Analyzing the Effects of Variable Performance Based Pay Programs on Productivity: Evidence from a Field Experiment

Andrew Aicher, Colgate University
Advisor: Professor Carolina Castilla and Professor Takao Kato

Abstract

Using data from an experiment at two manufacturing plants in Yangzhou, China, I analyze the effects of both high and low team leader bonuses on team member productivity. Through a fixed effects regression, I estimate that team members whose leader receives a bonus above the median level perform roughly the same as compared to a previous gift exchange program, while those team members whose leaders received a bonus at or below the median decrease their productivity by roughly 29 percent. This conclusion both supports the hypothesis that a positive reward is more beneficial to productivity than a negative one and raises questions as to the total effect of the program as a whole. Further analysis to analyze the nature of these effects is warranted.

I. Introduction

Regardless of the context, there has always been a debate about the best way to motivate: the "carrot", or the "stick"? Is it positive reinforcement, or negative feedback that drives us the most? While this discussion takes place across all aspects of life, it has especially interesting implications when applied to the labor force. The focus of this paper is to analyze a nuanced version of employee motivation that incorporates both a "carrot" and, to some extent, a "stick" methodology. It is widely purported throughout economic literature that variable performance based systems are effective in increasing employee performance (see, for example, Bhattacherjee 2005). For the purposes of this paper, a variable performance based pay system is defined as one that pays different employees different amounts, which correspond to their relative levels of production. The focus of this paper is to analyze the mechanisms of such a system. What drives an increase in employee productivity? Do those employees getting paid the most drive continued productivity increases (the "carrot")? Do those employees getting paid the least work to increase their productivity in light of their lower relative productivity (the "stick")? Is it in fact the case that the system motivates equally across all levels of pay? Answering these questions will not only speak to the true effectiveness of a variable performance based pay system, but it may also define conditions under which these systems are optimal, and when they are not.

Specifically, this paper analyzes data from an eighteen month long experiment conducted at two manufacturing plants in China. In these plants, teams produce various products, and team leaders are responsible both for individual production and team organization. Team leaders in one of these plants were exposed to two different types of bonus systems for this experiment. The first was a form of gift exchange; all leaders received an equal bonus that was not tied to team performance. The second was a variable performance based bonus, and it is this treatment

that this paper will focus on. Additionally, there are eight months' worth of observations prior to any of the experimental treatments, which allows for individual fixed effects to be controlled for. With my analysis, I estimate the effect of relative performance pay of team leaders on the rest of the team members. I further examine whether or not there is a heterogeneous effect dependent upon positive and negative feedback. In particular, I analyze the effect on team performance of a leader earning a bonus at or below the median (a "stick" team, for the purposes of my analysis) as compared to a leader earning a bonus above the median (a "carrot" team). Additionally, given the nature of the experimental design, I am able to investigate the overall effect on team member performance of changing the team leader bonus program from one of gift exchange to a program of variable performance based pay.

My analysis has resulted in several findings. Firstly, the impact of the variable performance pay program has the same impact on productivity as the gift exchange program for "carrot" teams. For "stick" teams, however, the transition results in a significant decrease in productivity. Additionally, when "stick" and "carrot" teams are redefined in a more extreme sense (increasing the distance from the median a team's leader must be for them to qualify as either "carrot" or "stick"), the transition from gift exchange to variable performance pay has a positive impact on "carrot" team productivity as well as a negative impact for "stick" teams of an even larger magnitude. While the efficacy of both gift exchange programs (Romero 2010) as well as variable performance pay programs (Bhattacherjee 2005) is well documented by existing literature, my comparison of the two programs, as well as the decomposition of the variable pay performance program can add valuable insight to the ongoing conversation.

The rest of this paper will be organized as follows: Section II will be a review of relevant literature from both economics and psychology. Section III will incorporate both a detailed

description of the experimental design and the data utilized. Sections IV, V, and VI include a description of my specific econometric model, hypotheses and results, as well as discuss the results of several robustness checks. Section VII will discuss alternative interpretations of my results that may be valid and section VIII will conclude.

II. Literature Review

While this paper focuses on the economic implications of a variable performance based pay system, it is important to note that there is a vast literature in psychology surrounding performance based pay. Rynes (2005) addresses many of the psychological hypotheses that have been posed. Historically, psychologists have hypothesized that performance evaluation and extrinsic motivation would crowd out intrinsic motivation and ultimately be detrimental to performance. Empirical analyses, however, have consistently failed to support these hypotheses. Additionally, Rynes (2005) purports that on average, performance increases with negative feedback, a finding that is potentially of great importance to my analysis.

In a more in depth psychological analysis, Strombach (2015) finds that pay for performance systems result in increased productivity by affecting reward related regions of the brain, as opposed to task related neural representations. Additionally, Strombach (2015) notes the large decrease in productivity that arises when pay for performance systems are removed, implying that the longevity of a pay for performance system is a major determinant of its overall effectiveness.

Encinosa III (2006) analyzes group behavior and how different groups respond to different types of incentive pay programs. Two conclusions that are drawn in this sociological analysis that are echoed in much of the economics literature is that incentive strength decreases with

group size and that in general workers in a group view themselves as worse off if they make less money relative to other group members.

There is a vast literature in economics regarding variable performance based pay systems. This literature analyzes both group and individual scenarios, and it is widely (although not exclusively) found that the introduction of variable performance based pay systems leads to increased productivity. Lazear (2000) made significant contributions to this field of study by recognizing two different mechanisms by which these payment plans increase productivity. The first is through the direct effect of the incentive, and the second is sorting. In a study of the glass repair company Safelite's complete overhaul of their payment system, Lazear (2000) found that not only did an incentive based payment system increase average production, but it attracted a more capable workforce and a piece rate wage led to increased variation in individual production. The identification of sorting, or the idea that "pay-for performance" schemes attract a more capable workforce, fills a void left by the incentive explanation, and as is purported in Lazear (2003), explains the dynamics of incentive based pay when incentives alone cannot.

As my analysis involves a combination of both individual and group incentives, it is important not only to make the distinction between the two, but also to recognize the different implications of each. Like Lazear (2000; 2003), Levenson (2011) analyzes the impact of the introduction of a variable performance based pay system on individuals. In this study of Direct store delivery workers who are responsible for delivering their companies product, stocking retail shelves, and negotiating with management for more shelf space and the sale of new products, a new payment plan is introduced in which sales growth (as opposed to sales revenue) is rewarded with higher pay. As part of this plan, those workers at or above a certain rate of growth made the same or more than they did under the previous payment plan, and those below

the benchmark made less. The conclusions from this study are in line with those of Lazear (2000) and Lazear (2003) in that the introduction of a variable performance based pay system leads to increased productivity.

The existing literature regarding group incentive payment plans is vast and covers a wide variety of conditions and results. Bhattacherjee (2005) is a study of an Indian firm that employs three different types of group incentive pay programs. Two of these programs are forms of profit sharing, or firm wide bonuses based on overall firm production. The third, and that which is found to be most effective in increasing productivity, is a more decentralized payment system that rewards smaller groups for increased production. This finding is supported not only by the psychological literature cited above, but also by other studies, such as Zenger (2000) which found not only that incentive intensity and group size have a negative relationship, but also that management participation and longevity of payment plans have a positive relationship with incentive strength.

While the studies discussed above represent a significant literature that speaks to the positive impacts of group incentive payment plans, there exists also a literature with more ambiguous findings. In a study of 189 stores of a retail chain in which stores could earn a bonus for outperforming three other stores for four weeks, Deflagaauw (2010) found that stores who were close to others in relative productivity increased their productivity, but stores that lagged far behind did not. This study found that on average there was no increase in productivity as a result of this relative performance pay plan. Other studies that make similar findings include Shaw (2002), which found that on teams with high levels of interdependence, variable pay plans have no effect, and possibly even a negative effect, on overall performance.

As is stated above, my study is one that involves a combination of individual and group incentives. While team leaders are the only people receiving variable bonus payments, team member performance is the variable of interest. There is a vast literature that exists regarding incentive pay plans for teachers. This literature is an interesting point of comparison for my analysis, in that teachers are the only ones receiving bonus pay, but it is student performance that is the variable of interest. Muralidharan (2009) is a study of 300 schools in India in which both school wide and individual bonuses for increased student performance were introduced. Both types of programs resulted in increased student achievement, but the individual programs were found to have a more lasting effect. Imberman (2012) is a study that uses the share of students within a given department that a teacher interacts with as a proxy for incentive strength. This method is in line with the generally supported hypothesis that incentive strength decreases with group size. That is to say, as student share increases, an individual teacher has less of an effect on individual student performance. This study concludes that there is no significant relationship between teacher incentive plans and student achievement. Brehm (2015) is a study of a merit based pay tournament in which teachers receive awards when their students reach certain levels of achievement. The hypothesis in this study is that teachers who are closest to an award threshold will increase their effort the most and therefore their students will exhibit the greatest increases in achievement. The results of the study, however, find no relationship between the tournament and student achievement, and attribute this finding to an inability to accurately relay feedback about teacher ability and success, highlighting the need for precision in incentive pay system design.

Similarly to the labor literature, the education literature discussed above includes both positive and negative findings. Both areas of study put forth the conclusion that longevity of

performance pay systems is integral to the success of such a program, and they both support the idea that the positive effects of group incentives decrease with group size. Given the contradictory findings within the literature, continued analysis of both individual and group incentives is warranted.

In addition to the aforementioned literature, two additional studies were invaluable to my own, as both of them analyze the same data that I do, and their description and organization of the data was integral to my analysis. The first, Romero (2010) focuses solely on the gift exchange portion of the experiment, and finds not only that the program results in increased employee performance, but also that team leaders decrease their individual performance in order to better organize their teams for increased overall team production. The second, Pindiwe (2015), investigates the impact of the introduction of extrinsic motivation on intrinsic motivation, and finds no evidence of the destruction of intrinsic motivation due to the introduction of extrinsic motivation. My results support the conclusions of both of these previous studies, and adds to them with its more in depth analysis of the introduction of the variable performance pay system.

III. Data and Experimental Set-up

The data for this study is drawn from an eighteen month long experiment conducted at two manufacturing plants in Yangzhou, China. Observations were collected daily at an individual level and cover various aspects of day-to-day operations at the firm. During the first eight months of observations, from October 2008 until June 2009, no changes were made in the firm. During this time, bonus pay for team leaders was variable and averaged 100 RMB. These observations create a control period, and allow for individual fixed effects to be controlled for. On July 1, 2009, the CEO of the firm announced that for the next three months; July, August and September, team leaders at one of the plants (the treatment plant) would receive an additional

bonus of 200 RMB per month that was not tied to performance, while operations at the second plant (the control plant) continued as before. While this program was introduced on July 1, the payment of the bonus lagged its announcement by three months. Payment of bonuses began on October 1, 2009, and continued on a monthly basis, with each monthly bonus given out being payment for the month exactly three calendar months previous. On January 1, 2010, the CEO announced that the bonus payments for the months of October, November, and December 2009 (whose distribution began on January 1, 2010) would be variable and tied to team performance. That is to say, the bonus payment for the month of October, paid out in January, was one of five values, 180 RMB, 190 RMB, 200 RMB, 210 RMB or 220 RMB. The same is true for the bonus payments for the months of November and December 2009, paid out in February and March 2010.[1]

It is important to note potential flaws in this experimental design. Because the period where a variable bonus is assessed to each team leader takes place before the announcement of the program, there is a certain degree of deception of team leaders taking place. There is no reason why team leaders would not believe that during October, November, and December 2009 they are still under the gift exchange program, and that their bonuses paid out in January, February, and March 2010 will be indicative of that program. When they find out in January 2010, however, that the past three months had in fact involved a variable performance pay program, they may be surprised (either positively or negatively, depending on the level of bonus pay they receive relative to that of the gift exchange program), and their effort and performance may change accordingly.

[1] See tables 9 and 10 for detailed layouts of the experiment.

Despite this potential confounding issue, the implications of which will be discussed in detail in a later section, this study assumes that in January, February, and March 2010, team leaders are newly aware of the variable performance pay system in place and are acting accordingly. It is this three-month period where I aim to investigate the potentially dynamic effects of such a system, by analyzing the different impact the program has on "carrot" and "stick" teams.

Workers at each of the plants are divided into a total of 16 teams and generally have one or two team leaders. The plants produce five different products, and any employee can produce any of the products. Employees are paid a piece rate wage based on pre-determined earned hours for producing one unit of a given product. Team leaders receive bonus pay in addition to their piece rate wage for organizing team members and assigning employees to the production of certain products. Summary statistics of team leader performance during each of the three relevant time periods, as well as for those teams that are defined as either "carrot" or "stick" teams during the variable performance pay system support the conclusions of Romero (2010) that team leaders may potentially reduce their personal production during the gift exchange period in an effort to increase overall team production through better team organization.[2] Summary statistics of team member (leaders excluded) for each of the same periods are inconclusive, and provide additional motivation for further investigation. While mean team member productivity increases from 1.10 to 1.27 from the control period to the gift exchange program, it decreases to 1.08 as a result of the transition from the gift exchange program to the variable performance pay program, while "carrot" teams have an overall mean productivity of 0.99 and "stick" teams average 1.38.[3] Additionally, an analysis of a kernel density estimate for both "carrot" and "stick" teams during

[2] See table 6 for a detailed breakdown of team leader productivity.
[3] See table 5 for a detailed breakdown of team member productivity.

each of the treatment periods also motivates further investigation, as there is no obvious conclusion to be drawn from them.[4]

For the purposes of my analysis, the three teams at the control firm as well as 4 teams in the treatment firm were excluded from the dataset. The 4 teams that were removed from the treatment firm were removed because they did not have sufficient team leader data, so the impact of the bonus program on their performance could not be analyzed. The control firm was removed because its production was consistently and considerably higher than that of the treatment firm for the entirety of the experiment and since it was, for the entirety of the experiment, treated the same as treated teams in the control period, it had misleading impacts on the results of my analyses. The final dataset that I used for my empirical testing consisted of 214 workers on 9 teams in the treatment firm. There was a significant amount of switching between teams that took place over the observation period for certain workers, which allows team fixed effects to be controlled for.

IV. Econometric Model

The dependent variable in all of my analysis is a measure of team member productivity, which is calculated as the earned hours across all products for an employee in a given day divided by the amount of time they worked during that day.[5] Additionally, I use individual, time and team fixed effects to isolate the effect of the variable bonus payments on employee productivity. In order to analyze the overall impact of the variable bonus system, I needed to create three dummy variables to control for all of the bonus payment changes that occurred in the firm. The first variable, "Gift Exchange" is indicative of the transition from the control period to the first treatment period, or the point at which the 200 RMB bonus is added to the initial

[4] See tables 7 and 8.
[5] See table 4 for relevant summary statistics.

variable bonus. Chronologically, this variable takes the value zero for all dates up to and including June 30, 2009, and takes the value one for all subsequent dates. The second, "Performance Pay," indicates the time at which the second treatment period started, during which the 200 RMB bonus was removed and the variable bonus, which averaged 200RMB, was introduced. Chronologically, this variable takes the value zero for all dates up to and including December 31, 2009, and takes the value one for all subsequent dates.

The final variable that makes my analysis possible is called "carrot." This variable indicates whether or not an individual is on a team whose leader received a variable bonus above the median. While each team leader received a monthly bonus of on of 180 RMB, 190 RMB, 200 RMB, 210 RMB or 220 RMB, these bonuses were adjusted for days worked and as a result, I calculated the median bonus payment based on an adjusted daily leadership bonus calculated as the monthly bonus for a given leader divided by the number of days that leader worked in a given month. With these three variables, as well as controls for individual, time, and team fixed effects, I am able to estimate the impact of receiving positive feedback (the "carrot") in the form of a high performance-based bonus on productivity, as well as the impact of receiving negative feedback (the "stick") in the form of a low performance based bonus.

It is worth noting again that I am analyzing at the impact of a team leader's bonus on the productivity of their team. The assumption being made is that the impact of a change in a team leader's bonus will be visible in his team members' performance, and will be brought about by a change in effort on the part of the team leader. As Romero (2010) illustrates, team leaders are likely to decrease their personal output in favor of increasing their efforts in team organization

and management in response to an increased bonus, and for this reason team leaders are excluded from my productivity analysis.[6]

My regression analysis is of the form:[7]

$$Productivity_{it} = \beta_0 + \beta_1 Gift\ Exchange_t + \beta_2 Performance\ Pay_t + \beta_3 Carrot_{it} + \boldsymbol{\delta IndividualDummies_i} + \boldsymbol{\gamma WeekDummies_t} + \boldsymbol{\lambda TeamDummies_i} + \varepsilon_{it}$$

This was a panel data estimation with individual, team, and week fixed effects with robust standard errors.[8] The interpretation of the coefficients of this regression is as follows: β_1 captures the change in productivity resulting from the introduction of the gift in the first treatment period. β_2, the coefficient on "Performance Pay" captures the impact of going from the first treatment period to the second for a worker on a team whose leader received a bonus at or below the median for the variable bonus treatment period (the "stick" effect). Lastly, β_3 captures the difference in the change in productivity brought about by the transition to the second treatment period between a "carrot" team member and a "stick" team member. Given that β_2 represents the "stick" effect and β_3 represents the difference between the "carrot" and "stick" effects, the sum of β_2 and β_3 is equal to the effect on an individual transitioning from the first treatment period to the second on a team whose leader received a variable bonus above the median in the second treatment period (the "carrot" effect).[9]

Before discussing the results of the above analysis, it is prudent to address relevant hypotheses. This experiment provides a unique opportunity not only to analyze the dynamics of a

[6] Romero (2010)'s analysis only looks at the idea of gift exchange, using data from the control and first treatment periods, but the logic for excluding leaders holds throughout the observation period.
[7] Bold terms represent a vector of coefficients on a set of dummy variables.
[8] Standard Errors are clustered at the individual worker level.
[9] Professor Kato was of invaluable assistance in interpreting the results of the model.

variable performance based payment plan as I aim to, but it also allows for a comparison of two types of rewards. The fact that the first treatment period involves a gift exchange program and the second an incentive based scheme sets up an interesting juxtaposition of the two.

There is an immense literature on gift exchange, and the overwhelming finding is that it leads to increased productivity.[10] Given this fact, the obvious hypothesis is that the coefficient β_1 should be positive; the introduction of a gift exchange program should lead to increased productivity of workers. The value that β_2 should take, however, is unclear. While both gift exchange and incentive pay programs have been shown to increase productivity, it is not immediately intuitive what the impact of introducing an incentive pay program after introducing (and subsequently removing) a gift exchange program should be. As is discussed in Strombach (2015), the removal of incentives can have detrimental effects on productivity, so it is possible that the switch from a gift exchange program to an incentive based one could decrease overall productivity through the impact of the original bonus plan. Additionally, given the slightly nuanced interpretation of this specific model, the difference between the "carrot" and "stick" effects must be considered as well. Given the evidence from some of the psychological literature regarding negative feedback, there is a plausible explanation for the "stick" effect on productivity to be positive. It has been shown that individuals respond positively to negative feedback[11], and given the transition from a gift exchange program to one where team leaders now receive less of a bonus based on poor team productivity, it seems reasonable to assume that "stick" team leaders would increase their effort so as to continue to make as much for their leadership as possible. Alternatively however, it is also valid to think that given their teams' poor performance and thus their lower bonus, they might decrease their effort in leadership so as to

[10] See Romero (2010) for analysis and references.
[11] See Rynes (2005).

increase their effort in production, in an attempt to raise their earnings based on their piece rate wage. This decreased effort in leadership could very well lead to decreased team member productivity.

Like that of the "stick" teams, the analysis of "carrot" teams is not straightforward. It is possible that the extrinsic motivation resulting from an increased bonus, "carrot" team leaders could further increase their effort in response and their teams' productivity could rise.[12] Additionally, and similarly to the potential case of the "stick" team leaders, "carrot" teams could see increased productivity in light of their leaders increasing their frontline production. Being further rewarded for their leadership duties, they may feel that more of their time would be better spent producing, and raising their earnings that they gain via their piece rate wage. There are also several plausible explanations, however, for a decrease in the productivity of "carrot" teams as a result of transitioning from the first treatment to the second. One such explanation is that income effects are at play. Given that tying performance to team leader bonuses has resulted in either the same or higher bonuses for "carrot" teams, it is possible that "carrot" team leaders are less motivated to perform their leadership duties in light of their higher income. There is also a potential explanation for a decrease in "carrot" team productivity that is attributable to the experimental design.[13]

V. Results

One element of this regression analysis that warrants discussion prior to that of the results, is a minor collinearity issue between the week dummies used to control for overarching time effects and those variables that represent the periods of gift exchange and variable performance

[12] While the bonus for a "carrot" team leader does not necessarily increase from the first treatment period to the second, it is higher relative to their peers, where it previously was not.
[13] This will be discussed in a later section.

pay. Five of the week dummies are omitted from my regressions due to collinearity (the rest remain, however, since the time dimensions of the time and treatment dummies are vastly different). For this reason, any plant wide shocks that may have taken place during these five weeks are unobserved and these five weeks form a baseline for time related shocks that the coefficients on the other time dummies are relative to. However, the time dummies are included as controls in this analysis and the interpretation of their coefficients is not relevant. Still, it is necessary to note that these potentially unobserved time shocks could impact the values of the "Gift Exchange" and "Performance Pay" coefficients. The values of the coefficient on "carrot", however, are unaffected by this issue as this variable analyzes changes for a given worker within a given time period.

The results of my regression analysis are detailed in the appendix, but a brief summary is as follows: The transition, for all workers, from the control period to the period of gift exchange resulted in a 19 percent increase in productivity (relative, as all of the following results are, to the overall mean value of productivity for the entire experiment).[14] The transition from the period of Gift Exchange to the period of variable performance pay resulted, for "stick" teams in a 29 percent decrease in productivity.[15] The effect on "carrot" teams of the transition from the gift exchange period to the period of variable performance pay (equal, in the regression outlined in the previous section, to the sum of $\beta 2$ and $\beta 3$), was found not to be statistically significant from zero.[16] In other words, the two programs had the same effect on productivity for "carrot" teams. In addition to the regression above, I ran a very similar one with a slightly different definition of productivity. Instead of regressing the independent variables on the raw observed values for productivity, I regressed them instead on a z-score of productivity (defined as the number of

[14] For a detailed breakdown of productivity, see the Appendix.
[15] Both of these results were statistically significant at the 1% level
[16] An F-Test concluded that the sum of $\beta 2$ and $\beta 3$ is not significantly different from zero.

standard deviations from the mean each observation is). This analysis was done not only in an effort to standardize my variable of interest, but also potentially to shed more light on the magnitude and significance of my results. The results of this analysis are detailed in the appendix; however they do not change the magnitude or direction of any of the results discussed above.[17]

These results, while found through a somewhat complicated interpretation of regression, are intuitive and in line with existing literature. The idea that a "carrot" team had roughly the same productivity as it did during the gift exchange program while the "stick" teams decreased their productivity significantly is in line with existing literature regarding the positive effects of not only gift exchange, but of positive reinforcement overall.[18] While these results are interesting in and of themselves, I performed several robustness checks in an effort to look deeper into the dynamics of the variable performance-based pay program, which will be discussed in the following section.

VI. Robustness Checks

While several of my hypotheses were confirmed by my initial regression analysis, further investigation is warranted. One effort I made in order to strengthen some of my conclusions was to better define "carrot" and "stick" teams. As discussed previously, I originally defined a "carrot" team as one whose leaders' median adjusted daily leadership bonus was above the overall median for the relevant period, and a "stick" as one whose leaders' median adjusted daily leadership bonus was at or below the median.[19] This definition, however, is subject to the valid criticism that those teams whose leaders' median adjusted daily leadership bonus was at or

[17] See table 1 for detailed regression results.
[18] See Strombach (2015).
[19] The period during which variable leadership bonuses were decided was from October 1, 2009 until December 31, 2009.

around the median would not experience true "carrot" or true "stick" motivation. This idea is especially concerning since the variable bonus averages 200 RMB, so these teams in the middle are receiving roughly the same bonus as they did in the gift exchange period. Additionally, an argument could be made that the original selection method for "carrot" and "stick" teams lacked exogeneity, as the "carrot" and "stick" teams are not randomly defined since they are the selected based on productivity levels in prior periods. In an effort to combat these potential issues, I created three groups of teams where I previously had two. For this analysis, I use a z-score of each leaders adjusted daily leadership value. In order to find the teams with the most "extreme" leadership bonus values in either direction, I looked at the average z-score for each team. The new "carrot" group is made up of the three teams with the highest average z-score of their adjusted daily leadership bonus values, and the new "stick" group is made up of the three teams with the lowest such values. The three remaining teams serve as a neutral group and can be used as a point of reference. In order to create these three groups I defined two new dummy variables, "Extreme Carrot" which takes a value of one during the final three months of the observation period for the top three teams and a value of zero for all others, and "Extreme Stick" which takes a value of one for the bottom three teams in this period and a value of zero for all others. In order to analyze the impacts of these new variables, I ran the following regression:

$$Productivity_{it} = \beta_0 + \beta_1 Gift\ Exchange_t + \beta_2 Performance\ Pay_t + \beta_3 Extreme\ Carrot_{it} + \beta_4 Extreme\ Stick_{it} + \boldsymbol{\delta IndividualDummies_i} + \boldsymbol{\gamma WeekDummies_t} + \boldsymbol{\lambda TeamDummies_i} + \varepsilon_{it}$$

This regression provided interesting results of its own which not only support, but also strengthen the conclusions drawn after analyzing my initial regression. The transition from the

period of gift exchange to that of variable performance pay for the newly redefined "carrot" teams (equal, in the above regression, to the sum of β_2 and β_3) resulted in a 16 percent increase in productivity. "Stick" teams, as defined in this new manner, experience a 34 percent decrease in productivity (with their total effect being equal in the above regression to the sum of β_2 and β_4).[20] The results of this regression further the results of my initial regression, concluding that "carrot" teams, as defined in a more extreme capacity, increased their performance as a result of the variable performance-based pay scheme beyond that which they sustained during the gift exchange program.[21]

The redefinition of "carrot" and "stick" teams as described posed one confounding issue. While most teams the were initially "carrot" teams either stayed "carrot" teams or entered the neutral group, one team switched from initially being a "carrot" team to being a "stick" team in the redefinition of the variables. In order to make sure that this team was not skewing my results, I ran both my initial regression and the regression with the redefinition of "carrot" and "stick" with the team in question excluded, and found that there were no significant changes to my results.[22]

A second robustness check that I performed tested for heterogeneity in the impact that the variable performance-based pay program had on teams with different initial levels of productivity. In order to do this, I created a baseline productivity for each team defined as that team's average productivity during the control period. I then interacted that variable with the "carrot" variable and ran the following regression:

[20] An F-Test of both the sum of β_3 and β_3 as well as of β_3 and β_3 concluded that the values were statistically different from zero.
[21] For detailed regression results, see table 2.
[22] For these detailed regression results, see tables 1 and 2.

$$Productivity_{it} = \beta_0 + \beta_1 Gift\ Exchange_t + \beta_2 Performance\ Pay_t + \beta_3 Carrot_{it} + \beta_4 Carrot*AverageProductivity_{it} + \boldsymbol{\delta IndividualDummies_i} + \boldsymbol{\gamma WeekDummies_t} + \lambda AverageProductivity_i + \varepsilon_{it}$$

The results of this regression are detailed in the appendix, but can be summarized as follows.[23] This analysis implies that "carrot" teams with higher baseline productivity will have a greater increase in productivity as a result of the variable performance-based pay program than those with lower baseline productivities. One plausible explanation for this is that teams with higher baseline efficiencies have higher quality team leaders and members, and therefore are more responsive to the bonus program.[24]

As with my initial regression, one final robustness check that I employed was a redefinition of my dependent variable, worker productivity, as a z-score. I then repeated all of my robustness checks with this z-score as my variable of interest. In all cases, the z-score regressions did not change the direction or significance of the effects, and the specific results of all of these regressions are detailed in the appendix.[25]

VII. Discussion of Alternative Interpretations

While the results of my analysis provide support of existing literature and add to the discussion of both incentive pay schemes in general and variable performance pay schemes specifically, some of the issues created by the experimental design leave these results exposed to alternative interpretation. The alternative interpretation I would like to focus on is centered on the fair wage-effort hypothesis as formulated by Akerloff and Yellen (1990). This concept is

[23] See table 3.
[24] Included in the appendix is table 4a, the results of a regression of the same form, but with the "Extreme Carrot" and "Extreme Stick" dummies included. The results support those discussed above.
[25] See tables 1, 2, 3, and 4.

popular throughout both economics and psychology literature and, in short, attributes effort of workers to their assessment of their wage with respect to fairness. If you apply this theory to the transition from the gift exchange period to that of variable performance pay in this experiment, the conclusions implied can explain the results I have found. Due to the fact that team leaders were not aware that there bonuses for October, November, and December 2009 were variable and based on team performance until January 2010, they were likely surprised by the finding. "Carrot" team leaders were positively surprised, as they were now receiving a higher bonus than they had previously. It is unlikely that they would asses this increase in wage as "unfair", and as such it follows that there was not a significant change in performance for the "carrot" teams as a result of the transition. The "stick" teams, on the other hand, were negatively surprised, and may very well have viewed their new, lower bonus as "unfair". If this in fact was the case, then the fair wage-effort hypothesis supports the claim that these leaders would decrease their effort in the face of an "unfair" wage and could explain the decrease in "stick" team performance as a result of the transition.

Additionally, although it is not an explicit conclusion of my analysis, my results suggest that overall productivity decreased as a result of the transition from the gift exchange period to the period of variable performance pay. As "carrot" team productivity was unchanged and "stick" team productivity fell, it follows that total productivity fell as well. This result is also supported by the fair wage-effort hypothesis, as it is plausible that a shift from a gift exchange program to one of variable performance pay (especially given the delayed manner in which the shift was announced) could be viewed as unfair by all and lead to decreased productivity across the board.

While this alternative interpretation provides a plausible explanation of what is happening in this experiment, the conclusions I have drawn based on the incentive effects (as opposed to a measure of fairness) are still valid, especially given the results of my robustness checks. When redefining the "carrot" and "stick" teams in a more extreme manner, "carrot" teams do significantly increase their productivity, which points to something other than just a perception of fairness at play.

VIII. Conclusion

There is a vast literature, both economic and psychological, that discusses the positive impacts of variable performance based pay programs. The aim of this paper is to add to this literature by providing a potential explanation of the mechanism behind these positive impacts. Through a two-way fixed effects model, I analyzed the difference in the changes in productivity of team members at a Chinese manufacturing plant based on the relative level of bonus received by their team leaders.

After controlling for the various changes in team leader bonus structure the plant underwent, in addition to individual, time, and team fixed effects, I was able to conclude that the "carrot" effect of a variable performance based pay program, or that experienced by those team members whose leaders' median bonus value was above the median value for the period, was positive and statistically significant at the one percent level. Additionally, I found that the "stick" effect, or that experienced by those team members whose leaders' median bonus value was at or below the median level for the period, was negative and statistically significant. While this is only one experiment, the idea that it is the "carrot" that drives productivity increases when a variable performance based pay system is employed has significant implications, and supports a vast literature that purports that positive feedback and reinforcement lead to increased effort and

productivity in the workplace (Strombach 2015). While my initial theory was that a variable pay system my not be optimal if it is driven by the "carrot" effect, the results of this experiment do in fact make a case in the program's favor. When using the more extreme definitions of "carrot" and "stick," it becomes apparent that carrot teams not only continue to work at the same capacity that they did during the gift exchange program, but they increase their productivity past that point. Although the "stick" effect was negative, without it there would be no "carrot" effect to be had, so it appears, through the results of analyzing this experiment, that if the combined effect of a variable performance based pay program is positive, that it is effective due to its impact on those receiving higher pay. This is an idea that is supported in psychological literature, and matching those psychological hypotheses with empirical results is an interesting conclusion, and should definitely play a role in the decision making process when firms choose what type of incentive program to employ (Strombach 2015).[26]

While the results of this analysis are interesting, they do pose certain issues. While the experiment allows for an analysis in the difference between the "carrot" and "stick" effects, it does not allow for an explicit comparison of these effects, or of the variable performance-based pay program as a whole, to the control period. Because of this, further analysis into the dynamics of such programs is warranted. Additionally, the fair-wage hypothesis as formulated by Akerloff and Yellen (1990) provides a valid alternative interpretation of my results; however it lacks explanatory power when considering the results of my robustness checks. Despite these issues, the results discussed above should undoubtedly add to the ongoing discussion of optimal incentive plans and the impact that variable performance plans have on productivity, and ultimately profitability for firms.

[26] As is noted by much of the literature discussed in section II, different settings can have different optimal incentive plans.

References

Akerlof, George A., and Janet L. Yellen. "The Fair Wage-Effort Hypothesis and Unemployment." The Quarterly Journal of Economics 105.2 (1990): 255-83. Print.

Bhattacherjee, Debashish. "The Effects of Group Incentives in an Indian Firm: Evidence from Payroll Data." Labour 19.1 (2005): 147-73. Print.

Brehm, Margaret, Scott A. Imberman, and Michael F. Lovenheim. Achievement Effects of Individual Performance Incentives in a Teacher Merit Pay Tournament. Working Paper 21598 ed. National Bureau of Economic Research, 2015. Print.

Delfgaauw, Josse, et al. "Dynamic Incentive Effects of Relative Performance Pay: A Field Experiment." Labour Economics 28 (2014): 1-13. Print.

Encinosa, William E., Martin Gaynor, and James B. Rebitzer. "The Sociology of Groups and the Economics of Incentives: Theory and Evidence on Compensation Systems." Journal of Economic Behavior & Organization 62.2 (2007): 187-214. Print.

Imberman, Scott A., and Michael F. Lovenheim. "Incentive Strength and Teacher Productivity: Evidence from a Group-Based Teacher Incentive Pay System." Review of Economics and Statistics 97.2 (2015): 364-86. Print.

Lazear, Edward P. Output-Based Pay: Incentives, Retention Or Sorting?. IZA Discussion Paper, 2003. Print.

Lazear, EP. "Performance Pay and Productivity', American Economic Review, 90 (5), December, 1346-61." INTERNATIONAL LIBRARY OF CRITICAL WRITINGS IN ECONOMICS 207.2 (2007): 216. Print.

Levenson, Alec R., et al. "Optimizing Incentive Plan Design: A Case Study." Available at SSRN 1925770 (2011)Print.

Muralidharan, Karthik, and Venkatesh Sundararaman. Teacher Performance Pay: Experimental Evidence from India. NBER Working Paper 15323 ed. National Bureau of Economic Research, 2009. Print.

Pindiwe, Bothwell. "Hidden Costs of Reward? Investigating the Destruction of Intrinsic Motivation." Colgate University, 2015. Print.

Romero, Julia. "Reciprocity and the Manipulation of Effort Allocation: Evidence from a Field Experiment." Colgate University, 2010. Print.

Rynes, Sara L., Barry Gerhart, and Laura Parks. "Personnel Psychology: Performance Evaluation and Pay for Performance." Annu.Rev.Psychol. 56 (2005): 571-600. Print.

Shaw, Jason D., Nina Gupta, and John E. Delery. "Pay Dispersion and Workforce Performance: Moderating Effects of Incentives and Interdependence." Strategic Management Journal 23.6 (2002): 491-512. Print.

Strombach, Tina, Marco Hubert, and Peter Kenning. "The Neural Underpinnings of Performance-Based Incentives." Journal of Economic Psychology 50 (2015): 1-12. Print.

Zenger, Todd R., and CR Marshall. "Determinants of Incentive Intensity in Group-Based Rewards." Academy of Management Journal 43.2 (2000): 149-63. Print.

APPENDIX:

Table 1:

Regression of productivity and/or a z-score of productivity on the following variables and individual, time and team fixed effects. (columns 3 and 4 are the results of the regressions in columns 1 and 2 with the potentially problematic team removed).

VARIABLES	productivity	z_productivity	productivity	z_productivity
Gift Exchange	0.223***	0.0647***	0.198***	0.0574***
	(0.0621)	(0.0180)	(0.0723)	(0.0210)
Performance Pay	-0.335***	-0.0972***	0.0726	0.0211
	(0.105)	(0.0306)	(0.121)	(0.0351)
Carrot	0.235*	0.0684*	0.321**	0.0934**
	(0.137)	(0.0398)	(0.141)	(0.0410)
Constant	0.937***	-0.0685	0.661***	-0.149***
	(0.243)	(0.0706)	(0.146)	(0.0423)
Observations	32,826	32,826	28,121	28,121
R-squared	0.041	0.041	0.047	0.047
Number of label	214	214	196	196

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 2:

Regression of productivity and/or a z-score of productivity on the following variables and individual, time and team fixed effects. (columns 3 and 4 are the results of the regressions in columns 1 and 2 with the potentially problematic team removed).

VARIABLES	productivity	z_productivity	productivity	z_productivity
Gift Exchange	0.223***	0.0648***	0.190***	0.0552***
	(0.0622)	(0.0181)	(0.0720)	(0.0209)
Performance Pay	-0.121*	-0.0351*	0.191*	0.0556*
	(0.0689)	(0.0200)	(0.0999)	(0.0290)
Extreme Carrot	0.313***	0.0908***	0.339***	0.0984***
	(0.0501)	(0.0145)	(0.0530)	(0.0154)
Extreme Stick	-0.272***	-0.0790***	-0.250	-0.0727
	(0.0901)	(0.0262)	(0.171)	(0.0496)
Constant	0.970***	-0.0590	0.752***	-0.122***
	(0.253)	(0.0734)	(0.142)	(0.0412)
Observations	32,826	32,826	28,121	28,121
R-squared	0.041	0.041	0.047	0.047
Number of label	214	214	196	196

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 3:

Regression of productivity and/or a z-score of productivity on the following variables and individual, time and team fixed effects.

VARIABLES	productivity	z_ productivity	productivity	z_ productivity
Gift Exchange	0.217***	0.0631***	0.225***	0.0652***
	(0.0624)	(0.0181)	(0.0623)	(0.0181)
Performance Pay	-0.297***	-0.0863***	-0.108	-0.0313
	(0.105)	(0.0305)	(0.0684)	(0.0199)
Extreme Carrot			1.206**	0.350**
			(0.542)	(0.157)
Extreme Stick			-1.194***	-0.347***
			(0.350)	(0.102)
Extreme Carrot*AvgProd			-0.724*	-0.210*
			(0.423)	(0.123)
Extreme Stick*AvgProd			0.844***	0.245***
			(0.279)	(0.0809)
Carrot	-0.234	-0.0679		
	(0.143)	(0.0416)		
Carrot*AvgProd	0.464***	0.135***		
	(0.0956)	(0.0278)		
Constant	-1.807	-0.866*	-1.463	-0.766
	(1.571)	(0.456)	(1.609)	(0.467)
Observations	32,826	32,826	32,826	32,826
R-squared	0.040	0.040	0.041	0.041
Number of label	214	214	214	214

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 4: Summary statistics of relevant variables

Variable	Description	Mean	SD	Min	Max
productivity	output1/worktime.	1.17	3.44	0	143.84
plantime1	Total earned minutes per unit of product 1	425.15	1502.17	0	97048
plantime2	Total earned minutes per unit of product 2	55.67	155.44	0	9000
plantime3	Total earned minutes per unit of product 3	42.00	1525.41	0	97048
plantime4	Total earned minutes per unit of product 4	9.58	52.02	0	1220
plantime5	Total earned minutes per unit of product 5	3.16	25.77	0	1449
output1	The sum of product1-product5	535.57	2154.45	0	97808
worktime	Total minutes worked per day	449.48	237.61	0	1420

Table 5: A breakdown of productivity by period for team members

Period	Mean	SD	Min	Max
Control	1.10	0.96	0	43.69
Gift Exchange	1.27	5.14	0	143.84
Performance Pay	1.08	0.59	0	13.13
Only Carrot	0.99	0.55	0	13.13
Only Stick	1.38	0.59	0	4.51

Table 6: A breakdown of productivity by period for team leaders

Period	Mean	SD	Min	Max
Control	1.08	1.34	0	43.69
Gift Exchange	1.08	0.50	0	3.13
Performance Pay	1.21	0.80	0	12.78
Only Carrot	1.20	0.90	0	12.78
Only Stick	1.23	0.58	0	3.87

Table 7: A kernel density estimate for "carrot" teams during both the gift exchange and variable performance pay periods

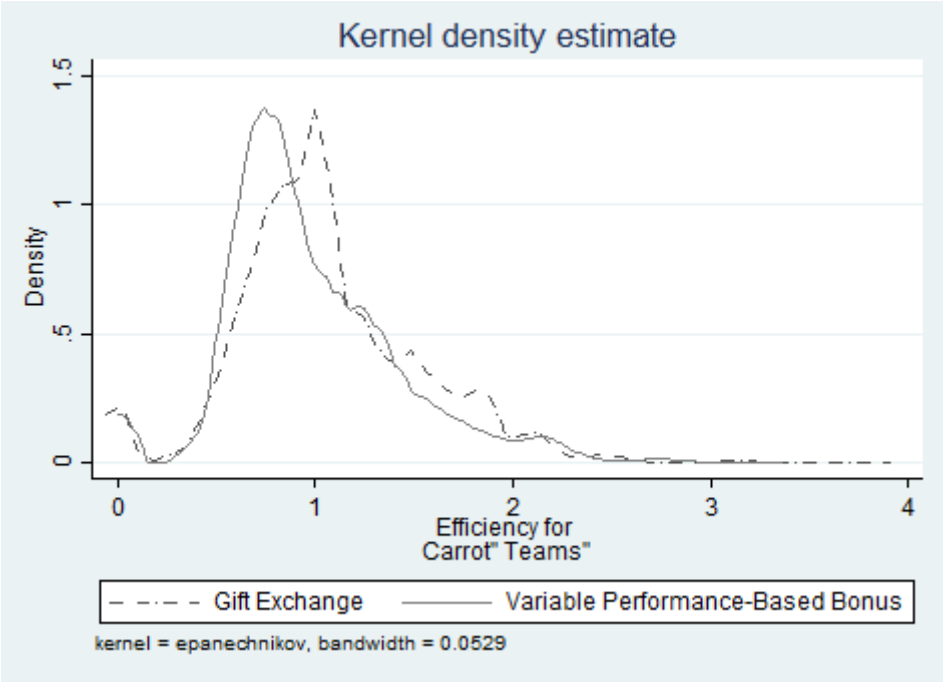

Table 8: A kernel density estimate for "stick" teams during both the gift exchange and variable performance pay periods

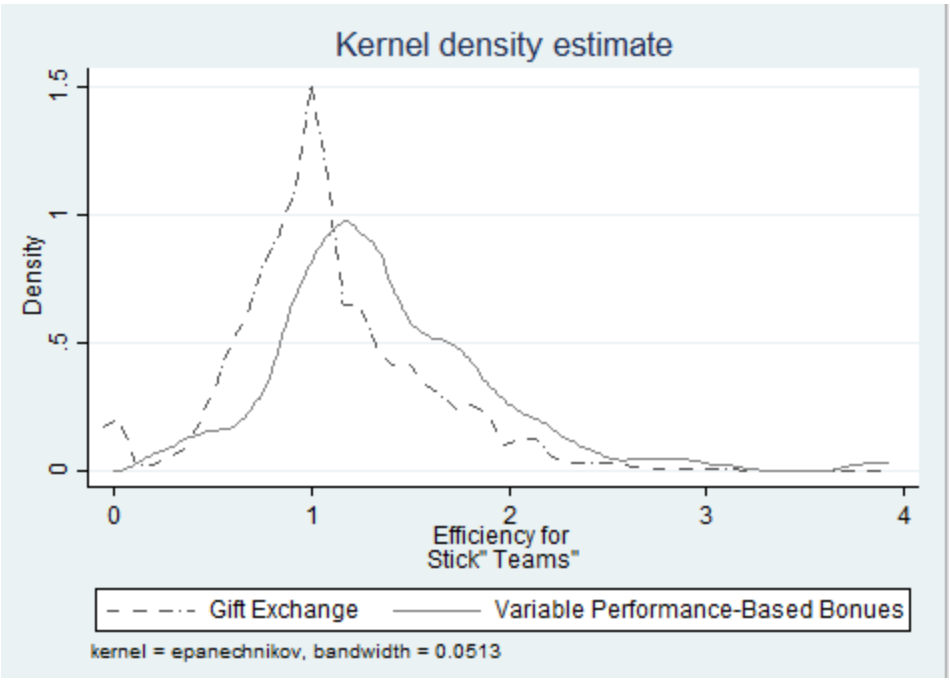

Table 9: A summary chart describing the experimental process:

Date	Front Office Action	Employee Environment
October 1, 2008	Observation period begins with no changes made by upper level management. Team leaders receive their original variable bonus averaging 100 RMB.	There are no changes to employee beliefs at this point, and this period acts as a control.
July 1, 2009	Upper level management announces that it will, for the next three months, be giving team leaders a gift of 200 RMB per month. This bonus is in addition to their original leader salary payments and is not tied to performance. The payment of this gift will be delayed by one fiscal quarter. That is to say, the payment for July will be made in October, the payment for August will be made in November, and the payment for September will be made in December.	Employees are now acting under the assumption that team leaders are receiving a gift from management in addition to their normal team leader salary.
October 1, 2009	The gift payments commence as described above.	Employees continue to act under the same assumptions as above.
January 1, 2010	Upper level management announces that bonuses for the months of October, November, and December (to be paid out in January, February, and March) are tied to team performance. Teams are split into five groups by monthly performance and their team leaders receive a corresponding leadership bonus, in addition to their original leadership salary (Note: the gift exchange program is removed). The five payment levels are: 180 RMB, 190 RMB, 200 RMB, 210 RMB and 220 RMB.	Employees are now acting under the assumption that team leaders are receiving a variable, performance based leadership bonus in addition to their original leadership salary.
March 31, 2010	End of observation period.	

Table 10: A timeline depicting the experimental process.