

Incentive Effects of Inequality and Economic Development*

Hyejin Ku[†]
Florida State University

Timothy C. Salmon[‡]
Florida State University

May 7, 2009

Abstract

The effect of inequality on economic development is often debated in the development literature with many finding that inequality among different groups in a society has a negative impact on economic growth while others find the opposite result. Our study investigates a behavioral phenomenon which may not only be associated with the existence of such inequality but may also significantly exacerbate any adverse consequences of inequality for growth. In particular we investigate whether or not individuals exhibit a discouragement effect in the face of inequality that leads to lower work effort. If such an effect exists it provides a mechanism for converting even idiosyncratic inequality into sustained inequality with adverse consequences for the individuals being affected by the inequality and the economy as a whole. We investigate this phenomenon using an economic experiment to allow us to cleanly vary the nature of inequality and to allow us to directly observe several characteristics of the workers. We find robust support for the existence of a discouragement effect lending credibility to the claims that such an effect would exist in external situations among workers confronted with disadvantageous inequality.

JEL Codes: C90, D61, D63, J24, O15, O40

Key Words: inequality, efficiency, productivity, experiment

1 Introduction

Economists have established many ways in which initial inequality may influence the paths of economic development. In a series of historical studies, Engerman and Sokoloff (1997, 2002 and 2005) argue that the initial differences in factor endowments between the North and South America contributed to the emergence of different institutions between the two regions, which in turn led to the divergence in the rates of growth. While the typical finding in the growth literature may be a negative correlation between inequality and growth in a cross-section of countries, there appear to be no conclusive findings on the relationship

*The authors would like to thank the National Science Foundation for research support in funding these experiments.

[†]Florida State University, Department of Economics, 113 Collegiate Loop, Room 263, PO Box 3062180, Tallahassee, FL 32306-2180, hku@fsu.edu. Phone: 850-644-7208 Fax: 850-644-4535.

[‡]Florida State University, Department of Economics, 113 Collegiate Loop, Room 263, PO Box 3062180, Tallahassee, FL 32306-2180, tsalmon@fsu.edu. Phone: 850-644-7207 Fax: 850-644-4535.

between growth and inequality from recent studies based on panel data analysis.¹ Due to the importance of this relationship between inequality and economic growth in the design of a broad range of economic development policies there is a need to provide a better understanding of the mechanics behind some of the linkages between growth and inequality. There is a substantial theoretical literature which proposes a number of different linkages between growth and inequality including Loury (1981), Galor and Zeira (1993), Banerjee and Newman (1993), Durlauf (1996), Benabou (2000) and Mookherjee and Ray (2003). This literature is aimed at doing two things. Some of these papers attempt to explain a link between inequality and growth while others focus mainly on attempting to explain how initial inequality may lead to persistent inequality that remains in the long run. By and large the linkage between inequality and either growth or persistent inequality is modeled as being based on differential investment opportunities of agents typically in the form of human capital and occupational choice.

We propose an additional linkage between inequality and growth that can also be useful in explaining the persistence of inequality which is based on a behavioral response individuals may exhibit when confronted with inequality. If there are members of a group who experience lesser opportunities for earnings and advancement than members of other more advantaged groups then one might expect that the disadvantaged individuals will exhibit some degree of discouragement from their situation and exhibit a variety of responses ranging from decreased on the job work effort, decreased human capital acquisition, decreased contributions to public goods and in general engaging in a lower level over a range of other activities that are important for individual as well as societal advancement. In the context of a wage inequality, Akerlof and Yellen (1990) propose a similar phenomenon as a theoretical construct which they refer to as the “fair-wage effort hypothesis” and it is described by stating that “workers have a conception of a fair wage; insofar as the actual wage is less than the fair wage, workers supply a corresponding fraction of normal effort.”

If such a behavioral response to inequality exists then there are a number of consequences associated with it. First, this effect *ceteris paribus* would generate a negative relationship between inequality and economic growth. In general there are a number of forces at play in an aggregate economy so this is not to claim that the overall effect must be negative if this effect exists but the partial effect of inequality on growth due to this behavioral effect would be negative. Second, separately or in addition to the theoretical models cited above this effect could generate persistent inequality from initial idiosyncratic inequality. The reason is that even if the initial inequality were idiosyncratic, if a worker responds by exerting lesser effort or investing less in human capital, then this response has the potential to legitimize and then perpetuate the initially low opportunities. Thus initial idiosyncratic inequality can be converted into longer term inequality. Any potential discouragement effect is, of course, likely to be worsened if the inequality in opportunity is related to a more generalized social phenomenon of race, ethnic, gender or caste based discrimination in which individuals from these disadvantaged groups expect to be confronted with unequal treatment in many aspects of their lives.

The goal of the present study is to determine if a discouragement effect is likely to exist. We present the results of a laboratory based experimental study designed to allow

¹Benabou (1996) and Banerjee and Duflo (2003) provide a comprehensive review of the literature on the issue of inequality and growth.

clean inference on whether or not individuals become discouraged in the face of inequality. The environment we construct in our experiment should be a very strong test of this effect because many of the aspects of discrimination and long run experience with the inequality will not be present in the experiment. Consequently, if we find a discouragement effect in our simplified laboratory environment then it will be robust evidence that such an effect could exist in a broader context in which the inequality is based on overt discrimination and individuals would have long term experience with feeling the effects of the inequality.

There is prior evidence on the issue but both the sign and robustness of any effect is unclear. One of the main reasons for this lack of clarity is due to the difficulty of identifying the discouragement effect in naturally occurring data. Much of the work in this area has been focused on identifying an effect due to wage inequality in workplaces. While there may be unfairness in naturally occurring wage schedules it is difficult to separate cleanly between a wage that is unfairly low and one that is deservedly low due to skill differences or to traits which would be unobserved by a researcher. The summary of the attempt to do this for identifying race based discrimination in wages against blacks in the US provided in Heckman, Lyons, and Todd (2000) explains many of the inherent problems in such attempts. Proper inference on the effects of inequality on effort though requires cleanly severing the link between wages and characteristics of a worker to find situations in which it is clear that “unfair inequality” is present. Were we able to find situations in which wages are clearly unequal based on something other than traits of workers related to their productivity then there is also an empirical difficulty in observing work effort. Effort itself is rarely observable and the measurement of most proxies for work effort in common use (rate of promotion, turnover etc. . .) are potentially related to any discrimination that generated the wage inequality which adds structural bias to the data and limits inference. Even were a viable proxy for effort to be available, there is still a problem of separating out effort differences due to pure wage effect and effort differences due to the existence of inequality. Separating these effects requires observing workers exerting effort when faced with the same wage in similar environments that differ only in regard to whether or not wage inequality is present. Due to these difficulties in inference from naturally occurring data, we present evidence from a laboratory experiment in which we will be able to observe effort directly while implementing exogenously set wage schedules with and without inequality to overcome these limits on inference.

In our experiment workers will be exogenously assigned wages as a proxy for general opportunity and they will then engage in a real effort task (encoding random sequences of letters into numerical code) for piece rate earnings. The subjects will also have an alternative task which they can engage in for very low earnings to represent an outside option which might be very loosely interpreted as a “shirking” option. This eliminates a potential confound present in many laboratory real effort tasks as with nothing else to do subjects could well engage in the main task out of boredom. Our experiment addresses another important confound which is necessary to distinguish between a wage effect and an inequality effect. Simply observing that lower paid workers worked less would not be conclusive in demonstrating an inequality effect as it could simply be due to the incentive effect of lower wages. Consequently our design will include control sessions in which all subjects receive equal wages so that we can identify the inequality effect by comparing the performance of workers receiving a particular wage in sessions with no inequality to that of

workers receiving the same wage in the presence of inequality.

Our experimental design will also address a less commonly investigated issue regarding how the response to inequality might vary with the relative sizes of the advantaged and disadvantaged groups. The relative group size may be important because the existence of an incentive effect hinges on workers perceiving a low wage as being unfair. That perception of unfairness may be diminished if the majority of workers are receiving the same low wage and may be heightened if only a small number do. To examine this possibility our experiments will vary the size of the high and low wage groups to determine if there is a systematic response of individuals to the status of their groups as being a minority or majority.

In examining the data from the experiments we will be interested in the overall effects on productivity as this will inform us on the efficiency aspects of inequality but we will also be looking at distributional issues. As argued in Ravallion (2001) it is important to look “beyond the averages” to understand the effects on subpopulations. Consequently, in addition to the overall average response we are also interested in determining if there are any specific groups who may respond more strongly to the existence of inequality than others. Towards that end we will be obtaining some basic demographic information on our subjects as well as using a few simple ability tests to measure various forms of cognitive ability.

In the end we do find that a discouragement effect exists and it is of non-trivial size. One key aspect of the experiment is that wages are assigned to subjects randomly and evidence provided in Bolton, Brandts, and Ockenfels (2005) suggests that some individuals will judge the fairness of an allocation based on the fairness of the procedure used to generate the allocation rather than simply on the actual allocation alone. This is an important point to understand in interpreting our results because our wage inequality is based on pure randomization which is typically deemed a fair procedure, the effect we find could be downwardly biased compared to external situations in which the wage inequality might be based on some form of overt discrimination which is more likely to be judged unfair. We do find evidence of a larger behavioral response for those subjects that we can establish should be more likely to view the allocation as unfair which supports that idea. Overall, the fact that we find any measurable discouragement effect in this laboratory setting which should be considered *a priori* as unlikely to generate much of a response should lend credibility to the claim that a similar effect is likely to exist in external situations in which discrimination is present and persistent.

In section 2 we will provide an overview of our experiment design. Section 3 will present a series of hypotheses regarding what one might expect to observe in the experiments based on prior literature. In section 4 we will present our results and we will provide a concluding discussion in section 5.

2 Experimental Design

The base task in this experiment involves subjects taking random strings of 4-letter “words” and using a code key to translate those letters into a numerical code. The subjects were shown the encoding key and the string of letters on a screen and would enter the encoded version below the word. The same code key was used for the entire session. They had a button which would allow them to submit a word and after doing so they would be

	Common Low	Common High	Disadvantaged Minority	Disadvantaged Majority
Blue Group (6 members)	\$0.03	\$0.09	\$0.09	\$0.03
Green (2 persons)	\$0.03	\$0.09	\$0.03	\$0.09
Number of Cells	4	2	8	8

Table 1: Summary of experimental design.

immediately given a new word. For every word they encoded correctly they were paid a piece rate wage which was exogenously set by the experimenters at the beginning of the session. This production task is very similar to the one used in Erkal, Gangadharan, and Nikiforakis (2008).

Each session consisted of 16 subjects and these 16 subjects were divided into two cells of 8 at random. Inside of those cells of 8, subjects were further randomly assigned into what we labeled “blue” and “green” groups. The subjects were informed of the color of their group before anything else happened in the experiment as an attempt to make it clear that group assignment was exogenous. The meaning of the group assignment was only explained later to the subjects and the actual meaning referred to the wage rate they would be assigned. Table 1 provides a summary of the experimental treatments as well as the number of observations of cells for each treatment. There are a total of four treatments contained in this design. There are two control treatments called Common Wage - Low and Common Wage - High in which all subjects received the same common wage regardless of group assignment. For consistency with the other treatments subjects were still divided such that there were 6 members of the blue group and 2 of the green in each treatment but the wages across groups were constant. The high wage was \$0.09 per correctly encoded word while the low wage was \$0.03. The other two treatments will be called the Disadvantaged Minority treatment (Minority) and the Disadvantaged Majority treatment (Majority). These two treatments introduce inequality by having one group possess the high wage and the other the low such that in the Minority treatment, the members of the 2-person group are assigned the low wage while in the Majority treatment the members of the 6-person group are assigned the low wage.

The experiment was programmed using z-Tree, Fischbacher (2007). After subjects saw an initial screen indicating the color of the group to which they were assigned, they were presented with a sample of the main screen for the experiment showing them the encoding task. In the course of explaining this screen to the subjects, they were explicitly informed of the wage rates that would be in effect for both groups. Also, before each round of production, subjects would see a screen which included a table showing them a column for each subject in their 8-subject cell indicating their group color and corresponding wage rate. The idea of stating this information to them repeatedly was to ensure that they clearly understood both the wage rate differential as well as the relative size of both groups. After the first round, this screen also showed them information on their own past earnings. They were not shown the earnings or production levels of other individuals in the experiment at any time. The only information they see about other subjects is their wage rates. As such there are no interactions between members of a group or members of a cell making each

subject independent of the others. This choice of feedback was made so that the only things that should be salient to the subjects that might affect their behavior are the treatment variables themselves and any session or group effects should be minimal.

In order to give subjects an outside option should they wish to avoid the production task, we included another task on their screen. This other task was the option of playing Tic-Tac-Toe (TTT) against the computer. The computerized opponent was programmed to be moderately difficult but beat-able. This task was only minimally incentivized in that it paid a subject \$0.01 per win. Due to the difference between this wage rate as well as the time it would take to win a game and the wage rate and time to correctly encode a word for the encoding task, it should have been quite clear to subjects that TTT would never compete in financial terms with the main task. It was designed mainly to be at least mildly more “fun” than the encoding task and allow subjects who did not want to engage in encoding another activity so that they would not just have to stare at the screen in boredom.

There were 12 rounds of production in each session with each round lasting four minutes. Subjects were not instructed on which task to engage in but rather they were told how both worked, the wage rates of both and then told they were free to allocate their time between the two tasks as they wished.² At the end of each of the 12 rounds, each subject was presented with the screen summarizing her output in the encoding task, earnings from the encoding task, earnings from TTT, and cumulative earnings. On the practice screen they were also allowed to practice TTT as well as the encoding task for a few times before moving on to the first paying round.

At the end of the 12 rounds we had subjects fill out a short demographic questionnaire and complete two short sets of questions intended to measure various aspects of cognitive decision making. The first of these is the Cognitive Reflection Test (CRT) described in Frederick (2005). This three question test is designed to determine the degree to which subjects engage in thoughtful and reflective versus quick and impulsive decision making. For example, one of the questions is:

“A bat and a ball cost \$1.10. The bat costs \$1.00 more than the ball. How much does the ball cost?”

The most common quick answer one might come up with is \$0.10 but this is clearly wrong as upon further reflection the correct answer is \$0.05. We used this measure because those who score low on it may be more inclined to make impulsive decisions and that impulsiveness might well lead to exhibiting greater effects due to the treatments.³

We also used a second set of questions involving pattern matching problems in which we asked subjects to fill in the number that fits best in sequences such as: 3 6 9 12 (). We had 10 sequences of this sort that varied in difficulty. This test should measure a general facility with numerate tasks (thus we will refer to this as measuring “numeracy”) and therefore

²Full instruction scripts are available from the authors upon request.

³Oechssler, Roeder, and Schmitz (2008) investigate whether the CRT is good overall measure of cognitive ability and find that it is correlated with other aspects of decision making such as risk and time preferences as well as likelihood of making certain types of reasoning mistakes. While they do find some correlation with these other elements of decision making, our focus is on using the CRT for the purpose it was designed which is to discriminate between impulsive and reflective decision makers.

might correlate with performance on tasks external to the lab. The importance of this will be explored in the next section. Subjects received no payment for completing these questions and we gave them 100 seconds per test to complete as many of the questions as they could in the time frame. For the CRT the majority of the subjects finished well before the time limit while for the pattern matching test the time constraint was binding. Since speed was an issue in the main production task the fact that the constraint was binding for the pattern recognition task is not a problem since our goal was to measure performance under time pressure. For the CRT one could imagine that a time constraint could generate even less reflective thinking than normal but since the constraint was so rarely even close to binding this should not be a problem.

We have conducted 11 sessions with 176 subjects generating the number of cells per treatment as noted in table 1. All subjects received \$10 for showing up to the session and sessions last a little over an hour. Subjects earned on average \$28.89 (\$39.72 for high wage workers and \$19.86 for low wage workers) including their show-up fee.

3 Hypotheses

As a way of providing a framework for evaluating the results of the experiments and explaining the relevant insights from prior literature on these issues we will provide a series of hypotheses regarding what one might expect to observe from these experiments. The most basic hypothesis serving as a motivation for these experiments is derived from the statement of the Fair Wage-Effort hypothesis as described in Akerlof and Yellen (1990) and we will use this as our first hypothesis:

Hypothesis 1 (Fair Wage–Effort / Discouragement Hypothesis) - *The productivity of low wage workers should be lower in the two treatment conditions involving inequality than in the control treatment without inequality.*

While Akerlof and Yellen (1990) state this hypothesis mostly on theoretical grounds there are a number of prior results which might be thought of as suggestive that this effect will exist. An immediate thought might be that the extensive literature on trust and reciprocity in the form of a trust game, Berg, Dickhaut, and McCabe (1995), or gift-exchange game, Fehr, Kirchsteiger, and Riedl (1993), would support this hypothesis but on a deeper investigation that is less clear. In the standard gift exchange game the only players involved in a transaction are the worker and employer with the gift-exchange effect occurring when an employer gives the worker a high fixed wage and the worker responds with high effort as a reciprocal response to that gift. Any labor supply response in this relationship is a pure wage effect while the discouragement hypothesis is based on the idea that there is an inequality effect. For an inequality effect to exist the worker needs to see the wages of others in order to base a judgement of the fairness of the wage. There is at least one paper, Charness and Kuhn (2006), which does allow workers to see the wages of others in a gift exchange game and the authors find no effort difference between making the wages of others public versus private. These results can not really be interpreted as providing inference on this discouragement hypothesis, though, because the wage differentials subjects saw in those experiments could have been reasonably interpreted as fair. The reason is that the wages

of the workers were set endogenously by a subject in the role of employer and the workers were heterogeneous with respect to the earnings their effort generated to the employer. So if another subject was receiving a higher wage there was a clear explanator for that which is that the other subject may be one who generates more money per unit of effort to the employer and if so the higher wage is justified. Burchett and Willoughby (2004) present data from a related experiment which is a more direct test of the discouragement effect in that they had subjects engage in a real effort task and the subjects received randomly assigned wage structures for performing the same task. The researchers found that subjects generated quite different work efforts depending on whether information about the wages of others was public with those earning a low per unit or fixed wage decreasing their output while those with the high wage increased their output. This study included only a single period of productivity though which leads one to wonder about the robustness of the result to longer time frames.

In addition to the laboratory evidence in favor of this hypothesis there is also field evidence regarding the claim but the results appear mixed. Brown, Gardner, Oswald, and Qian (2008) provides evidence that worker happiness is potentially negatively related to wage inequality. That study showed in part that the rate at which individuals quit their jobs is correlated with the level of positive skewness in the pay distribution which is consistent with the claim that employees at the low end of the wage distribution will quit (an extreme form of effort reduction perhaps) in response to wages that are low in a relative sense to others in their firm even though they might be less likely to quit given the same wage at a firm with a less skewed distribution. On the other hand, Carr (2008) uses PSID data to examine these issues and finds little support for a claim that unequal earnings in a community impact any of his measures of workplace performance. The inference in both of these studies may be limited though due to many of the issues explained before regarding why we are pursuing a laboratory approach.

There is a separate strand of research involving laboratory data, naturally occurring data and data from field experiments all suggesting the group identity can effect work effort which also allows for the possibility of wage inequality to do so if there is a connection between group identity and wage inequality. Hoff and Pandey (2006) conduct a study in which members of high and low castes perform a production task with a common piece rate wage. When caste is not mentioned prior to the task, members of both castes perform equivalently but when caste membership is made salient members of the low caste generate lower production. The authors argue that the explanation for this result is that low caste members may have believed that their effort would be rewarded less than that of others due to prior experience with discrimination even though the rewards were in fact common.⁴ This result is part of a larger literature on how social identity and organizational structure can effect behavior as explored in Akerlof and Kranton (2000) and Akerlof and Kranton (2005). There are multiple ways in which this result is important for the current study. The first important implication is that this result seems to confirm the claim that members of lower castes who feel discriminated against will exhibit a strong discouragement effect and in fact it can be so strong that even when the differential treatment is removed the

⁴The production task involved the subjects solving mazes and there were graders who determined if a maze had been correctly solved. So the expectation of differential returns could have been based on expected differential treatment from the grader eventhough caste status was never communicated to the graders.

effects may still persist. The additional implications will be explored in regard to two later hypotheses.

Taken as a whole, these various strains of literature build a compelling case that there is at a minimum an expectation that wage inequality could have a significant impact on productivity even if the effect has not been robustly demonstrated. In addition to the *discouragement* effect derived from the Fair Wage-Effort Hypothesis there is also the potential for an opposite *encouragement* effect which might influence the behavior of those earning high wages. Though this effect is certainly discussed much less in these literatures, the gift-exchange literature discussed above provides one justification for this hypothesis in that workers receiving a high wage may perceive it to be a “gift” and respond with high effort. Again, one must be careful to separate the pure wage effect from the inequality effect but if the existence of inequality enhances the view of the high wage as a gift then this could lead to the observance of an encouragement effect among the high paid workers. We mention this hypothesis mainly because of the economic efficiency implications which are that even if the discouragement effect exists then if this encouragement effect also exists it could counterbalance the discouragement effect leaving any negative efficiency consequences from the discouragement effect ameliorated or possibly even counterbalanced completely. This leads to our second hypothesis:

Hypothesis 2 (Encouragement Hypothesis) - *The productivity of the high wage earners should be higher in the two treatment conditions involving inequality than in the control treatment without inequality.*

In discussing the possibility for this discouragement effect it is important to understand what forces might be generating it. One source that might at first seem like a possibility for generating it can be found in the literature on social preferences in which there are common findings supporting the claim that individuals dislike inequality as demonstrated theoretically in the model of inequality aversion proposed by Fehr and Schmidt (1999) and by the ERC model in Bolton and Ockenfels (2000) among others. To determine if inequality aversion can explain the discouragement effect we need to think about how to extend these models into the production domain because the existing models of inequality aversion deal only with preferences over final allocations and not decisions over wealth production. The most direct way to extend these models is to assume that the inequality aversion manifests itself as a reduction in utility when final earnings among members of a group are unequal. Consider the utility function specified in equation 1 where w_i is the wage received by individual i , n_i is the productivity of individual i , $c(n_i)$ represents the cost to individual i of producing at the level of n_i , $h()$ measures the level of inequality in final wealth among all the relevant agents, and β_i is a parameter indicating the degree to which agent i is averse to inequality.

$$u(w_1n_1, \dots, w_In_I) = w_i * n_i - c(n_i) - \beta_i h(w_1n_1, \dots, w_In_I) \quad (1)$$

Cost of effort should clearly be increasing in effort, or $c' > 0$, while the effect of one individual’s effort on the overall level of inequality is contingent on their earnings relative to the overall average. If we consider $h()$ to just be a measure of the variance in the earnings distribution, which is one reasonable way of measuring the degree of inequality,

then any individual whose earnings are below the average will decrease h by increasing their own production, n_i , while anyone with earnings above average earnings will increase h by increasing n_i . If we let if \overline{WN} be the average earnings then the properties of $h()$ are summarize in equations 2 and 3.

$$h' < 0 \text{ if } w_i n_i < \overline{WN} \quad (2)$$

$$h' > 0 \text{ if } w_i n_i > \overline{WN} \quad (3)$$

The effect of adding $h()$ to the standard utility function is that for low wage earners higher productivity decreases expected inequality and this could be seen as offsetting the effort cost of increased productivity. This should lead to low wage earners producing at a higher level than if they were unconcerned about inequality in final earnings. High wage earners who are averse to inequality would face an additional cost from high production above a certain level in addition to any effort cost because increased productivity would worsen inequality in final earnings. Consequently they should be predicted to decrease work effort when faced with inequality in order to make the final wealth distribution more equitable. Predicting exactly how efforts would be chosen according to this model would require solving for a Nash equilibrium among all of the workers but since we have no control over the functional form of $c()$ or $h()$ in this experiment our focus will be only on the comparative static effects of moving from a condition with no wage inequality to one with wage inequality.

For any approaches to modeling inequality aversion similar⁵ to equation 1, we will find behavioral predictions opposite to those in hypotheses 1 and 2 which demonstrates that this way of extending inequality aversion to production decisions is actually not consistent with the discouragement effect. The key point that this demonstrates is that where the concern for inequality is focused is important for the direction of the predicted effect of inequality. In the motivation behind the discouragement effect the concern is placed at the intermediate wage level while here the focus is on final wealth allocations. The fact that these constructions yield conflicting directional predictions is quite interesting and may help later in interpreting any behavioral response we may see due to inequality. Towards that end we can state a third hypothesis based on this model of inequality aversion.

Hypothesis 3 (Inequity Aversion Hypothesis) - *The productivity of high (low) wage earners should be lower (higher) in the two treatment conditions involving inequality than in the control treatment without inequality.*

The previous hypotheses deal only with the effects of inequality itself and not with the group size treatments. Our intention with the group size treatments was to provide different contexts through which subjects might view the inequality. From the point of view of the low wage subjects, one might think that if they were almost singled out to be in the minority of the subjects in the experiment to receive a low wage that this might be viewed as less

⁵We note that ours is a primitive specification of inequality aversion which does not allow for such things as individuals caring more about inequality when their earnings are below average than when they are above. Augmenting the model to include aspects like these would not change the fundamental comparative statics so long as utility is decreasing in inequality and inequality is measured based on the variance in final earnings.

fair than were they just one among many who received the low wage. Consequently, if the group size differential matters we are led to make our fourth hypothesis:

Hypothesis 4 (Group Status Hypothesis) - *The productivity will be more responsive to the existence of inequality in the Disadvantaged Minority Treatment than the Disadvantaged Majority Treatment.*

The main motivation behind this hypothesis is to help understand the nature of any discouragement effect we might find and the specific prediction if the discouragement effect is operative is that if low wage workers view their position as less fair when they are in the minority then there will be a larger discouragement effect in that treatment. Due to the differential effects of a high/low wage workers effort on the overall level of earnings inequality in the different group configurations one can also construct an argument in which the inequality aversion effects are similarly magnified.

In addition to these broad treatment effects, we also envision that different groups should be more likely to exhibit a response to the presence of inequality than others. In particular, groups who might be more familiar with expecting that their returns from effort would be rewarded less than that of others should be more likely to exhibit a response to inequality than those less familiar with that position. This potential effect is the second implication we can draw from the results in Hoff and Pandey (2006) which showed that low caste members exhibited a discouragement effect in their study potentially related to their prior experience with discrimination. This suggests that in our setting in which we are inducing differential wages that members of groups who have experienced work force or wage discrimination may be more likely to exhibit responses to the existence of those differential wages. Due to the size and make-up of our sample we are unable to test for this effect along ethnic or racial lines which might be more likely to generate an effect but we can test for it along gender lines which leads to our fifth hypothesis:

Hypothesis 5 (Gender Hypothesis) - *Women will exhibit a discouragement effect while men should not.*

The basis for hypothesis 5 is the presumption that many women may feel as though women are compensated less than men for equal effort as this is reported quite often in the popular news media. The existence of a pay gap for women is much debated in the academic literature though because although there is clear evidence that on average women receive lower wages than men, this gap has been declining, Blau and Kahn (2000), and much of the gap can be attributed to occupational choice rather than overt discrimination, Macpherson and Hirsch (1995). Further, more recent data show that more women are now attending college than men with this at least partially attributed to greater expected career prospects for women now than in decades past, Goldin, Katz, and Kuziemko (2006). Consequently, the college age women in our sample may be unlikely to have internalized much of the past discrimination and they may therefore be unlikely to exhibit this greater sensitivity to inequality.

We can generate a related hypothesis based upon the results from the pattern matching questions that were designed to measure the facility of our subjects with numerate tasks. To the extent to which this task can be claimed to measure something about the cognitive

skills of our subjects, the scores on this task can be used as a general measure of cognitive ability. Subjects who are generally less able may also be used to facing situations in which they expect less reward for a similar level of effort as others and this too may predispose them to exhibiting a discouragement effect. This is the basis for our sixth hypothesis:

Hypothesis 6 (Numeracy Hypothesis) - *Subjects who score lower on the numeracy measure should exhibit a discouragement effect while those subjects scoring higher on the measure should not.*

Our final hypothesis is related to the performance on the CRT measure of the degree to which our subjects engage in reflective versus impulsive decision making. The tendency of a person to think deeply about matters related to the experiment is important to the potential judgement of whether or not the wage assignment is unfair. The wage rate assignments are unequal which could lead to an initial response by impulsive decision makers of judging that inequality to be “unfair.” More reflective decision makers though might think more deeply about the procedure used to assign those wage rates, pure randomization, and view the procedure itself as fair and judge the outcome that procedure to also be fair. This split judgement regarding the fairness from an allocation procedure like this is discussed in detail in Bolton, Brandts, and Ockenfels (2005) in which the authors provide evidence that some individuals will view the outcome of a procedurally fair process as fair even if the outcome of the process is inherently unfair while others will judge the fairness solely on the basis of the outcome itself. Since looking past the allocation itself to the process requires a more reflective thought process, it is a reasonable claim that subjects exhibiting more reflective decision making tendencies as measured by the CRT will be less likely to view the unequal wage allocations as unfair while those measured to be more impulsive will be more likely to see the allocation as unfair. Given that a behavioral response to the wage inequality is predicated on the judgement that the inequality is to some degree unfair, this reasoning leads to our seventh and final hypothesis.

Hypothesis 7 (CRT Hypothesis) - *Subjects who score lower on the CRT and are therefore measured to be more impulsive should exhibit a discouragement effect while the subjects scoring higher on the CRT should not.*

4 Results

4.1 Data Overview

As an initial look at the data from the experiment, table 2 displays some raw summary statistics regarding average per round output and average number of times per round workers chose to play TTT by treatment with low wage and high wage workers broken out separately. At this level of aggregation it appears that the general directional effects predicted by hypotheses 1 and 2 are present in that for low (high) wage workers the average productivity is lower (higher) in the two sessions with inequality. The other interesting point about aggregate productivity is that if there is a wage effect it appears to be negative because high wage workers produce less than low wage workers in the common wage treatments. We can also note that while neither low nor high wage workers play TTT with great frequency,

Table 2: Summary Statistics

		Common Wage	Disad Minority	Disad Majority
Low Wage Workers	Output	28.05 (8.65)	27.54 (8.37)	26.59 (8.36)
	TTT	1.58 (4.62)	1.35 (3.95)	1.31 (4.06)
High Wage Workers	Output	26.39 (9.25)	27.58 (8.16)	28.27 (6.56)
	TTT	1.49 (2.66)	0.99 (2.52)	0.43 (1.55)

Standard errors in parentheses.

there is some tendency for low wage workers to play more often (though with high variance) and high wage workers to play less when confronted with inequality.

In addition to the basic summary statistics we can examine figure 1 for another general overview of the data. This figure shows the average contributions per round separated out by treatment and by high wage versus low wage earners. The first characteristic to note is that there is definitely a learning trend by the subjects as output increases over time likely due to subjects gaining greater facility with the task and memorizing portions of the code key. Further, we also see that for low wage earners, the increase in productivity appears to be less in the treatments with inequality than in the control treatment without inequality while the reverse is true for the high wage workers. The combination of the summary statistics and the figure suggest that there may be a behavioral response to the inequality but the data at this level of aggregation does not yet allow us to formally test any of the above hypotheses.

Prior to engaging in the tests of the hypotheses though it is also useful to examine some baseline results on the determinants of productivity that will be important for framing the later analysis. Table 3 contains OLS regressions of average per period output on a dummy variable for the wage rate, a dummy variable for whether or not the subject is in the small group and then all of the demographic and cognitive control measures we will be using in later regressions. It also includes a similar set of OLS regressions with the dependent variable equal to the fraction of rounds in which a subject chose to play TTT.⁶ The only data considered in these regressions is derived from subjects in the common wage control treatments and each subject from those sessions enters into these regressions as a single data point. Note that the dependent variables are average output and average propensity to play TTT calculated over the entire session and so we have one observation for each of the 48 subjects in the common wage sessions. The question these regressions are intended to address is whether any of the demographic and cognitive measures we are using have any substantial impact on productivity or propensity to play TTT. Most of the variables should be self-explanatory but our cognitive measures warrant further explanation. Both CRT and Numeracy variables are constructed from the raw scores of how many questions

⁶We have conducted alternative specifications including probit in which we use a binary measure of whether or not a subject chose to play TTT in each round (using round by round data) and find the same results.

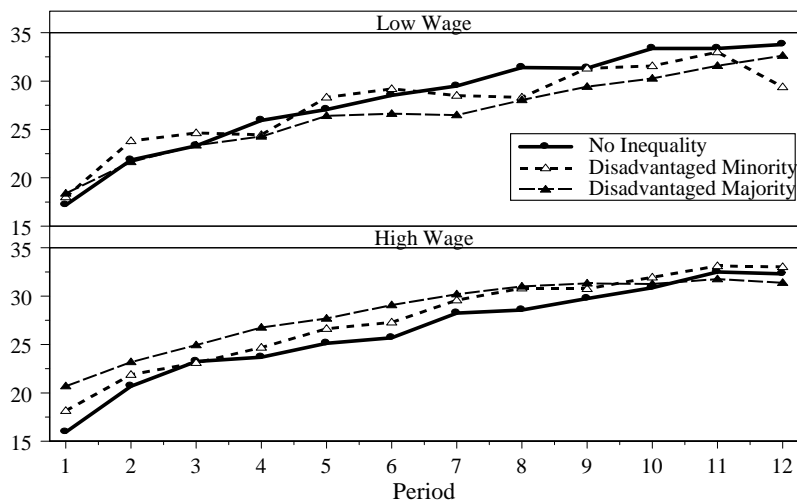


Figure 1: Average production per round.

a subject answered correctly on each set to produce binary dummy variables. The CRT variable is set equal to 1 if a subject answered any of the CRT questions correctly and 0 else. Out of our 176 subjects 57 got at least 1 answer correct on the CRT and the other 119 had 0 correct answers.⁷ In the case of Numeracy, the variable is set equal to 1 if the subject gave a number of correct answers on the pattern matching exercise which was above the median number correct (5) while the variable is set equal to 0 for those who gave a number of correct answers that was at or below the median. It should be quite clear that the CRT measure and this Numeracy measure are not measuring the same phenomenon because the correlation coefficient between these two constructed variables is 0.027.

The results from the regressions in table 3 demonstrate that for the most part, the demographic and cognitive measures have little impact on the performance of our subjects in the production task. There is evidence that NonWhite subjects may have been less likely to play TTT than White subjects but this difference had no significant impact on productivity. In particular we note that the three variables that we base hypotheses on, Female, CRT and Numeracy, are all uncorrelated with productivity. These results, of course, have no direct bearing on our stated hypotheses because the hypotheses concern how individuals of different types respond to inequality and the data from the sessions with inequality is not included in these regressions. What these regressions demonstrate is that none of these characteristics can be seen as proxies for ability on this specific task. Were any of our measures such as numeracy or gender correlated with performance on this task

⁷We note that this is lower performance than found in Frederick (2005), but in our case this test was administered after a long and tedious experiment and so some performance drop is to be expected. Since all we really wish to do is separate out the most deliberate thinkers from the others, this should not be much of a problem.

Table 3: Determinants of productivity and probability of playing TTT in common wage treatments.

	Output				Played TTT			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High Wage	-1.669 (2.066)	-2.270 (2.170)	-1.758 (2.036)	-2.111 (2.157)	0.086 (0.086)	0.068 (0.087)	0.091 (0.087)	0.072 (0.091)
Small Group	3.826** (1.751)	3.520* (1.915)	3.315 (1.994)	2.837 (2.154)	-0.125 (0.080)	-0.092 (0.086)	-0.134 (0.093)	-0.101 (0.105)
Age		3.430 (3.229)		4.505 (3.272)		-0.105 (0.117)		-0.074 (0.131)
Agesq		-0.079 (0.062)		-0.101 (0.064)		0.002 (0.002)		0.002 (0.003)
Female		1.534 (2.081)		1.117 (2.420)		0.073 (0.079)		0.072 (0.092)
NonWhite		1.234 (1.795)		1.369 (1.814)		-0.173** (0.075)		-0.169** (0.078)
Upperclass		0.491 (1.934)		-0.154 (2.248)		-0.118 (0.081)		-0.130 (0.085)
CRT			-1.697 (1.858)	-0.502 (2.252)			0.014 (0.079)	0.015 (0.090)
Numeracy			0.747 (1.950)	1.845 (2.203)			0.037 (0.084)	0.035 (0.099)
Constant	27.098*** (1.206)	-10.852 (39.301)	27.565*** (1.467)	-23.289 (39.340)	0.305*** (0.053)	1.555 (1.431)	0.284*** (0.071)	1.183 (1.602)
Obs	48	48	48	48	48	48	48	48
R^2	0.08	0.16	0.10	0.18	0.06	0.20	0.07	0.20

Robust standard errors in parentheses

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

then should we see a greater discouragement effect for low numeracy or female subjects and these groups also had lower productivity then it would not be clear if the discouragement effect was related to a lack of skill or facility with this task or to the other traits we are attempting to measure. Since productivity is in general not correlated with any of our cognitive or demographic characteristics this means that this task is well suited to our purposes.

4.2 Tests of Hypotheses

To present the formal results from the analysis we will proceed by presenting a series of regressions which allow us to test each of the seven hypotheses we previously developed. In each set of regressions we will use data from low and high wage earners separately and we will also present multiple specifications which will include either the demographic control variables contained in the regressions in table 3 (Age, Agesq, Female, NonWhite and Upperclass) the cognitive controls (CRT and Numeracy) or both. All of the regressions conducted use the output of each subject in each period as the dependent variable and are conducted as panel regressions using a subject level random effects specification with the standard errors clustered by subject. All regressions include a linear trend (i.e. the current period) to account for the general trend observed in figure 1 and a constant term. While the observed trend is not exactly linear, all of our results are robust to using separate dummy variables for each period to account for the trend and so we present the more parsimonious specification. Due to the fact that the coefficients on the trend variable, the constant and on the demographic/cognitive measures are of no intrinsic interest, we will suppress the actual values of those coefficients to conserve space but we will note when they are or are not included in the specifications. This allows us to concentrate on the key variables of interest for inference which are the dummy variables for the two inequality treatments, a dummy variable equal to one in the second half of a session and then interaction terms which interact the two treatment dummies with the second half dummy. Our inference regarding our hypotheses will be mainly based on those interaction terms. The reason for that is that the overall treatment dummies can be thought of as measuring two things. They provide a measurement of the initial response of the subjects to the treatments but they also provide a measure of any general group level differences of subjects across treatments. Due to the nature of the hypothesized effects, the discouragement effect in particular, it is reasonable to expect that the true impact of the existence of inequality and the treatments in general may take time to develop and so it is the behavior in the second half of the experiments, which is captured by the interaction terms, where we should expect to see the effect of the treatments. Because we have the treatment dummy and these interaction terms both, the interaction terms serve as difference in differences estimates of the treatment effect which allows us to clearly separate out any subject level heterogeneity across treatments and identify a purer version of the treatment effect.

The two base hypotheses underlying this study concern how low wage and high wage workers alter their productivity in the presence of inequality. Table 4 contains two sets of regressions using the structure outlined above examining this issue for the overall sample. As noted, one set of regressions includes only data from subjects receiving the low wage rate while the other set includes only data from subjects receiving the high wage rate.

Table 4: Response of Low and High wage workers to inequality.

	Dependent: Variable Output							
	Low Wage Workers				High Wage Workers			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Disad Minority	0.750 (1.376)	0.577 (1.417)	0.749 (1.399)	0.751 (1.426)	1.226 (1.637)	0.904 (1.800)	0.378 (1.470)	0.210 (1.662)
Disad Majority	-0.538 (1.171)	-1.005 (1.148)	-0.531 (1.150)	-0.843 (1.151)	2.990* (1.736)	2.721 (1.761)	2.018 (1.542)	1.900 (1.575)
SecondHalf(SH)	-0.316 (0.671)	-0.316 (0.673)	-0.316 (0.672)	-0.316 (0.673)	0.753 (1.041)	0.753 (1.044)	0.753 (1.042)	0.753 (1.045)
DisadMin*SH	-2.536* (1.373)	-2.536* (1.376)	-2.536* (1.374)	-2.536* (1.377)	-0.066 (1.330)	-0.066 (1.333)	-0.066 (1.331)	-0.066 (1.335)
DisadMaj*SH	-1.856* (1.003)	-1.856* (1.005)	-1.856* (1.004)	-1.856* (1.006)	-2.219 (1.485)	-2.219 (1.489)	-2.219 (1.486)	-2.219 (1.490)
Control Variables:								
Linear Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	No	Yes	No	Yes	No	Yes	No	Yes
Cognitive Scores	No	No	Yes	Yes	No	No	Yes	Yes
Obs (Groups)	1152 (96)				960 (80)			

Notes: Robust standard errors clustered by subject in parentheses

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

All regressions include constant.

Cognitive Measures: CRT and Numeracy Demographics : age, agesq, NonWhite, Upperclass, Female

Examining the two interaction terms DisadMin*SH and DisadMaj*SH leads to our first two results which correspond to tests of hypotheses 1 and 2.

Result 1 (Fair Wage–Effort / Discouragement Hypothesis) - *Workers receiving a low wage in the presence of others receiving a high wage generate statistically significantly lower output than workers receiving a low wage without the presence of high wage workers.*

Result 2 (Encouragement Hypothesis) - *The productivity of high wage workers shows no sensitivity to the presence or absence of inequality.*

The support for both of these claims is seen in table 4 in which we find that across all specifications of the regressions, the coefficients on the interaction terms are negative and significant for the low wage workers but insignificant for the high wage workers. Consequently we find support for hypothesis 1, the Fair Wage-Effort Hypothesis, but no support for the hypothesis 2, the Encouragement Hypothesis. There are a few important details about these estimates which are important for understanding how they should be interpreted. The first issue to note is that the coefficients on the interaction terms remain unchanged across all four specifications per data set which add and subtract control variables to the regression. This is to be expected since those control variables combined with the main treatment dummies are measuring any differences across treatments due to any differences in the subject population while the interaction terms measure the treatment effect net of those population differences. Consequently, adding and subtracting control variables should not be expected to have any impact on these interaction terms and so these estimated effects should be interpreted as strongly robust.

We should note, however, that the significance results indicated in table 4 reflect two sided tests. For the overall efficiency results that this paper is chiefly concerned with that is the appropriate test specification for determining if there is an overall effect on productivity associated with inequality. Given the stated hypotheses, one might reasonably argue that testing the explicit behavioral predictions requires one sided tests. Since the discouragement effect is significant for low wage workers on a two sided test, it should be clear that it will also be significant in a one sided test and in fact the significance level improves to the 5% level (p -value=0.0325). Since the sign of the relevant coefficients to test the encouragement effect are all negative, there is no test necessary to confirm that the encouragement hypothesis is not supported.

As an ancillary point we can also examine the main treatment effects as measured by the treatment dummies. These are mostly insignificant in table 4 and the coefficient estimate does vary as different control variables are added and subtracted. Since the coefficient estimates are varying depending on which set of control variables are included this indicates that there is some heterogeneity in subjects across treatments and those control variables are able to explain some of those differences. The overall interpretation though is that while there is no immediate response to the presence of inequality, the significance of the interaction terms for the low wage workers suggests that there is an effect after sustained experience with the differential wages.

We can also test hypothesis 3 based on the results in table 4 and the implications for it are clear since hypothesis 3 generated predictions exactly opposite of hypotheses 1 and 2.

This leads to our third result which is a rejection of hypothesis 3.

Result 3 (Inequity Aversion Hypothesis) - *The productivity of low wage earners declines in the presence of inequality while that of high wage earners is unchanged.*

This result is again based on two-sided significance tests but here we do see a small difference if we use one-sided tests instead. For low wage workers the effect seen is negative which clearly contradicts hypothesis 3 and no one sided test will change that. The predicted and observed effect for high wage workers are both negative, though, and a one-sided significance test does show that the effect is significant for the Disadvantaged Majority treatment at the 10% level (p -value=0.0685) while it is still not significant for the other treatment (p -value=0.4805). So there may be some evidence in favor of inequality aversion as modeled in equation 1 for high wage earners in the Disadvantaged Majority treatment. This is the treatment expected to be more likely to yield such an effect for high wage earners because they are in the minority and so one of them dropping their productivity has a greater effect on decreasing overall inequality than when there are a number of high wage earners. The key point for the purposes of this paper though is, regardless of what is driving the behavioral response, that the effect of inequality on production is not positive for either low or high wage workers.⁸

The next hypothesis that we can check is the group status hypothesis which states that the productivity response should be larger in the Disadvantaged Minority treatment than in the Disadvantaged Majority treatment. Since there are no significant productivity effects for the high wage workers, this hypothesis is not applicable to them but our fourth result is based on testing this hypothesis for the low wage earners.

Result 4 (Group Status Hypothesis) - *While the coefficient on the response to inequality for low wage earners is larger (in absolute value) in the Disadvantaged Minority treatment than the Disadvantaged Majority treatment, this difference is not statistically significant.*

The test of this hypothesis requires checking whether or not the coefficient $|DisadMin * SH| > |DisadMaj * SH|$ and we are only really interested in the test for the low wage workers. While the base predicted relationship does hold for them, the difference is not significant under any of the specifications (p -value=0.604). The failure of this hypothesis suggests that individuals judge the fairness of relative wages without consideration for the relative sizes of the advantaged and disadvantaged groups.

To test the remaining hypotheses, we have to conduct similar regressions to those found in table 4 for different subsets of the data. The first of these that we will test concerns

⁸We also note that the tests of hypotheses 1-3 presented are based as tests of models which do not allow both encouragement/discouragement and inequality averse effects simultaneously. One could construct a hybrid model with both forces present but this makes identification in the data difficult since all one can identify is the net effect from these two countervailing forces. Under that assumption we can still show clear evidence that a discouragement effect exists for low wage workers while there is still the possibility that both encouragement and inequality aversion effects exist but are cancelled out. Because the central concern of this paper is the overall efficiency consequences of inequality, we do not pursue separating these effects more finely.

whether or not there is a differential response to the inequality according to the gender of the subjects. Table 5 contains the results of regressions with the same specification as before but conducted separately for low/high wage females and low/high wage males. Hypothesis 5 suggested that female subjects might exhibit more of a discouragement effect than the male subjects but our the results in table suggest otherwise leading to our fifth result.

Result 5 (Gender Hypothesis) - *The productivity of females displays no statistically significant response to the presence of inequality.*

For both low wage and high wage females, we see no statistically significant response to the presence of inequality which leads to a clear rejection of the Gender Hypothesis. We do see that male subjects display a statistically significant discouragement effect in the Disadvantaged Majority treatment which suggests that if anything we find a reversal of the Gender Hypothesis, but it is almost certainly not the case that the reason for the observed response for males is due to a gender specific experience of prior discrimination. Overall we see little reason to suggest that women are responding any more to the inequality than men. This finding suggests that women do not see themselves as members of a low caste as the subjects in the Hoff and Pandey (2006) experiment and one might argue that this is consistent with the arguments in Goldin, Katz, and Kuziemko (2006) regarding the increase in college enrollments of women.

The next hypothesis we will test is whether or not the low numeracy subjects exhibit a differential response to the presence of inequality. Table 6 contains sets of regressions for low and high wage subjects with the data set split into low numeracy (i.e. those who gave a number of correct responses at or below the median number correct) and high numeracy (i.e. those who gave above the median number of correct responses) subjects and those regressions lead to our next result.

Result 6 (Numeracy Hypothesis) - *The productivity of low numeracy subjects receiving a low wage exhibits a negative and significant response to the presence of inequality while the high numeracy subjects display no response to inequality.*

With the data set split into low and high numeracy, the only sub-sample for which there is a negative and significant response to the presence of inequality in the low numeracy subjects receiving the low wage. The responses to the existence of inequality are not significant for the high wage / low numeracy subjects or any of the high numeracy subjects. This provides clear support for hypothesis 6 which predicted that the low numeracy subjects, perhaps due to their prior experience with receiving differential returns to effort, would exhibit a response to inequality while the high numeracy subjects would not. It is useful to reiterate that the results in table 3 shows that numeracy is not related to productivity in the baseline treatments and so when we segregate the sample by numeracy we are not also segregating the sample by productivity. Consequently, whatever quality of our subjects that our numeracy score is measuring is at least not simply a proxy for performance on this specific task.

Given that we find a discouragement effect among the low numeracy subjects we can also test hypothesis 4, the Group Status Hypothesis, again with this population. Again we

Table 5: Response of Male and Female workers to inequality.

	Dependent Variable: Output							
	Female Subjects				Male Subjects			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Low Wage Workers								
Disad Minority	0.201 (1.966)	0.318 (2.260)	0.655 (1.912)	1.099 (2.235)	1.158 (1.857)	1.010 (1.810)	1.179 (1.879)	1.048 (1.791)
Disad Majority	-1.464 (1.753)	-0.966 (1.789)	-1.168 (1.686)	-0.142 (1.657)	0.093 (1.635)	-0.265 (1.622)	0.129 (1.754)	-0.147 (1.765)
SecondHalf(SH)	0.233 (1.141)	0.233 (1.146)	0.233 (1.143)	0.233 (1.148)	-0.630 (0.824)	-0.630 (0.827)	-0.630 (0.825)	-0.630 (0.828)
DisadMin*SH	-2.216 (1.730)	-2.216 (1.736)	-2.216 (1.733)	-2.216 (1.740)	-2.651 (1.880)	-2.651 (1.886)	-2.651 (1.883)	-2.651 (1.889)
DisadMaj*SH	-1.550 (1.692)	-1.550 (1.699)	-1.550 (1.696)	-1.550 (1.702)	-2.561** (1.243)	-2.561** (1.247)	-2.561** (1.245)	-2.561** (1.249)
Panel B: High Wage Workers								
Disad Minority	-0.594 (2.534)	0.722 (2.544)	-0.463 (2.186)	0.453 (2.452)	2.322 (2.340)	1.627 (2.468)	1.581 (1.914)	0.327 (2.101)
Disad Majority	3.267 (2.492)	3.859* (2.318)	3.580 (2.254)	4.185* (2.293)	2.715 (2.502)	2.708 (2.536)	1.310 (1.905)	0.686 (1.952)
SecondHalf(SH)	1.224 (1.334)	1.224 (1.340)	1.224 (1.337)	1.224 (1.343)	0.012 (1.470)	0.012 (1.476)	0.012 (1.473)	0.012 (1.479)
DisadMin*SH	0.156 (1.874)	0.156 (1.883)	0.156 (1.879)	0.156 (1.888)	0.250 (1.786)	0.250 (1.793)	0.250 (1.790)	0.250 (1.797)
DisadMaj*SH	-3.067 (2.021)	-3.067 (2.031)	-3.067 (2.026)	-3.067 (2.036)	-1.090 (2.090)	-1.090 (2.098)	-1.090 (2.094)	-1.090 (2.102)
Control Variables:								
Linear Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	No	Yes	No	Yes	No	Yes	No	Yes
Cognitive Scores	No	No	Yes	Yes	No	No	Yes	Yes
Obs (Groups)	A: 528 (44) B: 432 (36)				A: 624 (52) B: 528 (44)			

Notes: Robust standard errors clustered by subject in parentheses

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

All regressions include constant.

Cognitive Measures: CRT and Numeracy Demographics : age, agesq, NonWhite, Upperclass, Female

Table 6: Response of Low and High Numeracy workers to inequality.

	Dependent Variable: Output							
	Low Numeracy Subjects				High Numeracy Subjects			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Low Wage Workers								
Disad Minority	-0.581 (1.672)	-0.642 (1.672)	-0.603 (1.709)	-0.586 (1.723)	2.728 (2.307)	2.682 (2.262)	2.736 (2.248)	2.920 (2.192)
Disad Majority	-1.291 (1.555)	-1.238 (1.637)	-1.301 (1.560)	-1.239 (1.642)	0.575 (1.719)	-0.271 (1.744)	0.588 (1.710)	-0.169 (1.833)
SecondHalf(SH)	-0.022 (0.821)	-0.022 (0.824)	-0.022 (0.821)	-0.022 (0.824)	-0.564 (1.024)	-0.564 (1.030)	-0.564 (1.026)	-0.564 (1.031)
DisadMin*SH	-2.811** (1.288)	-2.811** (1.293)	-2.811** (1.289)	-2.811** (1.294)	-2.333 (2.870)	-2.333 (2.887)	-2.333 (2.874)	-2.333 (2.890)
DisadMaj*SH	-2.174** (1.087)	-2.174** (1.091)	-2.174** (1.088)	-2.174** (1.092)	-1.632 (1.787)	-1.632 (1.797)	-1.632 (1.789)	-1.632 (1.799)
Panel B: High Wage Workers								
Disad Minority	1.833 (2.161)	1.938 (2.387)	1.747 (2.209)	1.869 (2.435)	-1.617 (1.852)	-2.170 (2.263)	-1.490 (2.022)	-3.364 (2.440)
Disad Majority	3.467 (2.193)	4.511* (2.342)	4.185* (2.349)	5.196** (2.369)	0.370 (1.967)	-0.410 (2.432)	0.074 (2.072)	-1.938 (2.611)
SecondHalf(SH)	0.468 (1.204)	0.468 (1.211)	0.468 (1.205)	0.468 (1.212)	1.412 (1.849)	1.412 (1.858)	1.412 (1.851)	1.412 (1.860)
DisadMin*SH	0.567 (1.737)	0.567 (1.747)	0.567 (1.739)	0.567 (1.749)	-1.006 (2.154)	-1.006 (2.165)	-1.006 (2.156)	-1.006 (2.167)
DisadMaj*SH	-2.205 (1.768)	-2.205 (1.778)	-2.205 (1.770)	-2.205 (1.780)	-2.685 (2.437)	-2.685 (2.449)	-2.685 (2.439)	-2.685 (2.451)
Control Variables:								
Linear Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	No	Yes	No	Yes	No	Yes	No	Yes
Cognitive Scores	No	No	Yes	Yes	No	No	Yes	Yes
Obs (Groups)	A: 708 (59) B: 444 (37)				A: 444 (37) B: 516 (43)			

Notes: Robust standard errors clustered by subject in parentheses,

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

All regressions include constant.

Cognitive Measures: CRT and Numeracy Demographics : age, agesq, NonWhite, Upperclass, Female

find that the DisadMin*SH coefficient is greater in absolute value than the DisadMaj*SH coefficient but this difference is insignificant. We also note once again that the uninteracted treatment effects are typically insignificant again suggesting that there is not a strong initial response to inequality among these populations either.

The final hypothesis we will check is the CRT hypothesis and to test it we have separated the population into two groups according to their performance on the CRT. We have constructed one set of data including all of the subjects who gave 0 correct answers on the CRT and then all subjects who gave at least one correct answer have been included in the second data set. The results from conducting our standard set of regressions on these two sub-samples with again the low and high wage workers being analyzed separately are contained in table 7. The results from this table form the basis of our seventh result.

Result 7 (CRT Hypothesis) - *Subjects in the CRT=0 sub-sample receiving a low wage exhibit a negative and statistically significant response to the presence of inequality. Subjects in the CRT>0 sample receiving a low wage exhibit no statistically significant response to the presence of inequality.*

We find strong support for the CRT hypothesis. The discouragement effect found for the low wage / CRT=0 subjects is around twice the magnitude of the response from the overall sample. We find that for the high CRT subjects earning a low wage there is a statistically significant increase in productivity in the inequality treatments relative to the no inequality control treatment which is what offsets the effect from the other subjects to decrease the magnitude of the response in the overall sample. We can once again check to see whether or not the group status hypothesis applies to this subsample and we again find that on an absolute basis the discouragement effect is greater in the Disadvantaged Minority treatment than the Disadvantaged Majority treatment as before but again it is insignificant. Further, the base treatment effects are again mostly insignificant indicating a lack of an immediate response to the inequality.

In interpreting these results it is important to be clear about two elements of the data. First, as shown in table 3 there is no correlation between CRT score and productivity in the control treatments which means we are not simply separating our sample by productivity. Second, while the CRT and numeracy measures are both measures of the cognitive capacities of our subjects, there is a correlation of .027 between these two measures which should make it clear that they are measuring different aspects of the decision making of the subjects. This supports the claim that the CRT measure could be a measure of the degree to which our subjects engage in reflective thinking. Since it is the more reflective thinkers who are more likely to consider the fairness of the allocation procedure in judging the fairness of the allocation, as did some of the subjects in Bolton, Brandts, and Ockenfels (2005), it seems the explanation for this differential effect is due to the low CRT subjects being more likely to view the differential wage assignment as unfair. This result is quite useful in helping to understand how our results may generalize outside of this laboratory environment.

4.3 Inequality and Efficiency

The previous section focussed on testing the hypotheses regarding the behavioral responses of individuals to the presence of inequality. The ultimate reason that those responses are

Table 7: Response of Low and High CRT workers to inequality.

	Dependent Variable: Output							
	CRT=0 Subjects				CRT>0 Subjects			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Low Wage Workers								
Disad Minority	-1.103 (1.622)	-0.882 (1.620)	-1.098 (1.647)	-0.833 (1.652)	4.644** (2.333)	4.474 (2.761)	4.661** (2.259)	4.067 (2.665)
Disad Majority	-1.561 (1.434)	-1.715 (1.382)	-1.557 (1.423)	-1.675 (1.385)	0.885 (2.005)	-0.621 (2.544)	0.987 (1.897)	-0.621 (2.519)
SecondHalf(SH)	0.694 (0.792)	0.694 (0.795)	0.694 (0.793)	0.694 (0.795)	-1.995* (1.081)	-1.995* (1.089)	-1.995* (1.083)	-1.995* (1.091)
DisadMin*SH	-4.455*** (1.380)	-4.455*** (1.384)	-4.455*** (1.381)	-4.455*** (1.385)	2.090 (2.640)	2.090 (2.659)	2.090 (2.644)	2.090 (2.663)
DisadMaj*SH	-3.248*** (1.067)	-3.248*** (1.071)	-3.248*** (1.068)	-3.248*** (1.071)	0.987 (1.891)	0.987 (1.905)	0.987 (1.894)	0.987 (1.908)
Panel B: High Wage Workers								
Disad Minority	0.835 (2.132)	0.477 (2.380)	0.296 (2.030)	0.184 (2.304)	2.367 (2.379)	2.113 (2.253)	0.887 (2.039)	0.130 (1.665)
Disad Majority	4.074* (2.405)	3.007 (2.425)	3.163 (2.464)	2.341 (2.429)	1.885 (2.247)	2.008 (2.096)	1.580 (1.813)	2.195 (1.561)
SecondHalf(SH)	0.430 (1.221)	0.430 (1.226)	0.430 (1.222)	0.430 (1.227)	1.492 (1.978)	1.492 (1.994)	1.492 (1.981)	1.492 (1.997)
DisadMin*SH	0.545 (1.576)	0.545 (1.582)	0.545 (1.577)	0.545 (1.584)	-1.497 (2.453)	-1.497 (2.473)	-1.497 (2.457)	-1.497 (2.477)
DisadMaj*SH	-0.331 (1.767)	-0.331 (1.774)	-0.331 (1.769)	-0.331 (1.776)	-4.304* (2.589)	-4.304* (2.609)	-4.304* (2.593)	-4.304* (2.613)
Control Variables:								
Linear Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	No	Yes	No	Yes	No	Yes	No	Yes
Cognitive Scores	No	No	Yes	Yes	No	No	Yes	Yes
Obs (Groups)	A: 792 (66) B: 636 (53)				A: 360 (30) B: 324 (27)			

Notes: Robust standard errors clustered by subject in parentheses

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

All regressions include constant.

Cognitive Measures: CRT and Numeracy Demographics : age, agesq, NonWhite, Upperclass, Female

interesting is in how those individual behavioral responses aggregate into overall productivity. Due to the fact that we generally find support for the discouragement effect and fail to find support for the encouragement effect it should be clear that the overall effect we find from inequality on productivity and efficiency is negative.

To help interpret the magnitude of the effect of inequality on efficiency we can provide a few calculations of the efficiency drop implied by the coefficients from the regressions. In the overall sample we find that production in the second half of the experiment declines by 1.9-2.6 encoded strings per round for the low wage subjects. Given that in the Common Wage - Low sessions average productivity over the second half of the experiment was 32.13 strings per round per person, the decrease implied by the coefficients represents a 5.9-8.1% drop in productivity. Were we to have found support for the encouragement effect among the high wage workers then this drop might have been offset for by increased productivity among high wage earners but we found that inequality had a neutral to negative impact on the output of high wage workers. So at a minimum the productivity of the high wage subjects does not counteract the productivity drop by low wage subjects.

In the results from the low CRT subjects we have a reasonable argument that these subjects were more likely to see the wage inequality as unfair and since what we really want to measure is the consequences on efficiency from inequality viewed as unfair then it is in the results from this sub-sample which might provide a more reliable benchmark for the efficiency consequences of inequality. The coefficients on the two interaction variables for the low wage / low CRT subjects range from -3.25 to -4.46. In the Common Wage - Low treatment, the subjects in the same CRT=0 classification produced on average 33.95 encodings per period and so those coefficients suggest a 9.6-13.1% drop in productivity. Again, the response from the high wage / low CRT does not offset this productivity decline. The high wage / low CRT subjects evidence no response to the inequality.

We do want to be clear that we are not proposing that the approximately 10% decline in productivity we find among low wage workers be interpreted as a reliable estimate of the magnitude of the discouragement effect that might exist among, for example, members of a low caste in India. The results from this experiment can certainly not provide an answer that specific or that accurate. What we claim is that since we were able to generate a discouragement effect of non-trivial size in this setting, then it provides strong support that there should exist a discouragement effect, for example, among members of a low caste in India and this is an important issue to consider in the design of policies aimed at enhancing economic development in countries with a substantial amount of discrimination based inequality as left alone the inequality will be persistent, possibly even worsening, and lead to negative consequences for economic growth and efficiency.

5 Conclusion

The question that motivated this study concerns whether or not individuals who are faced with unequal earnings opportunities will respond to any perceived unfairness by those receiving lesser opportunities decreasing work effort. We find evidence that in our overall sample that while there is no initial response to the inequality, after our subjects have experienced the inequality for several rounds those receiving a lower wage begin exerting less effort than their counterparts in a control group with no inequality. For the subset of our

sample who are measured to be impulsive decision makers through their score on the CRT, we find a much stronger discouragement effect which is consistent with the claim that those who are more likely to see the wage allocation as unfair exhibit a stronger discouragement effect.

These results represent clear and compelling evidence that in external situations when there are individuals faced with unequal opportunities which they deem unfair that they will also exhibit lower effort and the consequences of this behavioral response could be substantial. This discouragement effect could show up not only in on the job work effort but also in human capital investment and other activities which help a person advance into higher earnings groups. If the inequality in opportunity is not transient, then this discouragement effect could grow and combined with the forces described in Mookherjee and Ray (2003) and related studies can lead to persistent inequality and poverty. Thus this short term behavioral effect could have long term consequences for the initial generation of workers exhibiting the response as well as their descendents.

At the aggregate level, if there are populations of workers exhibiting lower work effort, investment in human capital and so on then there will certainly be negative impacts on economic growth and development. In our experiment we found output decreases of around 10% and while we will certainly not claim that this number is an accurate estimate of the magnitude of the discouragement effect in any other situation, it does suggest that the effects will be of non-trivial magnitude. The reason is that given the setup in this experiment, there is every reason to suspect that subjects should be immune to a discouragement effect building up over such a short time horizon from purely idiosyncratic wage assignment for performing a trivial task. Given the strength of the response that we find, it seems reasonable that in situations in which opportunity inequality is permanent and tied to an individual's race or caste then a discouragement effect should be even more likely to occur and could well be stronger than what we measure here. Given the long run consequences to the individuals and the aggregate economy, this issue should be of concern to those designing policies to foster economic growth in countries in which such inequality exists.

An immediate follow-up question to the demonstration of this effect, though, is how robust it is to the inclusion of other elements which exist in labor markets such as the opportunity for advancement. The first point in response to that question is that the main environment to which our study is meant to apply are to situations of ethnic, racial, gender or caste based discrimination in which such inequality is (at least perceived to be) permanent and outside the control of any single worker to effect. For these situations, the lack of opportunity for the disadvantaged workers to be promoted to the advantaged group is entirely appropriate. Still there will be other related situations in which there may be some possibility of moving from the disadvantaged to the advantaged group and/or one may consider policies which specifically allow that as a means of overcoming the discouragement effect. This leads to a question of how much of an advancement opportunity is necessary to overcome any discouragement effects? This is not an issue considered in this paper but it is an important question to be investigated in future work.

References

- Akerlof, G. A. and R. E. Kranton (2000). Economics and Identity. *The Quarterly Journal of Economics* 115(3), 715–753.
- Akerlof, G. A. and R. E. Kranton (2005). Identity and the Economics of Organizations. *Journal of Economic Perspectives* 19(1), 9–32.
- Akerlof, G. A. and J. L. Yellen (1990). The Fair Wage-Effort Hypothesis and Unemployment. *The Quarterly Journal of Economics* 105(2), 255–283.
- Banerjee, A. V. and E. Duflo (2003). Inequality and growth: What can the data say? *Journal of Economic Growth* 8, 269–299.
- Banerjee, A. V. and A. F. Newman (1993). Occupational choice and the process of development. *Journal of Political Economy* 101(2), 274–298.
- Benabou, R. (1996). Inequality and growth. *NBER Macroeconomics Annual* 11, 11–74.
- Benabou, R. (2000). Unequal societies: Income distribution and the social contract. *American Economic Review* 90(1), 96–129.
- Berg, J., J. Dickhaut, and K. McCabe (1995). Trust, Reciprocity, and Social History. *Games and Economic Behavior* 10, 122–142.
- Blau, F. D. and L. M. Kahn (2000, Fall). Gender differences in pay. *Journal of Economic Perspectives* 14(4), 75–99.
- Bolton, G. and A. Ockenfels (2000). ERC: A Theory of Equity, Reciprocity and Competition. *American Economic Review* 90, 166–193.
- Bolton, G. E., J. Brandts, and A. Ockenfels (2005). Fair procedures: Evidence from games involving lotteries. *Economic Journal* 115, 1054–1076.
- Brown, G. D., J. Gardner, A. J. Oswald, and J. Qian (2008). Does Wage Rank Affect Employees' Well-Being? *Industrial Relations* 47(3), 355–389.
- Burchett, R. and J. Willoughby (2004). Work productivity when knowledge of different reward systems varies: Report from an economic experiment. *Journal of Economic Psychology* 25(5), 591–600.
- Carr, M. (2008). Income Inequality and the Overworked American. Working Paper U Mass-Amherst.
- Charness, G. and P. Kuhn (2006). Does Pay Inequality Affect Worker Effort? Experimental Evidence. Working Paper UCSB.
- Durlauf, S. N. (1996). A Theory of Persistent Income Inequality. *Journal of Economic Growth* 1, 75–93.

- Engerman, S. L. and K. L. Sokoloff (1997). Factor endowments, institutions, and differential paths of growth among new world economies: A view from economic historians of the united states. In S. Haber (Ed.), *How Latin America fell behind: Essays on the economic histories of Brazil and Mexico, 1800-1914*. Stanford: Stanford University Press.
- Engerman, S. L. and K. L. Sokoloff (2002). Factor endowments, inequality, and paths of development among new world economies. *conomia: Journal of the Latin American and Caribbean Economic Association* 3(3), 41–88.
- Engerman, S. L. and K. L. Sokoloff (2005). Colonialism, inequality, and long-run paths of development. NBER Working Paper.
- Erkal, N., L. Gangadharan, and N. Nikiforakis (2008). Competition, luck and social preferences: A real-effort experiment. Working Paper, University of Melbourne.
- Fehr, E., G. Kirchsteiger, and A. Riedl (1993). Does Fairness Prevent Market Clearing? An Experimental Investigation. *Quarterly Journal of Economics* 108, 437–459.
- Fehr, E. and K. Schmidt (1999). A Theory of Fairness, Competition and Cooperation. *Quarterly Journal of Economics* 114(3), 817–868.
- Fischbacher, U. (2007). z-Tree: Zurich Toolbox For Readymade Economic Experiments. *Experimental Economics* 10(2), 171–178.
- Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives* 19(4), 25–42.
- Galor, O. and J. Zeira (1993). Income distribution and macroeconomics. *Review of Economic Studies* 60(1), 35–52.
- Goldin, C., L. F. Katz, and I. Kuziemko (2006, Fall). The homecoming of american college women: The reversal of the college gender gap. *Journal of Economic Perspectives* 20(4), 133–156.
- Heckman, J. J., T. M. Lyons, and P. E. Todd (2000). Understanding black-white wage differentials, 1960-1990. *The American Economic Review Papers and Proceedings* 90(2), 344–349.
- Hoff, K. and P. Pandey (2006). Discrimination, Social Identity and Durable Inequalities. *American Economic Review: Papers & Proceedings* 96(2), 206–211.
- Loury, G. C. (1981). Intergenerational transfers and the distribution of earnings. *Econometrica* 49(4), 843–867.
- Macpherson, D. A. and B. T. Hirsch (1995, July). Wages and gender composition: Why do women’s jobs pay less. *Journal of Labor Economics* 13(3), 426–471.
- Mookherjee, D. and D. Ray (2003). Persistent inequality. *Review of Economic Studies* 70, 369–393.

Oechssler, J., A. Roider, and P. Schmitz (2008). Cognitive abilities and behavioral biases. Working Paper, University of Heidelberg.

Ravallion, M. (2001). Growth, inequality and poverty: Looking beyond averages. *World Development* 29(11), 1803–1815.