ARE BONUS POOLS DRIVEN BY THEIR INCENTIVE EFFECTS? EVIDENCE FROM FLUCTUATIONS IN GAINSHARING INCENTIVES

ALAN M. BENSON AND SIMA SAJJADIANI*

Bonus pools, in which a worker's realized bonus depends both on a worker's share of the pool (which serves as the incentive) and on the size of the pool (which is largely outside of the worker's control), are a common method for distributing incentive pay. Using data on the variation in the size of the bonus pool generated by a US manufacturing plant's gainsharing plan, which varies incentives for quality and worker engagement, the authors evaluate the conditions under which such bonuses have incentive effects. Overall, results are cautionary: The evidence suggests gainsharing's benefits operate outside of the incentive channel, and incentives may backfire if they are too small or too diluted by group performance metrics. The authors illustrate how random variation in the size of bonus pools offers researchers a powerful, readily available, and underused tool for studying how workers respond to the availability and strength of incentives.

Academic work has sharply focused on piece rates and other "bottom-up" incentives in which the incentive schedule is fixed. Incentives also come from "top-down" bonus pools, however, and their sizes may fluctuate based on factors outside of workers' control. For example, firms may allocate a corporate-wide budget for raises, bonuses, or stock options based on company performance, and that budgeted pool may then be divided among workers or work units in proportion to their productivity, work quality, or other evaluated performance. From the perspective of the incentives

literature, these arrangements are puzzling: Does work quality matter less in quarters where corporate bonus pools are small? Nonetheless, such arrangements are ubiquitous.

Using variation in the size of bonus pools, this study tests whether such bonus pools operate specifically through their incentive effects. Indeed, when we observe that key performance indicators improve after bonus pools are introduced, it may be natural to attribute these improvements to the incentive effect. The effectiveness of such schemes might, however, instead result from the sense of shared ownership that they impart, the simultaneous adoption of other practices, the mere act of tracking these performance metrics, or some other reason hypothesized in the literature (Rynes, Gerhart, and Parks 2005; Bartel, Freeman, Ichniowski, and Kleiner 2011).

To identify the incentive effect apart from these other potential mechanisms, we exploit variation in the eligibility and the size of a shared bonus pool at a manufacturing plant that uses a gainsharing plan. This plan, which is typical of manufacturing plants, funds a bonus pool from the plant's quarterly operating budget. The operating budget is determined by the plant's orders, and a formula is used for projecting the costs associated with any order, minus the realized variable costs of capital, supply, freight, and other expenses outside workers' control. If operating costs in a given quarter are sufficiently below a budget threshold, then the bonus pool funds incentives for meeting plant-wide quality targets and worker engagement targets. This feature of gainsharing plans creates frequent, unexpected, and largely random variation in the availability and magnitude of a bonus pool; at our plant, we observe 16 instances of the plant crossing the eligibility threshold and 52 changes in the incentive's magnitude over 173 weeks. We exploit this variation to examine a preeminently practical question: Does plant-wide quality and worker engagement improve when workers are told they are currently eligible for these bonuses? In addition, how does this improvement depend on the size of the incentive?

Why would such bonus pools have incentive effects? For one reason, they are very common: More than a quarter of the workers surveyed in the 2006 General Social Survey (GSS) reported that they receive gainsharing bonuses (Kruse, Blasi, and Park 2010). In 1998, 46% of Fortune 1000 firms used gainsharing or profit-sharing incentives, even though incentives distributed in these ways are so diluted they have virtually no direct incentive effects on individual workers (Dube and Freeman 2010). Second, using a variety of metrics, studies suggest that several plants have seen performance improve after adopting gainsharing and profit-sharing, and plants that use these programs tend to perform better than those that do not (for a review, see Weitzman and Kruse 1990). Such research, however, has typically relied on before-and-after studies of individual plants (e.g., Hamilton, Nickerson, and Owan 2003), or across-firm correlational studies using survey data (Freeman, Kruse, and Blasi 2010). Such associations are problematic

because group incentives are typically adopted simultaneously with other practices that are designed to improve performance (Osterman 1994), and indeed, a large empirical literature finds that these programs are effective when implemented as part of a bundle of complementary practices (Kruse 1994; Huselid 1995; MacDuffie 1995; Blasi, Conte, and Kruse 1996; Ichniowski, Shaw, and Prennushi 1997). Given that gainsharing may be adopted simultaneously with other practices, and given that little fieldwork has evaluated the mechanisms underlying gainsharing programs, we should be cautious about assuming that gainsharing plans are effective specifically because of their incentive effects.

Our main findings are as follows. First, plant-wide quality (one of the incentivized performance indicators) did improve after the plan's adoption, but this improvement was part of a secular time trend that is consistent with the broader scope of quality initiatives the plant was instituting around the same time the gainsharing plan was introduced. This first analysis illustrates the challenges of isolating incentive effects from a before-and-after study alone, given simultaneity and endogeneity concerns. The subsequent (main) tests instead exploit variation in the size of the bonus pool in the post-adoption period to study their incentive effects.

Second, and most crucially, quality did not improve more when the bonus pool happened to be large compared to when there was no bonus pool. This finding serves as evidence that the gainsharing plan did not have direct incentive effects, at least for quality. By contrast, workers were more likely to submit process improvement suggestions in order to meet engagement targets when the gainsharing plan made them eligible for incentives, but unlike the measure for quality, the engagement measure was more clearly within workers' control and had an individual measure in addition to the plant-wide measure. These features suggest how similar incentives can be designed to be most effective.

Third, the gainsharing plan sometimes yielded very small bonus pools, and the evidence suggests that small bonuses had the unintended consequence of demotivating workers compared to no incentive at all. Although this result is somewhat anomalous within the whole of the incentives and gainsharing literature, it is consistent with psychological studies that suggest that small incentives can be less motivating than no incentives at all, because small incentives provide little extrinsic motivation but still crowd out intrinsic motivation (Deci 1971, 1972; Gneezy and Rustichini 2000; Gagné and Forest 2008; Frey, Homberg, and Osterloh 2013; Cerasoli, Nicklin, and Ford 2014). It is also consistent with the puzzling survey results observed by Freeman, Kruse, and Blasi (2010) that suggest effort may be greater when there is no group-based incentive program or expected group-based bonuses are large, rather than when incentives are small.

This study and others consistently find evidence that gainsharing programs are associated with superior performance, and other studies have been nearly universal in their conclusion that incentives are effective. Our

results, however, suggest that gainsharing and other forms of shared bonus pools are not effective because of their direct incentive effects, but rather through other mechanisms, such as complementary and often simultaneously adopted engagement practices. Our results suggest that gainsharing and shared bonus pools can also dilute incentives, potentially falling victim to the free-rider problem. This outcome contrasts other work that has largely found that peer monitoring deters free riding (e.g., Knez and Simester 2001).

Background on Gainsharing Plans

We evaluate incentives distributed through a gainsharing plan. Gainsharing plans typically use regular cash bonuses to incentivize the achievement of specific key performance indicators (KPIs), such as employee engagement, safety, or quality. Under a gainsharing plan, these bonuses are not fixed; rather, they are "self-funded" from cost savings or some other variable budget. In this way, they are similar to corporate bonus pools in which the size of the bonus pool is determined by the business unit's profit and loss, and then the employee's share of the bonus pool is determined by the number of points they were awarded (e.g., from a performance evaluation or for meeting KPIs). In gainsharing plans, cost savings include the budgeted minus actual costs of production, such as hiring, training, compensation, utility, material, supply, and freight. Gainsharing plans are also convenient for budgetary reasons, since they only distribute bonuses when other costs are low (Hatcher and Ross 1991; Nalbantian and Schotter 1997; Arthur and Aiman-Smith 2001; Bloom 2008; Milkovich, Newman, and Gerhart 2013).

As is the case in the plant we study, gainsharing plans are often adopted simultaneously with other high-performance work practices. For example, in manufacturing operations, gainsharing is one component of a "Scanlon plan." Such plans use gainsharing not just as a vehicle for disseminating incentives but also as a vehicle for imparting a sense of shared ownership: They provide the rewards for the kinds of process improvement activities that other Scanlon plan components enable. Several studies and reviews have concluded that Scanlon plans are indeed associated with superior performance along a variety of dimensions (Jones, Kato, and Pliskin 1994; Drago and Heywood 1995; Arthur and Huntley 2005). Gainsharing also often accompanies other engagement programs that are cited in the human resource literature as complements (such as total quality management or Six Sigma); again, whereas these programs offer workers the forums, such as semi-autonomous teams and quality circles, in which to propose process improvements, gainsharing provides the incentives to do so. Gainsharing plans should not be confused with profit-sharing plans, which typically differ in how they are distributed (e.g., through pension contributions) and what they reward (profit and stock performance, rather than specific intermediary metrics while the plant is bonus-eligible). Both

gainsharing plans and profit-sharing plans are typically cited as hallmarks of US manufacturing's movement to shared capitalism (Kruse et al. 2010).

Gainsharing plans are fairly common in the United States (Dube and Freeman 2010; Kruse et al. 2010). Within manufacturing, gainsharing plans are most commonly used by plants with team-based production, where monitoring individual workers is challenging (Weitzman and Kruse 1990; Jones et al. 1994; Drago and Heywood 1995; Kruse 1996). In the 2006 GSS, about half of US manufacturing workers with long-term contracts were covered by profit-sharing and/or gainsharing plans (Bryson, Freeman, and Lucifora 2012). Lincoln Electric's bonus program, which is the subject of a widely used case study and book, is also an example of a gainsharing plan (Berg 1975; Koller 2011).[1] More generally, gainsharing is one of many bonus schemes in which incentives are determined multiplicatively: both by what they intend to incentivize and also by some largely exogenous determinant, such as a share of profit or saved costs.

Hypotheses

Prior studies have largely concluded that gainsharing plans are associated with greater productivity, greater quality, and other performance improvements. Methodologically, these studies usually come in two varieties: 1) single-firm "before-and-after" studies, and 2) multi-firm studies that use surveys to compare outcomes at firms with gainsharing to those without. Examples of the former include Hatcher and Ross (1991), who found a time trend in quality after implementing a gainsharing scheme in a unionized manufacturing company; Wagner, Rubin, and Callahan (1988), who found that productivity improved after group-level incentives were introduced to a foundry; Knez and Simester (2001), who found that Continental Airlines overcame the free-rider problems of company-level incentives through intragroup peer monitoring; and Hamilton et al. (2003), who found that productivity rose by 18% after a garment manufacturer implemented a gainsharing plan. Examples of multi-firm survey studies abound, and these studies have found that gainsharing is associated with greater productivity, lower absenteeism, and other outcomes (Petty, Singleton, and Connell 1992; Poole and Whitfield 1994; Fernie and Metcalf 1995).

This study takes a different approach. As a preliminary test, we perform a "before-and-after" test of gainsharing, but we are primarily concerned with testing the gainsharing underlying mechanism, and especially, whether gainsharing had an actual incentive effect. Indeed, based on prior

---

[1]Lincoln Electric distributes bonuses from a pool funded by profits. These bonuses are awarded in proportion to an individual workers' points, which are awarded for four key performance indicators: ideas, quality, productivity, and reliability. Note that in periods in which the bonus pool is small, so too would be incentives for these four key performance indicators. Unlike Lincoln Electric, in the plant we studied the incentives in the gainsharing plan are distributed for meeting plant-wide performance rather than individual performance.

literature, it is less clear what mechanism is responsible for the association between gainsharing and performance improvements. One possibility is that gainsharing is associated with superior performance because such shared bonus pools have direct incentive effects. Alternatively, prior studies may have found a positive association between the adoption of gainsharing and superior performance for spurious reasons: Performance may improve due to the independent effects of simultaneously adopted practices, or because gainsharing complements other practices in ways that do not specifically rely on the gainsharing plan's incentive effects. Disentangling the effect of such bonus pools and simultaneous practices has been difficult, since both before-and-after studies and survey methods are vulnerable to the concern that gainsharing may be adopted as part of a package of complementary practices specifically designed to address the same problems (Osterman 1994; Drago and Heywood 1995; Huselid 1995; MacDuffie 1995; Ichniowski et al. 1997; Gittleman, Horrigan, and Joyce 1998; Kim 2005; Dube and Freeman 2010).

Whether gainsharing operates through the incentive channel is also theoretically ambiguous. Because effort costs are largely private whereas bonuses are shared, gainsharing would seem to be a prime candidate for free riding and the dilution of incentives among a large number of coworkers. Yet, there are also theoretical and empirical reasons to be skeptical that the $1/N$ problem applies in real-world workplaces. Kandel and Lazear (1992) showed that semi-autonomous teams can keep free riding in check through peer pressure and mutual monitoring, a prediction empirically supported by Knez and Simester (2001). This view has since been echoed by a variety of studies and incorporated into a literature that has treated autonomous work groups and team incentives as complements (for reviews, see Kruse 1993; Blasi et al. 1996).

In part because of the theoretical ambiguity and methodological challenges of isolating gainsharing incentives, relatively little is known about whether the effectiveness of gainsharing incentives depends on their size. With regards to this question, our study is perhaps most closely related to Freeman et al. (2010). Using National Bureau of Economic Research (NBER) surveys and the General Social Survey, they estimated whether workers were more likely to confront shirking workers, conditional on whether their employer had a gainsharing plan and on the self-reported size of the bonus that workers usually expected. They found that workers were less likely to confront shirkers if they had a group-based incentive plan—unless the incentive was sufficiently large. Freeman et al. concluded that "it is the intensity [of bonus] rather than the presence of profit sharing that seems to matter . . . very low levels of profit sharing may have a negative effect on anti-shirking activity, but this reluctance is apparently overcome as the bonus grows larger" (2010: 90).

Given that the prior literature has been favorable to gainsharing, to incentives, and to the ability of organizations to overcome free-rider

problems, our hypotheses are that gainsharing incentives will be associated with gains to the key performance indicators:

Hypothesis 1: When the gainsharing program is adopted, the program's targeted KPIs (quality and engagement) will improve.

Hypothesis 2: Quality will be greater when workers are eligible for the quality bonus (2A), the size of the quality bonus increases (2B), and workers are both eligible for the quality bonus and also quarterly quality targets to date have met quality targets (2C).

Hypothesis 3: Engagement will be greater when workers are eligible for the engagement bonus (3A), the size of the engagement bonus increases (3B), and workers are both eligible for the engagement bonus and also quarterly engagement targets to date have met engagement targets (3C).

Finally, we test the hypothesis that small incentives are demotivating. This test is driven by self-determination theory (Deci 1972), which proposes that motivation is the sum of extrinsic motives (e.g., workers want bonuses for quality) and intrinsic motives (e.g., workers enjoy producing high-quality work). Because any amount of extrinsic motives "crowds out" intrinsic motives, sufficiently large incentives are required to overcome the intrinsic motivation forfeited by the introduction of incentives. This scenario yields the classic "inverted-U" relationship between extrinsic motivators and outcomes, in which small incentives are more demotivating than no incentives at all (Gneezy and Rustichini 2000; Weibel, Rost, and Osterloh 2007; Freeman et al. 2010; Gneezy, Meier, and Rey-Biel 2011).

More specifically, Freeman et al. (2010) also hypothesized that gainsharing plans may incidentally yield small, demotivating incentives. In a survey, they also found evidence consistent with the U-shaped relationship, noting that coworkers were least likely to discipline shirking coworkers when incentives were small, compared to when they were large or absent. Therefore, we expect to see a non-monotonic relationship between quality (engagement) and bonus size:

Hypothesis 4A: The relationship between quality and incentive size follows an "inverted U-shape," such that item returns are greatest when the bonus pool is small but positive.

Hypothesis 4B: The relationship between engagement and incentive size follows a "U-shape," such that engagement is lowest when the bonus pool is small but positive.

Data

Setting

Data come from a 60-year-old urban, nonunionized manufacturing plant that produces heavy industrial equipment. The plant is the flagship plant of a publicly traded multinational manufacturing company. The plant employs between 300 and 400 full-time equivalent production employees. These

employees are assigned to one of three eight-hour shifts (day, evening, night), one of three functional groups (assembly, fabrication, and support), and one of 14 semi-autonomous teams within any function (e.g., fabrication includes sheet metal, machine shop, welding, paint, fabrication, technical services, and material control teams). The plant uses lean manufacturing techniques for process and inventory control, which also allows downstream teams to monitor inputs provided by upstream teams.

Each team is led by a cell leader who is also a production worker, and each functional area is led by an operational engineer. Teams are relatively small and interdependent; each of the 14 semi-autonomous teams are nominally responsible for testing parts before sending them down the assembly line, and downstream teams can implicitly monitor inputs by upstream teams and return them for rework if needed. The plant relies on peer monitoring within teams and across teams, and monitoring by supervisors to deter free riding.

Workers are rarely terminated, and voluntary turnover is also low. In 2014, turnover was 6.3%, the mean worker had 17 years of tenure, and 75 workers had more than 25 years of tenure.

The industrial equipment produced by the manufacturing plant is made-to-order. Each week, the plant produces an average of 103 items in about 12 different models. Much of the plant's own equipment is highly flexible and must be manually operated by skilled technicians. As such, this plant looks more like other manufacturing plants engaging in what Piore and Sabel (1984) referred to as "flexible specialization," or the production of small batches, and less like highly automated plants that engage in mass production. As we discuss in our Institutional Benchmarks and Generalizability section, we interpret the labor management relations and organizational climate plan to be unexceptional among non-unionized, mid-sized, make-to-order plants.

As part of the plant's lean manufacturing initiative, workers could submit continuous improvement suggestions either by submitting cards or by posting suggestions on the work group's large poster boards. Workers were also encouraged to engage through weekly quality circles and meetings with management.

The plant management invited the researchers to evaluate the gainsharing plan that began on a trial basis in January 2012. Management was considering an extension of the gainsharing program into the company's other plants as part of an effort to standardize compensation in all plants and was specifically interested in whether incentives were working. To this end, management provided access to staffing data, operational data, and other material related to the gainsharing program from January 2010 to July 2015. Prior to the gainsharing plan, the data link quality-related returns to the month they were produced. After January 2012, data on returns are linked to the exact date they were produced. For our empirics, this change is important because the size of the gainsharing bonus was recalculated and

disclosed weekly; for all items, we know whether workers were told they were eligible for quality bonuses during that week, and we can track whether items produced that week were ultimately returned. Overall, we have 24 months of data before the gainsharing plan was implemented and 173 weeks of data afterward. Our data and empirical strategy are informed by these materials and by our conversations with the plant's senior operational managers and workers.[2]

The plant implemented the gainsharing plan specifically to address quality issues. The year prior to the gainsharing plan's implementation, 12% of items produced by the plant were returned to the manufacturer and its distributors within one month by customers citing workmanship issues. The plant was also in the midst of implementing other multi-year process improvement programs, specifically lean manufacturing, Six Sigma, and Total Quality Management. Rather than providing fixed incentives for quality, the plant management adopted the gainsharing plan to avoid large compensation expenses when cash flow was poor. In addition to the incentive for quality, the gainsharing plan funded incentives for worker engagement. Besides the gainsharing plan, the only other elements of the compensation plan were the base pay, standard benefits, and a profit-sharing plan that was awarded annually and matched employees' contributions into a 401(k), with incentives for taking contributions in the company's stock.[3] Crucially, there are no short-term incentives (such as piece rates or team incentives) aside from the bonus pool that is distributed by the plant's gainsharing plan. Base pay and other compensation programs were left unchanged at the introduction of the gainsharing program, though in principle, the bonus program could have substituted for larger raises.

Naturally, the simultaneous implementation of other process improvement and worker engagement programs would pose a problem for any before-and-after evaluation of the gainsharing plan. As we will explain, however, the method for our main tests (Hypotheses 2–4) will exploit fluctuations in the size of the gainsharing incentives to examine whether the gainsharing plan had any discernible incentive effects.

---

[2]The conversations were with the director of operations, the subdivisional heads of operations, the lean manager, plant controller, the finance director of operations, the HR operations managers, the warranty supervisor, as well as warranty and shop-floor workers and the IT specialists who helped extract the data.

[3]The gainsharing plan should not be confused with the profit-sharing plan. Gainsharing plans and similar bonus pools are typically thought of as short-term incentives, whereas profit-sharing plans are typically thought of as long-term incentives tied to retirement benefits; this is the case here. Specifically, the gainsharing bonus depends on the plant's cost savings, whereas the profit-sharing plan is budgeted from the entire company's operating profit. Second, the gainsharing plan is awarded as a quarterly bonus, whereas the profit-sharing plan is awarded annually into the employee's 401(k). Third, as we show, the gainsharing bonus fluctuates widely, whereas the profit-sharing plan varies from 2.35% to 2.7% of annual earnings.

How the Gainsharing Plan Works

The plant's gainsharing plan is archetypal among US manufacturing plants (e.g., those described by Wagner et al. 1988; Hatcher and Ross 1991; Arthur and Aiman-Smith 2001). It explicitly funds incentives for two key performance indicators (KPIs), quality and engagement, as a share of the plant's idiosyncratic variable cost savings. The gainsharing plan works as follows.

Bonus Eligibility and Size

The plant distributes bonuses when variable cost savings (actual expenses minus budgeted expenses) exceed a fixed threshold of $50,000. In these cases, the bonus pool will be at least 20% of the size of variable cost savings. This 20% bonus was designed to provide workers with a sense of shared prosperity, addressing morale concerns after the company enjoyed large profits but relatively small wage increases after emerging from the recession. Among the plant's management and workers, there is no pretense that workers control variable cost savings. For example, variable cost savings can change rapidly and unexpectedly when the plant makes a major capital purchase, or when the variable costs budget, which is calculated based on a two-year moving average, abruptly changes because of a major capital purchase that occurred 25 months ago.[4] Other factors that affect variable cost savings include idiosyncratic hiring, training, freight costs, and the prices charged by contractors and suppliers. As such, the gainsharing plan was adopted to avoid paying out large incentives in quarters in which the plant was not profitable. By pegging the size of the bonus pool to a share of cost savings, management and owners are assured that bonuses will not be "too large" even in quarters the plant was losing money. Although this may seem strange, it is not unlike stock options for low-level employees or other corporate-wide bonus pools with budgets that can vary widely and be largely disconnected from the effort of any worker.

Incentives for Quality

When workers are bonus-eligible, workers receive an additional 15% of cost savings if the plant meets a plant-wide quality target. When workers are not bonus-eligible (because of insufficient variable cost savings), they are not eligible for the quality bonus. Quality is measured by the share of shipped items that are returned to the manufacturer because of short-term quality (STQ) claims believed to be within the workers' control. Specifically, the quality measure consists of the share of items that meet three criteria: 1) the item must be returned to the manufacturer or a distribution partner within one month of receipt; 2) the item must have been operated for less than 10 hours; and 3) the returned item must require repairs on a component

---

[4]The plant's controller, who does these calculations, reports to workers only what the variable cost savings are and not how they are calculated. Likewise, interviewed workers report that they do not know how this is calculated.

produced at the plant. Items that are not considered as STQ claims for the purpose of workers' incentives have been returned for non-quality-related reasons, or because its quality issues relate to parts supplied by contractors (e.g., engines and batteries). Of the returns, 9.47% (about 1% of all manufactured items) are excluded from the quality measure for these reasons.

Incentives for Engagement

When workers are bonus-eligible, workers receive an additional 15% of cost savings when they meet a plant-wide engagement target. Engagement is measured by the number of continuous improvement card submissions, which are suggestions for operational improvements submitted by workers and reviewed by managers. For the purpose of distributing incentives, the quality of the submission is not considered. To receive the additional 15% bonus, the worker must personally submit a continuous improvement card. The plant must also meet the total continuous improvement card target, which is equal to 60% of full-time employees.

Shown formulaically, the bonus pool in quarter q is represented by:

$$(1) \qquad Bonus_q = \left(Savings_q\right) * \left(S_q\right) * \left(0.20 + 0.15Q_q + 0.15E_q\right)$$

where S_q, Q_q, and E_q are indicators that take the value 1 if savings, quality, and engagement thresholds are respectively met, and $Savings_q$ is the quarter's variable cost savings. We provide the equation to make clear that when variable cost savings are less than $50,000, no incentive is given for quality or engagement; by "bonus-eligible," we mean that the variable cost savings have exceeded this $50,000 threshold so that workers are eligible to receive 15% of variable cost savings for meeting quality and engagement targets. When cost savings are large, the bonus pool is as great as 50% of variable cost savings if quality and engagement targets are met.

In the plant, the bonus was between 0% and 6% of a full-time worker's annual salary. In a survey of 40,767 workers who received gainsharing incentives in a range of industries, Kruse et al. (2010) found that gainsharing incentives are on average 8.9% of annual pay. Because quality is measured one month after items are produced, workers do not receive bonuses until about five weeks after the quarter ends.

Note that when workers are bonus-eligible, all workers receive the quality bonus if the plant meets quality targets. Although this would seemingly make the plant vulnerable to free-rider problems, the plant did not have the necessary process control systems to track mistakes down to specific teams and to establish team-specific targets.[5] The plant's management

---

[5]It is not uncommon for plants to adopt process control and quality tracking systems simultaneously with gainsharing and other incentive plans. Two well-known examples, Lincoln Electric and Safelite Autoglass, each had such tracking systems, each had some form of individual-level quality incentives, and each required operators to fix mistakes on their own time (Berg 1975; Hall, Lazear, and Madigan 2000; Lazear 2000).

seemed unsure how important this was, noting that it is not uncommon for incentive programs to reward plant-wide goals, and free riding might be checked by peer monitoring and a sense of cohesion. By contrast, when workers are bonus-eligible and the plant-wide engagement target is met, only workers who participated in the continuous improvement plan will receive a share from engagement bonus. We use the differences between the quality and the engagement measures to examine whether free riding does indeed pose a problem for group incentives like those for quality.

How the Gainsharing Plan Is Disclosed

Although the bonus is distributed quarterly, workers receive weekly updates regarding: 1) whether workers are bonus-eligible and the size of the bonus pool, if applicable, given the quarter's current cost savings; 2) whether their quality targets are being met; and 3) whether engagement targets are being met. These updates come in the form of dashboards posted throughout the plant, a weekly newsletter, and daily team meetings. These dashboards use green and red color codes to represent whether workers are bonus-eligible. Workers are also trained on the gainsharing plan as part of their orientation.

Figure 1's top panel gives an example of a bonus status dashboard. These dashboards are posted around the plant, including the front entrances, restrooms, breakrooms, and plant floor. Note that, in this case, the workers were explicitly not eligible for the 15% quality and engagement bonuses in quarter 1, even though quality and engagement targets were met. The second panel provides an example from the weekly company newsletter. The newsletter dashboard explains that workers were eligible for the bonus in quarter 4 and achieved the engagement target but not the quality target. In addition, unit leaders would typically remind workers at the beginning of their shifts if the variable cost savings thus far imply they would be bonus-eligible. Some interviewed workers report paying attention to the magnitude of the bonus, and others just pay attention to whether the indicator is green.

Although the plant management tries to keep these dashboards updated and workers always observe whether they are bonus-eligible or not, the exact size of the bonus pool was missing in 13% of weeks, the share of items returned to the manufacturer was missing in 27% of weeks, and the count of continuous improvement cards was missing in 11% of weeks. These measures are compiled by different people (i.e., the controller, the head of operations, and the IT specialist, respectively), and data would be missing due to holidays, vacation, and random administrative delays. To address this, we use multiple imputation. If these indicators are missing at random, then multiple imputation yields unbiased estimates for coefficients and standard errors (Schafer and Olsen 1998; Rubin 2004; StataCorp 2013). The
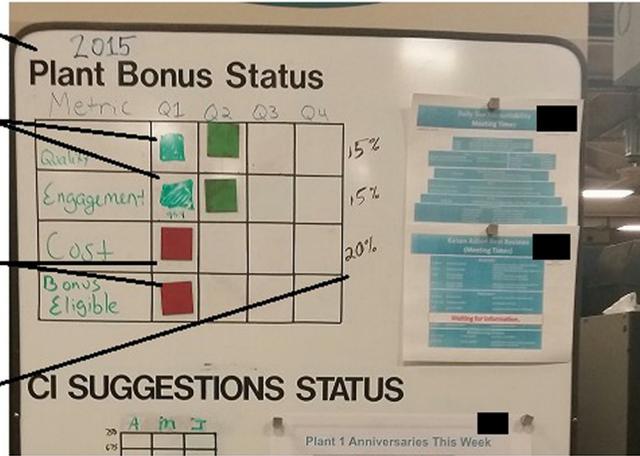
Figure 1. Bonus Eligibility Reporting

Dashboards like this are posted in highly trafficked locations.

Green indicates that engagement and quality targets are being met.

Red indicates that the cost target is not being met, and therefore workers are not bonus-eligible.

Reminder that met KPIs boost workers' share of savings.

Weekly newsletters also display dashboards.

Newsletters also provide performance versus target.

Text reminds workers whether they are eligible for quality and engagement bonuses.

2015
Plant Bonus Status

Metric	Q1	Q2	Q3	Q4	
Quality					15%
Engagement					15%
Cost					20%
Bonus Eligible					

CI SUGGESTIONS STATUS

Plant 1 Anniversaries This Week

Q4 Plant Bonus Final Results

Metrics	Goal	Actual
Variable cost budget	>$50k under	$81K
Quality (STQ)	<8.4%	8.6%
Employee engagement	368 pts	817 pts

Q4 Plant Bonus Results:

- **Variable Cost Budget** – Final results for the quarter were $81,000 better than our variable cost plan therefore eligible for the bonus.
- **Quality** – We ended the quarter with STQ missing plan for our Q4 quality performance.
- **Employee Engagement** – We ended the quarter with 817 CI board ideas and completions achieving the goal.

Main Results section describes the sensitivity of our findings to imputation and other specifications.

An Example with Timing

The timing of events can be confusing, so we give an example. Suppose that it is the fifth week of a quarter. Prior to the beginning of that week, the plant's management distributes the newsletter and updates the dashboards that are posted throughout the plant. These report whether variable cost savings since the beginning of the quarter have been sufficiently high such that the plant is bonus-eligible (yes or no). The dashboards and newsletters also report the size of variable cost savings (e.g., $60,000, beating the target of $50,000), whether the plant has been meeting quality targets (yes or no), how plant quality compares versus the target (e.g., 8.5% of items produced since the beginning of the quarter were returned, beating the target of 8.7%), whether the plant has been meeting engagement targets (yes or

no), and how plant engagement compares versus the target (e.g., 250 submissions, missing the goal of 300). Variable cost savings, which gives the potential size of the bonus pool, is updated every week, including the previous week's costs. The quality indicator is updated as items are returned to the plant; since items returned within one month for quality reasons count against the quality indicator, items produced in the first week are counted as either returned or not returned, and averages are used for those still at risk of being returned. The cell leaders also periodically remind workers about eligibility for bonuses for quality and engagement, especially if bonus pools are large.

After being prompted with their bonus eligibility, workers go about their shifts. In a typical week, the plant produces 105 items. These items are all at risk of being returned citing quality reasons for one month after they leave the plant floor. Later on, in the last week of the quarter, dashboards and newsletters will be updated until the second-to-last week of the quarter. Then workers will submit any remaining continuous improvement cards and produce the last items that will count toward the quality target. Because items produced in this week are still at risk of being returned for another month, the bonus for this quarter will not be paid until around the fifth week of the following quarter. If variable cost savings are greater than $50,000, up to 50% of savings will be distributed as a cash bonus if workers meet the quality and engagement goals.

The analysis uses what workers see in any one week (e.g., bonus eligibility given by cost savings) as the treatment. Given what workers saw, the dependent variables are 1) the share of items produced that week that were ultimately returned for quality reasons and 2) the number of continuous improvement cards submitted that week. Note that the language the plant uses to describe the bonus plan eligibility does not exactly capture what ultimately occurs. Every week the workers are told whether variable cost savings are such that workers are bonus-eligible for that week, and if so, the current size of that bonus pool. Strictly speaking, workers are bonus-eligible every week since variable cost savings in the remainder of the quarter are not yet realized; in expectation, very low variable cost savings more strongly suggest workers will not be bonus-eligible when the end of the quarter nears. We choose to use the plant's contemporaneous bonus eligibility and pool size rather than working with expectations for two reasons. First, the plant's communications literally tell workers that they are either currently eligible or not eligible for incentives, and what the current size of the bonus pool is, even though these generally do not reflect the final numbers. Second, working with the distribution of expectations rather than the contemporaneous numbers published by management necessarily introduces many statistical and behavioral assumptions. For these reasons, we feel that using what workers actually see in a given week is the most clear and transparent approach.

Can Workers Control Cost Savings?

The gainsharing plan was designed to distribute incentives for quality and engagement when the plant had cost savings, even though these cost savings were not intended to be within control of the workers. Likewise, the five plant workers interviewed for this study reported that although they were updated on all components of the gainsharing plan, they felt in control of only quality and engagement. Workers reported feeling aware of whether they were bonus-eligible but they did not feel in control of the underlying variable cost savings nor were they aware of how it is calculated.[6] The Descriptive Statistics and Preliminary Tests section also explores this issue statistically.

Empirical Strategy

Hypothesis 1: Time Trend

Our first hypothesis test reproduces a simple before-and-after test: We examine whether trends and jumps in quality occur before and after the gainsharing plan is adopted. These data draw from item return rates for 24 months before gainsharing begins and for 43 months after. We are able to examine changes in quality but not changes in employee engagement, because the plant did not keep track of engagement before the gainsharing program introduced their incentives. Furthermore, we do these analyses at the month level rather than at the week level, also because of data limitations addressed only after the gainsharing program was adopted. We emphasize that this test, like others before it, is vulnerable to simultaneity and confounding with other factors that are correlated with time. Indeed, in this case, since gainsharing was adopted at the same time as other process improvement programs, this replication exercise is simply meant to illustrate whether the same observational trends that we see in other studies are visible here.

Specifically, we perform an ordinary least squares (OLS) spline regression that allows for both instantaneous and gradual changes to quality after the gainsharing program is adopted.[7] These regressions take the form:

---

[6]For example, one worker reported that "you just hope [the bonus eligibility indicator] is green," and another noted he is "still waiting on an explanation two years later on how [variable cost savings] is calculated." Indeed, none of the interviewed workers knew what went into variable cost savings or how they could improve it, even though they were all well aware of how quality and engagement were measured. We did not hear of any workers specifically distrusting the cost numbers, just that their calculation was unclear.

[7]The spline regression is a flexible method of modeling the dependent variable (*quality*) over time. A simple alternative is to include in the model only the dummy term for the post-adoption period, which would reduce the test to essentially a two-sample *t*-test for equal means in the pre-period and post-period under the assumption of equal variance. Because this would involve estimating only one parameter, ($post_i$), this regression would have the greatest power but impose the strong assumption that quality would not otherwise change over time. Including a time trend term alone would test whether mean quality was different. The spline function, which further includes the interaction, allows gainsharing to affect the rate of the change in quality as well. This outcome is potentially important, as it is not clear whether gainsharing should affect quality instantly (e.g., due to the incentive effect) or over time (e.g., due to learning).

(2) $$Q_i = \beta_0 + \beta_1 G_i + \varepsilon_i$$

(3) $$Q_i = \beta_0 + \beta_1 G_i + \beta_2 pre_i + \beta_3(G_i * post_i) + \varepsilon_i$$

where Q_i is quality in month i, measured by quality-related item returns. In Equation (2), $\beta_1 G_i$ gives the mean difference in quality-related return rates after the gainsharing plan is adopted. In Equation (3), $\beta_1 G_i$ gives the instantaneous effect of gainsharing, $\beta_2 pre_i$ gives the monthly time trend before gainsharing was adopted, and $\beta_3(G_i * post_i)$ gives the monthly time trend after gainsharing was adopted.[8]

Hypothesis 1 is that the instantaneous effects and time trend after the adoption of the gainsharing program are negative, signifying that quality is improving.[9] Once again, this first test does not claim to identify the causal effect of gainsharing on quality or engagement. Rather, our aim is to show that the time trend is consistent with other studies. Any change in the trend may be caused by the incentive effects of gainsharing, but it may also be caused by the effects of gainsharing on engagement (or other non-incentive effects), the simultaneous adoption of other quality improvement programs (like lean manufacturing), or to mean reversion (if gainsharing is adopted when quality-related returns are particularly problematic).

Quality is only one of the two key performance indicators targeted by the gainsharing plan. Ideally, we would also test whether engagement improved after the introduction of gainsharing. Unfortunately, the gainsharing plan also introduced our engagement measure, continuous improvement cards. As such, we are only able to track changes in quality.

Hypotheses 2A, 2B, and 2C: Effect of Gainsharing Incentives on Quality

Having examined whether our time trends replicate the results of similar studies, we turn our focus to the main contributions of this study: evaluating the incentive mechanism of the gainsharing program. Hypotheses 2A, 2B, and 2C concern how quality depends on whether the item is produced when workers perceive they are bonus-eligible. (We will turn to the issue of crowding out intrinsic motivation in Hypotheses 4A and 4B.)

Recall from the data section that we are interested in this question: When workers observe they are eligible for quality bonuses, are they less likely to produce items that are ultimately returned for quality reasons? In other words, for all items that are produced at the plant, we see the date that they were produced, whether workers see that they were bonus-eligible in that week, and then whether those items were ultimately returned. Because the plant does not distribute bonuses until all items produced in that quarter are no longer eligible to be returned for quality-related

---

[8]That is, before gainsharing is adopted, only pre_i increments and $post_i = 0$. After gainsharing is adopted, pre_i stops incrementing and $post_i$ begins incrementing.

[9]Or, alternatively, the instantaneous effect is negative and the time trend is positive if the effect of gainsharing wears off.

reasons, we are directly measuring whether plant-wide quality improves when workers perceive an incentive for plant-wide quality.

The empirical models take the form:

$$(4) \qquad Y_i = \beta_1 eligibility_i + \mathrm{XB} + \varepsilon_i$$

$$(5) \qquad Y_i = \beta_1 size_i + \mathrm{XB} + \varepsilon_i$$

$$(6) \qquad Y_i = \beta_1 eligibility_i + \beta_2 quality_i + \beta_3 eligibility_i * quality_i + \mathrm{XB} + \varepsilon_i$$

where the parameters are as follows. Y_i is the share of items manufactured in week i that are eventually returned to the plant, citing workmanship issues. $eligibility_i$ corresponds to the bonus eligibility indicator (variable savings). This variable takes a value of 1 if the workers see that the bonus eligibility indicator is green in week i. The identification strategy relies on $eligibility_i$ being determined by factors largely outside of workers' control, switching the treatment on and off. The variable $size_i$ corresponds to the size of variable savings, with larger values meaning that the size of the bonus pool is larger in week i. This information is not conveyed on the simple dashboards posted throughout the plant, but it is available on the detailed dashboards posted near the entrance, at the center of the workplace, and in the weekly newsletter. Because there are no negative bonuses, this variable cannot be negative. The variable $quality_i$ corresponds to color of the quality indicator and takes a value of 1 if the dashboards report that the plant is meeting quality targets. This information is also conveyed by the detailed dashboards and newsletter. The variable $eligibility_i * quality_i$ takes the value 1 if the plant is bonus-eligible and also meeting the quality target. Last, XB includes control variables, including the log of worker hours and the percentage of overtime paid. As robustness checks, we perform an item-level probit and perform the regression adjusting quality-related returns for the composition of items produced that week.

Hypotheses 2A to 2C are that incentives work: Fewer items are returned to the plant for quality reasons when they are produced in weeks with an incentive for quality. Specifically, Hypothesis 2A is that the coefficient on $eligibility_i$ is negative, meaning that quality improves when employees are bonus-eligible. Hypothesis 2B is that the coefficient on $size_i$ is negative, meaning that quality improves with respect to the magnitude of the incentive for quality, and Hypothesis 2C is that the coefficient on $eligibility_i * quality_i$ is negative, meaning that quality improves when employees are bonus-eligible and quality target is met.

Hypotheses 3A, 3B, and 3C: Effect of Gainsharing Incentives on Engagement

Hypotheses 3A to 3C are the same as Hypotheses 2A to 2C, except they use engagement as the dependent variable in a Poisson regression. The empirical models take the form:

(7) $$E(Y_i) = exp(\beta_1 eligibility_i + \mathrm{XB})$$

(8) $$E(Y_i) = exp(\beta_1 size_i + \mathrm{XB})$$

(9) $$E(Y_i) = exp(\beta_1 eligibility_i + \beta_2 engagement_i + \beta_3 eligibility_i * engagement_i + \mathrm{XB})$$

where $E(Y_i)$ is the expectation under maximum likelihood for the number of continuous improvement card submissions in week i as predicted by the Poisson function, $engagement_i$ is an indicator for whether workers saw that they were meeting the engagement target, and all other variables and hypotheses tests are otherwise the same as in Hypotheses 2A to 2C. The variable $engagement_i$ corresponds to color of the engagement indicator and takes a value of 1 if the dashboards report that the plant is meeting engagement targets. Also, we control for the number of employees in the firm and the percentage of hours that were overtime.

Hypotheses 4A and 4B: The Effect of Small Incentives

Because the size of the bonus pool is funded as a fraction of variable savings, the bonus pool can be arbitrarily small, even if all quality and engagement targets are met. These small incentives may have adverse effects on morale or on workers' intrinsic motivation to provide high-quality work or to engage with the employer. Theoretically, this idea draws most from Titmuss (1970) and Deci (1972), who hypothesized that incentivizing a behavior undermines individuals' civic and intrinsic motives. Since then, social scientists have found that small incentives can be worse than having no incentives at all (for a review, see Frey and Jegen 2001; Gagné and Forest 2008; Frey et al. 2013; Cerasoli et al. 2014).

Interviewed workers did report being especially demoralized when cost savings was barely above $50,000, meaning that bonuses were very small.

To evaluate whether small incentives are demotivating, we begin by examining the plant's quality and engagement at separate $50,000 bins for the bonus pool's size. In particular, following prior work on crowding out, we are chiefly interested in whether we observe an inverted-U-shaped relationship between quality and the incentive size, where quality-related returns peak in weeks that workers observe small incentives for quality. For engagement, we are interested in whether we see a U shape, where new engagement card submissions drop when bonuses are small. After performing these semiparametric tests, we perform the statistical tests.

The empirical models take the form:

(10) $$Y_i = \gamma B + \varepsilon_i$$

(11) $$Y_i = \beta_1 size_i + \beta_2 size^2_i + \varepsilon_i$$

(12) $$Y_i = \beta_0 + \beta_1 large_i + \beta_2 none_i + \varepsilon_i$$

Table 1. Weekly Summary Statistics

Variable	Mean	SD
Dependent variable		
Actual quality (quality-related item return rate)	**0.085**	**0.036**
Engagement (new continuous improvement cards)	58.1	47.8
Treatments reported on simple dashboards		
Bonus eligibility indicator is green	0.563	0.498
Quality indicator is green	0.525	0.501
Engagement indicator is green	0.731	0.445
Bonus is at 0%	0.438	0.498
Bonus is at 20%	0.138	0.345
Bonus is at 35%	0.150	0.358
Bonus is at 50%	0.275	0.448
Treatments reported on detailed dashboards		
Variable cost savings ($1,000s)	56.125	108.251
Quality-related returns	0.071	0.013
Sum of engagement (continuous improvement cards)	402.763	294.33
Target variable cost savings ($1,000s)	50	0
Target quality-related returns	.083	0.674
Target engagement (continuous improvement cards)	263.337	129.056

Notes: Summary statistics report weekly means and standard deviations over 173 weeks. Simple dashboards report color codes for whether workers are bonus-eligible, and then whether quality and engagement targets are being met. Detailed dashboards include the size of the bonus pool and how quality and engagement measures compare with the targets. SD, standard deviation.

where Y_i includes quality-related returns (LPM) and engagement (Poisson), $size_i$ is the continuous term for the magnitude of the incentive, γ includes a vector of bins for the incentive size, $large_i$ is an indicator for whether the potential bonus pool was within the top quartile for size ($87,000 to $380,000), and $none_i$ is an indicator for there being no incentive. Hypothesis 4A is that the linear term on quality β_1 will be positive and the quadratic term β_2 will be negative for quality (and the opposite for engagement), where the derivative of returns with respect to the incentive size is small and positive, reflecting that quality-related item returns are greatest when incentives are small. Hypothesis 4B is that weeks with large and no incentives are associated with fewer returned items among items produced that week (greater quality) and greater continuous improvement card submissions (greater engagement) than weeks with small incentives.

Descriptive Statistics and Preliminary Tests

Before testing the main results, we examine whether the data are sufficiently well-suited to examine the incentive effects of gainsharing. This approach requires that the plan produced meaningful variation in the eligibility and the magnitude of the incentives, and ideally, that this variation is relatively uncorrelated with time.

Table 1 presents summary statistics for quality, engagement, and bonus eligibility across 173 weeks. The first row is the dependent variable, for

actual quality. This measure shows that, across weeks, a mean of 8.5% of items manufactured at the plant are ultimately returned due to workmanship issues. Note that the dependent variable comes from *actual* quality-related returns for items produced in week i, that is, from operational data collected at least two months after workers see whether they are bonus-eligible. An average of 58 new continuous improvement cards were submitted in any week.

The plant reported sufficient variable cost savings to be bonus-eligible in 56.3% of weeks. Likewise, quality targets were being met in 52.5% of weeks, suggesting that the management set reasonable "stretch" targets for quality, lending statistical power for both the treatment and the control. By contrast, the engagement target was being met in 73.1% of weeks.

The standard deviation of variable cost savings shows substantial variation in the magnitude of the incentive. Over 173 weeks, this variable has a minimum of negative $172,000 and a maximum of positive $430,000, which could yield a quarterly bonus of more than $1,000 for any worker.

Note that the actual rate of quality-related returns (the dependent variable) is slightly higher than the mean return rate reported on the dashboards (the independent variable). This finding is because the independent variable strictly reflects the measure that workers see when they choose their effort: orders returned within one month of shipment. By contrast, the dependent variable includes any quality-related returns, including those reported by the manufacturer's distributors. The numbers reported on the dashboard that workers see can also fluctuate as the plan's administrator replaces the estimated return rates for recently produced items with the actual return rates as items are either returned or pass at the one month mark; again, workers are ultimately compensated on the actual return rates that we use for the dependent variable.

Next, we examine whether variation in the eligibility of magnitude of incentives is sufficiently dispersed over time so that the test is not confounded by time trends, as before-and-after studies are. To examine this, we look at a scatterplot of quality and engagement over time, along with whether workers are incentive-eligible. Ideally, these measures are relatively uncorrelated and dispersed over time. We would be concerned, for example, if workers were always incentive-eligible after a certain date and never incentive-eligible before that date.

Figure 2 shows actual return rates over the study period by variable savings. The figure shows that return rates and variable savings are not strongly correlated with time. Figure 2 also could provide a visual approximation of the main tests; for example, if items are less likely to be returned if they were produced over weeks in which workers observe they are eligible for incentives, then the figure's solid points will systematically be lower on the y-axis than its hollow points.

Next, we test whether it is appropriate to treat residuals as independent, rather than using time lags in an autoregressive integrated moving average

Figure 2. Bonus Eligibility Status and Quality-Related Returns
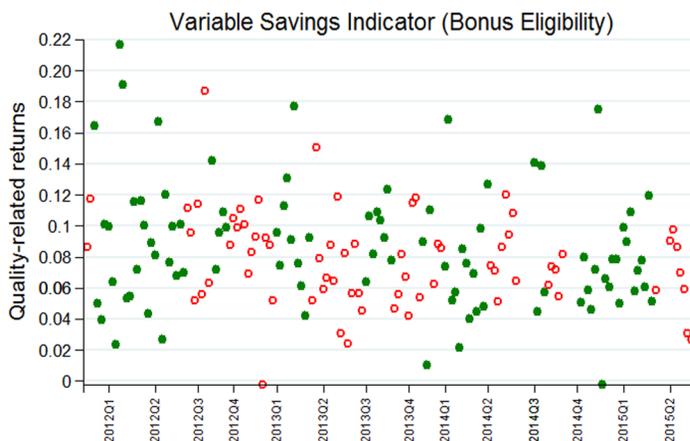
Notes: The scatterplot above relates weekly short-term quality-related return rates over time. Solid green dots represent weeks in which variable cost savings was green (being met), and the hollow red dots represent weeks in which the KPI was red (not being met). The former represent weeks in which the dashboards report that workers are eligible for the quality bonus, the latter represent those they do not.

(ARIMA) model. However, a Durbin-Watson alternative test (King 1981) finds no evidence for autocorrelation ($F = 3.178$, $p > 0.05$). We are then left to adjudicate between the more parsimonious OLS model and the fuller ARIMA model. To do so, we apply the Bayesian Information Criterion (BIC), and the Akaike Information Criterion (AIC) selection procedures. Both the BIC and the AIC test statistics favor the more parsimonious OLS over the ARIMA model.

Finally, because it is important for the identification strategy that workers cannot affect or anticipate cost savings, we examine whether changes in variable savings are consistent with workers' interests. To do so, we exploit that variable savings matter on the margin only when it crosses the $50,000 threshold, and matters most when workers are already meeting quality and engagement targets (because they are eligible for up to 50% of cost savings, rather than 20%). In other words, if workers can affect cost savings, then we would expect a positive correlation between changes in variable cost savings and being above the $50,000 threshold, and also with meeting quality and engagement targets. Contrary to workers' interests, however, we found that variable cost savings tended to fall (though not by a statistically significant amount) when incentives were in place.

Main Results

Time Trend

First, we examine whether this plant experienced an improvement in quality and engagement around the time that it implemented the gainsharing plan.

Table 2. OLS Spline Regression for Monthly Share of Items Returned, Before and
After Gainsharing Adopted

	DV: Share of items returned		DV: Log-share of items returned	
Variable	*(1)*	*(2)*	*(3)*	*(4)*
Gainsharing adoption	−0.03***	−0.006	−0.299***	−0.045
	(0.005)	(0.01)	(0.052)	(0.099)
Pre-gainsharing time trend		−0.0009*		−0.01*
		(0.0005)		(0.006)
Post-gainsharing time trend		−0.0005**		−0.006**
		(0.0002)		(0.0023)
Constant	0.114***	0.102***	−2.192***	−2.315***
	(0.004)	(0.008)	(0.0413)	(0.08)
R-squared	0.366	0.446	0.341	0.432
Observations	67	67	67	67

Notes: Standard errors in parentheses. The first two columns report the share of items produced in month i that were later returned for quality reasons. The latter two columns report the log-share. Time trend is in months. OLS, ordinary least squares; DV, dependent variable.

Table 2, column (1), because it has only one indicator variable, is functionally equivalent to a two-sample *t*-test. The constant shows that, prior to gainsharing's adoption, an average of 11.4% of items were returned to the plant for workmanship issues, and that this share subsequently declined by 2.98% to 8.42%. This improvement in quality-related returns is both statistically significant and substantive, perhaps suggesting that gainsharing may be associated with quality improvements. As shown in the time trends though, quality was improving even before gainsharing was adopted, and it continued to improve at a slightly slower rate (but not significantly slower rate) after gainsharing was adopted. Moreover, there was no significant discrete jump at the moment it was adopted, and the rate at which quality improved actually declined after gainsharing was adopted.

Overall, these results are merely meant to show that gainsharing in this setting was accompanied by improvements to quality, which other studies have found.

Incentives for Quality

We now turn to the key contribution of this article: using variation in the size of the gainsharing incentive to see whether quality and engagement improved most when gainsharing incentives were large.

Table 3 presents an OLS model for the share of items returned. Despite sufficient statistical power to detect even small differences in quality-related return rates, column (1) estimates that bonus eligibility does not yield a significant decline in item returns, and instead yields a modestly significant (at $\alpha = 10\%$) increase in quality-related item returns. Point estimates are precisely estimated, substantively very small, and in the opposite of the predicted direction: For ease of reading, the bonus size is in millions of dollars

Table 3. OLS for Weekly Share of Quality-Related Item Returns, by Incentive

Variable	(1)	(2)	(3)
Bonus-eligible	0.0082		0.0015
	(0.0053)		(0.0103)
Bonus size		−0.0375	
		(0.0459)	
Quality target met			−0.0001
			(0.0094)
Bonus-eligible*Quality target met			0.0103
			(0.0128)
Control: log worker-hours	−0.0644**	−0.0616**	−0.0680**
	(0.0206)	(0.0208)	(0.0210)
Control: percent overtime	0.0940	0.0881	0.0905
	(0.0925)	(0.0931)	(0.0931)
Constant	0.665***	0.645***	0.698***
	(0.185)	(0.186)	(0.188)
R-squared	0.086	0.071	0.110
Observations	173	173	173

Notes: Standard errors in parentheses. Bonus size is in millions. OLS, ordinary least squares.
*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

even though the range of observed bonus size over this period is only $380,000. This leads us to conclude that the bonus availability and size did not have substantive incentive effects, contrary to Hypotheses 2A to 2C.

Regarding the controls, log-hours worked and overtime go in opposite directions, and together, quality-related returns are minimized when workers are working full-time but not overtime.

To provide an intuition for the magnitude, note that workers were not eligible for a bonus in 83 weeks, over which 8,910 items were produced and 8% were returned. These weeks function as a control group in which workers are not eligible for bonuses. If we expect workers to be motivated by the bonus, then we should expect weeks in which workers were bonus-eligible to have lower return rates. However, in 90 weeks over which workers were told they were eligible for a bonus, of the 8,812 items produced, 8.75% were returned. Again, this illustrates that quality-related returns went up, rather than down, when workers had incentives for improving quality. This result is not statistically significant and is in the opposite of the hypothesized direction.

Next, we examine whether results may be driven by compositional effects. The plant produced a total of 12 different models of industrial equipment over the 3.5 years we examine, and these models vary in cost and complexity. As a result, differences in quality-related return rates may be partly due to the product mix across weeks. To correct for this, we calculate the mean return rate for each model from all weeks. Next, we calculate the expected return rate within any one week given the composition of items. Last, we subtract the actual number of returns from the expected number of returns. This yields a number that is approximately mean zero and that

represents overperformance or underperformance in quality-related returns for the week's given product mix. Results for the product-mix adjusted test are substantively the same as those in Table 3.

Robustness checks also include a probit at the level of 17,752 items for whether an item is returned using Ai and Norton's (2003) method for calculating standard errors over the support of the bonus, taking the log-percent returns, and permutations of including time trends; each yielded substantively similar results. In exploring potential morale effects, we found that whether the prior quarter ended in a bonus also did not have a statistically significant effect on quality. In exploring potential learning effects, for example, whether the effectiveness of the incentive either grew or declined over time, we found the interaction of the bonus eligibility and a time trend was again precisely estimated and nearly zero. Results are not sensitive to whether we use a multiple imputation procedure to fill missing data.[10]

To summarize, the manufacturing plant produced 17,752 items over the 173 weeks, and workers' eligibility for the quality bonus and the size of the bonus varied widely. These features allow us to produce precise estimates of the return rates when workers see that they are eligible for incentives and when they do not. Despite this power, and across specifications, we do not find evidence that quality-related item returns were lower when those items were produced when workers perceived they were eligible for incentives for quality. If anything, moderate evidence supports that quality-related returns rose. This finding leads us to reject Hypotheses 2A to 2C.

Incentives for Engagement

Next, we explore the incentive for engagement. Recall that the incentive for quality puts an additional 15% of cost savings into the bonus pool if the plant's workers meet the plant-wide quality target, whereas the incentive for engagement does so if the plant's workers meet the plant-wide engagement target *and also* the individual worker submits one continuous improvement card. As such, the incentive for engagement has both a plant-wide and an individual component and is clearly within a worker's immediate control.

As shown in Table 4, workers respond to incentives for engagement. The bonus size is positively and significantly correlated with continuous improvement submissions. The effect is magnified when the workers are bonus-eligible and the engagement target is met.[11]

To summarize these results so far, we find little evidence that quality improves when workers are eligible for the quality bonus, but we find greater evidence that continuous improvement suggestions go up when

---

[10]Results are available from the authors upon request.

[11]Once again, as a nonlinear model, the interaction term should be treated with caution; we confirm that the interaction term is statistically significant at the full support of the direct effects (i.e., from 0 to 1), and results are also robust to treating the 2×2 combinations of bonus eligibility and engagement goal attainment as a 4-category dummy.

Table 4. Weekly Count of Continuous Improvement Submissions by Incentive

Variable	Poisson			OLS
	(1)	(2)	(3)	(4)
Bonus-eligible	0.0552***		−0.432***	−0.846***
	(0.0205)		(0.0950)	(0.232)
Bonus size		0.653**		
		(0.234)		
Engagement target met			1.166***	1.353***
			(0.0571)	(0.186)
(Bonus-eligible)*			0.461***	0.827***
(Engagement target met)			(0.0975)	(0.261)
Control: Headcount	0.0318***	0.0314***	0.0257***	
	(0.0008)	(0.0008)	(0.0008)	
Control: percent overtime	−3.280***	−3.510***	−2.762***	
	(0.332)	(0.357)	(0.330)	
Constant	2.704***	2.751***	1.973***	2.738***
	(0.0434)	(0.0429)	(0.0688)	(0.164)
Pseudo *R*-squared	.329	.331	.514	.538
Observations	173	173	173	173

Notes: Standard errors in parentheses. Columns (1)–(3) report results for Poisson models. Column (4) uses OLS, where the dependent variable is the log-count of submissions. Bonus size is in millions. OLS, ordinary least squares.
*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

workers are eligible for the engagement bonus. Together, these results suggest that something about the engagement measure makes this measure more responsive to gainsharing incentives—perhaps that it is more within workers' direct control and that individual workers cannot free ride because they must participate to be eligible for the bonus.

Do Small Incentives Crowd Out Intrinsic Motivation?

We examine whether quality and engagement are consistent with crowding out. If so, we would expect to see that items produced in weeks when the bonus pool was small to be most likely returned for quality-related reasons, and continuous improvement card submissions to be fewest when the bonus pool was small. We begin with the semiparametric analysis of quality and engagement within discrete $50,000 bins.

Figure 3 shows the mean quality-related return rates and new continuous improvement card submissions at given variable cost savings bins (the solid dots), the standard errors of those estimates (the vertical bars), and the predicted relationship between those outcomes and variable cost savings using a quadratic OLS regression (the dashed lines). Quality-related return rates are highest when the variable cost savings is at its smallest but still positive. When variable savings were between $50,000 and $100,000, quality-related return rates were significantly higher (mean 9.1%, standard error 0.5%) than they were when variable savings were negative (mean 8.1%, standard

Figure 3. Quality-Related Returns and New Continuous Improvement Cards by
Variable Savings
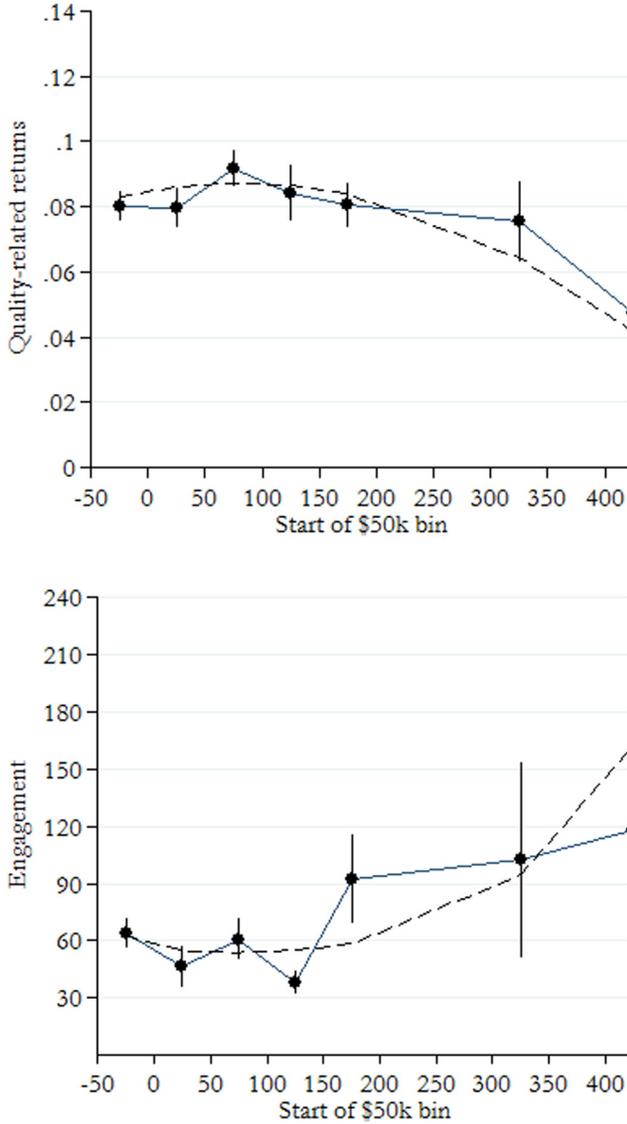
Notes: The figure relates quality-related return rates and engagement by variable savings bins, where bins are $50,000 wide. Vertical bars represent standard errors and the dashed line represents an OLS line where return rates and engagement is quadratic in variable savings.

error 0.3%) or when variable savings were greater than 150 (mean 7.8%, standard error 0.5%).

We also test for crowding out in a regression framework. Specifically, we first regress the rate of item returns against linear and quadratic terms of the bonus size. These regressions yield a positive and significant linear term and a negative and significant quadratic term. Taking derivatives, the

regression estimates that quality-related returns peak when variable cost savings are \$90,500. Note that incentives do not begin until variable cost savings reach \$50,000, and so incentives are smallest but positive when these variable cost savings are slightly above \$50,000. This test is consistent with crowding out and Hypothesis 4A. Similarly, we do the same for engagement, for which the linear term is negative and significantly different from 0, while the quadratic term is positive and significantly different from 0. Together, the estimated linear and quadratic coefficients estimate that engagement minimum is when variable cost savings are \$72,720. We find similar results for both quality and engagement when we discretize bonus pools into \$50,000 bins. Again, these results are consistent with Hypotheses 4A and 4B.[12]

Institutional Benchmarks and Generalizability

To evaluate external validity and scope conditions, we should distinguish how the gainsharing program at this plant compares to other gainsharing plans and other shared bonus pools. We should also be mindful to situate our results within the broader set of studies on gainsharing and shared bonus pools, especially as this study's distinct method and questions suggest lessons for the theoretical mechanisms and practical design of incentive plans.

Most aspects of the gainsharing program are familiar—bonuses are funded from variable cost savings, bonuses reward quality and engagement, and gainsharing was part of a larger set of practices. This plan's most notable design parameter is perhaps that it measures and rewards quality at the group level, and engagement requires both individual and group level targets to be met. This feature of the quality KPI leads us to extend the peer monitoring and team incentives literature to this setting, but it is unlike some other implementations of gainsharing, such as that at Lincoln Electric. Shared bonus pools more broadly also often vary with regard to the extent they reward individual or group performance metrics. For example, corporate bonus pools may be allocated to subdivisions based on group performance, and then those subdivisions may allocate a larger or smaller share of that subdivision's bonus pools to individuals based on their performance.

We should also distinguish how labor management relations at the plant are compared to others. Like 90.6% of US manufacturing workers, workers at this plant are not unionized (BLS 2016a). The mean turnover at our plant is 6.3%. In the United States, manufacturing has the longest tenure among major industry categories, which at 5.3 years is still far less than our setting (BLS 2016b).

---

[12]Results are available upon request.

More important, for the purposes of this study, we might consider whether workers' perceptions of compensation and motivation were exceptional. Fortunately, during the end of 2011, right before gainsharing was implemented, the company hired consultants to perform a climate study, which featured a survey to which 72% of the workers who were later covered by the gainsharing plan responded. On labor relations and compensation questions, workers' responses were similar to benchmarks for trust in management (5.21 compared to 5.30 out of 7.0), on the perceived fairness of compensation and rewards (4.81 compared to 4.80), and perceived motivation to provide discretionary effort (6.16 compared to 6.30). The plant performed more poorly on engagement (4.63 compared to 5.50) and workers' perception of the company's commitment to customer satisfaction (5.31 compared to 5.80), which contributed to the plant's later adoption of the gainsharing program. Ideally, we would like to know how workers' perceptions compare during the gainsharing program, but unfortunately, such data are not available. Given that quality did improve and that gainsharing programs are fairly common, we could only speculate that this plant was not too dissimilar from the "average" US manufacturing plant, or at least those that would also hire a consultancy to perform such a benchmark. If labor management relations at this plant were nonetheless too weak for workers to trust the credibility of the bonus program, then it suggests that many plants would need to overcome a high bar for credibility, perhaps starting with making the underlying mechanics behind bonus plans as simple and transparent as possible.

This brings us to our second point—the generalizability of the research method. Academic work has emphasized incentive regimes for which the size of the incentive is fixed; in practice, however, incentives that vary for factors outside of workers' control are rather common and offer ripe opportunities for future empirical work. For this study, the determinants of variable cost savings were sufficiently obfuscated so that we could not isolate an outside variable that is easily measured and clearly correlated with the size of the bonus pool. One might imagine that a future study could, for example, use the price of oil as an instrument for the size of the bonus pool at an energy company that apportions bonuses according to individual- or group-level performance indicators. Indeed, Bertrand and Mullainathan (2001) found that executive pay depends largely on factors, as with oil prices, that are outside of their control; similarly, this could be used as an instrument for the size of bonus pools and the incentives that they fund. Given that shared bonus pools could potentially operate through a variety of theoretical mechanisms with different practical implications, the research opportunity offered by exogenous variation in the size of shared bonus pools may be underappreciated.

Limitations and Future Directions

Although we hope further work uses this method to advance our understanding of shared bonus pools, this setting has some notable limitations.

First, although the workers and management we interviewed reported that variable cost was not within workers' control, and we could not find evidence that variable cost moves in ways consistent with their incentives, we may never know for certain whether workers redirect effort toward cost savings when they see they are incentive-eligible. This missing piece is a concern because, when workers observe they are eligible for incentives, they might shift their effort away from quality and toward boosting variable cost savings, thereby creating a complicated optimization problem in which the size of the potential bonus pool grows when they try to maximize cost savings, but their share of the bonus pool grows as they redirect effort toward quality. Again, this scenario is not consistent with anything we have seen in the plant, but it is nonetheless a concern that could be addressed if incentive sizes were set experimentally or by some observable and more clearly exogenous determinant. For example, if an oil company distributes incentives as points toward a corporate bonus pool, and the bonus pool budget depended on the price of oil, then a researcher could use the price of oil as an instrument for the magnitude of incentives to identify that bonus pool's incentive effect.

Second, gainsharing plans vary in their setting and their implementation. In a traditional Scanlon plan, workers should feel that cost savings are within their control. In a classic incentives model in which workers are more risk averse than their employer, employers should remove variable costs that are clearly outside workers' control (such as suppliers' prices) and align incentives to those that are within workers' control. Although this circumstance does not appear to be the case here, this plant's experience might simply provide further evidence for what other plants already consider a best practice.

Third, our study focuses on a rather limited question: Do gainsharing bonuses have incentive effects? Our results do not positively affirm what it is about gainsharing or other accompanying practices that makes them complementary. Rather, our study suggests that gainsharing's benefits operate primarily through other mechanisms, such as a sense of shared ownership.

Fourth, our study invokes crowding-out theory to explain why small incentives correspond to lower engagement and quality. In this setting, however, negative variable cost savings do not necessarily correspond to no bonuses, but rather a small probability of small bonuses. Self-determination theory, which is the central explanation for crowding out within psychology, hypothesizes that small rewards demotivate because the required measurement reduces workers' autonomy; however, the gainsharing plan is meant to engage workers and promote autonomy, and their performance is measured whether the bonus pool is zero, small, or large. Our invocation of crowding-out theory generally, and self-determination theory specifically, may strain what we really know about small probabilistic incentives. Although these results are statistically significant, the magnitudes are small and may not be substantively significant.

We hope that further studies will use variation in the size of shared bonus pools that fund specific incentives, like those typified in gainsharing, and will further shed light on these questions.

Conclusion

Whereas much of the work on incentives has focused on fixed incentive schedules, shared bonus pools with budgets that can vary based on outside factors are also ubiquitous. Because these outside factors are often beyond workers' control, our approach offers researchers and companies a simple method for examining how workers actually respond to the incentives that bonus pools fund.

We take this insight to a manufacturing plant's gainsharing program. Our institutional features and results highlight the general challenge of isolating the effectiveness of such a program, but illustrate how variation in bonus pools can be used to isolate and study their incentive effects.

Specifically, we find that quality did not statistically improve when workers were told they were eligible for quality bonuses. Rather, there may have been an increase in items returned to the plant for quality reasons when the bonus pool was particularly small. Workers did submit more continuous improvement cards when they were eligible for gainsharing incentives, and the plant never missed the total worker engagement target in quarters when the plant was bonus-eligible. In contrast to the quality incentive, this second incentive had an individual component (a worker had to submit a continuous improvement card to be eligible for the 15% bonus), and it was also clearly within workers' control.

After the analysis was complete, we presented these results to the plant's human resource officer, the operational engineers for all three functional areas, the chief operating officer, and the plant controller. In this meeting, the plant's management noted that the original intention of the gainsharing plan was to avoid paying incentives that could exacerbate a bad quarter, and to otherwise use its incentives to motivate workers to improve quality and engagement. Even prior to the meeting, the plant's managers were suspicious about whether the bonus itself was having its intended effect. We cautioned that the existence of the gainsharing plan could have had benefits (e.g., a sense of shared ownership) even if there was no evidence that workers were responding directly to the presence of incentives. Based on these results though, the management ultimately concluded that the gainsharing incentives were not meeting their intended objectives, giving pause to the planned roll out of gainsharing incentives at other plants to standardize compensation across their global manufacturing operations.

Overall, this plant's experience implies lessons for gainsharing plans, especially when taken in context of other studies. First, we may expect that the primary benefits of the gainsharing plan may not come directly from the incentives themselves but rather from the worker engagement and process

improvement programs that typically accompany these programs. Second, managers should be cautioned that gainsharing and other forms of shared bonus pools may incidentally distribute rewards that are trivially small, and this may have a demotivating effect compared to the results of offering no incentive at all. Third, the differences between engagement and quality suggest that performance measures should be tied to smallest measurement units (such as individuals and groups) rather than to plants as a whole. In some sense, this finding is vindication for the "$1/N$" problems with incentives, which have been challenged by other empirical work. It is also a notable difference between this plant and the well-known case of Lincoln Electric, which also featured a shared bonus pool that fluctuated with the firm's profitability, except that incentives were distributed on the basis of individual performance metrics. As we emphasized to this plant's management, we do not claim that gainsharing plans do not work, but rather that little evidence supports they work through the incentive effect. Gainsharing might instead be emphasized as a program for providing workers with a sense of shared governance rather than incentives, and management practice and communication should reflect this.

For academics, this study serves as an example of how we can use variation in the size of shared bonus pools to study incentives. In our plant, bonus eligibility would change quickly and unexpectedly in response to a large and complicated set of factors. This outcome could open the door for much more empirical work on how workers' respond to the precise implementation of incentive programs.

References

Ai, Chunrong, and Edward Norton. 2003. Interaction terms in logit and probit models. *Economics Letters* 80(1): 123–29.

Arthur, Jeffrey B., and Lynda Aiman-Smith. 2001. Gainsharing and organizational learning: An analysis of employee suggestions over time. *Academy of Management Journal* 44(4): 737–54.

Arthur, Jeffrey B., and Christopher L. Huntley. 2005. Organizational learning curve: Assessing the impact of deliberate learning on organizational performance under gainsharing. *Academy of Management Journal* 48(6): 1159–70.

Bartel, Ann P., Richard B. Freeman, Casey Ichniowski, and Morris M. Kleiner. 2011. Can a workplace have an attitude problem? Workplace effects on employee attitudes and organizational performance. *Labour Economics* 18(4): 411–23.

Berg, Norman A. 1975. Lincoln Electric Co. Harvard Business School Case 376-028. Boston, MA.

Bertrand, Marianne, and Sendhil Mullainathan. 2001. Are CEOs rewarded for luck? The ones without principals are. *Quarterly Journal of Economics* 116(3): 901–32.

Blasi, Joseph, Michael Conte, and Douglas Kruse. 1996. Employee stock ownership and corporate performance among public companies. *Industrial and Labor Relations Review* 50(1): 60–79.

Bloom, Matt. 2008. A century of compensation research. In Julian Barling and Cary L. Cooper (Eds.). *The SAGE Handbook of Organizational Behavior*: Volume One. *Micro Approaches*. Thousand Oaks, CA: SAGE Publications.

[BLS] U.S. Bureau of Labor Statistics. 2016a. Accessed at https://www.bls.gov/news.release/pdf/union2.pdf (July 14, 2017).

———. 2016b. Accessed at https://www.bls.gov/news.release/pdf/jolts.pdf (July 14, 2017).

Bryson, Alex, Richard Freeman, Claudio Lucifora, Michel Pellizzari, and Virginie Perotin. 2012. Paying for performance: Incentive pay schemes and employees' financial participation. CEP Discussion Paper. London: Centre for Economic Performance, London School of Economics.

Cerasoli, Christopher P., Jessica M. Nicklin, and Michael T. Ford. 2014. Intrinsic motivation and extrinsic incentives jointly predict performance: A 40-year meta-analysis. *Psychological Bulletin* 140(4): 980–1008.

Deci, Edward. 1971. Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology* 18(1): 105–15.

———. 1972. Intrinsic motivation, extrinsic reinforcement, and inequity. *Journal of Personality and Social Psychology* 22(1): 113–20.

Drago, Robert, and John S. Heywood. 1995. The choice of payment schemes: Australian establishment data. *Industrial Relations* 34(4): 507–31.

Dube, Arindrajit, and Richard Freeman. 2010. Complementarity of shared compensation and decision-making systems: Evidence from the American labor market. In Douglas L. Kruse, Richard B. Freeman, and Joseph R. Blasi (Eds.), *Shared Capitalism at Work: Employee Ownership, Profit and Gain Sharing, and Broad-Based Stock Options*, pp. 167–99. Chicago: University of Chicago Press.

Fernie, Sue, and David Metcalf. 1995. Participation, contingent pay, representation and workplace performance: Evidence from Great Britain. *British Journal of Industrial Relations* 33(3): 379–415.

Freeman, Richard, Douglas Kruse, and Joseph Blasi. 2010. Worker responses to shirking under shared capitalism. Published to Chicago Scholarship Online, University of Chicago Press. doi: 10.7208/chicago/9780226056968.001.0001. Also accessed at NBER Working Paper Series. 2008. URL: http://www.nber.org/papers/w14227.pdf. Cambridge, MA: National Bureau of Economic Research.

Frey, Bruno, and Reto Jegen. 2001. Motivation crowding theory. *Journal of Economic Surveys* 15(5): 589–611.

Frey, Bruno S., Fabian Homberg, and Margit Osterloh. 2013. Organizational control systems and pay-for-performance in the public service. *Organization Studies* 34(7): 949–72. Also accessed at http://bournemouth33.eprints-hosting.org/20879/4/Organization_Studies-2013-Frey-949-72.pdf (December 26, 2016).

Gagné, Marylène, and Jacques Forest. 2008. The study of compensation systems through the lens of self-determination theory: Reconciling 35 years of debate. *Canadian Psychology* 49(3): 225–32.

Gittleman, Maury, Michael Horrigan, and Mary Joyce. 1998. "Flexible" workplace practices: Evidence from a nationally representative survey. *Industrial and Labor Relations Review* 52(1): 99–115.

Gneezy, Uri, and Aldo Rustichini. 2000. Pay enough or don't pay at all. *Quarterly Journal of Economics* 115(3): 791–810.

Gneezy, Uri, Stephan Meier, and Pedro Rey-Biel. 2011. When and why incentives (don't) work to modify behavior. *Journal of Economic Perspectives* 25(4): 191–210.

Hamilton, Barton H., Jack A. Nickerson, and Hideo Owan. 2003. Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *Journal of Political Economy* 111(3): 465–97.

Hatcher, Larry, and Timothy L. Ross. 1991. From individual incentives to an organization-wide gainsharing plan: Effects on teamwork and product quality. *Journal of Organizational Behavior* 12(3): 169–83.

Huselid, Mark A. 1995. The impact of human resource management practices on turnover, productivity, and corporate financial performance. *Academy of Management Journal* 38(3): 635–72.

Ichniowski, Casey, Kathryn Shaw, and Giovanna Prennushi. 1997. The effects of human resource management practices on productivity: A study of steel finishing lines. *American Economic Review* 87(3): 291–313.

Jones, Derek C., Takao Kato, and Jeffrey Pliskin. 1994. Profit sharing and gainsharing: A review of theory, incidence and effects. Economics Working Paper No. 125. Annandale-on-Hudson, NY: Levy Economics Institute, Bard College.

Kandel, Eugene, and Edward P. Lazear. 1992. Peer pressure and partnerships. *Journal of Political Economy* 100(4): 801–17.

Kim, Dong-One. 2005. The choice of gainsharing plans in North America: A congruence perspective. *Journal of Labor Research* 26(3): 465–83.

King, Maxwell L. 1981. The alternative Durbin-Watson test: An assessment of Durbin and Watson's choice of test statistic. *Journal of Econometrics* 17(1): 51–61.

Knez, Marc, and Duncan Simester. 2001. Firm-wide incentives and mutual monitoring at Continental Airlines. *Journal of Labor Economics* 19(4): 743–72.

Koller, Frank. 2011. Spark: *How Old-Fashioned Values Drive a Twenty-First-Century Corporation: Lessons from Lincoln Electric's Unique Guaranteed Employment Program.* Philadelphia, PA: PublicAffairs, Perseus Books Group.

Kruse, Douglas. 1993. Does profit sharing affect productivity? NBER Working Paper Series No. 4542. Cambridge, MA: National Bureau of Economic Research.

———. 1994. Profit sharing and public policy. *Journal of Economic Issues* 28(2): 439–48.

———. 1996. Why do firms adopt profit-sharing and employee ownership plans? *British Journal of Industrial Relations* 34(4): 515–38.

Kruse, Douglas L., Joseph R. Blasi, and Rhokeun Park. 2010. Shared capitalism in the US economy: Prevalence, characteristics, and employee views of financial participation in enterprises. In Douglas L. Kruse, Richard B. Freeman, and Joseph R. Blasi (Eds.), *Shared Capitalism at Work: Employee Ownership, Profit and Gain Sharing, and Broad-Based Stock Options*, pp. 41–75. Chicago: University of Chicago Press.

MacDuffie, John Paul. 1995. Human resource bundles and manufacturing performance: Organizational logic and flexible production systems in the world auto industry. *Industrial and Labor Relations Review* 48(2): 197–221.

Milkovich, George T., Jerry M. Newman, and Barry Gerhart. 2013. *Compensation*, 11th edition. New York: McGraw-Hill Education.

Nalbantian, Haig R., and Andrew Schotter. 1997. Productivity under group incentives: An experimental study. *American Economic Review* 87(3): 314–41.

Osterman, Paul. 1994. How common is workplace transformation and who adopts it? *Industrial and Labor Relations Review* 47(2): 173–88.

Petty, M. M., Bart Singleton, and David W. Connell. 1992. An experimental evaluation of an organizational incentive plan in the electric utility industry. *Journal of Applied Psychology* 77(4): 427–36.

Piore, Michael J., and Charles F. Sabel. 1984. *The Second Industrial Divide: Possibilities for Prosperity.* New York: Basic Books.

Poole, Michael, and Keith Whitfield. 1994. Theories and evidence on the growth and distribution of profit sharing and employee shareholding schemes. *Human Systems Management* 13(3): 209–20.

Rubin, Donald B. 2004. *Multiple Imputation for Nonresponse in Surveys.* Hoboken, NJ: John Wiley & Sons.

Rynes, Sara L., Barry Gerhart, and Laura Parks. 2005. Personnel psychology: Performance evaluation and pay for performance. *Annual Review of Psychology* 56: 571–600.

Schafer, Joseph L., and Maren K. Olsen. 1998. Multiple imputation for multivariate missing-data problems: A data analyst's perspective. *Multivariate Behavioral Research* 33(4): 545–71.

StataCorp LP. 2013. Stata multiple-imputation reference manual. Accessed at https://www.stata.com/manuals13/mi.pdf (July 2016). College Station, TX: A Stata Press Publication.

Titmuss, Richard. 1970. *The Gift Relationship.* London: George Allen.

Wagner, John A., Paul A. Rubin, and Thomas J. Callahan. 1988. Incentive payment and non-managerial productivity: An interrupted time series analysis of magnitude and trend. *Organizational Behavior and Human Decision Processes* 42(1): 47–74.

Weibel, Antoinette, Katja Rost, and Margit Osterloh. 2007. Crowding-out of intrinsic motivation: Opening the black box. Unpublished Working Paper. Accessed at http://dx.doi.org/10.2139/ssrn.957770.

Weitzman, Martin, and Douglas Kruse. 1990. Profit sharing and productivity. In Alan S. Blinder (Ed.), *Paying for Productivity: A Look at the Evidence.* Washington, DC: The Brookings Institution.