.560	-13.480
			(7.39)*	(7.47)*
*SC * High*			38.830	40.349
			(10.56)***	(10.47)***
Controls	No	Yes	No	Yes
Constant	114.693	96.028	135.970	120.768
	(4.22)***	(7.99)***	(5.06)***	(8.82)***
R-squared	0.002	0.023	0.061	0.087
Observation	600	600	600	600

Notes: Robust standard errors, clustered at the group level, reported in parentheses. Included controls are the variables included in Table 1. ***Coefficient significant at 1%, **Coefficient significant at 5% and *Coefficient significant at 10%.