Output Restriction and the Ratchet Effect: Evidence from a Real-Effort Work Task*

Eric Cardella†
Texas Tech University

Briggs Depew‡
Utah State University
IZA

April 21, 2017

Abstract

The "ratchet effect" refers to a phenomenon where workers who are compensated based on productivity strategically restrict their output because they rationally anticipate that high levels of productivity will be met with increased or "ratcheted-up" expectations in the future. While there is ample anecdotal evidence suggesting the presence of the ratchet effect in real workplaces, it is difficult to empirically identify strategic output restriction among workers. In this study, we implement a novel experimental design using a real-effort work task and a piece-rate incentive scheme to investigate the presence of the ratchet effect using two different methods for evaluating worker productivity: (i) when productivity is evaluated based on the output of each individual worker, and (ii) when productivity is evaluated collectively based on the output of a group of workers. We find strong evidence that workers restrict their output when productivity is evaluated at the individual-level. However, we find very little evidence of output restriction when productivity is evaluated collectively at the group-level. We attribute the latter result to the free-riding incentive that emerges when productivity is evaluated at the group-level. Although, we find that output restriction among workers re-emerges if workers are able to communicate. Furthermore, by using a real-effort work task, we are able to examine an important dynamic productivity implication of the ratchet effect that has not yet been explored in the literature. Particularly, our results indicate that output restriction among workers reduces future productivity through reduced learning-by-doing.

Keywords: ratchet effect, output restriction, piece-rate pay, real-effort task, learning-by-doing

JEL Codes: J30, J40, D70, D01, C92

---

* We thank Luke Boosey, Danny Brent, Jeffrey Butler, Gary Charness, David Cooper, Martin Dufwenberg, Sebastian Goerg, Mark Isaac, Taylor Jaworski, Cark Kitchens, Charles Noussair, Alex Roomets, Alec Smith, Todd Sorensen, and Brock Stoddard for helpful comments and suggestions. We also thank seminal participants at Florida State University and the University of Arizona, and conference participants at the 2015 Texas Experimental Association Symposium, the 2015 North-American Economic Science Association meetings, and the 2015 Southern Economic Association meetings. We are grateful to Don Johnson and the Texas Tech Alumni Association for their support and research resources provided. A previous version of the paper circulated under the title "Testing for the Ratchet Effect: Evidence form a Real-Effort Work Task".

† Corresponding author; Rawls College of Business, Texas Tech University, Lubbock, TX 79409; Telephone: (806) 834-7482; Email: eric.cardella@ttu.edu

‡ Department of Economics and Finance, Utah State University, Logan, UT 84322-3500; Email: bdepew@lsu.edu

"In theory, piecework was simple. The company set a fair price for each unit of completed work and workers were paid according to their output…In practice, piecework never worked this way since employers always cut the price they paid workers." – *(Clawson, 1980, p. 169)*

1 Introduction

It is well documented that performance-pay jobs play a large role across many different industries in the economy (Lemieux et al., 2009).[1] In a static setting, a primary motivation for implementing performance-pay is to mitigate the agency problem and incentivize effort provision by workers (Stiglitz, 1975; Lazear, 1986; Gibbons, 1987; Lazear, 2000; and Prendergast, 1999 for a review).[2] However, in a dynamic setting, a potential drawback of performance-pay is that workers may have an incentive to "shirk" by strategically restricting the amount of output they produce. The reason is that if management is unable to commit to a multi-period compensation schedule, workers may rationally anticipate that management will respond to high output levels with increased quotas or lower performance-pay (e.g., piece-rates, commissions, bonuses) in the future. Thus, workers would be resigned to exerting higher levels of effort in the future for a similar level of overall compensation. This phenomenon where workers strategically restrict their output relative to their true capability is known as the "ratchet effect" (e.g., Laffont & Tirole, 1988).

The primary motivation of this paper is to evaluate the presence of the ratchet effect in a simulated work environment that involves *real-effort* and the potential for group dynamics. Specifically, we develop a novel experimental design where participant workers complete a real-effort work task over two work periods under a piece-rate incentive scheme. Our design enables us to test if workers strategically restrict their output in the 1st work period when there is a rational expectation that their 2nd period piece-rate will be reduced if they are *too* productive in the 1st period. After establishing the presence of the ratchet effect – output restriction among workers – when productivity is measured based on individual output, we test for the existence of the ratchet

[1] For example, Lemieux et al. (2009) show that the overall proportion of performance-pay jobs in the U.S. has increased from about 3 percent in the late 1970s to approximately 45 percent in the 1990s. The significant presence of performance-pay across various industries has also been documented by Skelton & Yandle (1982), who note that "piece-rate plans are included in at least 75 percent of the contracts in rubber, textiles, fabricated metal and the stone and glass industries…Furthermore, farm workers, watermen, and commissioned salesmen are often paid on a piece-rate" (pp. 201-202). More recently, Kuhn & Lozano (2008) document evidence, by way of Lawler et al. (2001) that the incidence of incentive pay across fortune 1000 firms has increased over the latter part of the 20th century.

[2] Empirically, several papers have documented increases in productivity under piece-rates, compared to fixed-pay schemes, including: Seiler (1984), Banker et al. (1996), Fernie & Metcalf (1999), Lazear (2000), Paarsch & Shearer (2000), Shearer (2004), Bellemare et al. (2010), and Carpenter & Gong (2016).

effect when productivity is measured collectively, based on output of the group. Under this group-level setting, we further analyze the role of group communication in workers' decisions to restrict output. Lastly, by using a real-effort work task rather than a chosen-effort task, we are able to exploit the natural variation in worker ability to explore important learning dynamics that are possibly linked to the ratchet effect; namely, we investigate the extent to which output restriction can reduce future productivity through reduced learning-by-doing.

Since the development of formal principal-agent models in the 1980s, theoretical models of the ratchet effect have been extensively studied under various contexts (e.g., Freixas et al., 1985; Lazear, 1986; Baron & Besanko, 1987; Gibbons, 1987; Ickes & Samuelson, 1987; Laffont & Tirole, 1988; Dearden et al., 1990; Kanemoto & MacLeod, 1992; Olsen & Torsvik, 1993; Dalen, 1995; Meyer & Vickers, 1997; Carmichael & MacLeod, 2000; Choi & Thum, 2003; Puller, 2006; Bhaskar, 2014).[3] Within labor markets, the general theoretical structure involves privately informed workers (the agents) choosing effort levels over multiple work periods. If management (the principal) is unable to commit, ex-ante, to a multi-period compensation/output schedule, then what typically results is a pooling equilibrium where the high ability (or low cost of effort) workers mimic the low ability (or high cost of effort) workers by choosing low effort, thus concealing their true high ability.[4] The incentive for high ability workers to conceal their true ability arises because they know that high levels of output will signal high ability, which would then induce management to set a less favorable compensation scheme (or more demanding output schedule) in the future. The existence of pooling equilibria in these dynamic principal-agent models provides the theoretical foundation for the emergence of the ratchet effect.

The theoretical implications of the ratchet effect are consistent with substantial qualitative anecdotal evidence suggesting the presence of the ratchet effect in real workplaces. In particular,

[3] Besides effort provision in the workplace, other contexts where ratchet effect dynamics have been explored include: (i) input allocations and output targets in centrally planned economics or multi-divisional firms, where high productivity firms or divisions within a firm may produce less efficiently to avoid lower input allocations or high output targets in the future; (ii) compliance with environmental regulation, where firms may be less inclined to innovate more environmentally friendly technology for fear of more stringent regulation in the future; (iii) regulation of natural monopolies, where the monopolist may be less inclined to invest in cost-reducing technologies for fear of more stringently regulated prices in the future, and (iv) sales targets, where salespersons may reduce effort and sales in the current period if they anticipate sales targets in the future will be based on current sales.

[4] Even if management attempts to commit, ex-ante, to not raising worker expectations in the future, there are ways for management to, ex-post, renege on such commitments, as discussed by Gibbons (1987), Ickes & Samuelson (1987), Dearden et al. (1990), and Carmichael & MacLeod (2000). For example, management can assign new workers to the job (at a lower piece-rate), re-assign old workers to new jobs (with a new piece-rate scheme), or re-classify jobs with a new, less-favorable piece-rate scheme (Clawson, 1980).

the works by Mathewson (1931), Lawler (1971), Edwards (1979), Montgomery (1979), and Clawson (1980), among others, provide numerous industry accounts, case studies, worker narratives, and discussions of apparent output restriction by workers under piece-rate incentive schemes (see Levine, 1992 for a thorough review of this literature). For example, Edwards (p. 99) writes that "the second, more serious difficulty [with piece-rate incentive schemes] was that piece-rates always contained an incentive for workers to deceive employers and restrict output." Clawson (p. 175) notes that "although workers generally knew that they could have produced substantially more, they understood it was not in their interest to do so." These anecdotal accounts often suggest that the reason for output restriction among piece-rate workers was anticipation of future piece-rate reductions if they were too productive and earned too much, which is consistent with the theoretical rationale underpinning the ratchet effect.

Despite the abundant theoretical work modeling the ratchet effect and the anecdotal evidence suggestive of the presence of the ratchet effect in workplaces, there is surprisingly little in the way of empirical research aimed at formally testing for the ratchet effect. In line with the arguments put forth by Charness et al. (2011), this is likely a result of the significant challenges associated with identifying the ratchet effect in real workplaces. The most obvious of these is the difficulty observing the *true* ability of workers, which, consequently, renders it difficult to identify if, and to what extent, workers are strategically restricting their output. In addition, the emergence of the ratchet effect typically hinges on specific contractual and informational features of the interaction between workers and management (e.g., private information about ability, management's inability to perfectly identify output restriction, and management's inability to commit to a long term compensation scheme), which are also difficult to verify in practice. However, by implementing a simulated work experiment featuring a real-effort task, we are able to identify the distribution of true ability among the sample of workers, as well as control for requisite informational and contractual features for the ratchet effect to *possibly* emerge.

In our experimental design, participant workers complete a real-effort work task under piece-rate compensation for two work periods. Importantly, the design enables us to first recover an estimate of the true distribution of output capability for our sample of participant workers in both the 1st and 2nd work periods. We do this by considering a "baseline" condition where participant workers work for both periods under a fixed piece-rate scheme (i.e., where there is no scope for strategic output restriction in the 1st work period). Given this estimate of true output capability, we are able to directly identify the ratchet effect (in the aggregate) by evaluating if workers restrict their output in the 1st work period when faced with the consequence of a reduction in their piece-rate in the 2nd period if they are, individually, too productive in the 1st period.

Motivated by anecdotal evidence that firms often evaluate productivity at the group-level rather than the individual-level (Lawler, 1971; Edwards, 1979; Clawson, 1980), we extend our empirical analysis of the ratchet effect in two important ways. First, we test if participant workers strategically restrict their output in the 1st period when they face the consequence of their piece-rate being reduced in the 2nd period if the group of workers (to which they are exogenously assigned) is *collectively* too productive in the 1st period. By tying the piece-rate reduction to group productivity, the strategic structure within the group of workers is transformed in a way that shares similar properties to threshold public goods games.[5] As such, under this group condition, a "free-rider" problem arises where each individual worker has an incentive to work at full capacity and free-ride off the output restriction of others in the group, which can then potentially eliminate the ratchet effect. Second, we extend the group-level setting by incorporating a "pre-play" communication stage where the workers can discuss the work task prior to commencing work. Allowing groups to communicate can foster cooperation, thus potentially mitigating the above-mentioned free-rider problem and facilitating the emergence of the ratchet-effect. Communication among workers regarding output restriction seems plausible, in practice, and is alluded to anecdotally, as suggested by Clawson (1980, p. 177): "in order to enforce output quotas it was definitely necessary for some workers to pressure and coerce others."

Studying the ratchet effect in a group context can have important implications, especially in real-world settings where many workers perform a similar task but management is imperfectly informed about the difficulty of the task, the value added of an individual employee, and/or the true productivity of the firm's technology. Such circumstances may cause management to opt for compensation that is based on group-level measures of output. In this setting, if workers do collectively restrict output to avoided future rate cuts or quota increases, this reduces the incentive among the workers to innovate and reveal productivity enhancing information to management,

[5] See Palfrey & Rosenthal (1984) and Bagnoli & Lipman (1989) for a more detailed theoretical presentation of threshold public goods games, and van de Kragt et al. (1983), Dawes et al. (1986), Rapoport & Eshed-Levy (1989), Isaac et al. (1989), and Palfrey & Rosenthal (1991) for some of the early experimental investigations into such games. Similar to the general allocation rule by which the threshold public good is provided, in the group context we explore, the piece-rate is not reduced in the 2nd period if *enough* of the workers in the group restrict output in the 1st period.

(Dearden et al., 1990; Carmichael & MacLeod, 2000), which is an essential component of productivity growth for the firm (Carmichael & Macleod, 1993).

We find strong evidence of aggregate output restriction among workers in the 1st work period when productivity is evaluated at the individual-level, consistent with the ratchet effect. However, when we investigate group dynamics, we find very little evidence of aggregate output restriction in the 1st period when productivity is evaluated collectively at the group-level without communication; in this setting, workers appear to be trying to free-ride off the output restriction of others in the group, which essentially results in full output production across workers. In contrast, when we allow pre-play communication among the group of workers, we document significant output restriction consistent with the reemergence of the ratchet effect, which suggests that communication can facilitate coordination of output restriction among workers and play and important role in the existence of the ratchet effect in the workplace.

We are aware of four experimental papers that have directly investigated the ratchet effect – Chaudhuri (1998), Cooper et al. (1999), Charness et al. (2011), and Bellemare & Shearer (2015) – all of which have considered individual level productivity/output.[6] The first three studies consider relatively stylized experimental designs with *chosen* effort/output. In particular, Chaudhuri (1998) considers a setting where the principle chooses an output quota, while the agent chooses an output level. Chaudhuri finds that most agents played naively by signaling in the 1st period whether they were a high or low productivity type (via their output choice), thus finding little empirical support for the presence of the ratchet effect.[7] The experimental design of Cooper et al. (1999) builds from some early literature of the ratchet effect within centrally planned

[6] We are also aware of two empirical papers aimed at indirectly identifying the ratchet effect. Specifically, Allen & Lueck (1999) use contract data between landowners and tenant agriculture workers to analyze several predictions based on implication of the ratchet effect. In their analysis, they find little evidence that supports their prediction and, hence, conclude that their data reveal little evidence of the presence of the ratchet effect. The null finding of Allen & Lueck may be because the ratchet effect is not important in modern agriculture or the regression analysis suffers from a lack of identifying variation as the model makes the assumption that the landowner's contractual match between a new tenant farmer and existing tenant farmer is exogenous. A recent paper by Macartney (2016) investigates possible ratchet effects in effort provision of teachers when they are faced with the possibility of receiving bonuses if student performance exceeds specific targets. Using student performance data, the author documents evidence of decreases in student performance, which is consistent with predicted reductions in teacher effort arising from the ratchet effect.

[7] Charness et al. (2011) speculate, with regard to the lack of evidence of the ratchet effect found by Chaudhuri (1998), that "possible explanations for this result include the relative complexity of the game and the lack of context provided to the subjects that might have impeded the learning process" (p. 516). Furthermore, the ratchet effect hinges on the principle increasing expectations in the future when they know the agent is a high-ability type. However, Chaudhuri finds little evidence that principles actually set more stringent quotas in the 2nd period when interacting with a high-ability agent. As a result, given that high output levels in the 1st period are seldom met with increased quotas in the 2nd period, it is not surprising that agents did not restrict output in the 1st period.

economies (Berliner, 1976; Weitzman, 1980). Cooper et al. use both students and actual Chinese firm managers as subjects (firms and planners), and, among other results, they find significant evidence of the ratchet effect; namely, high productivity *firms* tend to choose lower output levels in the 1st period, compared to what would be statically optimal, to avoid a more demanding production target in the 2nd period. Charness et al. (2011) consider a labor market setting where a worker must choose high/low output and the *firm* chooses a high/low rental rate to charge the worker. Like Cooper et al. (1999), Charness et al. do find evidence of the ratchet effect – a substantial number of high ability workers choosing low output in the 1st period to avoid facing the higher rental fee in the 2nd period. Based on the theoretical insights of Kanemoto & MacLeod (1992), Charness et al. also introduce market competition (on the side of both workers and firms) and find that the ratchet effect is virtually eliminated under the presence of competition. The study by Bellemare & Shearer (2015) uses a natural field experiment to test for the ratchet effect at a tree planting firm. The authors find empirical evidence that workers restrict their output in a trial period (compared to a control period) when faced with the possibility of receiving a *higher* piece-rate in the future if their productivity is low enough in a trial period (compared to the control period).

We contribute to the literature on the ratchet effect along a number of important dimensions. First, we document evidence of output restriction among workers under the credible threat of reduced future piece-rates using a real-effort work task. In doing so, we establish robustness of the emergence of the ratchet effect with real-effort, and add further credibility to the very limited extant literature documenting direct empirical evidence of the ratchet effect (Cooper et al., 1999; Charness et al., 2011). By using a real-effort work task embedded within a controlled lab experiment, combined with the incorporation of a substantial degree of “field” context in the experimental protocol, our study helps bridge the gap between the results documented in prior chosen-effort lab experiments and the anecdotal accounts of the ratchet from real workplaces that tend to be more qualitative and not well-identified. Second, we provide an empirical examination into the emergence of the ratchet effect when productivity is evaluated collectively at the group-level (with and without group communication), which we believe is important for two reasons: (i) it can change the economic incentives of workers in ways that can potentially mitigate the ratchet effect, and (ii) it seems plausible that, in practice, group-level evaluation and compensation is more representative of how management evaluates the productivity of its workers, especially at firms where many workers are preforming a similar task. Third, by using a real-effort task and the associated natural variation in worker ability, we are able to analyze possible dynamic productivity implications of the ratchet effect and, importantly, show that strategic output restriction by workers reduces their future productivity. A probable mechanism for this finding is that output restriction carries the negative externality of reduced learning-by-doing. More broadly, we view our study as contributing to the growing body of literature aimed at deepening our understanding of how workplace incentives impact employee productivity via the use of controlled experiments.[8]

2 Experimental Design

We conducted an experiment involving a real-effort work task designed to test for the ratchet effect by identifying strategic output restriction by workers. All sessions were conducted at the Rawls College of Business at Texas Tech University. Participants were recruited from a college maintained, subject-pool database. At the time of invitation, participants were told that participation in the study would involve working on a simple task monetary compensation. In total, 35 experimental sessions were conducted and 229 participant workers partook in the study (this includes participants and sessions from an additional follow-up condition that is described in Section 5); 53% were female, and the average age was 21.6 and the age range was 18 years to 44 years. We used a between-subjects design where each participant took part in only one session of a given experimental condition. The average session lasted 45 minutes, and the average earnings were $14 USD. A full copy of the experiment instructions can be found in the Appendix.

2.1 Real-Effort Work Task and the Work Environment

In collaboration with the Texas Tech Alumni Association (TTAA), we organized a real-effort work task that consisted of stuffing and sealing TTAA donor solicitation mailers.[9] Assembling a mailer required the participant worker to: (i) stuff a mailer into the mailing envelope (with the address facing through the clear plastic window on the front of the envelope); (ii) stuff in a return envelope behind the mailer; and, (iii) seal the envelope. In total, approximately 17,500 TTAA mailers were assembled over the course of this study. For the remainder of the paper, the output level of a participant worker will be in reference to the number of completed, assembled mailers.

[8] In lieu of attempting to cite all papers in the area of experimental labor research, we instead refer readers to the survey article on "lab labor" by Charness & Kuhn (2011) and the references therein for a review of the literature.

[9] We refer readers to Gill & Prowse (2015) for a discussion of some advantages of using a real-effort task compared to chosen effort. While the discussion of Gill and Prowse is focused on the "slider task" they develop, the authors note that the beneficial attributes of the slider task are also shared by the envelope stuffing task (p. 4). Other prior studies that have used an envelope stuffing task as the real-effort component of the experimental design include: Konow (2000), Falk & Ichino (2006), Carpenter & Gong (2016), and DellaVigna et al. (2016).

The mailer task is particularly well-suited for the purposes of studying the ratchet effect for several reasons. First, the task is simple, straightforward, and easy to understand, which essentially eliminates the possibility of inaccurate/incorrect completion of the task.[10] Second, the mailer task is not analytically intensive (e.g., anagrams, word unscrambling, puzzle-solving, or 2-digit multiplication), which implies there is not a strong cognitive component, and that output is an increasing function of effort. Third, there is little scope for substantial differences in the quality of assembled mailers, which is important because this mitigates possible tradeoffs that could arise between the quality of work and quantity of output when workers restrict output.[11] Taken together, this mailer task enables us to cleanly identify the ratchet effect; namely, if workers are strategically restricting their output by reducing effort. Furthermore, the use of a real-effort task, the partnership with the TTAA, and the legitimacy of the mailers provide a substantial degree of "field" context that is in line with a real piece-rate job, which positions our study into the domain of what Charness et al. (2013) refer to as an "extra-laboratory" experiment. As such, we feel that the real-effort task, in combination with the embedded field context, increases the external validity of our results (Friedman & Sunder, 1994; Falk & Fehr, 2003; Charness & Kuhn, 2011; Gill & Prowse, 2015).

Every participant worker assembled mailers for two 10-minute work periods. The piece-rate compensation scheme for each work period, which varied based on the experimental condition (described in detail below), was clearly stated to the participant workers in the instructions. After completing the 1st 10-minute period, participant workers had an approximate 10-minute break where they filled out a short questionnaire containing some general demographic questions (e.g., age, gender, work experience, etc.), some personality measures, and the 3-question cognitive reflection test (Frederick, 2005). During this time, the experimenter privately counted the number of completed mailers for each participant worker and indicated on a "compensation record" sheet how many mailers the worker had completed, their total compensation for the 1st period, and their

[10] In fact, during all sessions, an experimenter observed (ex-post) that each participant worker correctly assembled the mailers as instructed. There were no instances where a worker's mailers were not correctly assembled. Similarly, there was never any indication throughout the study by the TTAA that any of the mailers completed as part of the study were assembled in an unsatisfactory manner, thus indicating a vast majority of mailers were assembled correctly.

[11] In particular, piece-rate schemes have the potential to induce substitution between quality for quantity, as discussed theoretically by Stiglitz (1975) and Lazear (1986) and documented empirically by Paarsch & Shearer (1999) and Bellemare et al. (2010). If there is a substantial quality dimension to the work task, then workers who restrict their output may produce higher quality. We are not suggesting that such a tradeoff between quality and quantity is not interesting and potentially important. Rather, in this paper we focus specifically on restriction of output, and thus, we want to isolate production quantity and abstract away from the quality dimension of the task. Investigating if workers produce higher-quality work when they are restricting output is an interesting topic for future research.

piece-rate for the 2nd period. Participant workers then assembled mailers during the 2nd 10-minute period. After the 2nd period, the experimenter privately counted the completed assembled mailers, which concluded the session. Each worker was privately paid their total earnings, which was the sum of their piece-rate compensation from the 1st and 2nd work periods.

All experimental sessions were conducted in a conference room that was set up to resemble a simulated mailer assembly workplace. In the room, there were seven workstations equipped with all the necessary materials for assembling mailers. Each work station was separated by a privacy carrel, and in both work periods, participant workers assembled mailers within the confines of their privacy carrel. Thus, there was little scope for peer effects influencing productivity (Falk & Ichino, 2006), as participant workers were unable to observe the progress of the other workers or the output level of other workers.[12] An experimenter read the instructions aloud and provided a visual demonstration of how to properly assemble a mailer. Participant workers were informed in the instructions that the experimenter would not be continually monitoring their progress throughout the work periods so they were "free to work at [their] own pace and complete as many mailers as [they] can or choose to do in each work period."[13] All participant workers were informed that mailers were part of a TTAA campaign and that the mailers would actually be mailed.[14] A picture of the workplace environment and a sample mailer are provided in the Appendix.

[12] We acknowledge that having workers work in private may not be a representative characteristic of all workplaces, and the inability to observe other workers' output abstracts away from the possible channel of peer-effects influencing the decision to restrict output. That said, there are many piece-rate jobs (that might give rise to the ratchet effect) where worker productivity would not be perfectly observed by other workers including: sales, agricultural planting and harvesting (where there is sufficient geographic distance between workers), and manufacturing and textiles (where there is sufficient separation and visual obstructions between workers). Moreover, we would expect that peer-effects (e.g., peer pressure) would only amplify the emergence of the ratchet effect; therefore, in the conditions where we document evidence of output restriction, we would expect our findings to be strengthened if we also layered in an implicit peer-effect component arising from an openly observable workplace.

[13] An experimenter did make one cursory pass through the room right as the 1st work period began to ensure that each participant worker was not having any issues with their mailer assembly materials or additional questions. In each instance, there were no issues that involved the experimenter to temporarily delay the experiment or provide feedback to participants. After this initial pass through the room, the experimenter did not walk through the room during any of the remaining work time for either the 1st or 2nd work period.

[14] It is plausible that because the work task involved stuffing mailers for a charitable organization – the TTAA – participants may have been compelled to exert more effort, especially since students at Texas Tech University are likely in support the overall *mission* of the TTAA (see Besley & Ghatak, 2005 and Prendergast, 2007 for discussions and models of workers being motivated by the mission of the organization). Such an effect would be consistent with the findings documented by Carpenter & Gong (2016), where workers are more productive at a politically motivated mailer task when the mission of the mailer matches their political preferences, and the findings documented by DellaVigna et al. (2016) where workers are more productive at a mailer task for a charity when the mailers are actually mailed out, compared to when the mailers are thrown out. While it is possible that the charitable nature of the mailer task may have impacted effort levels compared to a more abstract work task, this impact would likely result in an overall level effect across all treatments. As a result, the validity of our identification of the ratchet effect across

2.2 Measuring True Output Capability

Since the ratchet effect entails workers restricting their output, relative to their capability, it is necessary to know the distribution of the true output capability across workers in order to empirically test for the presence of the ratchet effect. To identify this distribution of the true output capability in our sample of participant workers, we conducted an initial BASELINE condition using a fixed piece-rate scheme. The BASELINE condition is an integral aspect of our experimental design as it establishes a benchmark for the true output capability of the workers in our sample, which then enables us to test for strategic output restriction. In the BASELINE condition, all participant workers received a piece-rate of $.20 (20 cents) per assembled mailer in *both* the 1st and 2nd work period.[15] Importantly, participant workers were informed that the 2nd period piece-rate *did not* depend on 1st period output; hence, there was no strategic reason for workers to restrict their output in the 1st work period. As a result, we maintain that the observed distribution of output in the 1st work period in the BASELINE condition provides an estimate of the true output capability of our participant worker sample in the 1st period, conditional on a $.20 piece-rate. This approach of using a piece-rate scheme to measure an individual's true productivity has been similarly implemented by Abeler et al. (2011), Kube et al. (2013), and Gneezy et al. (2017). The BASELINE condition consisted of 42 participant workers (7 sessions).

Before introducing the main experimental conditions used to test for the ratchet effect, it is pedagogical to first present the aggregate output data for the 42 participant workers in the BASELINE condition. We present the BASELINE data first because the specification of the main ratchet effect conditions depends, in part, on the observed distribution of output in the BASELINE, which will be made evident in Section 2.3 below. Figure 1 displays the output distribution for both the 1st and 2nd 10-minute work periods. In terms of summary statistics of 1st period output, the average output was 34.6 mailers, the median was 34, the minimum was 17, and the maximum was 54. An important observation, as revealed in Figure 1, is that there is substantial variation in output levels of our participant workers. This variation is important because it establishes the presence of heterogeneity of different ability "types" of workers in our sample. Having a range of ability types

treatments remains intact. Moreover, the fact that the charitable nature of the task may provide added non-pecuniary motivations for participant workers implies that any observed reductions in effort would be a lower bound; this would make it less likely that we observe workers restricting output, and thus, harder to identify the ratchet effect in the data.

[15] Prior to the study, the authors and a few kind colleagues performed a crude productivity assessment regarding the number of assembled mailers that could be completed in a 10-minute period. Based on these output levels, the piece-rate of $.20 per assembled mailer was chosen to target an acceptable, ex-ante, average earnings level.

under a real-effort task will provide a more robust test of the ratchet effect, compared to the limited prior literature that has considered only two types of workers under chosen-effort.

In terms of summary statistics for output in the 2nd work period in the BASELINE condition, the average output was 44.7 mailers, the median was 44.5, the minimum was 23, and the maximum was 63. From the data and Figure 1, it is clear that our participant workers exhibited a significant increase in productivity from the 1st period to the 2nd period (Wilcoxon signed rank test: $p < .001$). We suspect that this is likely attributed to learning-by-doing, wherein the participant workers become more efficient at completing the task. DellaVigna et al. (2016) also find significant increases in productivity in a mailer assembly task over multiple work periods, which they similarly attribute to learning-by-doing. For example, increases in productivity in assembling mailers could result from participant workers implementing more efficient assembly methods including: (i) re-arranging the 3 components of the mailer within the carrel to facilitate quicker stuffing of the mailer, (ii) stuffing both the mailer letter and the return envelope together into the mailing envelope, as opposed to each piece separately, and (iii) implementing a quasi-assembly line approach of stuffing many mailers (without sealing them) and then sealing a stack of mailers. In Section 5 we explore the dynamic implications of the ratchet effect on productivity by analyzing 2nd period output and by establishing a learning-by-doing effect.

2.3 Experimental Conditions to Test for the Ratchet Effect

We implement three additional experimental conditions to test for the ratchet effect. Similar to the BASELINE condition, participant workers in these three ratchet effect conditions receive a piece-rate of $.20 in the 1st 10-minute work period. Unlike the BASELINE condition, these ratchet effect conditions differ in terms of the piece-rate in the 2nd 10-minute work period, as well as how the 2nd period piece-rate is determined. The general structure of each of these ratchet effect conditions is that workers will face the consequence of working for a reduced piece-rate of $.10 in the 2nd work period if productivity is *too high* in the 1st work period.[16]

As the criteria for evaluating whether productivity was *too high* in the 1st period (to warrant the piece-rate reduction in the 2nd period), we exogenously set a productivity threshold, denoted as **T,**

[16] For clarity and consistency with the underpinnings of the ratchet effect, when describing the compensation scheme we use the phrasing that piece rates will be reduced when 1st period productivity is *too high.* However to minimize possible experimenter demand effects, we do not use this descriptive phrasing in the experimental instructions. Rather we provide a very abstract piece-rate scheme that states their 2nd period piece-rate is a function of 1st period output. There is no mention in the instructions of 2nd period piece-rates being reduced if output is *too high.*

that is determined from the 1st period distribution of output in the BASELINE condition, which is made known to participant workers. In order for the investigation into the ratchet effect to be salient, the following two conditions are necessary when choosing a value of **T**: (i) **T** be set *low enough* that it is binding for most of our participant workers (i.e., the worker's true output capability is higher than **T**), which enables us to test for deliberate output restriction, and (ii) **T** be set *high enough* such that high ability workers have an incentive to restrict output; namely, their payoff is higher when they restrict output not to exceed **T** in the 1st period and receive the $.20 piece-rate in the 2nd period, compared to producing at full capability in the 1st period and receiving a piece-rate reduction to $.10 in the 2nd period. To satisfy these two conditions, we set the value of **T** = 29 mailers, which is effectively the 25th percentile of the 1st period output distribution from the BASELINE condition. By setting **T** equal to the 25th percentile of the distribution, **T** will be binding, in expectation, for approximately 75% of the participant workers, satisfying condition (i). At the same time, given the maximum observed output levels in the BASELINE condition of 54 and 63 mailers in the 1st and 2nd work periods, respectively, a worker of this capability would earn a higher payoff restricting output to 29 in the 1st period, compared to producing at full capability of 54 in the 1st period, thus satisfying condition (ii).[17]

While the implementation of this productivity threshold (around which the piece-rate is determined) is a stylized component of the experimental design, it is important for two reasons. First, it ensures common knowledge among the workers of the potential for the piece-rate to be reduced, and ensures that workers *rationally* anticipate a piece-rate reduction if they are too productive in the 1st period, which is essential for the emergence of the ratchet effect. This element of the design operationalizes what would be management's inability to commit to a piece-rate scheme over time, which implies that they would lower piece-rates in response to high output levels by workers. As such, our design essentially automates the manager's decision, and *forces* a piece-rate reduction in response to high observed output levels, which is how management is expected to dynamically respond in theory; this feature enables us to simplify the workplace setting by focusing specifically on how dynamic incentives impact the productivity of workers and their propensity to restrict output. Second, it eliminates any ambiguity for workers regarding how much

[17] To see this note that 29*($.20) + 63*($.20) = $18.40 > 54*($.20) + 63*($.10) = $17.10. We acknowledge that setting **T** = 29 is somewhat arbitrary; however, we see little reason to think that our results are specific to **T** = 29 and would not generalize to other values of **T** that satisfy the two stated necessary conditions.

output would be deemed as *too productive*, which enables us to more clearly identify output restriction in relation to the productivity threshold. Moreover, there is anecdotal evidence suggesting that workers may infer an explicit productivity threshold based on day rate equivalents (Clawson, 1980; Mathewson, 1931). For example, Clawson (p. 171) reports that "from cumulative experience they [workers] learned that if their earnings exceeded what they would have earned on a day rate by more than a certain percentage, they could expect their rate to be cut." Workers essentially become aware of the "maximum" management will pay per day, and then are able to deduce the maximum output that would generate that equivalent day rate. Thus, implementing an ex-ante productivity threshold is informationally equivalent to workers having a common understanding of an effective maximum day rate, which seems plausible in many circumstances.

The first of the three main experimental conditions we implement, which we denote as our INDIVIDUAL condition, is designed to test for the ratchet effect when productivity is evaluated at the individual-level. Namely, do participant workers strategically restrict their output if they rationally anticipate that their piece-rate will be reduced in the 2^{nd} period if, individually, they are *too* productive in the 1^{st} period. In the INDIVIDUAL condition, participant workers are informed, prior to starting work in the 1^{st} period, that if their 1^{st} period output level exceeds **T** = 29 mailers, then their 2^{nd} period piece-rate will be *reduced* in half to $.10. Hence, the difference between the BASELINE and INDIVIDUAL conditions is that in the INDIVIDUAL condition workers face the outcome of a piece-rate reduction in the 2^{nd} work period if they are too productive in the 1^{st} period.

The second experimental condition we implement, which we denote as our GROUP condition, tests for the ratchet effect when productivity is evaluated collectively for a group of workers. Namely, do participant workers strategically restrict their output if they rationally anticipate that their piece-rate will be reduced in the 2^{nd} period if the group of workers is, collectively, *too* productive in the 1^{st} period? All GROUP condition sessions consisted of a group of 7 workers. In the GROUP condition, the piece-rate in the 2^{nd} work period for participant workers depends on the 1^{st} period output levels of all 7 workers in the group. Specifically, workers are informed, ex-ante, that if 4 or more of the 7 workers in the group produce more than **T** = 29 mailers in the 1^{st} period, then the 2^{nd} period piece-rate will be reduced in half to $.10 for all 7 workers. The difference between the GROUP and BASELINE conditions is that participant workers face the outcome of a reduced piece-rate in the 2^{nd} work period if too many workers in the group are *collectively* too productive in the 1^{st} period.

The third experimental condition we implement, which we denote as the GROUP COMM condition, tests for the ratchet effect when productivity is evaluated collectively at the group-level, while *additionally* allowing workers to communicate with each other about the work task. The GROUP COMM condition is equivalent to the GROUP condition, except there is a 3-minute, pre-work communication phase. During the 3 minutes, the 7 workers were informed that they could collectively discuss "anything related to the study and the associated mailer assembly task." The group discussion was face-to-face, and during the discussion period the experimenter left the room to ensure privacy.[18] All communication among the group was restricted to the 3-minute discussion, and no additional communication between workers allowed during the work periods. After the 3-minute discussion period ended, the experimenter re-entered the room, and the remainder of the session proceeded in the same way as the GROUP condition. Thus, the only difference between the GROUP and GROUP COMM conditions is the ability of the group of 7 participant workers to openly communicate with each other for 3 minutes prior to starting work in the 1st period.

Importantly, with the exception of the compensation scheme, all treatments featured the same instructions and work conditions. Moreover, all participant workers were randomly drawn from the same database; hence, we maintain the assumption that the distribution of worker output capability observed in the BASELINE condition is representative and remains stable over all the experimental conditions. Thus, differences in productivity across treatments can be attributed to differences in how the 2nd period piece rate is determined based on 1st period output. For these additional conditions, our sample consists of 45 participant workers in the INDIVIDUAL condition (7 sessions), 42 in the GROUP condition (6 sessions), and 49 in the GROUP COMM condition (7 sessions). A summary of the experimental conditions and the corresponding piece-rates in both the 1st and 2nd work period is presented in Table 1.

3 Behavioral Hypotheses

Our first objective is to show that workers strategically restrict output and to establish the emergence of the ratchet under a piece-rate pay scheme when productivity is evaluated at the

[18] Although face-to-face communication is a strong form of communication, compared to anonymous chat, we chose this form for two reasons. First, given that our design was not computerized and is simulating a real workplace environment, there was no practical way to seamlessly integrate an anonymous computer chat. Second, face-to-face communication is likely the most prominent and natural way in which workers communicate with each other in the workplace. We did not record the group discussion (and informed participants of this) because we wanted to reduce experimenter demand effects and we wanted participants to feel comfortable discussing output restriction and possible cooperation or collusion in private, as would likely be true of real workplace discussions among workers.

individual-level. Given the assumption that the distribution of true output capability of the participant workers in the INDIVIDUAL condition is consistent with the BASELINE condition, the following requisite conditions are present for the ratchet effect to emerge in the INDIVIDUAL condition: (i) high ability workers (i.e., those workers capable of producing more output than the productivity threshold **T** = 29 mailers) are aware that if they produce at their full capability in the 1st work period, then their piece-rate will be reduced in the 2nd period, and (ii) the high ability workers have an incentive to restrict their output in the 1st period at or below **T** = 29; namely, they receive a higher payoff and have to exert less effort if they restrict their output in the 1st period. Thus, we expect to see the emergence of the ratchet effect in the INDIVIDUAL condition – high ability workers will respond rationally and strategically restrict their output at or below **T** = 29 mailers in the 1st period, leading to the following testable hypothesis:

HYPOTHESIS 1: *Average 1st period output in the INDIVIDUAL condition* ***will be*** *lower than in the BASELINE condition, and the proportion of workers producing 29 or fewer mailers in the 1st period* ***will be*** *larger in the INDIVIDUAL condition than in the BASELINE condition.*

Our second research question is whether we see the emergence of the ratchet effect when productivity is evaluated collectively at the group-level, which is important for two reasons. First, this may be representative of how management actually evaluates the productivity of their workforce. The narratives and discussions provided in Mathewson (1931), Edwards (1979), and Clawson (1980) point toward uniform piece-rates across equivalent types of workers, as well as productivity being evaluated at the group-level. For example, Edwards (1979, p. 99) states, "if all or most workers responded to the piece-rate with enough production to raise their wages substantially, then the expected job completion time would fall, and the piece-rate would be adjusted accordingly." Similarly, Clawson (1980, p. 170) discusses how "unless workers collectively restrict output they were likely to find themselves working much harder, producing much more, and earning only slightly higher wages." We assert that the use of a uniform piece-rate across workers and group-level productivity evaluation is especially likely to be implemented in workplaces employing many workers who are completing similar tasks, which is the type of simulated workplace environment we consider in our study.[19]

[19] Evaluating productivity at the group-level may be realistic in workplaces where it is difficult or costly for management to observe individual-level output. Management may also prefer uniform piece-rates across equivalent classes of workers to avoid any hostility and negative attitudes that may result from differential piece-rates.

Second, from an economic standpoint, evaluating productivity at the group-level can change the incentives of the individual workers in the group (see Prendergast, 1999 for a discussion). Within the ratchet effect context, it becomes *possible* for some workers to work at their full capability and still not have their piece-rate reduced if *enough* other workers restrict their output. That is, there is an incentive for workers to "free-ride" off the output restriction of other workers in the group. This creates a tension between what is optimal for the individual worker and what is collectively optimal for the group of workers, akin to a social dilemma (Dawes, 1980; Samuelson et al., 1984). The account by Clawson (1980, p. 174) speaks to workers' incentives to free-ride and the corresponding social dilemma that can arise, as he notes: "whereas restriction of output was in the interests of workers as a class, each individual worker had a large incentive to exceed the quota." While not specifically in the context of group-level productivity, there is extensive literature documenting at least some degree of free-riding across a range of social dilemmas, e.g., public goods games and common pool resource games.[20] More relevant to our study, the potential for free-riding on effort provision in the workplace when compensation is, at least in part, determined by group-level performance has been discussed (e.g., Kandel & Lazear, 1992; Hamilton et al., 2003; and Prendergast, 1999 for a review) and documented empirically (e.g., Weiss, 1987; Nalbantian & Schotter, 1997; Van Dijk et al., 2001; Bandiera et al., 2013). That being said, it strikes us as quite plausible that free-riding behavior may be salient in the context of effort provision and output production in workplaces where productivity is evaluated at the group-level.

In the GROUP condition, we create a setting where the piece-rate in the 2nd period is reduced only if the majority of the participant workers (4 of 7) are too productive in the 1st period, based on producing more than **T** = 29 mailers in the 1st period. Recall, **T** = 29 was approximately the 25th percentile of the output distribution for the BASELINE condition, so we expect 1 to 2 workers in the group of 7, on average, to have a true capability less than or equal to 29. Thus, in order to avoid having the piece-rate reduced for the entire group, it is likely that at least two high ability worker (and possibly as many as four) would need to restrict their output at or below 29 mailers. Therefore, even in the extreme case where 4 high ability workers need to restrict their output to avoid the piece-rate reduction, there is a clear opportunity for some of the workers to produce at full capacity

[20] See Dawes et al. (1977), Kim & Walker (1984), Isaac et al. (1984), and Isaac et al. (1985), the survey by Ledyard (1995) and the reference therein for examples of some of the early experimental studies illustrating evidence of free-riding behavior, as well as a more recent study by Fischbacher & Gachter (2010) and the survey by Chaudhuri (2011) for thorough reviews of the more recent experimental literature on public goods games.

and try to free-ride off the workers who restrict their output in the 1st period. However, if all of the high ability workers in the group attempt to free-ride, then this will induce full effort provision and output production across workers in the group. As a result, we expect the ratchet effect will be mitigated in the GROUP condition – high ability workers will be less likely to restrict their output at or below **T** = 29 mailers in the 1st work period, which leads to the following testable hypothesis:

HYPOTHESIS 2: *Average 1st period output in the GROUP condition* ***will not be*** *lower than in the BASELINE condition, and the proportion of workers producing 29 or fewer mailers in the 1st period* ***will not be*** *larger in the GROUP condition than in the BASELINE condition.*

At this point, it is informative to note that our GROUP condition features a strategic structure that is very similar to a binary-choice threshold public goods game (e.g., van de Kragt et al., 1983; Dawes et al., 1986; Palfrey & Rosenthal, 1991). In these games, individuals in a group face a binary decision between contributing nothing or their entire endowment to the public good. If *enough* people contribute – to meet or exceed the pre-defined threshold – then the public good is provided for everyone. Importantly, in these games, individuals have an incentive to not contribute and free-ride off the contribution of others. In a similar manner, in our GROUP condition if *enough* people restrict their output (so that at least 4 of 7 have a 1st period output of 29 or less) then the piece-rate is not reduced for the entire group (akin to having the public good provided). Although, as we have already established, each individual worker has an incentive to produce at full output capacity and free-ride off the output restriction of others. It is well established (e.g., Palfrey & Rosenthal, 1984; Cadsby & Mayes, 1999) that in threshold public goods games two primary types of equilibria emerge: (i) a "strong free-riding" equilibrium where nobody contributes and the public good is not provided, and (ii) multiple asymmetric equilibria where exactly the necessary number of people need to provide the public good contribute. Given that the GROUP condition features similar strategic properties to a discrete-choice threshold public goods game and the corresponding tension between what is individually optimal and socially optimal, Hypothesis 2 essentially corresponds to a prediction of the emergence of the strong free-riding equilibrium.

Our third research question is whether communication among workers promotes the emergence of the ratchet effect when worker productivity is evaluated at the group-level. Our motivation for studying the effect of communication is twofold. First, from a practical perspective, it seems reasonable that in many workplaces workers have the opportunity to discuss their work, their pay scheme, and the possible implications of their effort and productivity on future pay. For example, Mathewson (1931, p. 57) documents a case where "a bench worker fitting brass plates in a woodworking plant found he could easily exceed the customary number which the other men finished. His fellow-workmen observed this fact also and warned him that the whole group would have to reach the same point, if the boss noticed his higher production, and the rates would be cut." More generally, Clawson (1980, p. 175) notes that "numerous incidents of this kind [management ratcheting-up expectations] led workers to develop a class awareness of the need to restrict output…The concept of a class means that workers shared such experiences, and they developed a common viewpoint and approach, a common consciousness, as a basis from which to confront experiences or proposals." These anecdotal accounts suggest that communication among workers can promote output restriction and engender the ratchet effect along, at least, two dimensions: (i) by increasing the collective understanding in the group of workers that high output levels will likely be met with piece-rate reduction (or quota increases) in the future, and (ii) by helping coordinate the output restriction among the group of workers.[21]

In terms of pure economic incentives, the non-binding communication stage in the GROUP COMM condition does not alter the incentives structure and, hence, should have no differential impact on productivity compared to the GROUP condition. In particular, the incentive to produce at full capability and free-ride off the output restriction of other workers is still present in the GROUP COMM condition. Moreover, because participant workers assembled mailers within privacy carrels and the mailers were counted be the experimenter in private, there is no scope for post-experiment reputational consequence from workers being able to identify who free-rode and violated a cooperative agreement. That said, there is ample prior experimental literature documenting how non-binding communication can foster cooperation (see Dawes et al. 1977; Isaac & Walker, 1988; Bornstein & Rapoport, 1988; Orbell et al., 1988; Bornstein, 1992; Cooper et al., 1992; Charness, 2000; Duffy & Feltovich, 2002; Blume & Ortmann, 2007; Chaudhuri et al., 2009; and Sutter & Strassmair, 2009 for notable examples). Thus, if group communication can increase the collective understanding and the cooperative tendencies among workers, we would

[21] Possible mechanisms by which worker communication could facilitate coordination of output restriction include: collective informal agreements on production levels, non-binding commitments by workers to restrict output, increased peer-pressure to adhere to the group norm of output restriction, or possibly even coercion. For example, Clawson (1980, p. 177) reports that "in order to enforce output quotas it was definitely necessary for some workers to pressure and coerce others."

expect to observe output restriction and the emergence of the ratchet effect in the GROUP COMM condition, leading to the following testable hypothesis:[22]

HYPOTHESIS 3: *Average 1st period output in the GROUP COMM condition* ***will be*** *lower than in the BASELINE condition, and the proportion of workers producing 29 or fewer mailers in the 1st period* ***will be*** *larger in the GROUP COMM condition than in the BASELINE condition.*

Continuing with our parallel of the group setting to a discrete-choice threshold public goods game, Hypothesis 3 essentially corresponds to the prediction that pre-play communication among workers will enable workers to coordinate on the socially efficient equilibrium outcome where enough workers restrict output in the 1st period to ensure the piece-rate is not cut in the 2nd period, which benefits all workers in the group.

4 Results

We proceed by analyzing each of the three main experimental conditions – INDIVIDUAL, GROUP, and GROUP COMM – to identify if participant workers are restricting their output in the 1st period as evidence of the ratchet effect. Table 2 reports the aggregate 1st period output statistics for the INDIVIDUAL, GROUP, and GROUP COMM condition, compared to the BASELINE condition. Figure 2 plots the CDFs of 1st period output for each of the 4 treatments.

4.1 Establishing the Ratchet Effect in the INDIVIDUAL Condition

We consider 1st period output data from the INDIVIDUAL condition, in comparison to the BASELINE condition, to show that participant workers in the INDIVIDUAL condition are restricting their output below the productivity threshold, **T** = 29. Table 2 reveals that the average output in the 1st period in the INDIVIDUAL condition was 29.4 mailers, which is significantly lower than the BASELINE average of 34.6 (Mann-Whitney test: $p < .001$). Similarly, the median

[22] Recall that we evaluate whether a group of workers is too productive based on whether 4 or more of the workers produce more than **T**=29 mailers in the 1st period. We note that this method for evaluating group productivity, rather than using an aggregate measure of overall group productivity like the total or average number of completed mailers, may make it easier for the group of workers to collude and collectively coordinate output restriction when there is communication among the group; as a result, any effects that we find of group communication facilitating output restriction are likely and upper bound. That said, we expect the ratchet effect would be less likely to emerge if group productivity is evaluated using an aggregate measure where it may be more difficult for workers to coordinate, even when the group of workers is able to communicate.

output in the INDIVIDUAL condition was 28 mailers, which is significantly lower than the median of 34 in the BASELINE condition (K-sample medians test: $p < .001$).[23]

To provide more precise evidence of output restriction by participant workers in the INDIVIDUAL condition, we look at the proportion of workers producing an output level at or below the threshold **T** = 29. From Table 2, we see that 32/45 (71%) participant workers in the INDIVIDUAL condition completed 29 or fewer mailers, compared to 10/42 (24%) in the BASELINE condition, which is significantly different (Fisher's exact test: $p < .001$). More specifically, we can also look at just the proportion of participant workers completing 28 or 29 mailers. In the INDIVIDUAL condition, 16/45 (36%) workers complete 28 or 29 mailers, compared to 1/42 (2%) in the BASELINE condition, which is significantly different (Fisher's exact test: $p < .001$).[24] In terms of the distribution of 1st period output, the CDF plots in Figure 2 confirm a shift in the INDIVIDUAL condition from output levels above 30 mailers to levels below 30, especially toward 27-29; the distribution of 1st period output in the INDIVIDUAL condition is significantly different from the BASELINE (Epps-Singleton test: $p < .001$).[25]

Taken together, the data strongly support H1. Specifically, in the INDIVIDUAL condition, participant workers (in the aggregate) respond rationally by producing significantly less output in the 1st period compared to the BASELINE condition; furthermore, a significantly larger proportion of workers in the INDIVIDUAL condition completed less than or equal to **T** = 29 mailers in the 1st period, compared to the BASELINE condition. This finding is summarized in Result 1:

---

[23] Taking a more conservative statistical approach, we can also compare the session-level average 1st period output levels across the INDIVIDUAL and BASELINE conditions. In the BASELINE condition, the average 1st period output levels for each of the 7 sessions, in order from highest to lowest, were: 38.0, 36.4, 35.5, 34.6, 34.3, 33.8, and 30.5; in the INDIVIDUAL condition, the corresponding session-level averages were: 33.0, 32.4, 30.3, 30.2, 28.1, 26.2, and 26.0. Comparing these session-level averages, the INDIVIDUAL condition is significantly different than the BASELINE condition (Mann-Whitney test: $p = .004$), and this is robust if we instead use session-level median output levels (Mann-Whitney test: $p = .003$).

[24] We include both 28 and 29 as output levels representing deliberate output restriction, as opposed to just the threshold level **T** = 29, to allow for possible misinterpretation of the instructions on the part of the participant workers. In particular, some participant workers may have deliberately stopped at 28 to avoid the risk that they misinterpreted the instructions thinking that producing 29 would actually result in the piece-rate reduction (i.e., the old adage that it's better to be "safe than sorry"). Given that 9 of 45 participant workers in the INDIVIDUAL condition completed 28, while 0 of 42 completed 28 in the BASELINE, we feel confident asserting that 28 represented a deliberate choice for many of these participant workers in the INDIVIDUAL condition. However, our results are qualitatively robust if we consider only the proportion of workers completing 29; it is 7/45 (16%) in the INDIVIDUAL condition and 1/42 (2%) in the BASELINE condition, which is still significantly different (Fisher's exact test: $p = .059$).

[25] Because the distribution of completed mailers is discrete, we test for distributional differences across treatments using an Epps-Singleton test in lieu of the more commonly used Kolmogorov-Smirnov (KS) test (Goerg & Kaiser, 2009). However, the results from the distributional tests across treatments are all robust if a KS-test is used instead.

***RESULT 1** – We find strong empirical evidence of the ratchet effect in the INDIVIDUAL condition. A significant portion of participant workers in the INDIVIDUAL condition appear to be strategically restricting their output in the 1st period relative to their true capability.*

4.2 Testing for the Ratchet Effect in the GROUP Condition

Next, we test for the ratchet effect in the GROUP condition, where the productivity is evaluated collectively based on the output levels of each of the 7 workers in the group; namely, do participant workers restrict their output at or below **T** = 29 in the 1st work period. Table 2 reveals that the average output in the 1st period across the 42 participant workers in the GROUP condition was 33.5 mailers and the median was 32.5 mailers, compared to the BASELINE average and median of 34.6 and 34, respectively. Neither the average nor median output levels in the 1st period are significantly different between the GROUP and BASELINE conditions (Mann-Whitney test: p = .456; K-sample medians test: p = .827, respectively).[26]

Again, we can more precisely test for the presence of strategic output restriction in the GROUP condition by looking at the proportion of participant workers producing at or below the threshold **T** = 29 in the 1st period. Only 12/42 (29%) participant workers in the GROUP condition completed less than or equal to 29 mailers in the 1st period, which is not statistically different from the 10/42 (24%) workers in the BASELINE condition (Fisher's exact test: p = .804). Similarly, the proportion of participant workers who completed 28 or 29 mailers in the 1st period was 4/42 (10%), which is not significantly different from the 1/42 (2%) workers in the BASELINE (Fisher's exact test: p = .360). Lastly, the CDFs plots in Figure 2 reveal that the 1st period output distribution in the GROUP is very similar to that of the BASELINE condition, and these distributions are not statistically different (Epps-Singleton test: p = .913).

We can also directly compare 1st period output in the GROUP condition to the INDIVIDUAL condition to verify that there is significantly less output restriction among workers in the GROUP condition. Indeed, the average and median 1st period output levels of 33.5 and 32.5, respectively, in the GROUP condition are significantly larger than the average of 29.4 and median of 28 in the INDIVIDUAL condition (Mann-Whitney test: p = .002; K-sample medians test: p < .001,

[26] Similar findings emerge if we consider session-level output. The average output levels for each of the 6 sessions in the GROUP condition, from largest to smallest, were: 38.6, 33.7, 33.7, 32.9, 31.3, and 30.4, while in the BASELINE condition, the session-level averages were: 38.0, 36.4, 35.5, 34.6, 34.3, 33.8, and 30.5; these session-level averages are not significantly different (Mann-Whitney test: p = .153), and this result is robust if we instead use session-level median output levels (Mann-Whitney test: p = .282).

respectively). In addition, the proportion of workers in the GROUP condition who produced under 29 mailer is significantly less than in the INDIVIDUAL condition (Fisher's exact test: p < .001), as well as the proportion of worker who complete 28 or 29 mailers (Fisher's exact test: p = .004). Finally, comparing the CDF plots of 1st period output between GROUP and INDIVIDUAL conditions in Figure 2, there is a shift to larger values in GROUP (Epps-Singleton test: p = .004).

Overall, the data reveal significantly less output restriction in the GROUP condition compared to the INDIVIDUAL condition, which suggests that measuring productivity at the group level can mitigate the ratchet effect compared to when productivity is measured individually. Furthermore, in the environment we consider, the group dynamics appear to virtually eliminate output restriction among workers. In particular, there is very little difference in the average or median output levels in the 1st period between participant workers in the GROUP and BASELINE conditions; further, there is no significant difference in the proportion of workers who completed less than or equal to **T** = 29 mailers in the 1st period. This finding supports H2 and is summarized in Result 2:

***RESULT 2** – We find no empirical evidence of the ratchet effect in the GROUP condition. Participant workers in the GROUP condition do not appear to be restricting their output in the 1st period relative to their true capability.*

4.3 Testing for the Ratchet Effect in the GROUP COMM Condition

The last part of our main analysis is testing for output restriction of participant workers and the presence of the ratchet effect in the GROUP COMM condition, where the 7 workers were given 3 minutes to discuss the work task prior to commencing work. From Table 2, we see that the average 1st period output in the GROUP COMM condition was 31.7 mailers, which is marginally significantly lower than the BASELINE average of 34.6 (Mann-Whitney test: p = .093), and the median output of 29 mailers in the GROUP COMM condition is significantly lower than the median of 34 in the BASELINE condition (K-sample medians test: p = .002).[27]

[27] Because of the pre-play communication among workers in the GROUP COMM condition, each participant's output level may no longer be independent. As a result, the inferences from the statistical tests used to compare individual level output data between the BASELINE and GROUP COMM conditions ought to be interpreted with some degree of caution. That said, even if we take a conservative approach and compare the BASELINE and GROUP COMM conditions at the session level (where observations are independent), we still find evidence of lower 1st period output in the GROUP COMM condition. Specifically, the session-level average 1st period output levels in the GROUP COMM condition were: 34.7, 34.7, 34.4, 30.7, 30.0, 29.6, and 27.6, while in the BASELINE condition they were: 38.0, 36.4, 35.5, 34.6, 34.3, 33.8, and 30.5; these session-level averages border on being significant using a Mann-Whitney test (p = .109), and are significant if we use a more powerful Fisher-Pitman permutation test (p = .056). In addition, the session-level median output levels are significant different (Mann-Whitney test: p = .029).

Looking more specifically at the proportion of workers producing an output level at or below the threshold **T** = 29, 26/49 (53%) workers in the GROUP COMM condition completed 29 or fewer mailers compared to just 10/42 (24%) in the BASELINE condition, which is significantly different (Fisher's exact test: p = .005). Only considering the proportion of participant workers completing 28 or 29 mailers, there were 18/49 (37%) workers in the GROUP COMM condition compared to 1/42 (2%) in the BASELINE condition, which is also significantly different (Fisher's exact test: p < .001). Figure 2 confirms a leftward shift in the distribution of 1st period output toward 27-29 mailers in the GROUP COMM condition from the BASELINE condition; with the distributions being significantly different (Epps-Singleton test: p = .046).

A more direct method to examine the role of communication among the worker groups, and to verify that there is significantly more output restriction in the GROUP COMM condition, is to compare 1st period output in the GROUP COMM condition to the GROUP condition. The average and median 1st period output levels of 31.7 and 29, respectively, in the GROUP COMM condition are significantly lower than the average of 33.5 and median of 32.5 in the GROUP condition (Mann-Whitney test: p = .093; K-sample medians test: p = .053, respectively). Additionally, the proportion of workers in the GROUP COMM condition who produced under 29 mailer is significantly more than in the GROUP condition (Fisher's exact test: p = .015), as well as the proportion of worker who complete 28 or 29 mailers (Fisher's exact test: p = .002).

Taken together, the data reveal evidence of significantly more output restriction by workers in the GROUP COMM condition compared to the GROUP condition, indicating that communication among workers can help coordinate output restriction and facilitate the emergence of the ratchet effect. Moreover, when we compare 1st period output in the GROUP COMM condition to the BASELINE condition, we see clear evidence of output restriction in the GROUP COMM condition; namely, participant workers (in the aggregate) produce significantly less output compared to the BASELINE condition; additionally, a significantly larger proportion of workers in the GROUP COMM condition complete less than or equal to **T** = 29 mailers in the 1st period compared to the BASELINE condition. This finding supports H3 and is summarized in Result 3:

RESULT 3 *– We find strong empirical evidence of the ratchet effect in the GROUP COMM condition. A significant portion of participant workers in the GROUP COMM condition appear to be strategically restricting their output in the 1st period relative to their true capability.*

5 Dynamic Implications of the Ratchet Effect on Future Productivity

One of the advantages of using a real-effort task as the foundation for our experimental design is the ability to examine the possible dynamic implications of the ratchet effect. We hypothesize that a plausible channel through which current output restriction can impact future productivity is through learning-by-doing since workers accumulate less on-the-job learning.

5.1 Output Restriction and Reduced Learning-by-Doing

The underpinnings of learning-by-doing rest in the notion that agents become more adept and efficient at completing a task through the experience gained at completing the task.[28] In a seminal paper, Arrow (1962, p. 155) argues that "learning is the product of experience. Learning can only take place through the attempt to solve a problem and therefore only takes place during activity." Lucas (1988, p. 27) notes that "as many economists have observed, on-the-job-training or learning-by-doing appear to be at least as important as schooling in the formation of human capital." Within the context of this study, if workers restrict their output, then they are gaining less experience with the work task, which can reduce learning and future productivity. Relatedly, if workers are restricting their output, then there is less incentive for workers to develop innovative methods for completing the work task more efficiently since they will be unable to reap the full benefit of increases in production capability associated with the innovation. Thus, output restriction has the potential to stifle innovation (Dearden et al., 1990; Carmichael & MacLeod, 2000), which can further reduce productivity.

In the BASELINE condition, recall that we observed workers producing 29 percent more mailers in the 2nd period relative to the 1st period (an average increase from 34.6 to 44.6), which suggests that workers are potentially learning through experience and becoming more innovative in their production methods.[29] More specifically, if learning-by-doing effects are present, then we may expect a positive relation between the growth in productivity from the 1st period to the 2nd period and the amount produced in the 1st period (i.e., the amount of doing). Figure 3 plots the growth in output, calculated as 2nd period output minus 1st period output, against 1st period output

[28] We refer readers to Thompson (2010) for a thorough discussion of learning-by-doing and a comprehensive review of the literature finding support for learning-by-doing, as well as a recent paper by Levitt et al. (2013), who document empirical evidence of learning-by-doing in a car manufacturing plant.

[29] Although the task seems simple in nature, casual observation of workers by the experimenters (at the time of collecting and counting mailers after each period) revealed some degree of production innovation, as many workers appeared to have adopted an assembly line approach and/or re-configured their workspace to enhance efficiency.

for workers in the BASELINE condition. From Figure 3 we see that there is generally a positive trend (r = .229) between 1st period output and the growth in output. Moreover, a simple regression of the growth in output on 1st period output yields a positive and significant coefficient on 1st period output ($\beta = .129$; *p-value* = .054). The fact that we observe significantly higher average output levels in the 2nd period compared to the 1st period, combined with output growth in the 2nd period being positively related to 1st period output, suggests the potential presence of learning-by-doing. This effect is consistent with findings from DellaVigna et al. (2016) who also document increases in productivity in a mailer task attributed to learning-by-doing.

To broadly explore if, and to what extent, deliberate output restriction impacts subsequent productivity, we compare 2nd period output of participant workers in the INDIVIDUAL condition with the BASELINE condition. For this analysis we focus specifically on the INDIVIDUAL condition and do not consider the GROUP COMM condition for two primary reasons. First, we observe the largest degree of output restriction in the 1st period in the INDIVIDUAL condition, which makes it the most suitable in terms of power when testing whether output restriction impacts 2nd period productivity. Second, comparing 2nd period output levels between the INDIVIDUAL and BASELINE provides the cleanest analysis because we circumvent any possible confounds that could result from the pre-play communication, selection effects, or group-based evaluation that were present in the GROUP COMM condition. Given the observed empirical evidence of significant output restriction in the INDIVIDUAL condition, we hypothesize the following:

HYPOTHESIS 4: *Average 2nd period output in the INDIVIDUAL condition* ***will be*** *lower than in the BASELINE condition.*

Indeed, we do find that average 2nd period output in the INDIVIDUAL condition is significantly lower than in the BASELINE condition, which we present in more detail in Section 5.2. However, an important caveat to this result is that a subset of the workers in the INDIVIDUAL condition – those workers who produced more than 29 mailers in the 1st period – had a lower piece-rate in the 2nd period. Therefore, in addition to a potential learning-by-doing effect, the decrease in 2nd period output in the INDIVIDUAL condition may be due to a possible reduction in productivity by the subset of workers who earned a lower piece-rate in the 2nd period – a wage effect. In Section 5.3 we turn our attention to identifying the magnitude of the potential wage effect. In particular, we run a follow-up experimental condition, which we refer to as the BASELINE LOW condition, where all participants receive a $.20 piece-rate in the 1st period and $.10 piece-rate in the 2nd period.

Importantly, comparing the 2nd period output between the BASELINE LOW and BASELINE conditions provides us with an estimate of the wage effect. Previewing our findings, we document a relatively small wage effect in the BASELINE LOW condition. Importantly, in Section 5.4 we show that this small wage effect is not big enough to generate the observed reduction in 2nd period output in the INDIVIDUAL condition. Through some additional disaggregated data analysis, we are able to provide strong evidence that the reduction in 2nd period output in the INDIVIDUAL condition is being driven largely by the subset of workers in the INDIVIDUAL condition who deliberately restricted output, which is attributed to reduced learning-by-doing.

5.2 The Effect of Restricting 1st Period Output on 2nd Period Productivity

Figure 4 displays the average 2nd period output for the BASELINE and INDIVIDUAL conditions as well as the respective CDFs of 2nd period output. From the left panel of Figure 4, we see that average 2nd period output in the INDIVIDUAL condition was 39.8 compared with an average output of 44.6 in the BASELINE condition (an 11% reduction); this difference of 4.8 mailers is significantly different (Mann-Whitney test: p = .016). Similarly, the median 2nd period output was 38 in the INDIVIDUAL condition and 44.5 in the BASELINE, which is also significantly different (K-sample medians test: p = .071). The CDFs in the right panel of Figure 4 show a clear shift in the distribution toward lower 2nd period output levels in the INDIVIDUAL condition (Epps-Singleton test: p = .115).[30] The data suggest that, in the aggregate, 2nd period output is lower in the INDIVIDUAL condition compared to the BASELINE, which is summarized in Result 4:

RESULT 4 *– We document empirical evidence that current output restriction via the ratchet effect can reduce future productivity. Participant workers in the INDIVIDUAL condition (where output restriction was present in the 1st work period) completed significantly fewer mailers in the 2nd work period relative to the BASELINE condition.*

5.3 Identifying the Wage Effect

As previously discussed, the observed differences in 2nd period productivity between the INDIVIDUAL and BASELINE conditions documented above may be the result of two potential effects: (i) a learning-by-doing effect *and* (ii) a wage effect that reduces effort stemming from receiving a lower wage. Recall that in the INDIVIDUAL condition, participant workers either receive a piece-rate of \$.10 or \$.20 in the 2nd period (depending on 1st period productivity), whereas

[30] If these two distributions are compared using a KS-test, then there is a significant difference (KS-test: p = .014).

all participant workers in the BASELINE condition receive a piece-rate or $.20 in the 2nd period. As such, there is the possibility that the observed reduction in 2nd period productivity in the INDIVIDUAL condition is, at least partially, a result of lower productivity by the subset of participant workers earning the lower piece-rate of $.10; namely, the 28.9% of workers in the INDIVIDUAL condition who produced 30 or more mailers in the 1st period. All else equal, these workers may exert lower effort when receiving $.10 per mailer rather than $.20.[31]

5.3.1 Additional BASELINE LOW Condition

To rule out the possibility that the significant decrease in 2nd period output in the INDIVIDUAL condition is being driven *entirely* by the potential wage effect, we conducted a follow-up experimental condition, denoted as BASELINE LOW, to estimate the wage effect. In the BASELINE LOW condition, all participant workers receive a piece-rate of $.20 in the 1st period (as in the BASELINE condition); however, unlike in the BASELINE condition, the piece-rate in the BASELINE LOW condition is reduced to $.10 in the 2nd work period for *all* participant workers. We ran 8 additional sessions of the BASELINE LOW condition with 51 total participants.

Importantly, the empirical distribution of 2nd period output in the BASELINE LOW condition provides us with an estimate of the true output capability of our worker sample under a $.10 piece-rate. Hence, comparing the distribution of 2nd period output in the BASELINE LOW condition to the BASELINE condition isolates the potential wage effect by eliminating the learning-by-doing channel.[32] In particular, we are interested in identifying the potential wage effect for the subset of workers in the distribution who completed 30 or more mailers in the 1st period. The reason being is that if the wage effect is driving the reduction in 2nd period output between the INDIVIDUAL and BASELINE conditions, it must be coming exclusively from the subset of workers in the INDIVIDUAL condition who produced 30 or more in the 1st period (i.e., only those workers who

[31] There is mixed laboratory evidence on the effects of reduced piece-rates on output. Using a mailer assembly task, DellaVigna et al. (2016) find that output decreased by 12% when piece-rates were reduced from 20 cents to 10 cents. Although, Carpenter & Gong (2016) do not find any significant difference in worker productivity when the piece rate was reduced from $1.00 to $.50. Relatedly, Paarsch & Shearer (2009) and Bellemare & Shearer (2011) document significant decreases in worker output in tree planting resulting from lower piece-rates.

[32] In both the INDIVIDUAL and BASELINE LOW conditions, participant workers were informed about the piece-rate pay scheme for both periods before work began in the 1st period; hence, workers in both conditions were not surprised in the 2nd period by an unexpected piece-rate reduction. An implication of revealing the compensation scheme, ex ante, is that we minimize the potential for reductions in 2nd period output in either the INDIVIDUAL or BASELINE LOW conditions to be a result of behavioral reactions to changes in the wage (e.g., indignation, spite, retaliation, moral reaction to an unfair wage). This assertion is consistent with recent findings by Sliwka & Werner (2016) who document experimental evidence that productivity of workers does not dynamically respond to changes in the wage rate, when the changes in wages are ex-ante known and anticipated by workers.

had their piece-rate reduced to \$.10). By disaggregating the data and comparing 2nd period output between the BASELINE LOW and BASELINE conditions for the subset of workers who produced 30 or more in the 1st period, we can specifically identify an estimate of the *relevant* average wage effect on productivity for those workers in the distribution of 1st period output that is 30 or more.

Prior to presenting the results on 2nd period output, it is necessary that we first compare 1st period output between BASELINE LOW and BASELINE; notably, since BASELINE LOW sessions were conducted after the original sessions, we need to ensure that participant workers in the BASELINE LOW condition represent a similar sample of workers.[33] Importantly, it also ensures than any differences in 2nd period output between BASELINE and BASELINE LOW are not a result of differences in 1st period output, which could arise, for instance, if workers in the BASELINE LOW are responding strategically in the 1st period to the anticipated lower wage in the 2nd period. Namely, if workers anticipate the productivity gains associated with learning-by-doing, then the incentive to invest in learning in the 1st period (through increased effort) is weaker in the BASELINE LOW condition because of the lower return from learning in the 2nd period (\$.10 vs \$.20), which could result in lower 1st period output levels.

Recall, the mean and median 1st period output levels in the BASELINE condition were 34.6 and 34, respectively. In the BASELINE LOW condition, the mean and median 1st period output levels were 34.4 and 34, respectively, which are not significantly different from the BASELINE (Mann-Whitney test: p = .874; K-sample medians test: p = .941). Figure 5 displays the distribution of 1st period output in the BASELINE and BASELINE LOW conditions, which are not significantly different from each other (Epps-Singleton test: p = .798). The striking similarity in the distribution of 1st period output between BASELINE and BASELINE LOW suggest that there is no potential confounding effects from differences in worker ability when comparing 2nd period output between BASELINE LOW and the other conditions. Additionally, there appears to be no suggestive evidence that workers in the BASELINE LOW condition are strategically responding in the 1st period to the anticipated lower piece-rate in the 2nd period. As a result, any observed difference in 2nd period output between BASELINE LOW and BASELINE can confidently be attributed purely to the wage effect.

[33] Importantly, participants from the BASELINE LOW condition were recruited from the same subject-pool database as the four original sessions, so there is no reason to suspect, ex-ante, any systematic differences in the types of workers who participated in the BASELINE LOW condition.

5.3.2 Estimating the Wage Effect

We now turn our attention to the 2nd period output to provide an estimate of the wage effect. The mean and median 2nd period output levels in the BASELINE LOW condition were 43.6 and 44, respectively, compared to 44.6 and 44.5, respectively, in the BASELINE condition. Thus, the reduction in mean 2nd period productivity was 1.0 mailer in the BASELINE LOW condition compared to the BASELINE, which is not statistically significant (Mann-Whitney test: p = .529), and there is similarly no significant difference in the median 2nd period output (K-sample medians test: p = .941). This finding suggests that, within our experimental setting, over the entire distribution of 1st period output, the wage effect is minimal. This is consistent with findings documented by Carpenter (2016) and Carpenter & Gong (2016) in a mailer assembly task.[34]

That said, there is substantial heterogeneity in worker ability, and it is possible that the wage effect is not constant over the entire distribution of 1st period output. As discussed earlier, of particular interest is identifying the potential wage effect for those workers in the INDIVIDUAL condition who had a 1st period output of 30 or more, i.e., those workers who were actually treated with a piece-rate reduction. To estimate the wage effect for this subset of workers, we disaggregate the data and compare 2nd period productivity in the BASELINE and BASELINE LOW conditions for only the subset of workers who produced 30 or more in the 1st period. Looking specifically at this subsample, the mean and median 2nd period output levels in the BASELINE LOW condition were 45.9 and 45, respectively, compared to 48.6 and 49, respectively, in the BASELINE condition. The mean reduction in output of 2.7 in the BASELINE LOW condition, relative to the BASELINE condition, is larger than when comparing the effect on the entire distribution, although the difference is not statistically significant (Mann-Whitney test: p = .122).

We can obtain a more robust estimate of the wage effect, conditional on 1st period output, through a simple regression analysis. Particularly, we flexibly regress 2nd period output on 1st period output by using a fifth degree polynomial that is specified as follows:

[34] In the aggregate, we observe a small wage effect. We suspect that in our setting, the lack of effort response to the decrease in piece-rates could have possibly been attributed to a few things. First, the charitable nature and the mission of the TTAA could have, at least partially, crowded out the monetary incentives. Although, Sliwka & Werner (2016) also document experimental evidence that worker output does not respond to changes in wages (when the wages changes are known ex-ante and fully anticipated) in an abstract real-effort task with no charitable component. Second, the process by which the piece-rate was reduced could have weakened the wage effect. Specially, the piece-rates in our setting are effectively generated from an external process and not from an actual human principal, which could have mitigated reciprocal motivations of the agents to changes in the piece-rate, as discussed and shown experimentally by Charness (2004) and Charness & Levine (2007).

$$y_i = \alpha + f(period1_i) + g(period1_i) \times BASELINELOW_i + \delta BASELINELOW_i + \varepsilon_i.$$

In the regression model, y_i represents 2^{nd} period output for individual i, $f(\cdot)$ represents a fifth degree polynomial in 1^{st} period output, $g(\cdot)$ represents a fifth degree polynomial in 1^{st} period output interacted with an indicator that takes the value of one if individual i is in the BASELINE LOW condition, and zero if individual i is the BASELINE condition. Note, the estimated wage effect is equal to the marginal effect of y with respect to the $BASELINELOW$ indicator variable.

Since we model the wage effect as depending on 1^{st} period output, we report our results graphically. Figure 6 plots the estimated wage and 95% confidence interval for the distribution of the subset of workers of interest – those workers who produce 30 or more mailers in the 1^{st} period. From Figure 6, it is evident that the magnitude of the estimated wage effect varies minimally over the relevant range of 1^{st} period productivity, and is rarely statistically different from zero. Moreover, the wage effect is at most 3.97 (when 1^{st} period output is 39) for this subset of workers.

5.4 Evidence of Reduced Learning-by-Doing from Output Restriction

Recall that average output in the 2^{nd} period decreased by 4.8 mailers in the INDIVIDUAL condition, relative to the BASELINE condition (Result 4). Yet, we documented in Section 5.3 that the overall wage effect, estimated as the difference in average 2^{nd} period output between the BASELINE and BASELINE LOW conditions, was only 1.0 mailers. It is evident that the magnitude of the wage effect across the entire distribution of workers is not sufficiently large to account for the *entire* reduction in 2^{nd} period productivity in the INDIVIDUAL condition.

The implausibility that the wage effect is driving the entire reduction in productivity is even more pronounced when we properly account for the fact that only a subset of the workers in the INDIVIDUAL condition – those 28.9% of workers that produced 30 or more mailers in the 1^{st} period – received the reduced piece-rate of \$.10 in the 2^{nd} period. Therefore, any potential wage effect in the INDIVIDUAL condition is coming from a non-random subsample of less than one-third of the distribution. We can determine the necessary size of the wage effect needed to justify a null learning-by-doing effect by conditioning on the fact that only 28.9% of the workers in the INDIVIDUAL condition were actually treated with a wage decrease. In particular, the local average treatment effect of the wage decrease on the treated workers in the INDIVIDUAL condition would have to be: 4.8/.289 = 16.6 mailers (a 37.2% reduction from the BASELINE mean) to justify a null learning-by-doing effect. When directly comparing the implied 16.6 mailer decrease to the actual observed 1.0 mailer decrease (BASELINE vs. BASELINE LOW), it overwhelmingly suggests that the decline in 2nd period productivity is not due to the wage effect.

Moreover, when we consider only the 28.9% of relevant workers in the INDIVIDUAL condition who received the wage reduction, we estimated that the average wage effect for this subset of treated workers was 2.7 mailers and at most 3.97 mailers (Section 5.3.2). This indicates that even if we consider only the subset of relevant workers in the INDIVIDUAL condition who were impacted by a lower wage in the 2nd period, the estimated wage effect is still not sufficiently large to account for the entire reduction in average 2nd period output in the INDIVIDUAL condition, relative to the BASELINE, and to justify a null learning-by-doing effect. Rather, we contend that the reduction in productivity is largely being driven by reduced leaning-by-doing among the workers that deliberately restricted output in the INDIVIDUAL condition.

An alternative method of disentangling the wage effect from the learning-by-doing effect is to provide a bound of the learning-by-doing effect. Under a weak assumption that the wage effect is monotonic, we can obtain a lower bound estimate of the magnitude of the learning-by-doing effect arising from output restriction in the INDIVIDUAL condition by comparing the 2nd period output in the INDIVIDUAL condition to the BASELINE LOW condition. In the BASELINE LOW condition, all workers received $.10 per mailer in the 2nd period, while in the INDIVIDUAL condition a subset of only 28.9% of the workers received $.10 per mailer in the 2nd period; as such, a comparison of the average 2nd period output between the INDIVIDUAL and BASELINE LOW conditions will provide a lower bound estimate (in magnitude) of the learning-by-doing effect. By way of explanation, note that average 2nd period output in the INDIVIDUAL condition would be expected to be even smaller under a counterfactual scenario in which all workers received the reduced $.10 piece-rate in the 2nd period. As a result, the magnitude of the difference in 2nd period output between the INDIVIDUAL and BASELINE LOW conditions represents a lower bound of the productivity decrease resulting from reduced learning-by-doing.

The mean and median 2nd period output levels in the BASELINE LOW condition were 43.6 and 44, respectively; compared to a mean and median of 39.8 and 38, respectively, in the INDIVIDUAL condition. Thus, workers produced 3.8 fewer mailers, on average, during the 2nd period in the INDIVIDUAL condition (an approximate 9% reduction in output), which is significant (Mann-Whitney test: $p = .019$). Also, a test for equality of 2nd period output distributions can be rejected (Epps-Singleton test: $p = .030$). The fact that we observe significantly lower 2nd period production in the INDIVIDUAL condition compared to the BASELINE LOW condition strongly suggests a large learning-by-doing effect, which accounted for *at least* a 9% reduction in average 2nd period productivity.

We conclude this section with some additional disaggregated analysis on 2nd period output across treatments to further substantiate the learning-by-doing effect. Specifically, if we first look at the subset of workers in the INDIVIDUAL condition who produced 27 or less in the 1st period, then this represents the workers who neither restricted output nor had their piece-rate reduced in the 2nd period; as such, there should be no selection effects or wage effects for these workers in the INDIVIDUAL condition. Therefore, the output profiles for this subset of workers in the INDIVIDUAL condition should be similar to the analogous subset of workers in the BASELINE condition. Considering only workers who produced 27 or less in the 1st period, the average output levels in the 1st and 2nd period were 24.2 and 32.3, respectively, in the INDIVIDUAL condition and 23.4 and 31.6, respectively, in the BASELINE condition; neither of which are significantly different across the two conditions. Hence, the virtually identical output profiles suggest that workers in the INDIVIDUAL condition who produced 27 or less did not contribute to the aggregate reduction in 2nd period output in the INDIVIDUAL condition.

Next, if we consider the subset of workers in the INDIVIDUAL condition who produced 30 or more in the 1st period, then this represents the workers who didn’t restricted output but did have their piece-rate reduced in the 2nd period. For these workers in the INDIVIDUAL condition, there may be a wage effect and, therefore, the output profiles for this subset of workers can loosely be compared to the subset of workers in the BASELINE LOW condition (where workers also received the reduced piece-rate in the 2nd period). A caveat of considering the distribution of this subset of workers from the INDIVIDUAL condition is that it may not be random due to the selection of workers who strategically restricted output. That said, for only the workers who produced 30 or more in the 1st period, the average output levels in the 1st and 2nd period were 36.9 and 45.7, respectively, in the INDIVIDUAL condition and 37.2 and 45.9, respectively, in the BASELINE LOW condition; neither of which are significantly different across the two conditions. Importantly, the fact that the output profiles for the subset of workers who produced 30 or more in the 1st period are similar between the INDIVIDUAL and BASELINE LOW conditions suggests that our estimated wage effect of 2.7 for treated workers in the INDIVIDUAL condition, which is identified from the relevant comparison group in the BASELINE LOW condition, is reasonable;

this provides further evidence that the relative reduction in aggregate 2nd period productivity in the INDIVIDUAL condition is largely due to learning-by-doing, and not a wage effect.

Overall, the results from the BASELINE LOW condition, combined with additional disaggregated data analysis of 2nd period productivity, suggest that the large reduction in 2nd period productivity in the INDIVIDUAL condition is attributed primarily to the 36% of workers who deliberately restricted their output in the 1st period. Given that these workers did not experience a piece-rate reduction, the relative reduction in 2nd period productivity must be an artifact of the restriction of output in the 1st period. We contend that a probable mechanism linking the output restriction of these workers with the reduction in future productivity is reduced learning-by-doing.

6 Concluding Remarks

Piece-rate incentive schemes are a common form of compensation in many workplaces, especially in the manufacturing, textile, agriculture, and sales sectors. The conventional economic wisdom behind the implementation of piece-rates is to incentivize effort provision by workers and, thus, mitigate shirking. However, in a dynamic setting, piece-rates can give rise to the *ratchet effect* – a phenomenon where workers strategically restrict their current output, relative to their true capability, because they rationally anticipate that high levels of output will be met by management with decreased piece-rates or higher quotas in the future. While there is a substantial amount of theoretical work supporting the emergence of the ratchet effect, as well as more qualitative anecdotal evidence suggestive of output restriction among workers under piece-rate schemes, there is little empirical research on the presence and implications of the ratchet effect.

In this study, we implement an experimental design where participant workers complete a real-effort task (assembling donor solicitation mailers) under a piece-rate pay scheme for two work periods. Importantly, we are able to recover an estimate of the distribution of true output capability for participant workers, which enables us to empirically test for output restriction and, therefore, identify the presence of the ratchet effect. After first empirically establishing the presence of the ratchet effect within our sample of workers when productivity is evaluated based on individual output, we then explore how group dynamics affect the emergence of the ratchet effect. In particular, we test whether output restriction among workers is reduced when productivity is evaluated collectively at the group-level, and whether communication among workers helps to facilitate group-level output restriction. Lastly, implementing a real-effort work task, in combination with the documented evidence of output restriction, enables us to examine how output restriction impacts future productivity through the channel of reduced learning-by-doing.

We find that workers appear to be rationally restricting output in the condition where they face the consequence of a reduced piece-rate if they are too productive in the 1st work period; as such, our results provide robust evidence for the emergence of the ratchet effect in a real-effort work task when productivity is evaluated at the individual-level. However, when productivity is evaluated at the group-level, we do not observe a decrease in 1st period output levels, indicating no evidence of output restriction. Our results suggest that in the workplace setting we consider, the free-riding incentive that arises within the group is strong enough to overcome the incentive to restrict output, which results in full effort provision by workers and output levels in line with true capability, thus mitigating the ratchet effect. However, when we augment the group-level condition to allow for pre-play communication among workers, we again document strong evidence of output restriction of workers as the ratchet effect re-emerges; communication among workers appears to foster cooperation and help coordinate collective output restriction among workers, thus overcoming the free-rider problem. [35] Our group-level results are particularly relevant for firms who employ many workers that perform a similar task and where management is likely to be imperfectly informed about production technology (e.g., sales, agriculture, manufacturing, assembly lines), since group-based compensation schemes would be more likely to occur under such circumstances. The ratchet effect can still emerge when productivity is evaluated at the group-level; namely, the successful coordination among workers to restrict output, facilitated through communication, which could have important productivity implications for the firm.

An advantage of using a real-effort task within a simulated workplace environment is the ability to investigate a possible dynamic productivity implication of the ratchet effect. We hypothesize that output restriction can reduce future productivity through reduced learning-by-doing. In our data, we find strong evidence that output restriction in the 1st work period leads to significantly

[35] By linking the piece-rate that all workers in the group receive in the 2nd period to group-level productivity in the 1st period, the strategic environment in the 1st period is essentially transformed into a threshold public goods game. Consistent with the possible equilibria that arise in such games, in the GROUP condition when workers aren't able to communicate we observe basically no output restriction, which is analogous to the emergence of strong free-riding equilibrium where nobody contributes to the public good. But, in the GROUP COMM condition where pre-play communication is permitted, we observe significant output restriction, which is in line with the cooperative equilibrium outcome where enough people contribute such that the public good is provided. While there is ample existing evidence that communication can promote cooperation in public goods games, our results suggest that communication is also instrumental in facilitating cooperation in a different domain: effort provision among workers and coordinating collective output restriction.

lower output levels in the 2nd work period. Importantly, based on the results of an additional follow-up experimental condition, we can rule out the possibility that the reduction in 2nd period productivity is being driven entirely by a potential wage effect from the subset of workers who receive a lower wage in the 2nd period. Furthermore, through some disaggregated data analysis that exploits the natural variation in worker ability, we show that the observed reduction in aggregate 2nd period productivity is largely attributed to the subset of workers who strategically restricted output in the 1st period. Because these workers didn't experience any wage reduction, our findings strongly suggest that a probable mechanism driving this reduction in productivity is reduced learning-by-doing. In fact, we estimate that the learning-by-doing effect experienced by the workers who restrict output in the 1st work period leads to a reduction in the 2nd period output of at least 9%. As such, our results provide evidence of a plausible indirect consequence of the ratchet effect in the workplace: deliberate output restriction by workers imposes a negative externality, through reduced learning-by-doing, that reduces future productivity.

Workers are often an important driver of innovation and productivity growth for firms. Carmichael & Macleod (1993) discuss how "firms in this situation [where workers know more about the fine details of production than the managers] need the enthusiastic cooperation of their workers if they are to innovate successfully" (p. 143). If workers are restricting output in anticipation of ratchet effect dynamics, then there is less incentive for workers to innovate and invest time and effort into developing productivity enhancing methods. Importantly, we find evidence supporting this claim, and our results suggest that output restriction can significantly reduce learning by-doing, which is an important channel through which innovation and production efficiency would occur. As a result, not only can the ratchet effect reduce firm productivity directly through current output restriction of its workers, but also indirectly by reducing innovation and stifling future productivity growth of the firm. Furthermore, we conjecture this detrimental effect on future productivity coming through reduced innovation and learning-by-doing would be even more impactful for firms, especially young firms, in new industries or industries that are labor intensive and where technology is rapidly evolving where there is a greater scope for productivity growth. For such firms, piece-rate pay or productivity laden compensation schemes, in addition to generating ratchet effect dynamics, could also hinder long term productivity growth through reduced innovation and learning-by-doing. This effect may be even more pronounced when groups of workers are able to collectively collude to conceal the true productivity of new technologies.

We acknowledge that our experimental design is stylized and may not fully represent naturally occurring workplaces. That said, we assert that the stylized features of the design allow for clean identification of output restriction among workers and the emergence the ratchet effect under both individual and group based productivity measures, the impact of communication among workers, and the impact of current output restriction on future productivity. Moreover, our design does feature a *natural* real-effort task that is likely to be perceived by participant workers as regular, economically valuable work (Falk & Ichino, 2006), as well as some degree of field context arising from the partnership with a university alumni association and the legitimacy of the mailers, which adds to the credibility of our main findings. That being said, the results from our study complement the prior chosen-effort experimental studies and qualitative field studies, and contribute more generally to our understanding of the ratchet effect in the workplace. Notably, output restriction among workers is a valid concern for firms implementing piece-rates (or performance-based pay more generally), when such firms are unable to dynamically commit to not revising the compensation scheme. At the same time, conditional on implementing a piece-rate pay scheme, firms should strongly consider evaluating productivity at the group-level (when possible); this allows firms to exploit the free-rider problem that arises from group-based productivity measures and possibly reduce the degree of output restriction among workers. Our study joins Charness et al. (2011) in identifying possible mechanisms that can mitigate the ratchet effect in the workplace. The likelihood of reducing the ratchet effect and inducing full-effort provision by measuring output at the group-level will be amplified in workplaces where communication among workers is more difficult and scarce (e.g., where workers may be physically separated from each other, work remotely, or have little interaction with other workers), as the lack of communication helps mitigate workers from collectively, coordinating output restriction.

While the focus of our study is on identifying output restriction of workers and the emergence of the ratchet effect in the workplace, our results may have important implications in other economic setting where the ratchet effect has been theoretically speculated to arise. For example, in the context of regulatory compliance by firms (e.g., emissions standards), firms have an incentive to under invest in efficiency improving technologies if they anticipate the regulator will set stricter standards in the future (Olsen & Torsvik, 1993; Dalen, 1995; Puller, 2006). In this setting, it seems unlikely that firms communicate and work collaboratively to jointly restrict innovation (for fear of possible anti-trust sanctions); assuming the regulator can evaluate outcomes at the group-level for all the firms in a given industry, then perhaps the ratchet effect as it relates to strategic under investment by firms is less severe than theoretically predicted. Another example where the ratchet effect might emerge is retail sales where sales managers have an incentive to reduce effort if they anticipate that high sales totals will be met with higher sales targets in the future (Bouwens & Kroos, 2011). In this context, if sales targets are determined at the group-level across multiple store locations, where communication among the different sales managers is likely to be minimal, sales ratcheting might be less likely to occur. Lastly, the ratchet effect have been recently explored in the context of teachers having an incentive to restrict effort when bonuses depend on student performance exceeding specific targets, where the targets can be revised upward based on prior performance of students (Macartney, 2016). If these performance targets are set based on prior group-level student performance from all classes at the district level, where rampant communication across teachers at different schools in the district is unlikely, then perhaps strategic under provision of effort by teachers would be less of a concern.

References

Abeler, J., Falk, A., Goette, L., & Huffman, D. (2011). Reference points and effort provision. *American Economic Review*, 101(2), 470-492.

Allen, D. W., & Lueck, D. (1999). Searching for Ratchet Effects in Agricultural Contracts. *Journal of Agricultural and Resource Economics, 24*(2), 536-552.

Arrow, K. J. (1962). The economic implications of learning by doing. *Review of Economic Studies, 29*(3), 155-173.

Bagnoli, M., & Lipman, B. L. (1989). Provision of public goods: Fully implementing the core through private contributions. *The Review of Economic Studies*, *56*(4), 583-601.

Bandiera, O., Barankay, I., & Rasul, I. (2013). Team Incentives: Evidence from a Firm Level Experiment. *Journal of the European Economic Association, 11*(5), 1079-1114.

Banker, R. D., Lee, S.-Y., & Potter, G. (1996). A field study of the impact of a performance-based incentive plan. *Journal of Accounting and Economics, 21*(2), 195-226.

Baron, D., & Besanko, D. (1987). Commitment and Fairness in a Dynamic Regulatory Relationship. *Review of Economic Studies, 54*(3), 413-436.

Bellemare, C., & Shearer, B. (2011). On the Relevance and Composition of Gifts within the Firm: Evidence from Field Experiments. *International Economic Review, 52*(3), 855-882.

Bellemare, C., & Shearer, B. (2015). Contracts, Commitment, and the Ratchet Effect: Evidence from a Field Experiment. *Mimeo*.

Bellemare, C., Lepage, P., & Shearer, B. (2010). Peer pressure, incentives, and gender: An experimental analysis of motivation in the workplace. *Labour Economics, 17*(1), 276-283.

Berliner, J. S. (1976). *The Innovation Decision in Soviet Industry.* Cambridge: MIT Press.

Besley, T., & Ghatak, M. (2005). Competition and incentives with motivated agents. *American Economic Review, 95*(3), 616-636.

Bhaskar, V. (2014). The Ratchet Effect Re-examined: A Learning Perspective. *Mimeo*.

Blume, A., & Ortmann, A. (2007). The effects of costless pre-play communication: Experimental evidence from games with Pareto-ranked equilibria. *Journal of Economic Theory, 132*(1), 274-290.

Bornstein, G. (1992). The free-rider problem in intergroup conflicts over step-level and continuous public goods. *Journal of Personality and Social Psychology, 62*, 597-606.

Bornstein, G., & Rapoport, A. (1988). Intergroup competition for the provision of step-level public goods: Effects of preplay communication. *European Journal of Social Psychology*, *18*(2), 125-142.

Bouwens, J., & Kroos, P. (2011). Target Ratcheting and Effort Reduction. *Journal of Accounting and Economics*, *51*, 171-185.

Cadsby, C., & Maynes, E. (1999). Voluntary provision of threshold public goods with continuous contributions: experimental evidence. *Journal of Public Economics*, *71*(1), 53-73.

Carmichael, H. L., & MacLeod, B. (1993). Multiskilling, Technical Change and the Japanese Firm. *The Economic Journal*, 104, 142-160.

Carmichael, H. L., & MacLeod, B. (2000). Worker Cooperation and the Ratchet Effect. *Journal of Labor Economics, 18*(1), 1-19.

Carpenter, J. (2016). The labor supply of fixed-wage workers: Estimates from a real effort experiment. *European Economic Review, 89*, 85-95.

Carpenter, J., & Gong, E. (2016). Motivating Agents: How Much Does the Mission Matter? *Journal of Labor Economics, 34*(1), 211-236.

Charness, G. (2000). Self-Serving Cheap Talk: A Test Of Aumann's Conjecture. *Games and Economic Behavior, 33*(2), 177-194.

Charness, G. (2004). Attribution and reciprocity in an experimental labor market. *Journal of Labor Economics*, 22(3), 665-688.

Charness, G., Gneezy, U., & Kuhn, M. (2013). Extra-Laboratory Experiments: Extending the Reach of Experimental Economics. *Journal of Economic Behavior and Organization, 91*, 93-100.

Charness, G., & Kuhn, P. (2011). Lab Labor: What Can Labor Economists Learn from the Lab? In O. Ashenfelter, & D. Card, *Handbook of Labor Economics* (Vol. 4, pp. 229-330). Elsevier.

Charness, G., Kuhn, P., & Villeval, M. C. (2011). Competition and the Ratchet Effect. *Journal of Labor Economics, 29*(3), 513-547.

Charness, G., & Levine, D. I. (2007). Intention and stochastic outcomes: An experimental study. *The Economic Journal*, 117(522), 1051-1072.

Chaudhuri, A. (1998). The ratchet principle in a principal agent game with unknown costs: an experimental analysis. *Journal of Economic Behavior and Organization*, *37*, 291-304.

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics, 14*(1), 47-83.

Chaudhuri, A., Schotter, A., & Sopher, B. (2009). Talking Ourselves to Efficiency: Coordination in Inter-Generational Minimum Effort Games with Private, Almost Common and Common Knowledge of Advice. *The Economic Journal, 119*(534), 91-122.

Choi, J. P., & Thum, M. (2003). The Dynamics of Corruption with the Ratchet Effect. *Journal of Public Economics, 87*, 427-443.

Clawson, D. (1980). *Bureaucracy and the Labor Process: The Transformation of U.S. Industry, 1860–1920.* New York: Monthly Review Press.

Cooper, D. J., Kagel, J. H., Lo, W., & Liang Gu, Q. (1999). Gaming Against Managers in Incentive Systems: Experimental Results with Chinese Students and Chinese Managers. *American Economic Review, 89*(4), 781-804.

Cooper, R., DeJong, D., Forsythe, R., & Ross, T. (1992). Communication in Coordination Games. *The Quarterly Journal of Economics, 107*(2), 739-771.

Dalen, D. M. (1995). Efficiency-Improving Investment and the Ratchet Effect. *European Economic Review, 39*(8), 1511-1522.

Dawes, R. (1980). Social Dilemmas. *Annual Review of Psychology*(31), 169-193.

Dawes, R. M., McTavish, J., & Shaklee, H. (1977). Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation. *Journal of Personality and Social Psychology, 35*, 1-11.

Dawes, R. M., Orbell, J. M., Simmons, R. T., & Van De Kragt, A. J. (1986). Organizing groups for collective action. *American Political Science Review, 80*(04), 1171-1185.

Dearden, J., Ickes, B. W., & Samuelson, L. (1990). To Innovate or Not to Innovate: Incentives and Innovation in Hierarchies. *American Economic Review, 80*(5), 1105-1124.

DellaVigna, S., List, J., Malmendier, U., & Rao, G. (2016). Estimating Social Preferences and Gift Exchange at Work. *NBER Working Paper No. w22043*.

Duffy, J., & Feltovich, N. (2002). Do Actions Speak Louder Than Words? An Experimental Comparison of Observation and Cheap Talk. *Games and Economic Behavior, 39*(1), 1-27.

Edwards, R. C. (1979). *Contested Terrain: The Transformation of the Workplace in the Twentieth Century.* New York: Basic Books.

Falk, A., & Fehr, E. (2003). Why Labour Market Experiments? *Labour Economics*(10), 399-406.

Falk, A., & Ichino, A. (2006). Clean Evidence of Peer Effects. *Journal of Labor Economics, 24*(1), 39-57.

Fernie, S., & Metcalf, D. (1999). It's Not What You Pay it's the Way that You Pay it and that's What Gets Results: Jockeys' Pay and Performance. *Labour, 13*(2), 385-411.

Fischbacher, U., & Gachter, S. (2010). Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments. *American Economic Review*, *100*(1), 541-56.

Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, *19*(4), 25-42.

Freixas, X., Guesnerie, R., & Tirole, J. (1985). Planning under Incomplete Information and the Ratchet Effect. *Review of Economic Studies, 52*(2), 173-191.

Friedman, D., & Sunder, S. (1994). *Experimental Methods: A Primer for Economists.* Cambridge University Press.

Gibbons, R. (1987). Piece-Rate Incentive Schemes. *Journal of Labor Economics, 5*(4), 413-429.

Gill, D., & Prowse, V. (2015). A Novel Computerized Real Effort Task Based on Sliders. *Mimeo*.

Gneezy, U., Goette, L., Sprenger, C., & Zimmermann, F. (2017). The limits of expectations-based reference dependence. *Journal of the European Economic Association*, forthcoming.

Goerg, S., & Kaiser, J. (2009). Non-Parametric Testing of Distributions - the Epps-Singleton two-sample test using the Empirical Characteristic Function. *The Stata Journal, 9*(3), 454-465.

Hamilton, B., Nickerson, J., & Owan, H. (2003). Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams. *Journal of Political Economy*, *111*(3), 465-497.

Ickes, B. W., & Samuelson, L. (1987). Job Transfers and Incentives in Complex Organizations: Thwarting the Ratchet Effect. *RAND Journal of Economics*, *18*(2), 275-286.

Isaac, M., & Walker, J. (1988). Communication and free-riding behavior: The voluntary contribution mechanism. *Economic inquiry, 26*(4), 585-608.

Isaac, M., McCue, K., & Plott, C. R. (1985). Public Goods Provision in an Experimental Environment. *Journal of Public Economics*, *26*(1), 51-74.

Isaac, M., Schmidtz, D., & Walker, J. (1989). The assurance problem in a laboratory market. *Public Choice*, *62*(3), 217-236.

Isaac, M., Walker, J., & Thomas, S. H. (1984). Divergent evidence on free riding: An experimental examination of possible explanations. *Public Choice, 43*(2), 113-149.

Kandel, E., & Lazear, E. P. (1992). Peer Pressure and Partnerships. *Journal of Policial Economy, 100*(4), 801-817.

Kanemoto, Y., & MacLeod, B. (1992). The Ratchet Effect and the Market for Secondhand Workers. *Journal of Labor Economics, 10*(1), 85-98.

Kim, O., & Walker, M. (1984). The free-rider problem: experimental evidence. *Public Choice, 43*(1), 3-24.

Konow, J. (2000). Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions. *American Economic Review, 90*(4), 1072-1091.

Kube, S., Maréchal, M., & Puppe, C. (2013). Do wage cuts damage work morale? Evidence from a natural field experiment. *Journal of the European Economic Association, 11*(4), 853-870.

Kuhn, P., & Lozano, F. (2008). The Expanding Workweek? Understanding Trends in Long Work Hours among US Men, 1979–2006. *Journal of Labor Economics, 26*(2), 311-343.

Laffont, J. J., & Tirole, J. (1988). The Dynamics of Incentive Contracts. *Econometrica, 56*(5), 1153-1175.

Lawler, E. (1971). *Pay and Organizational Effectiveness.* New York: McGraw-Hill.

Lawler, E. E., Mohrman, S. A., & Benson, G. (2001). *Organizing for high performance: Employee involvement, TQM, reengineering, and knowledge management in the Fortune 1000.* San Francisco: Jossey-Bass.

Lazear, E. P. (1986). Salaries and piece rates. *Journal of Business, 59*(3), 405-431.

Lazear, E. P. (2000). Performance Pay and Productivity. *American Economic Review, 90*(3), 1346-1361.

Ledyard, J. O. (1995). Public Goods: A Survey of Experimental Research. In J. H. Kagel, & A. E. Roth, *The Handbook of Experimental Economics* (pp. 111-194). Princeton, NJ: Princeton University Press.

Lemieux, T., MacLeod, B., & Parent, D. (2009). Performance Pay and Wage Inequality. *Quarterly Journal of Economics, 124*(1), 1-49.

Levine, D. I. (1992). Piece rates, output restriction, and conformism. *Journal of Economic Psychology, 13*(3), 473-489.

Levitt, S., List, J., & Syverson, C. (2013). Toward an Understanding of Learning by Doing: Evidence from an Automobile Assembly Plant. *Journal of Political Economy, 121*(4), 643-681.

Lucas, R. (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics, 22*(1), 3-42.

Macartney, H. (2016). The Dynamic Effects of Educational Accountability. *Journal of Labor Economics*, *34*(1), 1-28.

Mathewson, S. B. (1931). *Restriction of Output among Unorganized Workers.* New York: The Viking Press.

Meyer, M. A., & Vickers, J. (1997). Performance comparisons and dynamic incentives. *Journal of Political Economy*, *105*(3), 547-581.

Montgomery, D. (1979). *Workers control in America: Studies in the history of work, technology and labor struggles.* New York: Cambridge University Press.

Nalbantian, H. R., & Schotter, A. (1997). Productivity under group incentives: An experimental study. *American Economic Review, 87*(3), 314-341.

Olsen, T., & Torsvik, G. (1993). The Ratchet Effect in Common Agency: Implications for Regulation and Privatization. *Journal of Law, Economics, & Organization, 9*(1), 136-158.

Orbell, J., Van de Kragt, A., & Dawes, R. (1988). Explianing discussion-induced cooperation. *Journal of Personality and Social Psychology*, *54*, 811-819.

Paarsch, H., & Shearer, B. (1999). The Response of Worker Effort to Piece Rates: Evidence from the British Columbia Tree-Planting Industry. *Journal of Human Resources, 34*(4), 643-667.

Paarsch, H., & Shearer, B. (2000). Piece Rates, Fixed Wages and Incentive Effects: Statistical Evidence from Payroll Records. *International Economic Review*, *41*(1), 59-92.

Paarsch, H., & Shearer, B. (2009). The response to incentives and contractual efficiency: Evidence from a field experiment. *European Economic Review, 53*(5), 481-494.

Palfrey, T., & Rosenthal, H. (1984). Participation and the provision of discrete public goods: a strategic analysis. *Journal of public Economics*, *24*(2), 171-193.

Palfrey, T. R., & Rosenthal, H. (1991). Testing for effects of cheap talk in a public goods game with private information. *Games and Economic Behavior*, *3*(2), 183-220.

Prendergast, C. (1999). The Provision of Incentives in Firms. *Journal of Economic Literature, 37*(1), 7-63.

Prendergast, C. (2007). The Motivation and Bias of Bureaucrats. *American Economic Review, 97*(1), 180-196.

Puller, S. (2006). The strategic use of innovation to influence regulatory standards. *Journal of Environmental Economics and Management, 52*, 690-706.

Rapoport, A., & Eshed-Levy, D. (1989). Provision of step-level public goods: Effects of greed and fear of being gypped. *Organizational Behavior and Human Decision Processes*, *44*(3), 325-344.

Samuelson, C. D., Messick, D., Rutte, C., & Wilke, H. (1984). Individual and structural solutions to resource dilemmas in two cultures. *Journal of Personality and Social Psychology, 1*(47), 94-104.

Seiler, E. (1984). Piece Rate Vs. Time Rate: The Effect of Incentives on Earnings. *Review of Economics and Statistics, 66*(3), 363-376.

Shearer, B. (2004). Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment. *Review of Economic Studies, 71*(2), 513-534.

Skelton, B. R., & Yandle, B. (1982). Piece rate pay. *Journal of Labor Research, 3*(2), 201-209.

Sliwka, D., & Werner, P. (2016). How do agents react to dynamic wage increases? An experimental study. *Journal of Labor Economics,* forthcoming.

Stiglitz, J. E. (1975). Incentives, risk, and information: notes towards a theory of hierarchy. *The Bell Journal of Economics, 6*(2), 552-579.

Sutter, M., & Strassmair, C. (2009). Communication, cooperation and collusion in team tournaments-An Experimental Study. *Games and Economic Behavior, 66*, 506-525.

Thompson, P. (2010). Learning by Doing. In B. H. Hall, & N. Rosenberg, *Handbook of The Economics of Innovation, Vol. 1*. Elsevier.

Van de Kragt, A. J., Orbell, J. M., & Dawes, R. M. (1983). The minimal contributing set as a solution to public goods problems. *American Political Science Review, 77*(01), 112-122.

Van Dijk, F., Sonnemans, J., & Van Winden, F. (2001). Incentive systems in a real effort experiment. *European Economic Review, 45*(2), 187-214.

Weiss, A. (1987). Incentives and Worker Behavior: Some Evidence. In H. R. Nalbantian, *Incentives, Cooperation, and Risk Sharing* (pp. 137-150). Totowa, NJ: Rowman & Littlefield.

Weitzman, M. L. (1980). The ‘Ratchet Principle’ and Performance Incentives. *Bell Journal of Economics, 11*(1), 302-308.

Figure 1 – Distribution of Completed Mailers: BASELINE Condition

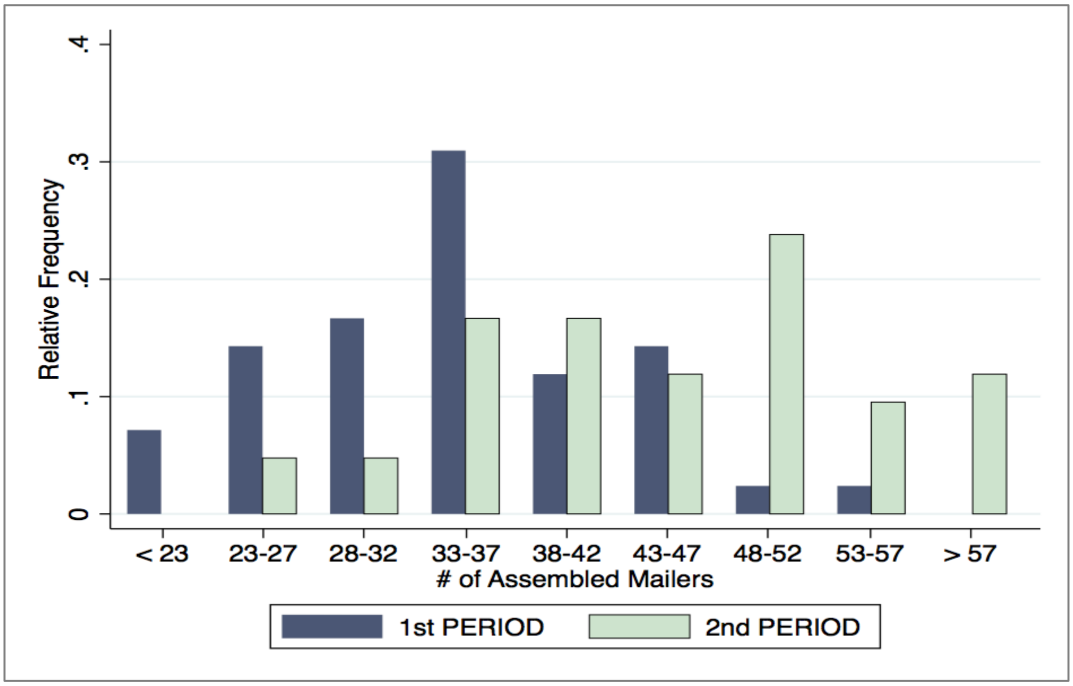

Figure 2 – Comparison of 1st Period Output Distributions across Conditions

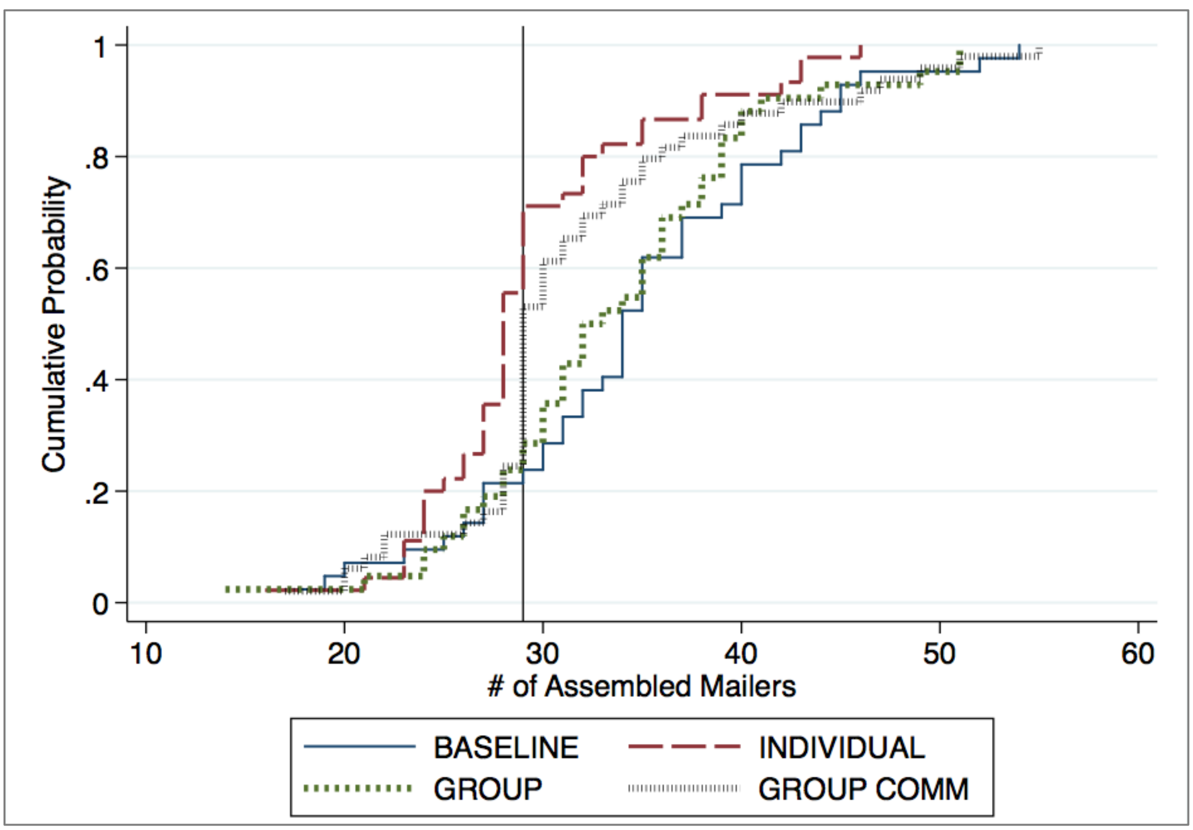

Figure 3 – Scatter Plot of 1st Period Output and Change in Output: BASELINE Condition

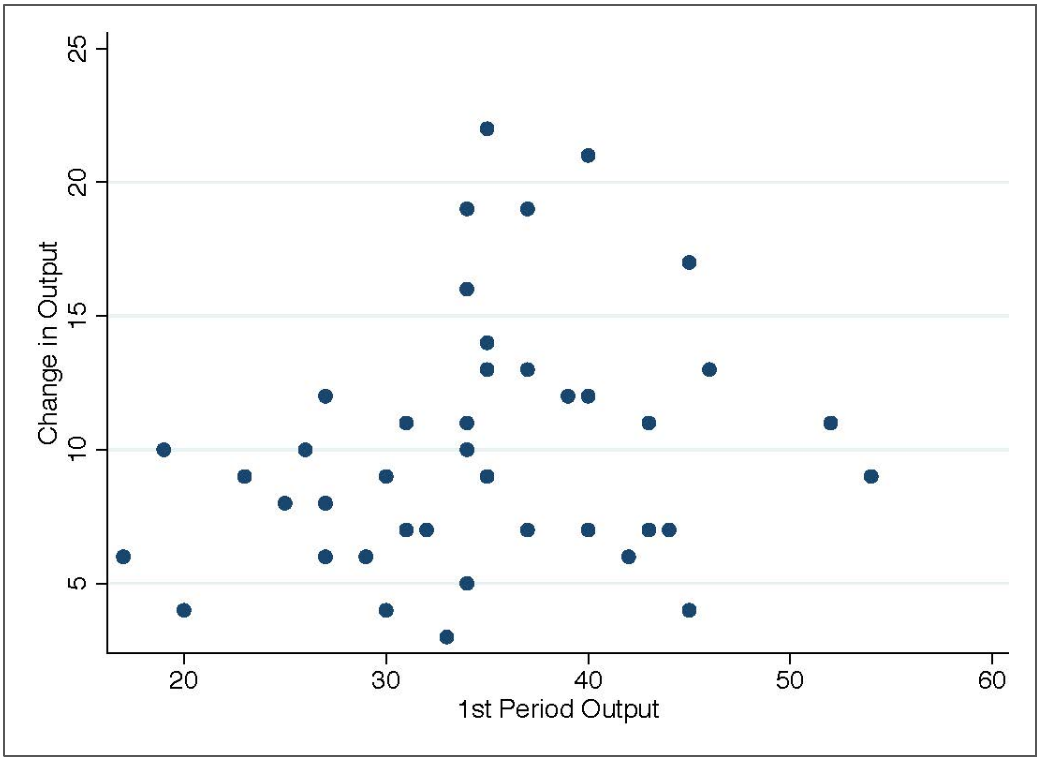

Figure 4 – Comparison of 2nd period output: INDIVIDUAL vs BASELINE

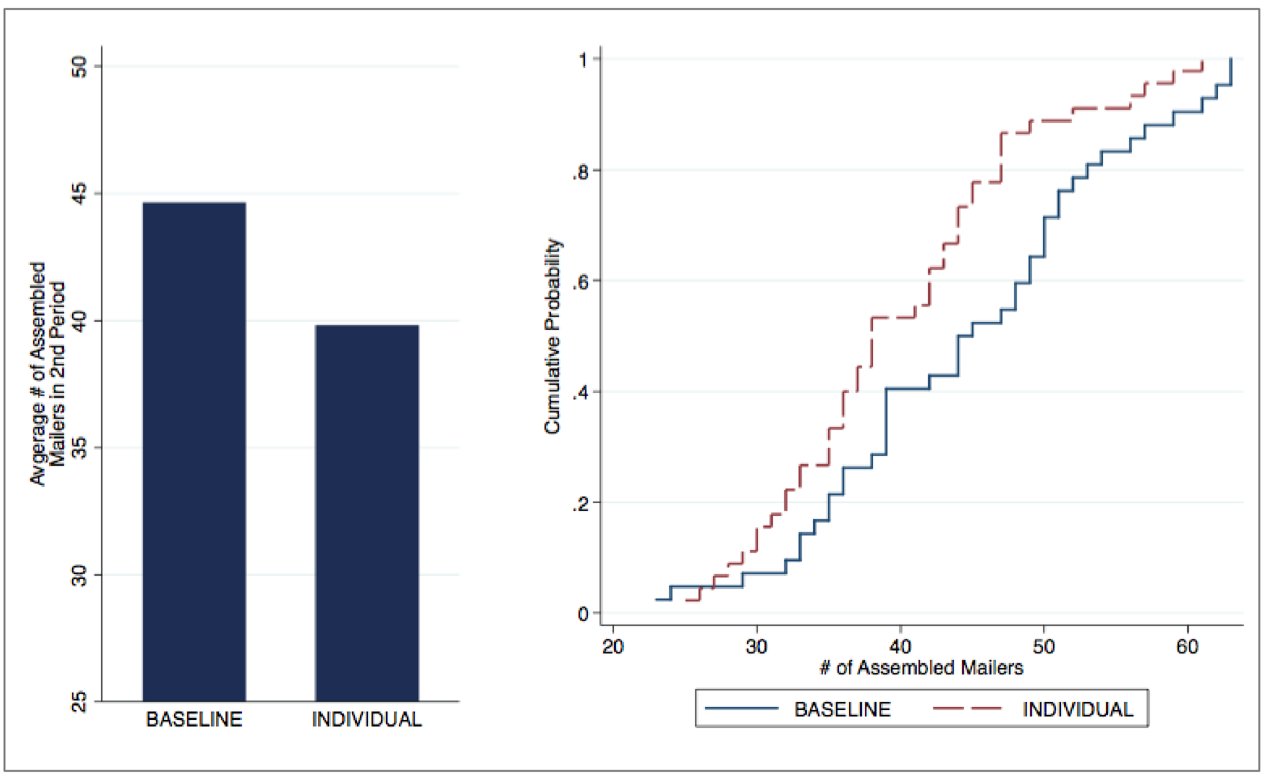

Figure 5 – Comparison of 1st period output: BASELINE vs BASELINE LOW

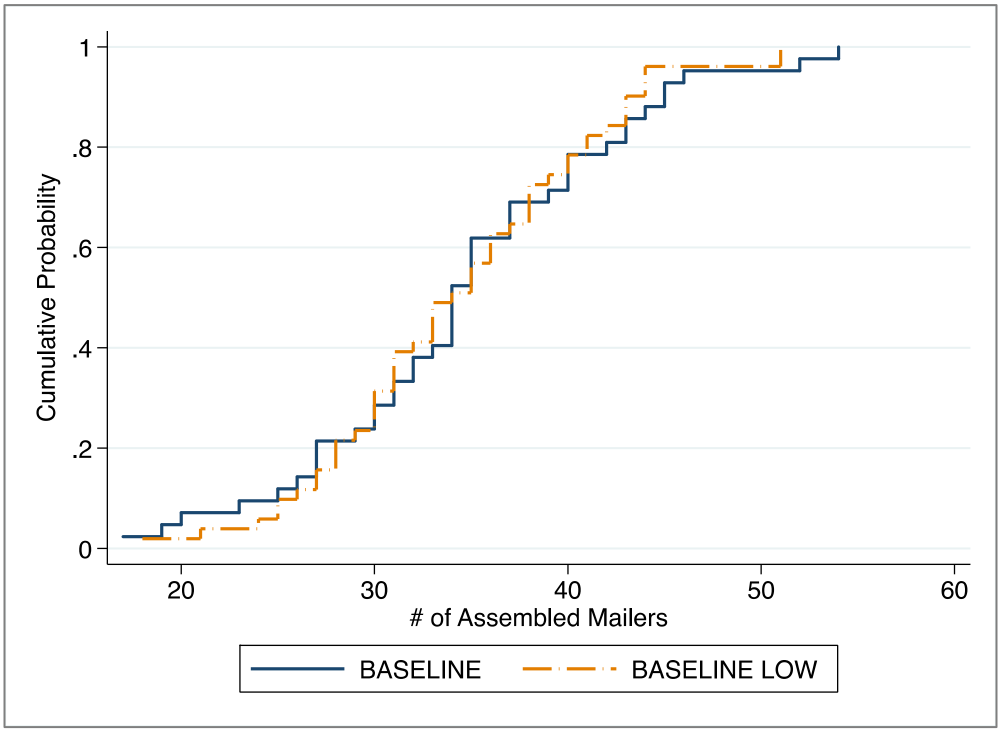

Figure 6 – Estimate of the Wage Effect: BASELINE LOW relative to BASELINE

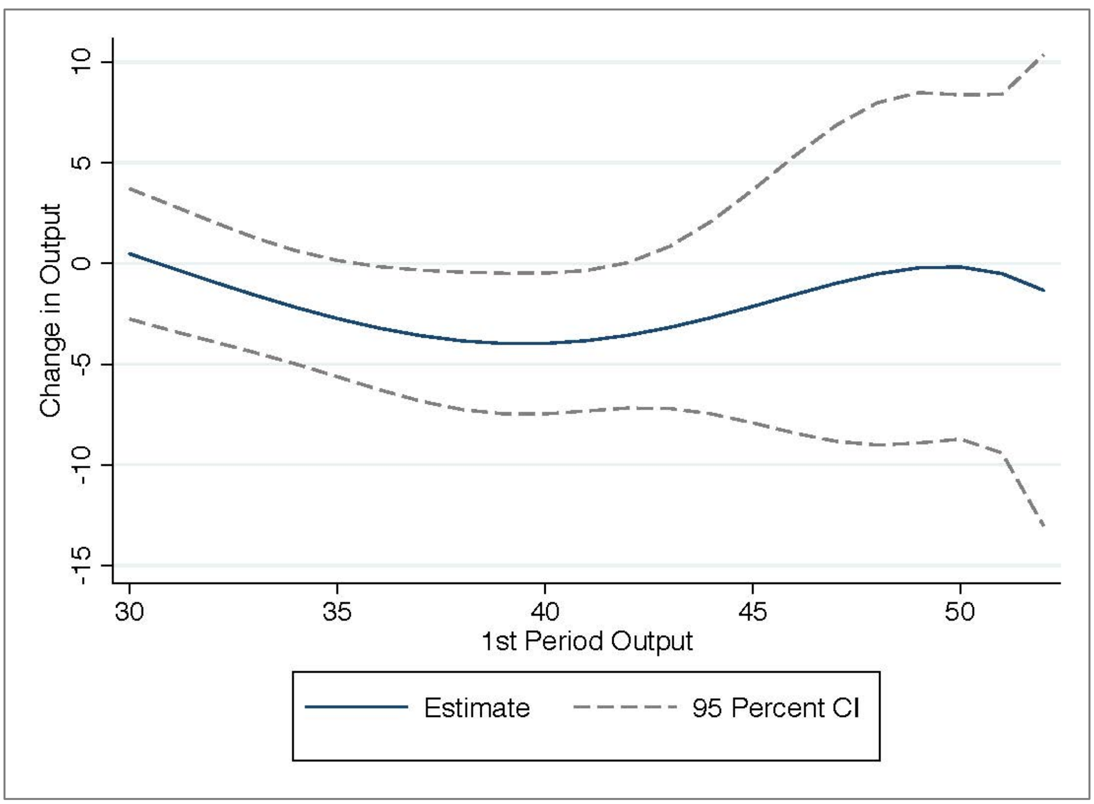

Table 1 – Summary of Experimental Conditions

	BASELINE n = 42 (7 sessions)	INDIVIDUAL n = 45 (7 sessions)	GROUP n = 42 (6 sessions)	GROUP COMM n = 49 (7 sessions)
1st Work Period **piece-rate**	$.20	$.20	$.20	$.20 *and* 3 minutes of pre-work discussion
2nd Work Period **piece-rate**	$.20 for all workers	$.20 for workers with output <= 29 in 1st work period $.10 for workers with output > 29 in 1st work period	$.20 if 4 or more of the 7 workers had output <= 29 in 1st work period $.10 if 4 or more of the 7 workers had output > 29 in 1st work period	$.20 if 4 or more of the 7 workers had output <= 29 in 1st work period $.10 if 4 or more of the 7 workers had output > 29 in 1st work period

Table 2 – Comparison of 1st Period Output across Experimental Conditions

Productivity Measure	Experimental Conditions			
	BASELINE (n = 42)	INDIVIDUAL (n = 45)	GROUP (n = 42)	GROUP COMM (n = 49)
Mean output	34.6	29.4***	33.5	31.7*
Median output	34	28***	32.5	29***
# of workers with output of 29 or fewer mailers	10/42 (24%)	32/45*** (71%)	12/42 (29%)	26/49*** (53%)
# of workers with output of exactly 28 or 29 mailers	1/42 (2%)	16/45*** (36%)	4/42 (10%)	18/49*** (37%)

Notes: This table reports the aggregate descriptive statistics for output in the 1st work period for each of the four experimental conditions. All statistical tests for the INDIVIDUAL, GROUP, and GROUP COMM conditions are pairwise comparisons in relation to the BASELINE condition. *Mean output* is tested using a Mann-Whitney U-test, *median output* is tested using a K-sample medians test, and the *# of workers with output of 29 or fewer* and *# of workers with output of exactly 28 or 29* are tested using a Fisher's exact test. *, ** and *** indicate statistically different from the baseline statistic at the 10%, 5%, and 1% levels, respectively.

Appendix A – Copy of Experimental Instructions

Participant Instructions
(BASELINE Condition)

Welcome and thank you for participating. Your participation is voluntary and you may leave at any time. The study is expected to take 45 minutes. Please remain quiet during the entire study. If you have any questions, please raise your hand and an Experimenter will come by and answer them privately. All actions during this experiment are to be completed individually, and verbal interaction with other participants is strictly **PROHIBITED.** Thank you for your cooperation.

In this study, you will have an opportunity to earn monetary compensation by assembling TTU Alumni Association mailers. You will have a total of 20 minutes to assemble mailers, and your total earning will depend on how many mailers you are able to assemble in the 20 minutes of allotted time. More detailed information about the mailer assembly task, the procedure and sequencing of the study, and the specific compensation scheme will be proved below.

The Mailer Assembly Task:
In your carrel, you will find: (i) a stack of envelopes on the left side of your carrel with a clear plastic "window", (ii) a stack of tri-folded TTU Alumni Association mailers, (iii) a stack of return envelopes on the right side of your carrel, (iv) an envelope moistener/sealer stick, (v) a tray labeled "completed mailers", and (vi) a compensation record sheet.

To assemble a mailer, you will need to: (step 1) stuff a tri-folded TTU Alumni Association mailer into the envelope with the clear plastic window. **The address on the lower left of the mailer must be facing forward through the clear plastic window of the envelop, so the address is visible through the envelope,** (step 2) stuff in a return envelope behind the tri-folded mailer, (step 3) seal the envelope (by using the moistener/sealer stick), (step 4) stack the completed mailer in the tray labeled "completed mailers". This completes 1 assembled mailer.

As you are assembling mailers throughout the 20 minutes, proceed through the stack of tri-folded mailers in sequence, from the top working your way through the stack. It is imperative that you stuff the mailers in this order, as the post office requires the completed stuffed envelopes be in the same sequence as the tri-folded mailer. To keep the mailers in order, please stack the completed mailers **face down** in the tray. When using the envelope moistener stick, you will need to apply slight pressure to ensure moister is being dispended onto the flap of the envelope.

Please turn your attention to the experimenter for a demonstration of the assembly process.

Procedure and Sequencing of the Study:
You will have **two 10-minute work periods** to assemble the mailers. In each of the 10-minute periods, you are free to assemble as many mailers as you are able to, or choose to. Note, however, that your monetary compensation (described in detail below) will depend on how many total mailers you assemble over both 10-minute periods. The sequencing of the study will be as follows: First, you will assemble mailers during the 1st 10-minute period. Next, an experimenter will come around to your carrel, collect your basket of completed mailers, count how many mailers you assembled, and mark that on the compensation record sheet. During that time, you will be asked to complete a short

questionnaire that will take approximately 8-10 minutes. After you finish the questionnaire, please remain quietly seated in your carrel and wait for the 2^{nd} period to begin. After all participants have finished the questionnaire, and the experimenter has finished counting the 1^{st} period mailers for each participant, you will then begin assembling mailers for the 2^{nd} 10-minute period. After the completion of the 2^{nd} period, an experimenter will again come around and collect your basket of completed mailers, count your assembled mailers, and mark this information on your compensation record sheet. Lastly, an experimenter will privately pay you your total earning in cash and you may leave.

Compensation:
Your earnings in this study will depend on how many total mailers you assemble over both 10-minute work periods.

1^{st} 10-minute period: You will be paid \$.20 in compensation **per** completed mailer you assemble in the 10 minutes of allotted time.

2^{nd} 10-minute period: You will be paid \$.20 in compensation **per** completed mailer you assemble in the 10 minutes of allotted time.

Your earnings in each of the two periods will be added together, and that will be your total compensation for the study.

General Final Remarks:
Throughout the work task, you will be assembling the mailers in private within the confines of your privacy carrel. As a result, the other participants will not be able to observe your progress throughout the work period, or the total number of mailers you assemble. Similarly, the experimenter will not be monitoring your progress throughout the work period, so you are free to work at your own pace and complete as many mailers as you can or choose to do in each work period. During each of the work periods, a timer will be displayed on the video screen so you will be able to keep track of how much time has elapsed in each work period.

At the conclusion of the study, you will be paid your compensation in cash. After you have been paid, you are free to quietly exit the room. As a reminder, there is to be no interaction or communication with any other participants throughout this study.

Participant Instructions
(INDIVIDUAL Condition)

Welcome and thank you for participating. Your participation is voluntary and you may leave at any time. The study is expected to take 45 minutes. Please remain quiet during the entire study. If you have any questions, please raise your hand and an Experimenter will come by and answer them privately. All decisions during this experiment are to be completed individually, and verbal interaction with other participants is strictly **PROHIBITED.** Thank you for your cooperation.

In this study, you will have an opportunity to earn monetary compensation by assembling TTU Alumni Association mailers. You will have a total of 20 minutes to assemble mailers, and your total monetary earning will depend on how many mailers you are able to assemble in the 20 minutes of allotted time. More detailed information about the mailer assembly task, the procedure and sequencing of the study, and the specific compensation scheme will be provided below.

The Mailer Assembly Task:
In your carrel, you will find: (i) a stack of envelopes on the left side of your carrel with a clear plastic "window", (ii) a stack of tri-folded TTU Alumni Association mailers on the right side of your carrel, (iii) a stack of return envelopes on the right side of your carrel, (iv) an envelope moistener/sealer stick, (v) a tray labeled "completed mailers", and (vi) a compensation record sheet.

To assemble a mailer, you will need to: (step 1) stuff a tri-folded TTU Alumni Association mailer into the envelope with the clear plastic window. **The address on the lower left of the mailer must be facing forward through the clear plastic window of the envelop, so the address is visible through the envelope,** (step 2) stuff in a return envelope behind the tri-folded mailer, (step 3) seal the envelope (by using the moistener/sealer stick), (step 4) stack the completed mailer in the tray labeled "completed mailers". This completes 1 assembled mailer.

As you are assembling mailers throughout the 20 minutes, proceed through the stack of tri-folded mailers in sequence, from the top working your way through the stack. It is imperative that you stuff the mailers in this order, as the post office requires the completed stuffed envelopes to be in the same sequence as the tri-folded mailer. To keep the mailers in order, please stack the completed mailers **face down** in the tray. When using the envelope moistener stick, you will need to apply slight pressure to ensure moister is being dispended onto the flap of the envelope.

Please turn your attention to the experimenter for a demonstration of the assembly process.

Procedure and Sequencing of the Study:
You will have **two 10-minute work periods** to assemble the mailers. In each of the 10-minute periods, you are free to assemble as many mailers as you are able to, or choose to. Note, however, that your monetary compensation (described in detail below) will depend on how many total mailers you assemble over both 10-minute periods. The sequencing of the study will be as follows: First, you will assemble mailers during the 1st 10-minute period. Next, an experimenter will come around to your carrel, collect your basket of completed mailers, count how many mailers you assembled, and mark that on the compensation record sheet. During that time, you will be asked to complete a short questionnaire that will take approximately 8-10 minutes. After you finish the questionnaire, please remain quietly seated in your carrel and wait for the 2nd period to begin. After all participants have finished the questionnaire and the experimenter has finished counting the 1st period mailers for each

participant, the experimenter will return the compensation record sheets and you will then begin assembling mailers for the 2nd 10-minute period. After the completion of the 2nd period, an experimenter will again come around and collect your basket of completed mailers, count your assembled mailers, and mark this information on your compensation record sheet. Lastly, an experimenter will privately pay you your total earning in cash and you may leave.

Compensation:
Your earnings in this study will depend on how many mailers you assemble in each of the 10-minute work periods.

1st 10-minute work period: You will receive $.20 in compensation per completed mailer you assemble in the 1st work period.

2nd 10-minute work period: Your compensation rate in the 2nd work period will depend on how many mailers you assemble in the 1st period. There are two possible scenarios for your compensation in the 2nd work period:

Scenario 1 – If you assemble **less than or equal to 29** mailers in the 1st period, then you will continue to receive **$.20 (20 cents)** per mailer you complete in the 2nd period.

Scenario 2 – If you assemble **more than 29** mailers in the 1st period, then your compensation will be reduced to **$.10 (10 cents)** per mailer you complete in the 2nd period.

After the 1st work period, the experimenter will be counting your completed mailers. **Depending on how many mailers you complete, the experimenter will check the appropriate box on your compensation record sheet indicating your per mailer compensation rate for the 2nd work period, based on the criteria above.** Thus, you will know whether your per mailer compensation rate in the 2nd period is 20 cents or 10 cents prior to starting the 2nd period. If you complete the desired number of mailers you want to assemble, you are free to stop working and quietly wait for the work period to end. Your earnings in each of the two periods will be added together, and that will be your total compensation for the study.

General Final Remarks:
Throughout the work task, you will be assembling the mailers in private within the confines of your privacy carrel. As a result, the other participants will not be able to observe your progress throughout the work period or the total number of mailers you assemble. Similarly, the experimenter will not be monitoring your progress throughout the work period, so you are free to work at your own pace and complete as many mailers as you can or choose to do in each work period. During each of the work periods, a timer will be displayed on the video screen so you will be able to keep track of how much time has elapsed in each work period.

At the conclusion of the study, you will be paid your total earnings in cash. After you have been paid, you are free to quietly exit the room. As a reminder, there is to be no interaction or communication with any other participants throughout this study.

Participant Instructions
(GROUP Condition)

Welcome and thank you for participating. Your participation is voluntary and you may leave at any time. The study is expected to take 45 minutes. Please remain quiet during the entire study. If you have any questions, please raise your hand and an Experimenter will come by and answer them privately. All decisions during this experiment are to be completed individually, and verbal interaction with other participants is strictly **PROHIBITED.** Thank you for your cooperation.

In this study, you will have an opportunity to earn monetary compensation by assembling TTU Alumni Association mailers. You will have a total of 20 minutes to assemble mailers, and your total monetary earning will depend on how many mailers you are able to assemble in the 20 minutes of allotted time. More detailed information about the mailer assembly task, the procedure and sequencing of the study, and the specific compensation scheme will be provided below.

The Mailer Assembly Task:
In your carrel, you will find: (i) a stack of envelopes on the left side of your carrel with a clear plastic “window”, (ii) a stack of tri-folded TTU Alumni Association mailers on the right side of your carrel, (iii) a stack of return envelopes on the right side of your carrel, (iv) an envelope moistener/sealer stick, (v) a tray labeled “completed mailers”, and (vi) a compensation record sheet.

To assemble a mailer, you will need to: (step 1) stuff a tri-folded TTU Alumni Association mailer into the envelope with the clear plastic window. **The address on the lower left of the mailer must be facing forward through the clear plastic window of the envelop, so the address is visible through the envelope,** (step 2) stuff in a return envelope behind the tri-folded mailer, (step 3) seal the envelope (by using the moistener/sealer stick), (step 4) stack the completed mailer in the tray labeled “completed mailers”. This completes 1 assembled mailer.

As you are assembling mailers throughout the 20 minutes, proceed through the stack of tri-folded mailers in sequence, from the top working your way through the stack. It is imperative that you stuff the mailers in this order, as the post office requires the completed stuffed envelopes to be in the same sequence as the tri-folded mailer. To keep the mailers in order, please stack the completed mailers **face down** in the tray. When using the envelope moistener stick, you will need to apply slight pressure to ensure moister is being dispended onto the flap of the envelope.

Please turn your attention to the experimenter for a demonstration of the assembly process.

Procedure and Sequencing of the Study:
You will have **two 10-minute work periods** to assemble the mailers. In each of the 10-minute periods, you are free to assemble as many mailers as you are able to, or choose to. Note, however, that your monetary compensation (described in detail below) will depend on how many total mailers you assemble over both 10-minute periods. The sequencing of the study will be as follows: First, you will assemble mailers during the 1st 10-minute period. Next, an experimenter will come around to your carrel, collect your basket of completed mailers, count how many mailers you assembled, and mark that on the compensation record sheet. During that time, you will be asked to complete a short questionnaire that will take approximately 8-10 minutes. After you finish the questionnaire, please remain quietly seated in your carrel and wait for the 2nd period to begin. After all participants have finished the questionnaire and the experimenter has finished counting the 1st period mailers for each

participant, the experimenter will return the compensation record sheets and you will then begin assembling mailers for the 2nd 10-minute period. After the completion of the 2nd period, an experimenter will again come around and collect your basket of completed mailers, count your assembled mailers, and mark this information on your compensation record sheet. Lastly, an experimenter will privately pay you your total earning in cash and you may leave.

Compensation:
Your earnings in this study will depend on how many mailers you assemble over both 10-minute work periods, as well as how many mailers each of the other six participants in the room assemble in the 1st period.

1st 10-minute work period: You will receive $.20 in compensation per completed mailer you assemble in the 1st work period.

2nd 10-minute work period: Your compensation rate in the 2nd work period will depend on how many mailers you assemble in the 1st period, as well as how many mailers each of the other six participants assemble in the 1st period. There are two possible scenarios for compensation in the 2nd work period:

Scenario 1 – If **4 or more participants** in this session assemble **less than or equal 29** mailers in the 1st period, then all seven participants will continue to receive **$.20 (20 cents)** per mailer they complete in the 2nd period.

Scenario 2 – If **4 or more participants** in this session assemble **more than 29** mailers in the 1st period, then the compensation for all seven participants will be reduced to **$.10 (10 cents)** per mailer they complete in the 2nd period.

After the 1st work period, the experimenter will be counting your completed mailers. **Depending on how many mailers you and the other six participants complete, the experimenter will check the appropriate box on your compensation record sheet indicating your per mailer compensation rate for the 2nd work period, based on the criteria above.** Thus, you will know whether your per mailer compensation rate in the 2nd period is 20 cents or 10 cents prior to starting the 2nd period. If you complete the desired number of mailers you want to assemble, you are free to stop working and quietly wait for the work period to end. Your earnings in each of the two periods will be added together, and that will be your total compensation for the study.

General Final Remarks:
Throughout the work task, you will be assembling the mailers in private within the confines of your privacy carrel. As a result, the other participants will not be able to observe your progress throughout the work period or the total number of mailers you assemble. Similarly, the experimenter will not be monitoring your progress throughout the work period, so you are free to work at your own pace and complete as many mailers as you can or choose to do in each work period. During each of the work periods, a timer will be displayed on the video screen so you will be able to keep track of how much time has elapsed in each work period.

At the conclusion of the study, you will be paid your total earnings in cash. After you have been paid, you are free to quietly exit the room. As a reminder, there is to be no interaction or communication with any other participants throughout this study.

Participant Instructions
(GROUP COMM Condition)

Welcome and thank you for participating. Your participation is voluntary and you may leave at any time. The study is expected to take 45 minutes. Please remain quiet during the entire study. If you have any questions, please raise your hand and an Experimenter will come by and answer them privately. All decisions during this experiment are to be completed individually, and verbal interaction with other participants is strictly **PROHIBITED.** Thank you for your cooperation.

In this study, you will have an opportunity to earn monetary compensation by assembling TTU Alumni Association mailers. You will have a total of 20 minutes to assemble mailers, and your total monetary earning will depend on how many mailers you are able to assemble in the 20 minutes of allotted time. More detailed information about the mailer assembly task, the procedure and sequencing of the study, and the specific compensation scheme will be provided below.

The Mailer Assembly Task:
In your carrel, you will find: (i) a stack of envelopes on the left side of your carrel with a clear plastic "window", (ii) a stack of tri-folded TTU Alumni Association mailers on the right side of your carrel, (iii) a stack of return envelopes on the right side of your carrel, (iv) an envelope moistener/sealer stick, (v) a tray labeled "completed mailers", and (vi) a compensation record sheet.

To assemble a mailer, you will need to: (step 1) stuff a tri-folded TTU Alumni Association mailer into the envelope with the clear plastic window. **The address on the lower left of the mailer must be facing forward through the clear plastic window of the envelop, so the address is visible through the envelope,** (step 2) stuff in a return envelope behind the tri-folded mailer, (step 3) seal the envelope (by using the moistener/sealer stick), (step 4) stack the completed mailer in the tray labeled "completed mailers". This completes 1 assembled mailer.

As you are assembling mailers throughout the 20 minutes, proceed through the stack of tri-folded mailers in sequence, from the top working your way through the stack. It is imperative that you stuff the mailers in this order, as the post office requires the completed stuffed envelopes to be in the same sequence as the tri-folded mailer. To keep the mailers in order, please stack the completed mailers **face down** in the tray. When using the envelope moistener stick, you will need to apply slight pressure to ensure moister is being dispended onto the flap of the envelope.

Please turn your attention to the experimenter for a demonstration of the assembly process.

Procedure and Sequencing of the Study:
You will have **two 10-minute work periods** to assemble the mailers. In each of the 10-minute periods, you are free to assemble as many mailers as you are able to, or choose to. Note, however, that your monetary compensation (described in detail below) will depend on how many total mailers you assemble over both 10-minute periods. The sequencing of the study will be as follows: First, you will assemble mailers during the 1st 10-minute period. Next, an experimenter will come around to your carrel, collect your basket of completed mailers, count how many mailers you assembled, and mark that on the compensation record sheet. During that time, you will be asked to complete a short questionnaire that will take approximately 8-10 minutes. After you finish the questionnaire, please remain quietly seated in your carrel and wait for the 2nd period to begin. After all participants have finished the questionnaire and the experimenter has finished counting the 1st period mailers for each

participant, the experimenter will return the compensation record sheets and you will then begin assembling mailers for the 2nd 10-minute period. After the completion of the 2nd period, an experimenter will again come around and collect your basket of completed mailers, count your assembled mailers, and mark this information on your compensation record sheet. Lastly, an experimenter will privately pay you your total earning in cash and you may leave.

Compensation:
Your earnings in this study will depend on how many mailers you assemble over both 10-minute work periods, as well as how many mailers each of the other six participants in the room assemble in the 1st period.

1st 10-minute work period: You will receive $.20 in compensation per completed mailer you assemble in the 1st work period.

2nd 10-minute work period: Your compensation rate in the 2nd work period will depend on how many mailers you assemble in the 1st period, as well as how many mailers each of the other six participants assemble in the 1st period. There are two possible scenarios for compensation in the 2nd work period:

Scenario 1 – If **4 or more participants** in this session assemble **less than or equal 29** mailers in the 1st period, then all seven participants will continue to receive **$.20 (20 cents)** per mailer they complete in the 2nd period.

Scenario 2 – If **4 or more participants** in this session assemble **more than 29** mailers in the 1st period, then the compensation for all seven participants will be reduced to **$.10 (10 cents)** per mailer they complete in the 2nd period.

After the 1st work period, the experimenter will be counting your completed mailers. **Depending on how many mailers you and the other six participants complete, the experimenter will check the appropriate box on your compensation record sheet indicating your per mailer compensation rate for the 2nd work period, based on the criteria above.** Thus, you will know whether your per mailer compensation rate in the 2nd period is 20 cents or 10 cents prior to starting the 2nd period. If you complete the desired number of mailers you want to assemble, you are free to stop working and quietly wait for the work period to end. Your earnings in each of the two periods will be added together, and that will be your total compensation for the study.

General Final Remarks:
Throughout the work task, you will be assembling the mailers in private within the confines of your privacy carrel. As a result, the other participants will not be able to observe your progress throughout the work period or the total number of mailers you assemble. Similarly, the experimenter will not be monitoring your progress throughout the work period, so you are free to work at your own pace and complete as many mailers as you can or choose to do in each work period. During each of the work periods, a timer will be displayed on the video screen so you will be able to keep track of how much time has elapsed in each work period.

At the conclusion of the study, you will be paid your total earnings in cash. After you have been paid, you are free to quietly exit the room. As a reminder, there is to be no interaction or communication with any other participants throughout this study.

Group Discussion Period:
Prior to the start of the 1st work period, the seven of you will have an opportunity to discuss the work task as a group. The group will be given 3 minutes for this discussion period. During these 3 minutes you are free to discuss anything related to this study and the associated mailer assembly task. This will be an open discussion amongst the group so we kindly ask that you be courteous and respectful of your fellow group members during the discussion. Before the discussion begins, I will ask you to all stand up and introduce yourself to the group. After that, the experimenter will leave the room, and you will have 3 uninterrupted minutes from that point for the group discussion. Also note, your discussion will remain private within the group and will not be recorded in any way. After the discussion period is up, the experimenter will return and you will be asked to quietly sit back down in your carrel. From that point forward, there is to be no more communication or interaction with other group members for the remainder of the study.

Participant Instructions
(BASELINE LOW Condition)

Welcome and thank you for participating. Your participation is voluntary and you may leave at any time. The study is expected to take 45 minutes. Please remain quiet during the entire study. If you have any questions, please raise your hand and an Experimenter will come by and answer them privately. All actions during this experiment are to be completed individually, and verbal interaction with other participants is strictly **PROHIBITED.** Thank you for your cooperation.

In this study, you will have an opportunity to earn monetary compensation by assembling TTU Alumni Association mailers. You will have a total of 20 minutes to assemble mailers, and your total earning will depend on how many mailers you are able to assemble in the 20 minutes of allotted time. More detailed information about the mailer assembly task, the procedure and sequencing of the study, and the specific compensation scheme will be proved below.

The Mailer Assembly Task:
In your carrel, you will find: (i) a stack of envelopes on the left side of your carrel with a clear plastic "window", (ii) a stack of tri-folded TTU Alumni Association mailers, (iii) a stack of return envelopes on the right side of your carrel, (iv) an envelope moistener/sealer stick, (v) a tray labeled "completed mailers", and (vi) a compensation record sheet.

To assemble a mailer, you will need to: (step 1) stuff a tri-folded TTU Alumni Association mailer into the envelope with the clear plastic window. **The address on the lower left of the mailer must be facing forward through the clear plastic window of the envelop, so the address is visible through the envelope,** (step 2) stuff in a return envelope behind the tri-folded mailer, (step 3) seal the envelope (by using the moistener/sealer stick), (step 4) stack the completed mailer in the tray labeled "completed mailers". This completes 1 assembled mailer.

As you are assembling mailers throughout the 20 minutes, proceed through the stack of tri-folded mailers in sequence, from the top working your way through the stack. It is imperative that you stuff the mailers in this order, as the post office requires the completed stuffed envelopes be in the same sequence as the tri-folded mailer. To keep the mailers in order, please stack the completed mailers **face down** in the tray. When using the envelope moistener stick, you will need to apply slight pressure to ensure moister is being dispended onto the flap of the envelope.

Please turn your attention to the experimenter for a demonstration of the assembly process.

Procedure and Sequencing of the Study:
You will have **two 10-minute work periods** to assemble the mailers. In each of the 10-minute periods, you are free to assemble as many mailers as you are able to, or choose to. Note, however, that your monetary compensation (described in detail below) will depend on how many total mailers you assemble over both 10-minute periods. The sequencing of the study will be as follows: First, you will assemble mailers during the 1st 10-minute period. Next, an experimenter will come around to your carrel, collect your basket of completed mailers, count how many mailers you assembled, and mark that on the compensation record sheet. During that time, you will be asked to complete a short questionnaire that will take approximately 8-10 minutes. After you finish the questionnaire, please remain quietly seated in your carrel and wait for the 2nd period to begin. After all participants have finished the questionnaire, and the experimenter has finished counting the 1st period mailers for each

participant, you will then begin assembling mailers for the 2nd 10-minute period. After the completion of the 2nd period, an experimenter will again come around and collect your basket of completed mailers, count your assembled mailers, and mark this information on your compensation record sheet. Lastly, an experimenter will privately pay you your total earning in cash and you may leave.

Compensation:
Your earnings in this study will depend on how many total mailers you assemble over both 10-minute work periods.

1st 10-minute period: You will be paid $.20 (20 Cents) in compensation **per** completed mailer you assemble in the 10 minutes of allotted time.

2nd 10-minute period: You will be paid $.10 (10 Cents) in compensation **per** completed mailer you assemble in the 10 minutes of allotted time.

Your earnings in each of the two periods will be added together, and that will be your total compensation for the study.

General Final Remarks:
Throughout the work task, you will be assembling the mailers in private within the confines of your privacy carrel. As a result, the other participants will not be able to observe your progress throughout the work period, or the total number of mailers you assemble. Similarly, the experimenter will not be monitoring your progress throughout the work period, so you are free to work at your own pace and complete as many mailers as you can or choose to do in each work period. During each of the work periods, a timer will be displayed on the video screen so you will be able to keep track of how much time has elapsed in each work period.

At the conclusion of the study, you will be paid your compensation in cash. After you have been paid, you are free to quietly exit the room. As a reminder, there is to be no interaction or communication with any other participants throughout this study.

Appendix B – Mailer Task and Workplace Environment

Sample of Mailer Assembly Task

Simulated Workplace Environment