

IZA DP No. 9981

**Testing for the Ratchet Effect:
Evidence from a Real-Effort Work Task**

Eric Cardella
Briggs Depew

June 2016

Testing for the Ratchet Effect: Evidence from a Real-Effort Work Task

Eric Cardella

Texas Tech University

Briggs Depew

*Louisiana State University
and IZA*

Discussion Paper No. 9981

June 2016

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Testing for the Ratchet Effect: Evidence from a Real-Effort Work Task*

The “ratchet effect” refers to a phenomenon where workers whose compensation is based on productivity strategically restrict their output, relative to their capability, because they rationally anticipate that high levels of output 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 actually empirically identify 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 directly test for the presence of the ratchet effect using two different methods for evaluating 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 of the ratchet effect when productivity is evaluated at the individual-level. However, we find very little evidence of the ratchet effect 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. Furthermore, we find the ratchet effect re-emerges if workers are able to communicate. Our experimental design, combined with using a real-effort work task, also allows us to shed light on an important dynamic implication of the ratchet effect that has not yet been examined in the literature – the role of the ratchet effect on future productivity via learning-by-doing.

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

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

Corresponding author:

Briggs Depew
Department of Economics
Louisiana State University
Baton Rouge, LA70803-6306
USA
E-mail: bdepew@lsu.edu

* We thank Luke Boosey, Danny Brent, 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 resources provided.

“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 in the US 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, one potential drawback with performance-pay is that workers may have an incentive to “shirk” by strategically restricting the amount of output they produce. The reason is that workers 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 exert higher levels of effort in the future in return 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 empirically test for the presence of the ratchet effect in a simulated work environment. Specifically, we examine whether workers strategically restrict output production, relative to their true capability, when they are faced with the consequence of reduced pay in the future if they are *too* productive. We develop a novel experimental design where participant workers complete a *real-effort* work task under a piece-rate incentive scheme over two work periods. 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. We test for the presence of the ratchet effect using two different criteria for evaluating productivity: (i) when productivity is measured based on individual

¹ For example, Lemieux et al. (2009) show that the overall proportion of performance-pay jobs 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).

output, and (ii) when productivity is measured, collectively, based on group-level output. Furthermore, by using a real-effort work task rather than chosen effort, we are able to evaluate an important possible dynamic implication of the ratchet effect, namely, the extent to which strategic output restriction impacts 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 to management, which would then induce management to set a more demanding output schedule (or a less favorable compensation scheme) in the future. The emergence of pooling equilibria in these dynamic principal-agent models provides the theoretical foundation for the ratchet effect.

The theoretical implications of the ratchet effect are consistent with substantial anecdotal evidence suggesting the presence of the ratchet effect in real workplaces. In particular, the works by Mathewson (1931), Lawler (1971), Edwards (1979), Montgomery (1979), and Clawson (1980),

³ Besides effort provision and output production in the workplace, other contexts that have been explored where the ratchet effect phenomenon could arise 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 have an incentives to 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).

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 very 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 worker 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 experiment featuring a simulated work environment and a real-effort work 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. That said, we contend our experimental environment provides a suitable platform for empirically testing for the presence of the ratchet effect in the workplace.

To our knowledge, there are four experimental papers that directly test for the ratchet effect: Chaudhuri (1998), Cooper et al. (1999), Charness et al. (2011), and Bellemare & Shearer (2015).⁵

⁵ We are also aware of two empirical papers aimed at indirectly identifying the ratchet effect. Specifically, Allen & Lueck (1999) use contract data (type of contract or terms of contract) between landowners and tenant agriculture workers to analyze several predictions based on implication of the ratchet effect. In their cross-section analysis, they find little evidence that supports their prediction and, hence, conclude that their data reveals little evidence of the

The first three studies consider relatively stylized experimental designs with two work periods, two types of workers (low or high ability), and *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.⁶ The experimental design of Cooper et al. (1999) builds from some early literature of the ratchet effect within centrally planned 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* effectively 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).

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.

⁶ 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 fact that the principle will increase 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.

We develop an experimental design using a simulated work environment where participant workers complete a real-effort work task under a piece-rate compensation for two work periods. Importantly, our 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 two periods under a fixed piece-rate pay 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 test for the ratchet effect (in the aggregate) by investigating if workers strategically 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 presence 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. 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 presence of 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. Furthermore, communication among workers regarding output restriction seems plausible, in practice, and is alluded to anecdotally, as evident by Clawson (1980, p. 177): “in order to enforce output quotas it was definitely necessary for some workers to pressure and coerce others.”

We find strong evidence of the ratchet effect when productivity is evaluated at the individual-level; namely, we observe significant output restriction by participant workers in the 1st work period in order to avoid facing a reduced piece-rate in the 2nd period. However, we find very little evidence of the ratchet effect when productivity is evaluated collectively at the group-level and workers are not allowed to communicate; workers appear to be attempting to free-ride off the output restriction of others in the group, which essentially results in full output production

aggregately across participant workers. In contrast, when we allow pre-play communication among the group of workers, the ratchet effect reemerges, suggesting that communication facilitates coordination of output restriction among workers.

We contribute to the literature on the ratchet effect along a number of important dimensions. First, we find evidence of output restriction among workers under the credible threat of reduced future piece-rates using a simulated workplace environment and real-effort work task; thus we add to the very limited extant literature documenting direct empirical evidence of the ratchet effect (Cooper et al., 1999; and Charness et al., 2011). By using a real-effort work task embedded within the controlled setting of a lab experiment, combined with incorporating a substantial degree of “field” context in the experimental protocol, our study helps bridge the gap between the results documented in prior lab experiments and the anecdotal accounts of the ratchet from real workplaces. Second, we empirically examine the ratchet effect when productivity is evaluated collectively at the group-level (with and without group communication), rather than at the individual-level, 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, this method is more representative of how management evaluates the productivity of workers. Third, by using a real-effort work task, we are able to analyze a possible dynamic implication of the ratchet effect and, importantly, we 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.⁷

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. Experimental sessions were conducted at the Rawls College of Business at Texas Tech University. Participants were recruited from a college maintained database to participate in an economics research study about productivity. At the time of invitation, participants were told that participation in the study would involve their working on a simple task for which they would earn monetary compensation based on their productivity. In

⁷ In lieu of attempting to cite all such 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.

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 4); 53% of the participants were female, and the average age was 21.6 years with a minimum age of 18 years and maximum of 44 years. We used a between-subjects design where each participant worker took part in one session of a given experimental condition. The average session length was approximately 45 minutes, and the average earnings were \$14 USD. A copy of the experiment instructions for all experimental conditions 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.⁸ There were three components of a mailer: the mailer, a return envelope, and a mailing envelope, which are depicted in Figure 1. 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); (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.

The mailer assembly 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.⁹ Second, the mailer task is not analytically intensive (e.g., an anagram task, word unscrambling task, puzzle-solving task, or multiplication task), which implies that output is a strictly increasing function of effort. Third, there is essentially no quality dimension associated with this task, which is important because it allows us to circumvent any possible tradeoffs between the quality of work and quantity of output that could arise if workers are restricting output.¹⁰ Taken together, this mailer task

⁸ 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).

⁹ In fact, during all sessions, an experimenter observed that each participant worker was correctly assembling the mailers as instructed and demonstrated by the experimenter. There were no instances where a worker was not correctly assembling mailers. Similarly, there was never any indication throughout the study by the TTAA that any of the mailers were being assembled in an unsatisfactory manner.

¹⁰ 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)

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 work task, the partnership with the TTAA, the legitimacy of the mailer task, and the simulated workplace environment 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 field context imbedded in the study, increases the external validity of our results (Friedman & Sunder, 1994; Falk & Fehr, 2003; Charness & Kuhn, 2011; Gill & Prowse, 2015).

All experimental sessions were conducted in a conference room that was set up to resemble a simulated mailer assembly workplace. In the room, seven workstations were set up, each of which was separated by privacy carrels. Each carrel was equipped with all the necessary materials for assembling mailers. A picture of the workplace environment is presented in Figure 2. Upon arrival for a session, each participant worker was assigned to a workstation. An experimenter read the experimental instructions aloud and provided a visual demonstration of how to properly assemble a mailer. All participant workers were informed that mailers were part of a TTAA campaign and that the mailers would actually be mailed.¹¹

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

and Bellemare et al. (2010). In the context of the ratchet effect, 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, and not possibly present in the workplace. Rather, in this paper we focus specifically on strategic restriction of output, and thus, we want to isolate output quantity and abstract away from the quality dimension of the task. Investigating if workers produce higher-quality work when they are producing less output, relative to their capability, is an interesting topic for future research but beyond the scope of the current paper.

¹¹ It is plausible that because the work task involved stuffing mailers for a charitable organization – the TTAA – participants may have been compelled to put in more effort and work harder, especially since students at Texas Tech University are likely in support of 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 certainly possible that the charitable nature of the mailer task may have impacted effort levels compared to a more abstract or neutral 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 treatments remains intact. Moreover, the fact that the charitable nature of the work 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, which would strengthen results providing evidence of the ratchet effect.

(described in detail below), was clearly stated to the participant workers in the instructions. After completing the 1st 10-minute work 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 piece-rate for the 2nd period. Participant workers then assembled mailers during the 2nd 10-minute work period. After the 2nd period, the experimenter again 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.

In both work periods, participant workers assembled mailers at their individual workstations within the confines of their privacy carrel. Thus, participant workers were unable to observe the progress of the other workers or the output level of other workers. In addition, participant workers were informed in the instructions that the experimenter would not be 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.”¹²

2.2 Measuring True Output Capability

Since the ratchet effect entails workers strategically restricting their output, relative to their capability, it is necessary to know the distribution of the true output capability among the workers in order to empirically test for the presence of the ratchet effect. To identify this distribution of the true output capability in our participant worker sample, we conducted an initial BASELINE condition using a fixed piece-rate scheme. 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.¹³ Importantly, the 2nd period piece-rate *did not* depend on 1st period output; hence, there was no strategic reason for participant workers to restrict their output and under-produce in the 1st work period. As a result, we maintain that the observed distribution of output in the 1st work period in

¹² An experimenter did make one pass through the room after about 1 minute into the 1st work period to ensure that each participant worker was not having any issues assembling mailers correctly. After this initial cursory pass through the room, the experimenter did not walk through the room during any of the remaining work time.

¹³ 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.

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 assess the true output capability was similarly used by Kube et al. (2013). The BASELINE condition consisted of 42 participant workers (7 experimental sessions).

Before introducing the main experimental conditions used to test for the ratchet effect, it is pedagogical to first present the aggregate data on the output of 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 3 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 3, is that there is substantial variation in output levels of our participant workers. This variation is important for several reasons. First, it establishes the presence of heterogeneity of different ability “types” of workers in our sample, which is necessary to test for the presence of the ratchet effect (i.e., if the “high” ability types are restricting their output). Second, having a range of ability types 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. When output between the 1st and 2nd work periods was compared, our participant workers exhibited a significant increase in productivity. We suspect that this is likely attributed to learning-by-doing, wherein the participant workers become more efficient at completing the task.¹⁴ 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 4 we explore the dynamic

¹⁴ DellaVigna et al. (2016) similarly find significant increases in productivity in a mailer assembly task over multiple work periods, and the authors attribute this increase in productivity to learning-by-doing.

implications of the ratchet effect on productivity by analyzing the 2nd period and decomposing differences across conditions into a learning-by-doing effect and a reduced-effort effect.

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 the ratchet effect via strategic output restriction by workers. Moreover, the participant workers in the BASELINE condition were randomly drawn from the same database as the participant workers used in the subsequent experimental conditions; 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 other experimental conditions.

2.3 Experimental Conditions to Test for the Ratchet Effect

We implement three main 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, the 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.

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**, that is determined from the 1st period distribution of output in the BASELINE condition, which is made known to participant workers. In order for our investigation of the ratchet effect to be salient, the following two conditions are necessary when choosing a value of **T**: (i) **T** be set *low enough* such that this threshold 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 strategic output restriction, and (ii) **T** be set *high enough* such that the high ability workers have an incentive to restrict their 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 assembling at full capability of 54 in the 1st period, thus satisfying condition (ii).¹⁵

While the implementation of this ex-ante productivity threshold is a stylized component of the experimental design, it is important for two reasons. First, it ensures common knowledge among the participant workers of the potential for piece-rates to be reduced and ensures that workers *rationally* anticipate the piece-rate reduction if they are too productive in the first period, which is a necessary condition for the ratchet effect to emerge. Second, it eliminates any ambiguity for workers regarding how much output would be deemed as too productive, which enables us to more clearly identify strategic output restriction in relation to the productivity threshold. We further note that, anecdotally, there is ample evidence suggesting that workers 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 for the given job, and then are able to deduce the maximum level of output that would generate that equivalent day-rate, conditional on the piece-rate. Thus, setting 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 workplaces implementing piece-rate payment schemes.

The first main ratchet effect condition we consider, 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 2nd period if, individually, they are *too* productive in the 1st period? In the INDIVIDUAL condition, the 2nd period piece-rate for participant workers depends on the worker’s 1st period output level. Prior to starting work in the 1st period, participant

¹⁵ To see this note that $29*(\$0.20) + 63*(\$0.20) = \$18.40 > 54*(\$0.20) + 63*(\$0.10) = \17.10 . Also, this calculation of the monetary incentive to restrict output in the first period does not account for the additional cost of effort association with producing the maximum output in the first period compared to restricting output. We acknowledge that setting $\mathbf{T} = 29$ is somewhat arbitrary; however, we see little reason to think that our results are specific to $\mathbf{T} = 29$ and would not generalize to other values of \mathbf{T} that satisfy our two necessary conditions.

workers are informed that if their 1st period output level exceeds $T = 29$ mailers, then their 2nd period piece-rate will be *reduced* in half to \$.10. In short, 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 2nd work period if they are too productive in the 1st period.

The second ratchet effect 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 2nd period if the group of workers is, collectively, *too* productive in the 1st period? All GROUP condition sessions consisted of a group of 7 workers. In the GROUP condition, the piece-rate in the 2nd work period for participant workers depends on the 1st period output levels of all 7 workers in the group. Specifically, workers are informed, prior to starting work in the 1st period, that if 4 or more of the 7 workers in the group produce more than $T = 29$ mailers in the 1st period, then the 2nd 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 2nd work period if too many workers in the group are *collectively* too productive in the 1st period.

The final ratchet effect 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 the discussion was private.¹⁶ After the 3-minute discussion period ended, the experimenter re-entered the room, and the remainder of the experimental session proceeded in the same way as the GROUP condition. Thus, the only difference between the GROUP and GROUP

¹⁶ Face-to-face communication is a strong form of communication among participant workers compared to say anonymous chat. That said, we chose to implement face-to-face communication 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 actually communicate with each other in the workplace. We did not record the group discussion (and informed participants of this) because we wanted participants to feel comfortable discussing output restriction and possible cooperation or collusion within the confines of private discussion, as would be true of real workplace discussions among workers.

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.

For these additional three conditions, our sample consists of 45 participant workers in the INDIVIDUAL condition (7 sessions), 42 participant workers in the GROUP condition (6 sessions), and 49 participant workers 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.

2.4 Behavioral Hypotheses

Our first primary research question is whether we see the emergence of the ratchet under a piece-rate pay scheme when productivity is evaluated at the individual-level. Our INDIVIDUAL condition allows us to test for the existence of the ratchet effect when productivity is evaluated at the individual-level. We assume that the distribution of true output capability of the participant workers in the INDIVIDUAL condition is consistent with the observed distribution from the BASELINE condition. Given this assumption, the following requisite conditions for the ratchet effect to emerge are present in the INDIVIDUAL condition: (i) there is heterogeneity in worker capability, (ii) the 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 (iii) 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 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 primary research question is whether we see the emergence of the ratchet effect when productivity is evaluated collectively at the group-level. The motivation for investigating the ratchet effect when productivity is measured at the group-level is twofold. First, from a practical standpoint, 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 measures 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.¹⁷

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). If piece-rates in the future are reduced only if the group is collectively *too* productive in the current period, then it is 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 and the ratchet effect, there is an extensive body of 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.¹⁸ 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;

¹⁷ 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.

¹⁸ We refer readers to 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 & Gächter (2010) and the survey by Chaudhuri (2011) for thorough reviews of the more recent experimental literature on public goods games.

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 are too productive in the 1st period, namely, if 4 or more of the group of 7 participant workers produce 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 and free-ride off the workers who restrict their output in the 1st period at or below 29 mailers. However, if all of the high ability workers in the group attempt to free-ride, then this will induce full effort provision and maximum 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 level 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.*

Our third primary 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.¹⁹

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 by 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 that 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 expect to observe output restriction and the emergence of the ratchet effect in the GROUP COMM condition, leading to the following testable hypothesis:²⁰

¹⁹ While identifying the exact mechanism by which worker communication could possibly facilitate coordination of output restriction is beyond the scope of our study, possible mechanisms 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.”

²⁰ 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 an upper bound. That said, we expect the ratchet effect would be less likely to emerge if group

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.*

3 Results

3.1 Testing for the Ratchet Effect in the INDIVIDUAL Condition

We first analyze the data to test for the ratchet effect in the INDIVIDUAL condition. Namely, we test whether participant workers in the INDIVIDUAL condition strategically restrict their output below the productivity threshold, $T = 29$. Table 2 presents the aggregate 1st period output statistics for the INDIVIDUAL condition and compares them with the BASELINE condition.

From Table 2, we see that for the 45 participant workers in the INDIVIDUAL condition, the average output in the 1st period was 29.4 mailers, which is significantly lower than the BASELINE average of 34.6 (Mann-Whitney test: $p < .001$). Similarly, the median 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: $p < .001$). 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: $p = .003$).

To further test for 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 in the INDIVIDUAL condition, 32/45 (71%) participant workers completed 29 or fewer mailers, compared to 10/42 (24%) in the BASELINE condition, which is strongly significantly different (Fisher's exact test: $p < .001$). More precisely, 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

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.

BASELINE condition, which is strongly significantly different (Fisher's exact test: $p < .001$).²¹ In terms of the distribution of output, Figure 4 presents the CDFs of 1st period output for the INDIVIDUAL and BASELINE conditions. The comparison of the CDFs confirms a shift in the distribution of output in the INDIVIDUAL condition from 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: $p < .001$).²²

Taken together, the data strongly supports H1. Specifically, in the INDIVIDUAL condition, participant workers (in the aggregate) produced 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 empirical finding is summarized in Result 1:

***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.*

3.2 Testing for the Ratchet Effect in the GROUP Condition

Next, we test for the ratchet effect in the GROUP condition; namely, do participant workers restrict their output at or below $T = 29$ in the 1st work period? Recall, in the GROUP condition, productivity is evaluated based on the collective output of the group of 7 participant workers, and the piece-rate is reduced to \$.10 in the 2nd period if 4 or more of the 7 workers produce more than $T = 29$ in the 1st period. The output statistics for the GROUP condition are presented in Table 2 and can be compared to the BASELINE condition.

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

²¹ 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: $p = .059$).

²² 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.

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). In terms of the session-level analysis, 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, recall that 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: $p = .282$).

Again, we can further 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, Figure 5 compares the CDFs of the 1st period output levels for the GROUP and BASELINE conditions. The 1st period output distribution in the GROUP is nearly identical to that of the BASELINE condition and is not statistically different (Epps-Singleton test: $p = .913$).

Overall, the data supports H2. Namely, 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 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.*

3.3 Testing for the Ratchet Effect in the GROUP COMM Condition

The last part of our main analysis is testing for the ratchet effect in the GROUP COMM condition. Recall that the GROUP COMM condition is identical to the GROUP condition with the exception that all 7 worker participants in the GROUP COMM condition were given 3 minutes to discuss the work task prior to commencing work. Table 2 presents the aggregate 1st period output statistics for the GROUP COMM condition and compares them to the BASELINE condition.

From Table 2, we can see that in the GROUP COMM condition the average output in the 1st period was 31.7 mailers, which is significantly lower than the BASELINE average of 34.6 (Mann-Whitney test: $p = .093$). Similarly, the median output in the GROUP COMM condition was 29 mailers, which is significantly lower than the median of 34 in the BASELINE condition (K-sample medians test: $p = .002$). Looking at the session-level data, for the GROUP COMM condition, the session-level average 1st period output levels were: 34.7, 34.7, 34.4, 30.7, 30.0, 29.6, and 27.6; recall that the BASELINE condition session-level averages were: 38.0, 36.4, 35.5, 34.6, 34.3, 33.8, and 30.5; comparing these session-level averages, the GROUP COMM condition borders on being significantly different than the BASELINE condition (Mann-Whitney test: $p = .109$).²³ However, if we instead compare the GROUP COMM and BASELINE conditions using session-level median output levels, the difference is statistically significant (Mann-Whitney test: $p = .029$).²⁴

Next we consider the proportion of workers producing an output level at or below the threshold $T = 29$. In the GROUP COMM condition, 26/49 (53%) participant workers completed 29 or fewer mailers compared to 10/42 (24%) in the BASELINE condition, which is strongly significantly different (Fisher's exact test: $p = .005$). Just considering the proportion of participant workers completing 28 or 29 mailers, 18/49 (37%) workers completed 28 or 29 mailers in the GROUP COMM condition compared to 1/42 (2%) in the BASELINE condition, which is also strongly significantly different (Fisher's exact test: $p < .001$). In terms of the entire distribution, Figure 6 presents the CDFs of 1st period output levels for the GROUP COMM and BASELINE conditions. The comparison of the CDFs confirms a shift in the distribution of output in the GROUP COMM condition toward 27-29 mailers; the GROUP COMM and BASELINE distributions are significantly different (Epps-Singleton test: $p = .046$).

Taken together, the data largely supports H3. Namely, in the GROUP COMM condition, participant workers (in the aggregate) produce significantly less output in the 1st period compared to the BASELINE condition; additionally, a significantly larger proportion of workers in the

²³ If we test for differences in these session-level averages using a more powerful Fisher-Pitman permutation test, the difference is significant ($p = .056$).

²⁴ 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 certainly independent), we still find significantly lower 1st period output in the GROUP COMM condition.

GROUP COMM condition complete less than or equal to $T = 29$ mailers in the 1st period compared to the BASELINE condition. This empirical finding 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.*

4 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. Specifically, we are interested in the extent to which current output restriction may reduce future productivity.

We hypothesize that a plausible channel through which output restriction can impact future productivity is through 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.²⁵ 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 given work task, which can reduce learning and future productivity. Relatedly, if workers are restricting their output in order to avoid exceeding some implicit productivity threshold, 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. The fact 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) in the BASELINE condition suggests that workers are potentially learning through experience and becoming more innovative in the production methods.²⁶

²⁵ 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.

²⁶ Although the task seems simple in nature, casual observation of workers by the experimenters (at the time of collecting and counting the completed mailers after each work period) revealed some degree of production innovation

To broadly explore if, and to what extent, strategic output restriction impacts subsequent productivity, we compare 2nd period output of participant workers in the INDIVIDUAL condition with that in the BASELINE condition. We are aware that in addition to a potential learning-by-doing effect, some of the workers in the INDIVIDUAL condition had a lower piece-rate in the 2nd period. In section 4.2 we turn our attention to identifying the learning-by-doing effect, separate from the possible decreased effort effect associated with the lower piece-rate. In addition, for this analysis we focus specifically on the INDIVIDUAL condition and not 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 of 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, which were not present in either the BASELINE or INDIVIDUAL conditions. 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.*

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

Figure 7 displays the average 2nd period output levels for the BASELINE and INDIVIDUAL conditions as well as the respective CDFs of 2nd period output. From the left panel of Figure 7, 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 in output), which is significantly different (Mann-Whitney test: $p = .016$). Similarly, the median 2nd period output is 38 in the INDIVIDUAL condition and 44.5 in the BASELINE, which is also significantly different (Medians test: $p = .071$). The CDFs in the right panel of Figure 7 show a clear shift in the distribution toward lower 2nd period output levels in the INDIVIDUAL condition compared to the

as many workers appeared to have adopted an assembly line approach and/or re-configured their workspace to enhance efficiency. Moreover, DellaVigna et al. (2016) similarly document evidence of increases in productivity in a mailer assembly task resulting from learning-by-doing.

BASELINE (Epps-Singleton test: $p = .115$).²⁷ The data suggests 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.*

4.2 Identifying the Effect of Reduced Learning-by-Doing

It is important to note that 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 a lower wage. Recall that the compensation scheme in the INDIVIDUAL condition pays participant workers either \$.10 or \$.20 in the 2nd period (depending on 1st period productivity), whereas all participant workers in the BASELINE condition earned a piece-rate or \$.20 in the 2nd period. As such, there is the possibility that the reduced 2nd period productivity observed 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. All else equal, workers may exert lower effort when receiving \$.10 per mailer rather than \$.20 per mailer.²⁸

To rule out the possibility that the significant decrease in 2nd period productivity in the INDIVIDUAL condition is being *entirely* driven by the possible wage effect, we conducted a follow-up experimental condition, denoted as BASELINE LOW, which is an augmented version of the BASELINE condition.²⁹ Similar to the BASELINE condition, in the BASELINE LOW

²⁷ If these two distributions are compared using a KS-test, then there is a significant difference (KS-test: $p = .014$).

²⁸ Such a claim would be consistent with recent findings documented by DellaVigna et al. (2016) where worker productivity on a mailer assembly task decreased, on average, by about 12% when piece-rates were reduced from 20 cents to 10 cents. Relatedly, Paarsch & Shearer (2009) and Bellemare & Shearer (2011) document significant decreases in worker output in tree planting resulting from lower piece-rates. Although, Carpenter & Gong (2016) consider a mailer assembly task, and they do not document any significant difference in worker productivity when the piece-rate is \$1.00 compared to \$.50.

²⁹ One might be inclined to think that it is possible to isolate the learning-by-doing effect by simply analyzing the 2nd period productivity of those participant workers in the INDIVIDUAL condition that restricted output, i.e., those who produced an output level of 28 or 29 in the 1st period. However, this will not yield a reliable estimate for two important reasons. First, there is a potential selection bias resulting from the participant workers who chose to restrict their output in the INDIVIDUAL condition. Second, because these participant workers restricted their output, we do not know their true capability; thus, we cannot identify the relevant comparison group of participant workers from the BASELINE condition for comparison of 2nd period productivity.

condition all participant workers receive a piece-rate of \$.20 in the 1st period; unlike in the BASELINE condition, in the BASELINE LOW condition all participant workers receive a piece-rate of \$.10 in the 2nd work period. Thus, 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. As suggested, the observed difference between the 2nd period output in the BASELINE and INDIVIDUAL conditions is an estimate of the potential joint learning-by-doing effect and wage effect. Importantly, the difference between the 2nd period output in the BASELINE and BASELINE LOW conditions eliminates the learning-by-doing channel and, thus, isolates the wage effect.

As a precursor to discussing the results from the BASELINE LOW condition, it is worth emphasizing that in the BASELINE LOW condition 100% of workers received the reduced piece-rate of \$.10 per mailer in the 2nd period, while in the INDIVIDUAL condition only 28.9% of workers received the reduced piece-rate of \$.10 in the 2nd period. Therefore, any potential wage effect in the INDIVIDUAL condition is coming from a (potentially) non-random subsample of less than one-third of the distribution. That said, if there is no learning-by-doing effect and the reduction in 2nd period output in the INDIVIDUAL condition is entirely being driven by the wage effect, then we should observe significantly lower 2nd period productivity in the BASELINE LOW condition compared to the INDIVIDUAL condition. Moreover, this is highly unlikely given that average output decreased by 4.8 mailers in the 2nd period in the INDIVIDUAL condition relative to the BASELINE condition; this would imply that the local average treatment effect of the wage decrease would have to be 16.6 ($4.8/.289 = 16.6$) mailers to justify a null learning-by-doing effect.

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

We ran 8 additional experimental sessions of the BASELINE LOW condition with 51 total participants. Prior to presenting the results on 2nd period output, it is necessary that we first compare 1st period output in BASELINE LOW to BASELINE; notably, since BASELINE LOW sessions were not part of the original research design and were conducted after the original sessions, we need to ensure that participant workers in the BASELINE LOW condition represent a similar sample of workers.³⁰ Recall, the mean and median 1st period output levels in the

³⁰ 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.

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; neither the mean nor median are significantly different (Mann-Whitney test: $p = .874$; K-sample medians test: $p = .941$). Similarly, a test for equality of distributions suggests that the 1st period output is essentially identical (Epps-Singleton test: $p = .798$). The striking similarities in the distribution of 1st period output between BASELINE and BASELINE LOW suggest that there is likely to be no confounding effects from differences in worker ability when comparing 2nd period output between BASELINE LOW and the previously run sessions.

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. Correspondingly, the mean and median 2nd period output in the BASELINE condition were 44.6 and 44.5, respectively. 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 different (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, the potential wage effect is minimal, and as a result, the observed average decrease of 4.8 mailers in the 2nd period in the INDIVIDUAL condition, relative to the BASELINE condition, is largely due to reduced learning-by-doing from output restriction.

Under a weak assumption that the wage effect is monotonic, we can obtain a lower bound estimate of the magnitude of the average 2nd period decrease in productivity in the INDIVIDUAL CONDITION, relative to the BASELINE condition, that is caused by reduced learning-by-doing. This is accomplished by comparing the 2nd period productivity in the INDIVIDUAL condition to the 2nd period productivity in 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 29% of the workers received \$.10 per mailer in the 2nd period; as such, a comparison of the average 2nd period productivity 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 participants received the reduced \$.10 piece-rate in the 2nd period. As a result, if average 2nd period output in the INDIVIDUAL condition is *lower* than in the BASELINE LOW condition, then the magnitude of

the difference represents a lower bound of the productivity decrease resulting from reduced learning-by-doing.

Looking at the 2nd period data, the mean and median output levels in the BASELINE LOW condition were 43.6 and 44, respectively. Recall that in the INDIVIDUAL condition, the mean and median 2nd period output levels were 39.8 and 38, respectively. Thus, workers produced 3.8 fewer mailers (an approximately 9% reduction in productivity), on average, during the 2nd period in the INDIVIDUAL condition than in the BASELINE LOW condition, which is significant (Mann-Whitney test: $p = .019$). Also, a test for equality of distribution of 2nd period output can be rejected (Epps-Singleton test: $p = .030$). The fact that we observe significantly lower 2nd period production in the INDIVIDUAL condition (where only 29% of workers received a lower piece-rate) compared to the BASELINE LOW condition (where 100% of workers received the lower piece-rate) suggests a large learning-by-doing effect. Namely, reduced learning-by-doing by workers in the INDIVIDUAL condition who deliberately restricted their output in the 1st period accounted for *at least* a 9% reduction in average group productivity in the 2nd period.

5 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 with decreased piece-rates or higher quotas set by management in the future. While there is a substantial amount of theoretical work supporting the emergence of the ratchet effect, as well as ample anecdotal evidence pointing toward likely output restriction of 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 among participant workers and thus identify the presence of the ratchet effect. To provide a more robust examination of the ratchet effect in the workplace, we test if it emerges using two different methods

for evaluating productivity: (i) at the individual-level based on each worker's own output level, and (ii) at the group-level based collectively on the output levels of all the workers in a group.

We find that output is significantly lower in the condition where workers face the consequence of reduced piece-rate if they are too productive in the 1st work period, which indicates that workers are restricting output; as such, our results provide strong support for the emergence of the ratchet effect 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 or the emergence of the ratchet effect. In the condition where productivity is evaluated at the group-level, a free-rider problem arises where workers can try to work at full capacity and free-ride off the output restriction of other workers. Our results suggest that in the workplace setting we consider, this free-riding incentive 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 find 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 by the workers, thus overcoming the above mentioned free-rider problem.

Because our design features a real-effort task, we are able to investigate a possible dynamic 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 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 a result of some workers receiving a lower wage in the 2nd period. Thus, output restriction in the 1st work period appears to reduce learning-by-doing among workers, which results in significant productivity losses (at least 9%) in the 2nd work period. 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 of reduced learning-by-doing that, consequently, reduces future productivity.

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 in testing for: output restriction 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, which 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. That being said, the results from our study can contribute more generally to our understanding of the ratchet effect in the workplace. Notably, our results suggest that 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 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 prevents 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, then 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

- 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.
- 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.

- Carmichael, H. L., & MacLeod, B. (2000). Worker Cooperation and the Ratchet Effect. *Journal of Labor Economics*, 18(1), 1-19.
- 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., & 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., 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., & Villeval, M. C. (2011). Competition and the Ratchet Effect. *Journal of Labor Economics*, 29(3), 513-547.
- 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.
- 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., & Gächter, 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*.
- 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., 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 Political 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. A., & 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). Explaining 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.
- 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.
- 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.
- 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 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 – Mailer Assembly Task

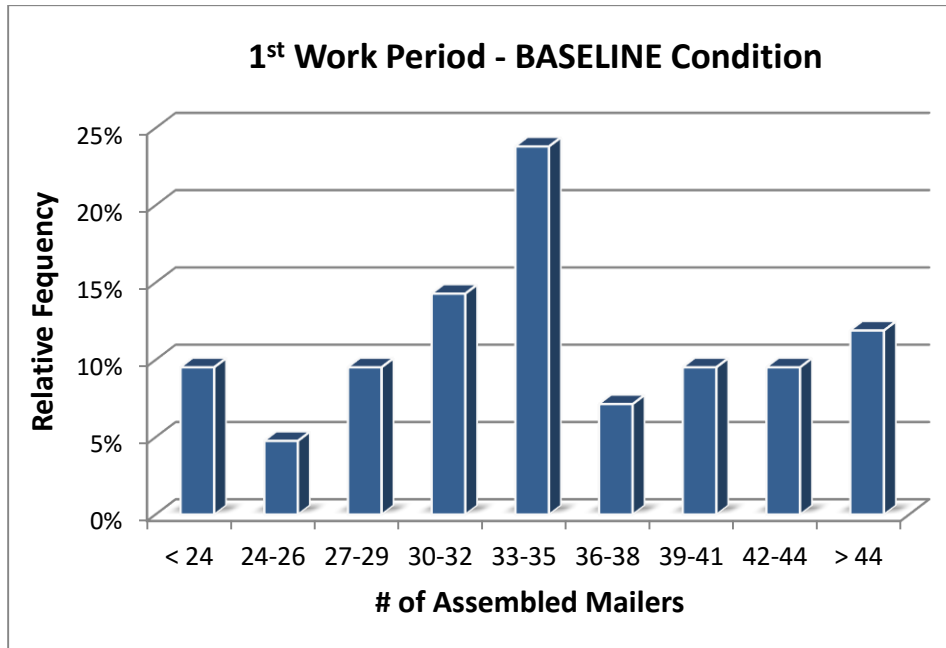


Figure 2 – Simulated Workplace Environment



Figure 3 – Distribution of Completed Mailers in the BASELINE Condition

Panel A – 1st Work Period



Panel B – 2nd Work Period

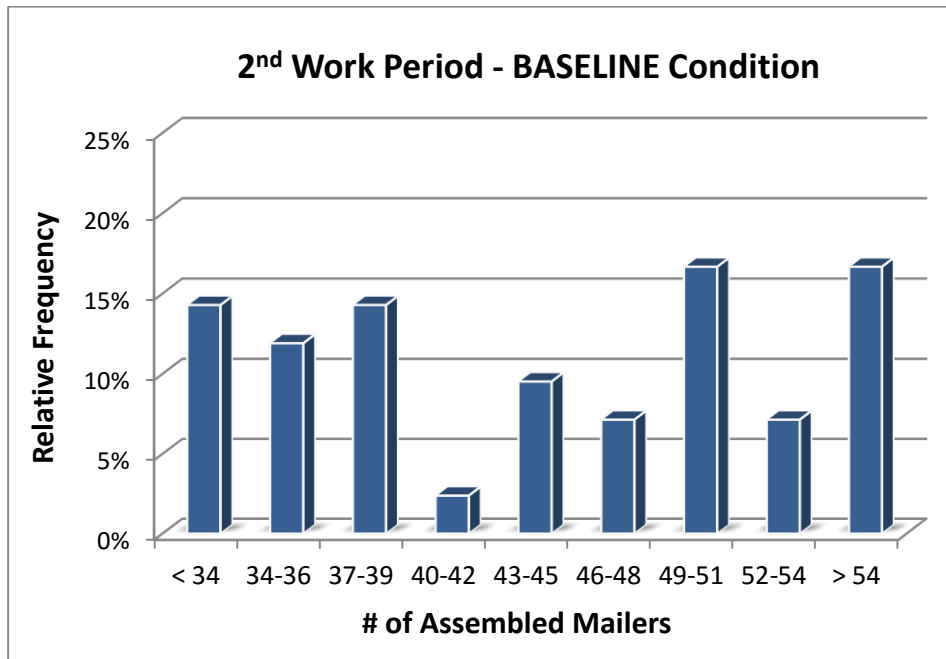


Figure 4 – Comparison of 1st period output distribution: INDIVIDUAL vs BASELINE

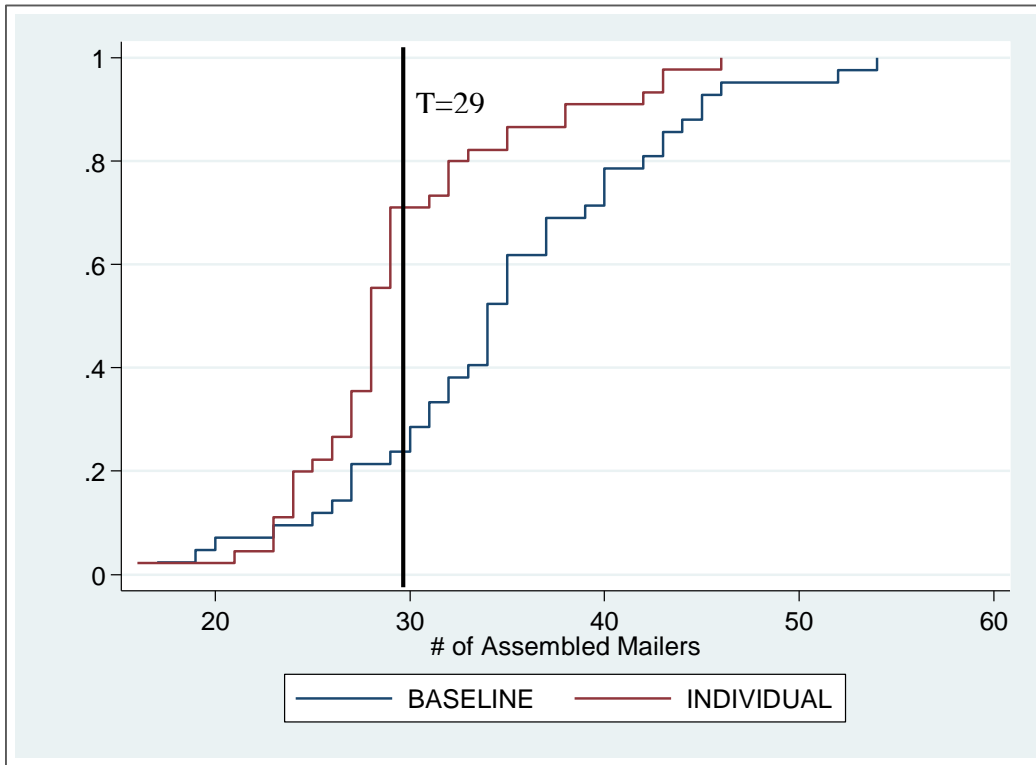


Figure 5 – Comparison of 1st period output distribution: GROUP vs BASELINE

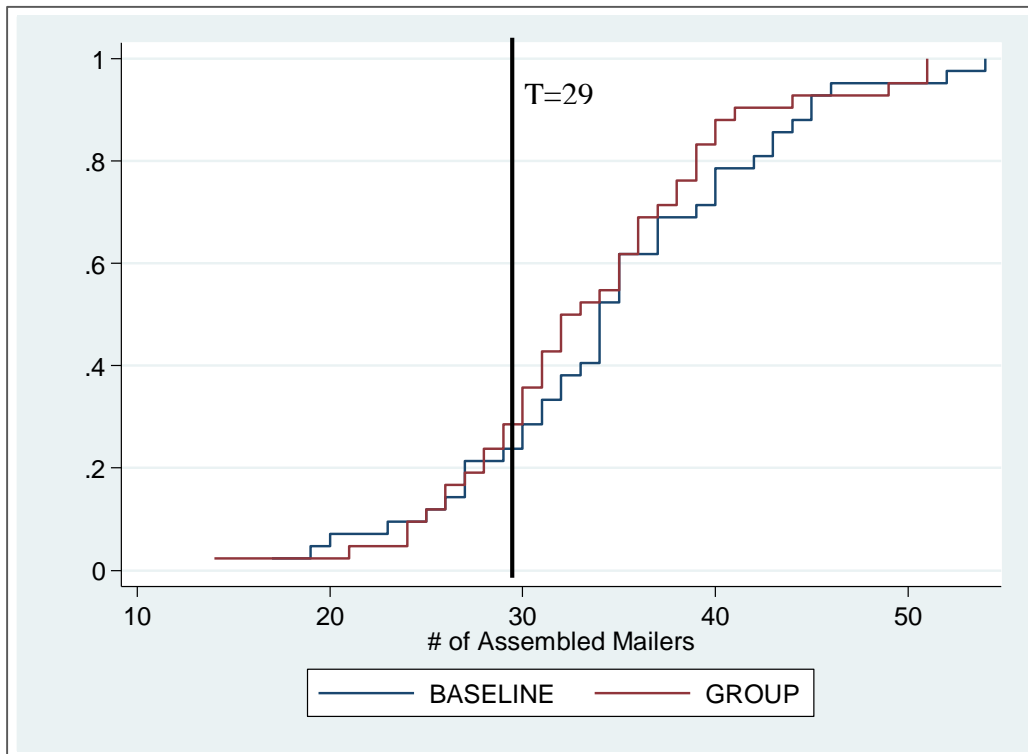


Figure 6 – Comparison of 1st period output distribution: GROUP COMM vs BASELINE

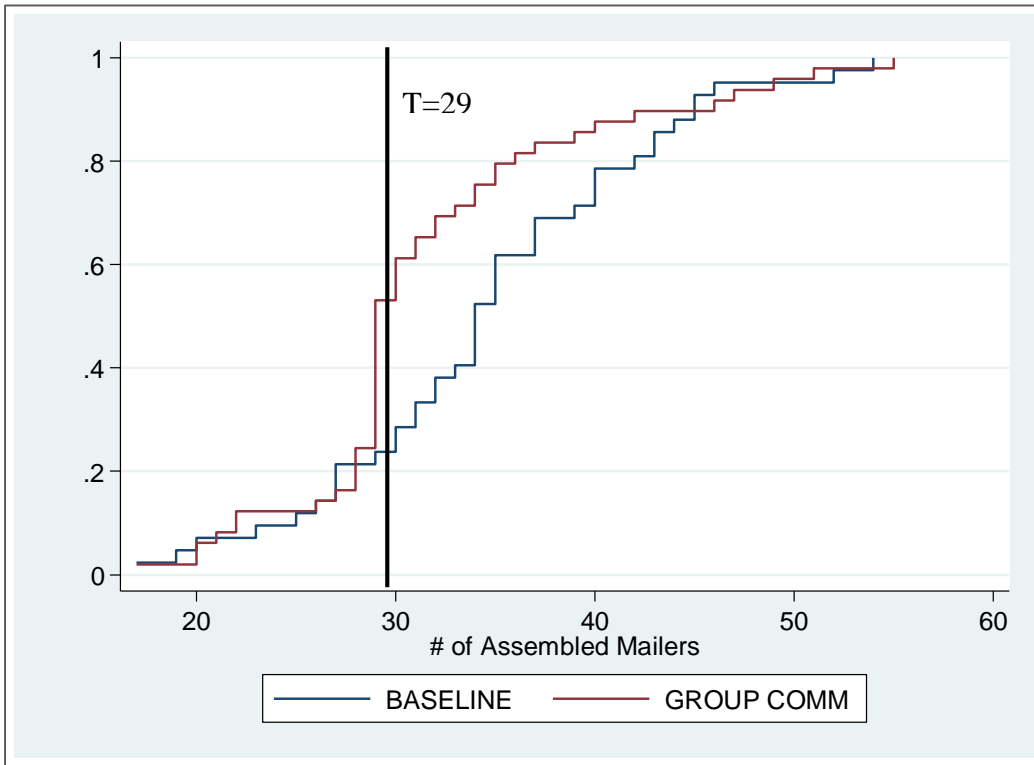


Figure 7 – Comparison of 2nd period output: INDIVIDUAL vs BASELINE

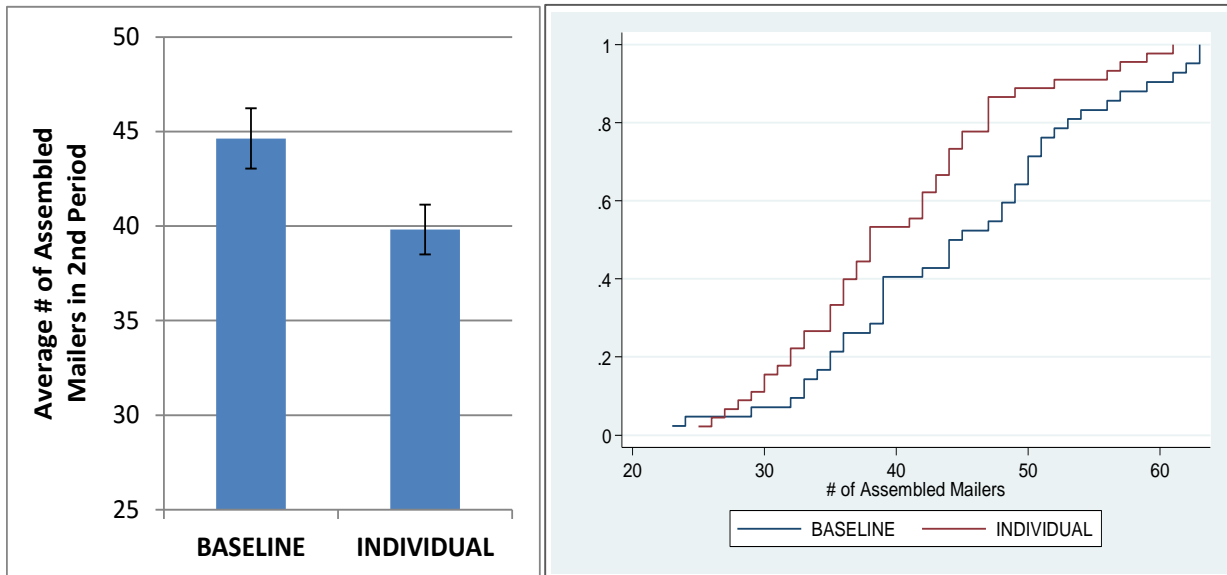


Table 1 – Summary of Experimental Conditions

	BASELINE n = 42 (7 sessions)	INDIVIDUAL n = 45 (7 sessions)	GROUP n = 42 (6 sessions)	GROUP COMM n = 49 (7 sessions)
<i>1st Work Period</i>				
piece-rate	\$.20	\$.20	\$.20	\$.20 and 3 minutes of pre-work discussion
<i>2nd Work Period</i>				
piece-rate	\$.20 for all workers	\$.20 for workers with output <= 29 in 1 st work period \$.10 for workers with output > 29 in 1 st work period	\$.20 if 4 or more of the 7 workers had output <= 29 in 1 st work period \$.10 if 4 or more of the 7 workers had output > 29 in 1 st work period	\$.20 if 4 or more of the 7 workers had output <= 29 in 1 st work period \$.10 if 4 or more of the 7 workers had output > 29 in 1 st 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)
<i>Mean output</i>	34.6	29.4***	33.5	31.7*
<i>Median output</i>	34	28***	32.5	29***
<i># of workers with output of 29 or fewer mailers</i>	10/42 (24%)	32/45*** (71%)	12/42 (29%)	26/49*** (53%)
<i># of workers with output of exactly 28 or 29 mailers</i>	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 – 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 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, (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 envelope, 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 moisture is being dispensed 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 in compensation **per** completed mailer you assemble in the 10 minutes of allotted time.

2nd 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 envelope, 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 moisture is being dispensed 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 envelope, 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 moisture is being dispensed 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 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 envelope, 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 moisture is being dispensed 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 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.