Deferred Compensation Vs. Efficiency Wages: An Experimental Test of Effort Provision and Self-Selection*

David Macpherson†
Trinity University

Kislaya Prasad‡
University of Maryland

Timothy C. Salmon§
Southern Methodist University

December 2013

Abstract

We compare the ability of two common compensation structures, efficiency wages (EW) and deferred compensation (DC), at inducing effort from workers. We test predictions on effort provision and elicit preferences between the two wage structures. The theoretical predictions on effort are generally well supported, although we find over-provision of effort with EW. In consequence, although the theoretical prediction that DC is more cost-effective is supported, the difference is small. We also find a marked preference for EW that cannot be explained by risk aversion. The two effects combine to largely dissipate any advantage that DC may have in inducing effort.

JEL Codes: D86, C90
Key Words: Incentive contracts, principal-agent model, self-selection, experiments.

*The authors thank the National Science Foundation for funding this research (SES-0920821). We also thank seminar participants at the University of Pittsburgh and at the Social Science Research Center Berlin (WZB) for comments and suggestions. We are especially grateful to two anonymous referees and an Associate Editor for their suggestions for improvement of the paper.

†Department of Economics, One Trinity Place, Trinity University, San Antonio, TX 78212-7200. David.MacPherson@Trinity.edu Phone: (210) 999-8112 Fax: (210) 999-7255

‡Robert H. Smith School of Business, University of Maryland, College Park, MD 20742. kprasad@rhsmith.umd.edu. Phone: 301-405-9637 Fax: 301-405-9655

§Southern Methodist University, Department of Economics, 3300 Dyer Street, Suite 301 Umphrey Lee Center, Dallas, TX 75275-0496. tsalmon@smu.edu, Phone: 214-768-3547, Fax: 214-768-1821

1 Introduction

Employers use many different techniques to induce their workers to exert high effort and two of the most commonly studied approaches involve using either an Efficiency Wage (EW) or Deferred Compensation (DC). The EW approach to inducing effort involves paying a worker a wage above their outside option with the threat of firing for subpar performance (see Shapiro and Stiglitz (1984)). The DC approach involves paying a worker low wages at the beginning of their career, possibly even below her outside option, but with a promise of high future wages if the worker successfully retains her job (see Lazear (1981) and Akerlof and Katz (1989)). The relative merits of the two are often debated in the literature, but both are considered as effective ways of inducing effort.[1]

While either method can induce effort from employees, there are a number of additional questions related to their effects on both employers and employees that warrant investigating. An obvious concern by employers is which wage profile generates the most effort for the least cost. Within each family of contracts there is the further question of how employee behavior responds to contract parameters such as the level of efficiency wages, or the extent of back-loading of compensation in DC schemes. Additionally, both types of contracts induce patterns of inter-temporal behavior that are of interest. For instance, Gibbons and Murphy (1992) have emphasized the role of "career concerns" as a form of motivation. As employees get closer to retirement the future gains from current effort also become smaller, weakening incentives. Pay that increases with tenure has the potential to offset this effect, yielding higher effort levels at dates closer to retirement. Such weakening would certainly be of concern in EW schemes which use uniform incentives over time. DC schemes would be expected to be more effective at inducing effort late in the career, although EW could be more effective early.

While these questions regarding relative effectiveness have existed for a while and are often discussed, providing answers to them using naturally occurring field data has been challenging. There is an initial difficulty of identifying the wage structure an individual is working under (see Gibbons and Katz (1992); Murphy and Topel (1990) and Lazear (2000b)). In addition, it is difficult to measure effort in a way which

[1] See Ritter and Taylor (1994) and Prendergast (1999).

will allow for a comparison of the efficacy of wage profiles across the different firms and possibly industries one would likely have to compare to obtain a representative spread of wage profiles. There are many studies which use various proxies for effort such as dismissals, Peter and Chauvin (1991), or absenteeism, Johansson and Palme (1996), but these proxies will be of little value when comparing across occupations and industries. There are some examples such as Lazear (2000a) which can test the effects of different payoff schemes on self-selection and effort provision but even in those cases the test can only cover short run incentives and not the long run incentive schemes involving deferred compensation. What is needed to test the difference between EW and DC schemes is the ability to observe an individual exerting effort on a common measurable dimension under different wage profiles for the duration of a wage profile or, failing that, to be able to observe comparable workers exerting effort in a comparable manner under different wage profiles, still for the duration of that wage profile. Due to the issues in observing effort and the nature of wage profiles as well as all of the additional endogeneity issues inherent to these relationships, clean identification of the impact of wage profiles on effort provision is likely to be impossible with naturally occurring data. Ideally one would want to observe a worker working the duration of his career under multiple different wage profiles so that one can observe his behavior in these different situations but such a thing is not possible in non-experimental data.

To address these fundamental behavioral questions regarding how individuals react to different wage profiles, we use a set of carefully designed laboratory experiments which isolate behavioral responses to different wage profiles. In our experiments, the subjects make effort decisions in an environment that is modeled as a standard life-cycle labor supply problem. The subjects make decisions under multiple wage profiles which allows examination of how they respond to different profiles. They also make decisions among profiles which allows us to observe self-selection and whether there are any systematic preferences among wage profiles. While there are certainly a large number of contextual factors that exist in real labor markets which do not exist in the experiment, the results from this experiment can help isolate some of the fundamental behavioral responses to the wage profiles.[2]

[2]One might be concerned about the subject pool we use for these experiments as the subjects are predominantly undergraduate students. There is a commonly voiced concern about such subject

There have been prior papers using experiments to examine EW and DC wage structures though with substantially different focus than our intent. There is a very large literature examining efficiency wages as a form of gift-exchange as envisioned in Akerlof (1982). Fehr, Kirchsteiger, and Riedl (1993) popularized this line of investigation and subsequent papers such as Rigdon (2002) refer to this gift-exchange effect interchangeably with an efficiency wage. In this version of the EW model, employers offer high wages to employees and the employees respond with high effort based on either an increase in morale or due to some sort of reciprocal altruism. Our study is based on the alternative specification of an Efficiency Wage model due to Shapiro and Stiglitz (1984). In this case, a worker is paid a wage above their outside option and faces the possibility of the employer firing them for insufficient work effort forcing them to return back to their (presumably less lucrative) outside option. Thus it is the employee's fear of being fired and losing their well paying job which induces effort rather than and feeling of gratitude from receiving a gift from their employer. To our knowledge there is only one prior experimental study that examines the ability of a DC profile to increase effort, Huck, Seltzer, and Wallace (2011). Those authors were interested in a different aspect of the DC model dealing with the credibility problem of an employer promising future wages and then reneging. They find that credibility can be achieved through reputation concerns of the employers. Our study abstracts away from these credibility concerns to achieve tighter theoretical predictions on behavior to allow for a detailed test of the underlying model. None of these studies were aimed at comparing the effort inducement properties of different profiles nor did they examine preferences people might have between profiles. One paper relevant to our study is Loewenstein and Sicherman (1991) who, drawing upon survey results, argue that workers prefer increasing wage profiles. This is quite contrary to our findings, and a likely reason is a critical difference in the environment. In their paper workers are certain to retain their job until the end of the life-cycle, where in our model there is uncertainty about retaining a job (so that the high late returns in a DC scheme may not be realized).

pools that they may not be representative of the subject pool of interest. There is substantial (and far too often ignored) data showing that this concern is generally misplaced, see Frchette (2009). Further, our subjects are ones soon to join the workforce and are the population of interest for understanding self-selection into wage profiles.

The experiments for this study consist of three parts. The first is a standard risk assessment module in which subjects make a choice among five different lotteries such that their choice indicates their degree of risk aversion. The second phase involves subjects making choices in a 10 period life cycle labor supply model under 8 different wage profiles which include both EW and DC profiles of varying levels of generosity. In this environment subjects receive a wage profile, either EW or DC, which defines their potential earnings for 10 period and they are asked to choose an effort level for each period which costs the worker some of his earnings but increases the probability of retaining the EW or DC profile instead of being "fired" and receiving a less generous outside option wage. The subjects do this for multiple different wage profiles allowing us to test how well their behavior corresponds to the standard theoretical predictions along multiple dimensions. Finally we have the subjects indicate their preferences between various EW and DC profiles to try to understand what sort of aspects of these wage profiles affect the preferences of individuals for wage profiles.

There are three primary results one observes from these experiments. The first is that despite the fact that the theoretical predictions regarding effort are quite complicated, the choices by the subjects are able to match several important dimensions of these predictions. This indicates that the standard theory is quite useful in helping to understand the behavior in this environment. The second key result though is that there is one key divergence between the predictions and the observations which is that subjects typically provide more effort than predicted. Finally we observe that in the self-selection phase of the experiment, our subjects exhibit quite strong preferences overall for the EW profiles instead of the DC profiles and that this preference is too strong to be explained by risk aversion.

Section 2 presents the theoretical model and hypotheses that underlie this study. Section 3 describes the design of the experiment. Section 4 presents an analysis of the results and Section 5 provides a concluding discussion.

2 Theory and Hypotheses

We determine how individuals respond to incentives over time when a compensation profile takes the form of either an Efficiency Wage (EW) or Deferred Compensation

(DC) profile. In the first case, a worker is paid a constant wage above his or her outside option. In DC, the wage profile has a positive slope, involving low initial wages but higher wages later. While many discussions of DC wage profiles involve schemes that are defined to make payments close to the outside option (with additional compensation for incremental effort costs), this is not a defining characteristic of a DC wage profile. A contract can both pay more than the outside option *and* involve back-loading. Since we wish to compare DC and EW contracts on fair terms, our DC contracts will involve both of these elements so that we can make potential payments to workers common between both types of profiles. In our classification structure, the key difference between a wage profile classified as an EW profile and one classified as a DC profile is the slope of the profile over time. Flat profiles will be referred to as EW profiles while upward sloping profiles will be labeled DC.

In each round of our experiment there are ten periods. In any given period, with full information of the future wage profile, the subject must choose an effort level $e_t \in [0, 10]$. Subjects incur a cost for using a higher effort level according to:

$$C(e_t) = \frac{1}{2} e_t^{2.25}. \tag{1}$$

The advantage in using higher effort is that the probability of retaining the job for another period is higher. In the experiment, subjects retain their job in period t according to the following probability distribution:

$$P(e_t) = \frac{2e_t}{2e_t + 1}. \tag{2}$$

The probability of losing ones job in this case is intended to include both the probability of getting fired and the probability of the firm closing which is why the probability of job retention is never 1. This is important in evaluating DC profiles and the likelihood of a firm going out of business is an important part of evaluating whether an individual will receive those late rewards from early effort. The base wage profiles used in the experiments are depicted in Figure 1. The profiles have been chosen so that *maximum potential* earnings (i.e. earnings in the event that the worker is not fired) are the same for each profile. Note that both *expected* earnings and *expected* effort will vary across wage profiles. Consequently, we do not claim that the contracts

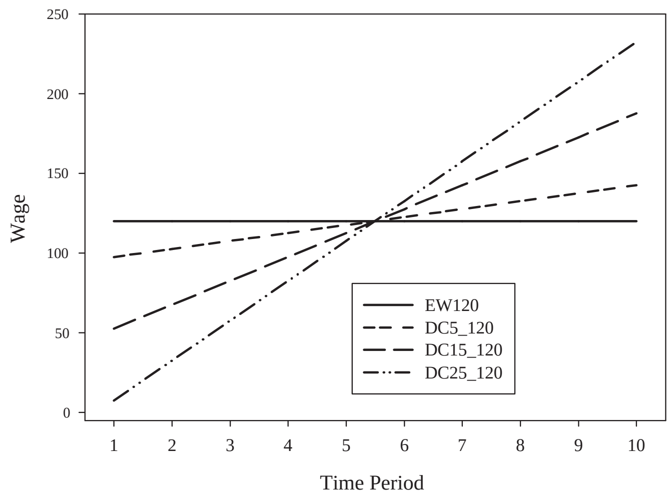

Figure 1: Wage Profiles

are equivalent on all dimensions. It is impossible to devise contracts of differing slope that are equivalent along all such dimensions and we chose to equalize potential rather than expected earnings. This does mean that we will not be able to do many cross profile comparisons in our analysis. Instead our strategy is to focus our attention on making comparisons of outcomes with the predictions of the theory. In the Figure, the line with zero slope depicts the base EW schedule – in each period the worker stands to receive a payment of 120 ECU, unless he gets fired. If he gets fired he gets the outside wage w_0 (we have $w_0 = 50$ ECU throughout the experiment). The remaining wage profiles in the figure involve DC, with slopes from the set $\{5, 15, 25\}$ and intercepts chosen to yield the same maximum potential earnings as in the EW case (*viz.* 1200 ECU).

We solve for equilibrium effort levels under different assumptions about risk preferences. We first consider the risk-neutral case, where workers are assumed to care only about expected earnings. We solve by backward induction, starting from the last period ($t = T \equiv 10$). If they reach this stage in the employed state, then the choice problem is:

$$\max_e V(e|T) \equiv P(e)w(T) + (1 - P(e))w_0 - C(e), \tag{3}$$

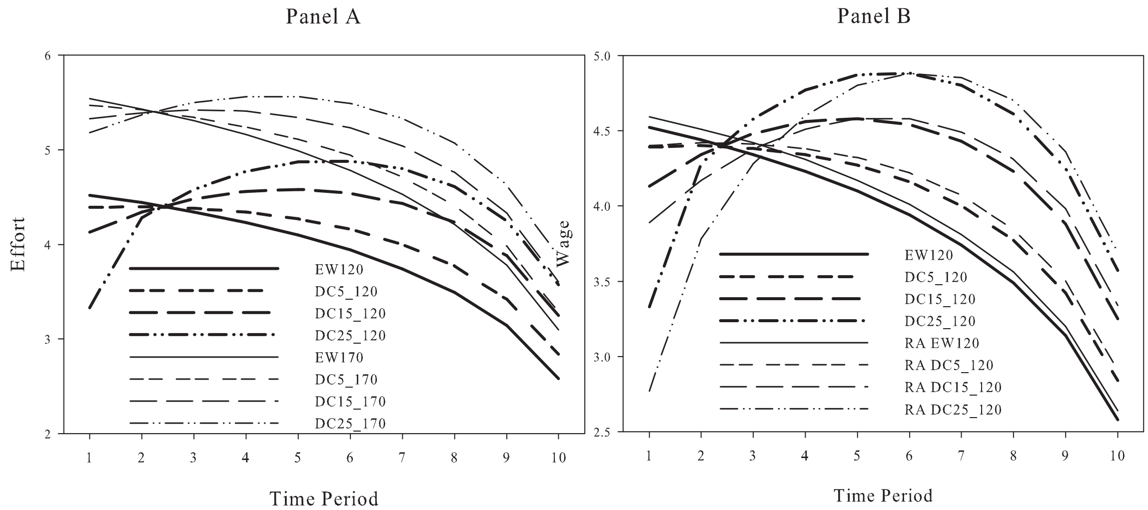

Figure 2: Predicted effort paths under different contract and risk specifications.

which is easily solved for the optimal effort e_T^*. Now we can solve a similar problem for $t = T-1$. If the subject is not fired he gets $w(T-1)$ and goes on to stage T in the employed state. If he is fired, he gets w_0 from $t = T-1$ onwards. In this case, his choice is to

$$\max_e V(e|T-1) \equiv P(e)[w(T-1) + V(e_T^*|T)] + (1-P(e))[w_0 + w_0] - C(e). \quad (4)$$

In this manner, we can compute the optimal effort for each stage recursively. Note that the subject can guarantee himself the outside option at zero cost by choosing $e = 0$. Hence, we do not need to separately consider the participation constraint. Equilibrium effort levels are depicted together in Panel A of Figure 2. Our first prediction is that:

Prediction 1 *On average, risk neutral subjects choose the effort levels depicted in Figure 2. In particular, efforts decline as we approach the end date. This tendency is offset by DC wage profiles with the profiles with the steepest slopes yielding the highest late stage effort.*

We have both qualitative predictions and quantitative predictions about exact levels of effort. Qualitatively, note that in early periods (i.e., periods 1 and 2) efficiency wages yield the highest effort, and effort declines for contracts with higher slope. This ordering is reversed for later periods, when higher slope contracts yield higher effort. This is because the first few periods of the DC profile are not very lucrative and the probability of making it past those periods leads to a discounting of the future payoffs. Thus retaining the DC profile does not seem as valuable as retaining the EW profile. For those that do survive in the DC profile after the first few rounds, the DC earnings surpass the EW earnings making the DC profile now more valuable and leading the worker to want to expend more effort to retain the DC profile than the EW profile. Overall, we find that contracts with higher slopes elicit higher effort levels. We verify this conclusion by computing total effort across the ten periods, as well as the expected total effort. The first approach measures effort conditional on retaining the job in each period. The total effort in the efficiency wage case is 38.52. The deferred compensation effort levels, in order of increasing slope, are $39.97, 42.42$, and 43.94 respectively. A better measure accounts for the likelihood that a worker will be fired and then computes the total effort expected over ten periods. In this case, the expected total efforts ordered by slope are $[25.66, 26.41, 27.69, 27.74]$. This second measure is the more useful and corresponds to the expected sample average of total effort expended by subjects. By either measure, we have higher efforts associated with steeper wage profiles. While the overall totals turn out to be similar to each other, it is important to note that the theory provides us multiple different predictions for tests of the theory other than simply total effort observed. For example, we will see below that expected earnings are very different for these contracts and Figure 2 already demonstrates that the time paths are quite different across wage contracts. The combination of these and other elements of the theoretical predictions will allow us clear tests of how well the model describes the behavior.

We next compute the expected total future earnings of a worker in period t assuming equilibrium effort in future periods. This is depicted in Panel A of Figure 3. Note that at the start of period 1 the expected earnings (at equilibrium effort) in the EW scheme are higher than for any of the DC schemes. Once subjects reach period 3, however, the DC schemes have higher expected earnings from that point forward. The characteristic shapes of the effort and earnings profiles result from the EW profile being initially more attractive. All profiles promise the same total amount conditional on never being fired (1200 ECU). For the DC profile, however, the high wages occur

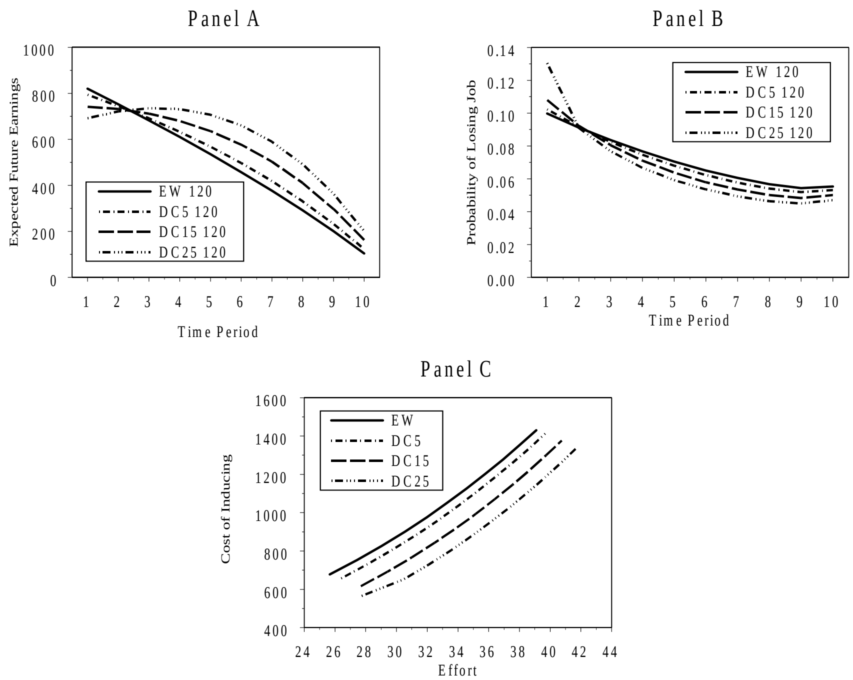

Figure 3: Theoretically predicted expected earnings, probability of job retention and cost of inducing various levels of effort all by contract form.

in later periods. Furthermore, even if equilibrium effort is provided, some likelihood exists of being fired early in the cycle. Consequently, at equilibrium effort levels, the EW scheme offers highest expected earnings.

Prediction 2 *Subjects choose higher total effort levels as slope gets steeper (EW having the lowest effort and DC25 the highest). In period 1, EW contracts offer the highest expected earnings, and expected earnings decline as the slope of the DC contract increases.*

Workers spend the least effort, and have greater expected earnings under EW, and so they would prefer this wage contract. Then, in successive order they prefer DC5_120, DC15_120, and DC25_120. The relative initial attractiveness of EW also explains the higher initial effort elicited by this scheme. Alternatively stated, a worker would least want to lose an EW job early in his tenure, and so is willing to work harder to retain this job at this stage. In later periods, the future earnings of the DC profiles dominate leading to the worker exerting more effort under those profiles as he would now least like to lose the DC profiles.

We now consider the effect on effort of increasing the total promised payments (from 1200 ECU to 1700 ECU). This involves changing the efficiency wage to 170 ECU, and shifting the intercepts of all DC wage profiles by an appropriate amount. The resulting effort is included in Panel A of Figure 2. We observe that

Prediction 3 *Subjects choose higher effort levels when the intercept of the wage profile are higher.*

A question that cannot be directly addressed by the predictions above relates to the employer's preference between EW and DC. For instance, consider an employer who can choose between an EW contract (varying the payment amount) and a DC contract with fixed slope (but varying the intercept). We know that any desired level of expected effort can be induced using either a contract from the EW family, or one from the DC family by adjusting the intercept accordingly. The key question is, if an employer wishes to induce some specific expected effort level of x, what sort of profile achieves that for the lowest expected cost? To answer this question, we compute for a range of potential average effort levels, the expected cost of inducing that effort level for an EW profile and then for DC profiles using the three slopes we use in the examples above ($\{5, 15, 25\}$) with the intercepts of all adjusted as necessary to achieve that effort level. The results are plotted in Panel C of Figure 3 which therefore shows the expected cost of inducing any particular level of effort using each of these classes of wage profiles. The line for the EW profile is clearly above all of the DC profiles indicating that it is the most expensive and the DC profiles are ordered such that the profile with the steepest slope is the least expensive. This leads to our next prediction.

Prediction 4 *DC profiles with a higher slope can be used to obtain each expected total effort level with lower expected payments. EW has highest expected payments.*

This prediction settles the question, raised at the beginning of this section, concerning whether EW or DC is more cost-effective for motivating workers. By increasing the slope of the contract we increase effort while simultaneously reducing expected earnings/payments. Thus, back-loading compensation is more effective at inducing effort than increasing the gap between expected earnings and the outside option.[3]

This relative ordering continues to hold for average cost – in other words, the cost per unit of effort is highest for EW and lowest for the DC profile with the steepest slope. The expected payments and efforts corresponding to the baseline contracts are presented in the top half of Table 3. Note that expected payments are different from the expected earnings depicted in Panel A of Figure 3 because the latter includes earnings from the outside option.

The computations account for workers being fired in equilibrium. The expected employment duration is calculated after computing the probability of getting fired in period t (and not before). So, for instance, if someone is fired in period 1, his duration is 0. The probabilities of getting fired in any period assuming equilibrium effort is provided are given in Panel B of Figure 3. Once again, the graphs are ordered (the highest at $t = 10$ is EW, the lowest is DC25_120, with DC5_120 and DC15_120 intervening). The predicted job durations are also included in table 3. Due to the fact that the contracts were constructed to all generate the same potential payments, these numbers have to be close together. Again though, we do find more variability in other predictions which allow us clean tests of how the wage profiles affect behavior.

Prediction 5 *Employment durations are longer for steeper DC profiles.*

Next, we consider the effect of risk-aversion on effort and on preference between wage profiles. We assume CRRA utility functions, $u(x) = x^\alpha$ (the coefficient of relative risk-aversion is $1 - \alpha$). Equilibrium efforts can be computed after minor modifications of the procedure used above. We solve by backward induction, starting from the last period ($t = T \equiv 10$). If they reach this stage in the employed state, then the choice problem is:

$$\max_e V(e|T) \equiv P(e)u(w(T) - C(e)) + (1 - P(e))u(w_0 - C(e)), \tag{5}$$

which is easily solved for the optimal effort e^*_T. Now we can solve a similar problem for $t = T - 1$. If the subject is not fired he gets $w(T - 1)$ and goes on to stage T in

[3] In principle we can, by increasing the slope of the DC contract to a sufficiently high value, make the expected earnings exactly equal to the outside option (plus effort costs). We did not do so because this entails negative payments in a large number of states, which we wanted to avoid. Negative payments are especially problematic for calculations involving CRRA utility functions.

the employed state. If he is fired, he gets w_0 from $t = T - 1$ onwards. In this case, his choice is to

$$\max_e V(e|T-1) \equiv P(e)[u(w(T-1)-C(e))+V(e_T^*|T)]+(1-P(e))[u(w_0-C(e))+u(w_0)]. \tag{6}$$

In this manner, we can compute the optimal effort for each stage recursively.

To examine the effect of risk-aversion on efforts we provide a comparison of effort choices for the base contracts when a subject is risk-averse. The results, for the case where the coefficient of relative risk-aversion has a value of 0.25, are depicted in Panel B of Figure 2.

Prediction 6 *At dates sufficiently close to the the end (T), risk-averse subjects with any contract choose higher effort levels than risk-neutral agents. Subjects with EW contracts always choose higher efforts, but subjects with DC contracts choose early date efforts that are equal or less than efforts of risk-neutral agents. At some point, they switch to choosing higher efforts, with the switch over point increasing with the slope of the DC contract.*

The total effort of risk-averse agents (conditional on never being fired) is higher for EW and DC5_120, but lower for DC15_120 and DC25_120. It should be noted that the magnitudes of the effort differences are quite small at later time periods.

A second set of questions addressed in our experiments relates to subjects' preference between contract types. Offered a choice between an EW contract and a DC contract, will all subjects make the same choice? If not, do these choices vary systematically with subject characteristics? To elicit preferences between contracts we had subjects express preferences between pairs of contracts (see Table 2).

The contracts we use for this task are different from the ones we use in the rest of the experiment. The reason is that we needed to vary the attractiveness of EW and DC contracts along an easy to understand property and so we varied the intercept of these profiles as that more clearly changes the attractiveness of a profile rather than the slope. To create the first of two sets of contracts we use, we fixed an EW contract and then offered a choice between this and one from a set of increasingly attractive DC contracts (see Table 2). The DC contracts all have a common slope of 15, but the intercept was allowed to vary. To create a second set of contracts, we fixed

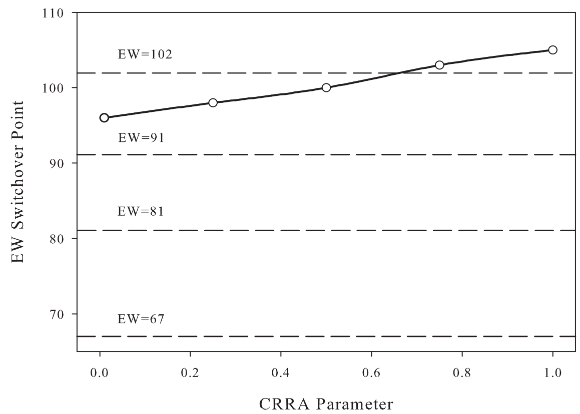

Figure 4: Predicted switchover EW profiles for risk averse decision makers in the DC fixed treatment.

a DC contract and asked for a comparison with one from a set of increasingly *less* attractive EW contracts. Both of these sets are constructed so that at the initial pair, the expected earnings are comparable though slightly higher for the DC contract. So initially we predict risk-neutral individuals will prefer the DC contract while sufficiently risk-averse individuals will prefer the EW contract. Then, as DC contracts are made more attractive, or as EW contracts are made less attractive, at some point the subject should be increasingly more likely to choose the DC. While it is the case that preferences among wage profiles like this can be affected by risk preferences, we constructed our contract sets to minimize the possibility that we should observe an impact of risk preferences on the observed decisions. Indeed, successive choice of EW indicates a willingness to forego significant amounts of money to avoid the DC contract. Figure 4 demonstrates this by showing lines for the four EW contract options, {102,91,81,67}, offered as alternatives to the fixed DC wage as well as points indicating the EW wages required to make individuals of varying degrees of risk aversion indifferent between that EW profile and the fixed DC profile. What can be clearly seen is that these indifference points are clustered tightly around the highest EW of 102 with none falling below the second highest alternative of 91. This means that at an EW of 102, individuals with higher degrees of risk aversion should be more willing to choose that EW compared to the DC alternative but no individuals regardless of risk aversion would prefer the other lower paying EW profiles over the DC profile. A similar figure could be shown for the other set of contracts. This analysis leads to the following prediction:

Prediction 7 *In the preference elicitation task, subjects should generally prefer the DC contracts offered though risk averse individuals might prefer the EW profile at the most generous EW option or least generous DC option.*

3 Experimental Design

The experiments were designed for two main purposes. First, to examine how people behave in the life-cycle labor supply model as described above. Second, to investigate whether and how people form preferences over different types of wage profiles. As controls for individual characteristics we also obtained data on various demographic characteristics of our subjects. The demographic variables we use were obtained through two methods. Standard demographic variables such as race/ethnicity and gender were self-reported through a questionnaire given at the end of the experiment. We also obtained information on the educational background of the subjects including their SAT/ACT scores and current GPA directly from the University Registrar's office which means this more detailed data is not self-reported. To obtain this information, subjects were asked to sign a secondary consent form which allows us access to the information and 266 out of our total sample of 290 or 92% of our sample did so.

The rest of the experiment is divided into two phases. In the first phase, subjects engaged in a simple risk preference assessment procedure and then in the second phase they went through the module on effort supply and self-selection. We will explain each in turn.

3.1 Risk Preference Assessment Phase

The risk preference assessment was done using a simple procedure similar to that used in Eckel and Grossman (2006), Binswanger (1980) and Binswanger (1981). The subjects were allowed to make a choice between receiving the returns from five different possible lotteries labeled A-E. Each lottery had two possible outcomes which could occur with equal probability. The computer interface mimicked a coin flip to convey that equal probability. For lottery A, the prize values for both a heads and a tails outcome were the same yielding a safe outcome but with relatively low expected value. Each additional lottery was constructed by subtracting off an amount of money from the heads outcome in lottery A to be the heads outcome in the new lottery while adding 2.5 times that amount to the tails of the A lottery for the new tails outcome. This structure delivers a stable pattern of the expected values of the lotteries increasing as the variance increases.[4]

Table 1 shows the lotteries used. It also shows the expected value of the lotteries and the ranges of risk aversion parameters using a standard CRRA utility function, $u(x) = x^{\alpha}$, that would be consistent with someone choosing that lottery as their most preferred. We report the information on the risk aversion parameters to help the reader understand the structure of the lotteries. For our purposes in the data analysis we will neither need nor use the risk aversion parameters as we require only a non-parametric measure for relative risk aversion among our population. This is why we chose to use this simple yet perhaps imprecise method for eliciting preferences over risk. Further, while risk aversion technically can matter in the decision environment of this experiment, given the functional forms used, large differences in risk aversion have relatively small effects on both average effort provision and on the self-selection task. This is actually an advantage for the main purpose of the study since our main interest in the analysis will be whether behavioral differences are correlated with the demographic characteristics of our subjects. It is therefore useful to have constructed an environment in which risk aversion is expected to matter little as that makes it less likely that any observed differences in behavior attributable to demographic differences are really just reflections of differences in risk preferences among demographic groups. We still want to control for risk aversion though in case it can explain some of the choice behavior, but as demonstrated it should not be expected to explain much about the behavior in this study by construction.

[4]The payoffs were constructed this way from base dollar values so lottery A involved $1 for sure. Lottery B involved getting $0.70 or $1.75 and so on. The amounts shown in table 1 are based on converting those dollar amounts into ECUs using the exchange rate for the main portion of the experiment. The minor deviations from the formula given are due to rounding differences in that conversion process.

Table 1: Lotteries used for risk aversion measure.

Lottery	Heads	Tails	Expected Value	CRRA Parameter Range
A	435	435	435.0	$\alpha \approx 0$
B	304	761	532.5	$0.24 \geq \alpha > 0$
C	217	978	597.5	$0.50 \geq \alpha > 0.24$
D	130	1196	663.0	$0.67 \geq \alpha > 0.50$
E	43	1413	728.3	$\alpha > 0.67$

3.2 Effort Supply and Self-Selection Phase

In the Effort Supply phase of the experiment, subjects made choices in multiple instances of the labor supply model previously described. In each 10-period round, subjects were endowed with either an EW or DC wage profile, see Table 2, which described the subject's potential earnings over the 10 periods. In each period, the subject would be asked to choose an effort level. Higher effort levels led to an increased chance of retaining the EW or DC profile though at higher cost. The functional forms for the cost of effort and probability of retaining the profile are as given in the previous section in equations 1 and 2 and are shown graphically in Figure 1. If in a period a subject fails to retain the EW or DC profile, he is "fired" and therefore receives his outside option wage for that period and all future periods. Subjects were shown their wage profile and these functions for probability of retaining the profile along with their effort cost using a graphical interface to make the consequences of their decisions easier to understand. A screenshot of the display is available in the Appendix which also contains the instructions.

In order to better understand how individuals respond to different elements of a wage profile we had subjects "work" under 8 different profiles as described in Table 2 under the section indicated as Rounds 1-8. There are two sets of four profiles used. In each set there is one EW profile and three DC profiles in which the DC profiles differ by both their intercept and slope. The second set of profiles is essentially the first set again but with the intercepts shifted upwards by 50. An important aspect of the design of this experiment is that it is impossible to make the wage profiles perfectly comparable. As described above the wage profiles were constructed so that potential earnings for all four profiles in a set are the same. This does not mean

Table 2: Wage profiles used in experiment.

	Rounds 1-8			
	Round 1 Wage		**Slope**
	120	**170**	
Outside Option	50	50	0
Efficiency Wage	120	170	0
DC5	97.5	147.5	5
DC15	52.5	102.5	15
DC25	7.5	57.5	25

	Self-Selection Phase			
	DC Fixed		**EW Fixed**	
	DC	**EW**	**DC**	**EW**
Pair 1	52.5/15	102	69.5/15	120
Pair 2	52.5/15	91	78.5/15	120
Pair 3	52.5/15	81	86.5/15	120
Pair 4	52.5/15	67	98.5/15	120

that expected earnings along the equilibrium effort paths are the same nor that the profiles are equally risky. The latter would have been impossible to guarantee and while we could have made the profiles such that equilibrium earnings would have been the same, such an equivalency would have been eliminated the first time a subject made an effort choice different from the theory. As will be seen in the results section, this means that we will not engage in much cross profile tests. Our tests will largely involve testing how behavior for each profile compares to the theoretical predictions. On the other hand it is important to know how much of any differences in behavior across profiles might be driven by the differences in generosity across them. This is why we included the second set of profiles as this allows us a clear test for how behavior changes by making a specific profile more generous.[5] In order to break up any ordering effects, half of the subjects saw the wage profiles in one order[6] and the other half saw them in the reverse order.

After subjects have completed all 8 rounds with the different wage profiles, they engaged in one final round with a self-selection component. Subjects were given a set of choices to make regarding profiles they might possibly use for the final round according to two treatments. In one treatment, the subjects saw four pairs of wage profiles and in each pair the DC profile was the same but the EW profile became increasingly less generous. We asked the subjects to make a choice for each pair. This treatment is referred to as the DC Fixed treatment. In the second treatment,

[5]As a side note we also point out that there were no humans in the role of employers in this experiment. It is certainly possible that the subjects could have responded differently to different profiles had there been humans in the role of meployers offering the profiles or even accruing benefits from the subjects' effort decisions. We decided to forego this element to acheive a cleaner design but we note the possibility that one might achieve different results had this element been different.

[6]$\{EW120,\ DC5\_120,\ DC15\_120,\ DC25\_120,\ EW170,\ DC5\_170,\ DC15\_170,\ DC25\_170\}$

EW Fixed, there is a fixed EW profile and the DC profile becomes increasing more generous over the four pairs. Subjects are asked to make a choice for each pair. The wage profile they will actually use in the last round will be chosen by selecting randomly from among the choices made.

As we will explore further in the results section, there are significant advantages and disadvantages in having the self-selection task always come after the first 8 rounds. The advantage is that by the time the subjects get to the self-selection task they should understand the relative payoffs achievable by different profiles. Without this experience it would not be clear that they would understand the relatively fine differences between profiles well enough to make sensible decisions on this task. Of course the downside is that the preferences we observe on the self-selection task could well be driven by prior experience rather than any underlying preferences. Given the complex nature of the task, there seems little reason to have expected sensible choices on the selection task if done first so the alternative design seemed unpromising. In the results section we will address these issues further.

The experiments for this study were conducted in the xs/fs Laboratory at Florida State University using a computer interface programmed in z-Tree, Fischbacher (2007), and the subjects were recruited using a system based on ORSEE, Greiner (2004). There were a total of 13 sessions conducted with a total of 290 subjects. Subjects received a participation fee of \$10 and the ECUs converted into US\$ at the rate of \$1=435ECUs. On average subjects received total compensation of \$28.11 and sessions lasted around 1.5-2 hours.

4 Results

We investigate two sets of questions in this section regarding the experimental data. We first investigate the degree to which the theoretical predictions can explain the effort choices and also try to determine whether there are systematic deviations from the theory overall and if any track with observed demographic differences. We then will proceed to determining any patterns of self-selection that emerge from the final element of the experiment.

4.1 Determinants of Effort

The first question to ask of the data is how closely did subjects conform to the predictions of the standard theoretical model. We can begin investigating this issue by comparing the observed data to some theoretical benchmarks using high level summary statistics. This comparison is presented in Table 3. For each wage profile we provide the predicted effort and earnings levels as well as the expected job duration and we then provide the observed values. The table includes the results of Wilcoxon tests comparing the observations to the theoretically predicted values. There are a few general points we can already observe from these statistics. First, it is clear that the observed effort levels are above the predicted levels except for the DC profiles with the steepest slopes. The excess effort provision leads to lower earnings and longer job durations. We don't intend these tests as our final tests of these issues because the theoretical predictions are much more detailed than simply predicting averages. The time path of labor provision and other aspects are also important. We will defer formal results regarding the test of the theory until later when we provide more detailed tests of the degree to which the data matches the theory.

Before presenting the econometric tests though there are two other looks at the data that are useful to see to obtain an intuitive understanding of the data. Figure 5 provides a histogram of the effort choices made for each wage profile. This figure shows that subjects rarely made effort choices at the boundaries of high or low effort but rather largely made interior choices and there was variation to those choices. There is variation that is predicted to occur temporally as predicted effort should change from period to period. Figures 6 and 7 provide a graphical comparison of the average effort chosen over time with the theoretical predictions. A careful inspection of these panels reveals that there is quite strong conformance between the theoretically predicted effort paths and the average of those observed. The main deviation appears to be an over provision of effort from the risk neutral prediction as already demonstrated in Table 3 but the shapes of the effort paths seem surprisingly consistent with the predicted paths.

To provide a deeper understanding of the relationship between actual and predicted effort as well as whether there are systematic deviations by demographic groups, we present table 4 which contains the results from multiple regressions aimed

Table 3: Predictions versus average observations on effort, earnings and employment duration.

		Contract Set 120			Contract Set 170		
	Contract	**Effort**	**Earnings**	**Duration**	**Effort**	**Earnings**	**Duration**
Predicted	EW	25.66	678.3	5.65	33.29	1050.7	6.18
	DC5	26.41	657.2	5.70	33.94	1031.3	6.21
	DC15	27.69	616.0	5.77	35.16	993.6	6.27
	DC25	27.74	565.4	5.80	36.22	956.6	6.31
Observed	EW	29.93***	571.68***	6.72***	36.72*	827.76***	6.89**
		(1.12)	(21.93)	(0.21)	(1.40)	(32.76)	(0.23)
	DC5	31.96***	529.38***	6.72***	37.15*	811.39***	6.94***
		(1.27)	(22.50)	(0.22)	(1.43)	(33.36)	(0.23)
	DC15	30.12	446.33***	6.41***	37.49	784.79***	7.11***
		(1.29)	(22.92)	(0.22)	(1.36)	(32.51)	(0.21)
	DC25	27.31	370.49***	5.96**	35.50	741.82***	7.00***
		(1.35)	(24.50)	(0.24)	(1.41)	(35.37)	(0.23)

Standard errors in parentheses.

*'s indicate p-value of Wilcoxon test compared to theoretical prediction. *** p<0.01, ** p<0.05, * p<0.1

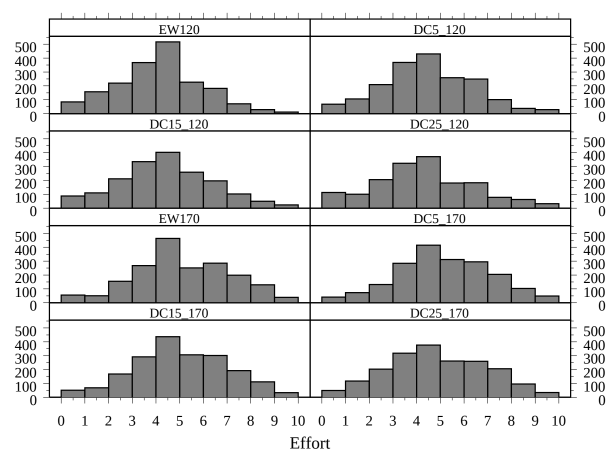

Figure 5: Histograms of effort choice for each wage contract.

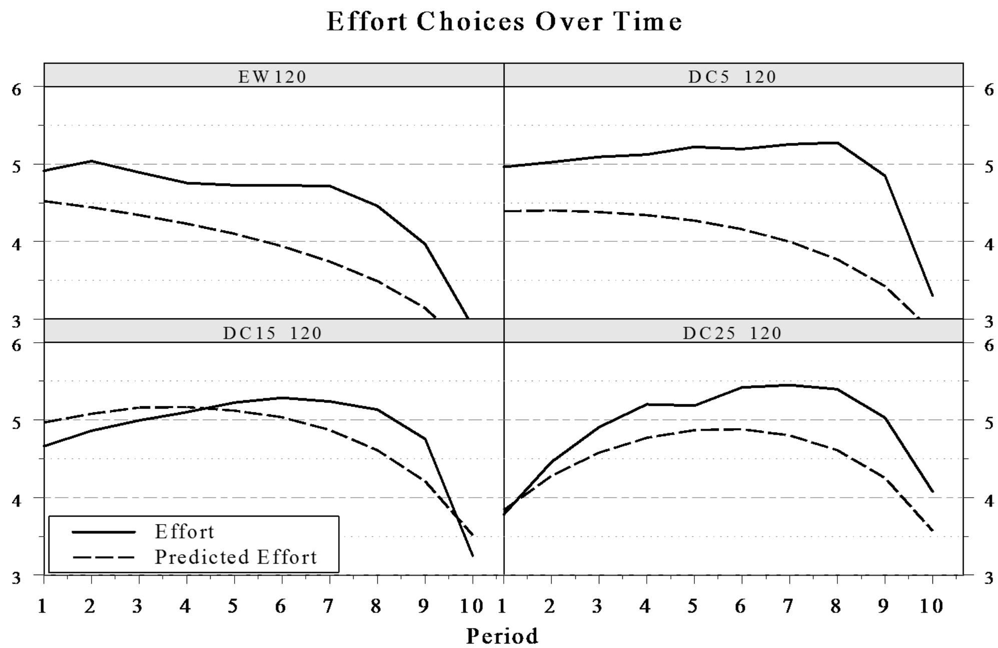

Figure 6: Comparison of theoretically predicted and observed effort over time by wage profile for the less generous wage set.

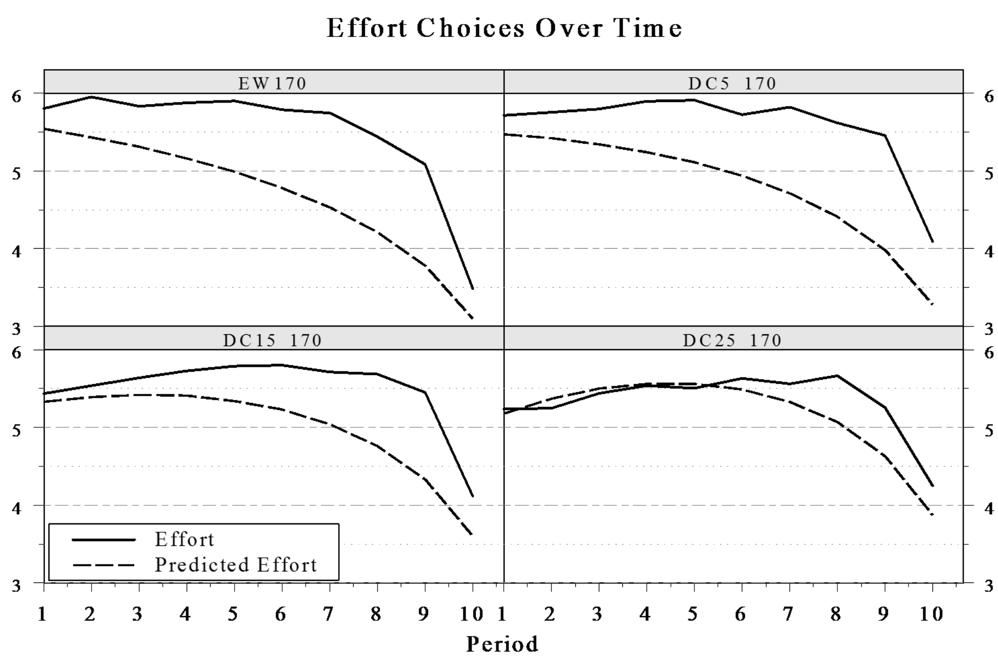

Figure 7: Comparison of theoretically predicted and observed effort over time by wage profile for the more generous wage set.

at addressing these issues. All of the regressions use the effort choice by an individual in a particular period of the decision task as the dependent variable with the key independent variable being the predicted level of effort for that period in that wage profile. All of the regressions use a random effects panel specification[7] with robust standard errors clustered on the individual level. Since there was no interaction between subjects in these experiments, each subject can be treated independently from all others. Due to the dynamic structure involving multiple rounds of the decision task and multiple periods inside of a round, we need to understand how effort might depend on both temporal dimensions. As a base method of controlling for these effects we include the ln of these values but we will also examine these temporal issues more directly with additional regressions later. Another important set of control variables involve dummy variables for each of the different wage profiles in use as they allow us to understand any differences in average effort provision by individual profiles. In some regressions, we also include variables corresponding to the demographic information and risk preferences we collected about the individuals.

Result 1 *The coefficient on predicted effort is close to 1 and in some specifications it is not significantly different from 1 indicating that observed effort tracks predicted effort. There is, however, systematic over provision of effort, most noticeably in the EW and DC5 cases. The systematic deviations can partly be explained by prior outcomes, time, and demographic variables.*

This result is of course foreshadowed by Figures 6 and 7, but the statistical analysis confirms this finding. In each of the regressions in table 4, the coefficient on the variable for Predicted Effort is near 1. In specifications (1), (4) and (5) the coefficients are significantly different from 1 but they are still quite near 1 while the coefficients on effort in specification (2) and (3) are not significantly different from 1 (p-values 0.005 and 0.006 respectively). This suggests that subjects adjusted their effort choices largely in the pattern suggest by the theoretical predictions. Were the theory precisely accurate we would see that coefficient be 1 while all others would be 0 and the latter prediction clearly does not hold as there are many other coefficients which

[7]Except for regressions (2) and (3) in table 4 which are fixed effects regressions for the purpose of allowing better comparability of the results in those two regressions as explained later.

are significantly different from 0. So while the subjects were adhering to an effort provision path similar to that predicted by the theory, there are systematic deviations.

The first systematic deviation to investigate is how effort might be changing across rounds. In regressions (1), (4) and (5) we find a positive and highly significant coefficient on the variable ln(*Round*) which suggests that effort is steadily increasing across rounds over and above what is predicted to occur due to the differences in wage profiles. Since profiles are changing across rounds, this is not the most direct test of whether effort is simply increasing due to time spent in the experiment. Given the structure of the experiment design, the best comparison for this purpose involves comparing effort levels in Round 1 with those of Round 8. Half of the subjects saw the wage profile EW120 in each of those rounds while the other half saw DC25_170. Since both profiles are in use in both rounds we can compare effort levels in both rounds to try to observe any differences due to dynamics or ordering. Columns (2) and (3) of table 4 contain regressions using only Round 1 and Round 8 data respectively and we observe that the coefficient on Predicted Effort in the Round 8 data is slightly higher though neither is significantly different from 1. To test if the two coefficients themselves are different we conducted a fully interacted version of a fixed effects panel regression nesting both of the regressions in columns (2) and (3) in the regression (i.e., all variables including the fixed effects dummies were interacted with a dummy variable for Round 8). Whether there was any difference between these rounds can be determined off of the coefficient on the interaction between Predicted Effort and Round 8 That coefficient comes out to be .018 with a $p-$value of 0.900. Thus there is in fact no significant difference in how subjects provided effort over the course of the experiment that appears to be primarily attributable to temporal effects. On the other hand we also included a variable for the job duration in the previous round and find that it is negative and significant indicating that subjects do respond to some degree to prior outcomes.

The dummy variable coefficients for the different wage profiles reveal a systematic trend in how subjects respond to EW and DC profiles. The base constant, which refers to EW120, is of indeterminate sign and sporadic significance and the coefficient on the EW170 dummy is generally insignificant. While the coefficients on the DC5_120 and DC5_170 dummies are insignificant, the coefficients for the DC profiles with

	(1)	(2)	(3)	(4)	(5)
Predicted Effort	0.866***	1.013***	1.031***	0.877***	0.866***
	(0.044)	(0.111)	(0.096)	(0.044)	(0.049)
RA				-0.081	-0.050
				(0.056)	(0.057)
RA * Period>4				0.048	0.059
				(0.035)	(0.037)
Period>4				0.010	-0.023
				(0.106)	(0.111)
White				0.486***	0.348**
				(0.162)	(0.175)
Asian				0.077	0.613
				(0.383)	(0.585)
Male				0.393**	0.278
				(0.161)	(0.171)
RA Phase Earn				<0.001	<0.001
				(<0.001)	(<0.001)
GPA					-0.194
					(0.149)
SAT					0.002**
					(0.0007)
$\ln(Round)$	0.427***			0.428***	0.405***
	(0.080)			(0.080)	(0.087)
$\ln(Period)$	0.134***	-0.003	-0.123*	0.056	0.043
	(0.042)	(0.086)	(0.067)	(0.042)	(0.045)
DC5_120	0.126			0.128	0.164
	(0.109)			(0.108)	(0.113)
DC15_120	-0.746***			-0.751***	-0.708***
	(0.125)			(0.125)	(0.132)
DC25_120	-0.461***			-0.461***	-0.426***
	(0.118)			(0.117)	(0.125)
EW_170	-0.043			-0.049	0.015
	(0.123)			(0.123)	(0.129)
DC5_170	-0.112			-0.118	-0.071
	(0.131)			(0.130)	(0.139)
DC15_170	-0.280**			-0.288**	-0.275**
	(0.130)			(0.130)	(0.140)
DC25_170	-0.322**			-0.334**	-0.288*
	(0.146)			(0.145)	(0.156)
Job Duration $t-1$	-0.033***			-0.032***	-0.034***
	(0.007)			(0.006)	(0.007)
Constant	0.840***	-0.232	0.856*	0.281	-0.773
	(0.246)	(0.552)	(0.476)	(0.365)	(0.896)
Obs (Clusters)	13,109 (290)	1,735 (290)	2,053 (290)	13,109 (290)	11,437 (252)

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Dependent variable is Effort. (2) uses only data from round 1 while (3) uses only data from round 8.

Table 4: Determinants of Effort.

the steeper slopes are consistently negative and significant. This demonstrates a consistent pattern that the DC profiles induce either less of an upward shift over the theoretically predicted effort profiles than the EW profiles or a slightly downward shift in a few cases.

The final two columns in table 4 replicate the first column but also include demographic variables about the subjects. These demographics variables were included as they were ones of key interest in examining whether they have implications for the self-selection phase of the experiment. Before understanding the effects on selection we first need to understand if they affect work effort. The effect of risk aversion is complicated as increases of risk aversion would lead to lower early effort and higher later effort thus we have included the Risk Aversion variable as well as an interaction term between it and a dummy variable for if the period is above the fourth. As expected given figure 2, there is not a significant difference over the latter part of the experiment as the predicted differences are small. We do find that the base Risk Aversion coefficient is negative and significant in one of the specifications perhaps indicating a tentative effect of risk aversion on the early rounds of effort provision. We do find though that the White and Male dummy coefficients are significant in specification (4) and those effects are positive. Specification (5) adds in our variables that are proxies for general ability and willingness to exert effort, i.e. the subject's SAT score and GPA, and we find that the SAT score is positive and significant but GPA is not. Further, in this specification the White dummy coefficient remains significant while the Male dummy coefficient falls out. This suggests that the Male result found in specification (4) was due to an omitted variable bias from leaving out this ability measure. Once this is corrected for, there is no significant difference by gender.

In addition to this general conformance with the theoretical predictions we can stress test the theory further by testing whether or not several of the comparative static predictions made by the theory are supported in the data. The first set of such predictions we will look at involve several different metrics that would be of interest to employers considering using either DC or EW profiles as a way to induce effort from employees. The questions of interest include what profile generates higher effort, lower cost, a better ratio of cost to effort and which might decrease job turnover. These comparative statics results are summarized as Prediction 2–5 (see also Figure 3). The next result provides a summary of this analysis.

Result 2 *Consistent with the theory we find that an upward shift in wages leads to an increase in effort, EW schemes generate the highest payment to workers, and the ratio of effort to gross earnings, though not earnings net of effort cost, is increasing in the slope of the wage profile. Contrary to the theoretical predictions, effort is not increasing with the slope of the profile and neither is job tenure.*

The support for this result is contained in the regressions shown in table 5. That table contains five regressions, all of the same form just with a different dependent variable. The regressions are fixed effects panel regressions with standard errors clustered on the individual subjects. As the theoretical predictions relevant to these questions made predictions regarding how the relevant dependent variable should change with respect to the slope of the profile and the overall generosity of the profile, the key independent variables are dummy variables for the three different slopes used in the DC profiles leaving the EW profile as the default and then another dummy variable for whether the profile is from the second, more generous set. The data for these regressions involve using the total effort, total earnings etc. for each subject in each round rather than period by period observations which yields the $290 * 8 = 2320$ observations.

The first regression in table 5 looks at the effect of the slope of the wage profiles on total effort. The prediction is that total observed effort should increase with the slope of the profile but we get no effect for two of the slope conditions and the effect of the steepest slope is negative and significant. Similarly the time workers spend at a job before losing it should also increase with the slope, as a consequence of the increased effort. The regression in column five uses Job Tenure or round in which the individual last held the preferred profile as the dependent variable and as one would expect we get identical results as with effort which is again counter to the theoretical prediction. We note that while we demonstrated that the magnitude of the predicted effort differences was not large as the slope varied thereby making the differences perhaps difficult to observe, our empirical results are not showing simply insignificant differences but differences opposite to the base prediction. Effort and Job Tenure are also predicted to increase with the generosity of the wage and in this case we find

Dep Var:	(1) Tot Effort	(2) Tot Earnings	(3) Effort / Earn	(4) Effort / Net Earn	(5) Job Tenure
Slope=5	0.511	-18.750	0.004**	0.004	-0.043
	(1.101)	(33.668)	(0.002)	(0.002)	(0.197)
Slope=15	-0.406	-61.740*	0.004**	0.007**	-0.092
	(1.101)	(33.457)	(0.002)	(0.003)	(0.195)
Slope=25	-2.459**	-103.100***	0.005**	0.008	-0.360*
	(1.153)	(33.208)	(0.002)	(0.007)	(0.191)
170 Set	6.268***	321.500***	-0.004***	-0.014***	0.416***
	(0.892)	(24.483)	(0.001)	(0.004)	(0.144)
ln($Round$)	4.657***	58.770***	0.005***	0.013***	0.314***
	(0.653)	(18.626)	(0.001)	(0.004)	(0.106)
Constant	24.550***	580.200***	0.0504***	0.055***	5.899***
	(1.178)	(31.786)	(0.002)	(0.004)	(0.208)
Obs (Clusters)	2320 (290)	2320 (290)	2320 (290)	2320 (290)	2320 (290)
R^2	0.060	0.093	0.019	0.017	0.011

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Analysis of comparative static predictions.

that these predictions supported as the relevant dummy variables are indeed positive and significant.

The other three columns in table 5 show regression results related to total earnings and then the ratio of effort to gross and net earnings. The theoretical prediction is that the earnings to the workers should be highest in the EW profiles and lower in the DC profiles. Column 2 contains the relevant regression and we see that while the dummy variable coefficient for the low slope is insignificant, the coefficients on the other two dummy variables are negative and significant. So the EW profiles at a minimum generate earnings no lower than any of the DC profiles. Since expected efforts are approximately similar, we can compute the predicted value of total earnings, which can be confirmed to be consistent with the ordering of expected earnings in Figure 3. The prediction on the effort to earnings ratio is that it should be increasing in the slope as effort should be rising while earnings are falling. While effort was not found to rise as expected, we still find that the earnings fall enough so that the coefficient on each of the slope dummies are positive and significant. They are, however, all of the same magnitude indicating that each of the DC profiles exerted the same level of effect rather than the effect being increasing in the slope. When using earnings net of effort cost – more of a measure of interest to workers rather than the employers – we find only the slope=15 dummy coefficient to be significant.

Our next result on the effort decisions involves a more careful look at the effort subjects chose by period and how this varies with the wage profile.

Result 3 *Consistent with the theory we find in the early periods, effort is higher in EW profiles but effort in later periods is higher in DC profiles.*

As previously demonstrated, the effort ranking between EW and DC profiles changes between early and late rounds with initial effort predicted to be higher in EW profiles and later effort predicted to be higher in DC profiles. Given the parameters used here, the switch over point should occur in period 3. To test this more detailed prediction table 6 provides a set of 10 regressions, one for each possible period of the decision task, in which the effort in that period is regressed on the same variables as the previous set of regressions. What we expect to see is that there should be negative coefficients on the slope dummies for periods 1 and 2 and then beginning in period 3 we should see the slope dummies become positive (Figure 2). This is a very detailed prediction which relies on subjects being able to work out very nuanced implications of the wage profiles. For the higher two slopes we observe results very close to the predicted relationship and those should be expected to be stronger than the results for Slope=5. In periods 1 and 2 the two dummy variables for Slope=15 and Slope=25 are indeed negative and significant. Beginning in period 4 those slope dummies become essentially 0 and then by period 6 they are positive and typically significant. This pattern is still seen to a lesser extent for the Slope=5 case.

4.2 Self-Selection

In looking at the results from the self selection task there are a few summary statistics that are important to examine prior to interpreting the results. The structure of the self-selection task involved giving the subjects four pairwise choices between EW and DC profiles. In the EW Fixed treatment, multiple DC profiles are compared to a common EW profile and the opposite was true in the DC Fixed treatment. In both treatments, at one end of the generosity of the varying profiles the EW should have

	Pd 1	Pd 2	Pd 3	Pd 4	Pd 5	Pd 6	Pd 7	Pd 8	Pd 9	Pd 10
Slope=5	-0.107	-0.156**	-0.020	0.078	0.128	0.083	0.190*	0.270**	0.456***	0.296*
	(0.076)	(0.066)	(0.071)	(0.074)	(0.079)	(0.097)	(0.103)	(0.110)	(0.122)	(0.156)
Slope=15	-0.415***	-0.360***	-0.112	-0.012	0.067	0.142	0.138	0.263**	0.357***	0.204
	(0.090)	(0.078)	(0.080)	(0.090)	(0.099)	(0.101)	(0.104)	(0.112)	(0.128)	(0.170)
Slope=25	-0.911***	-0.692***	-0.243**	0.020	-0.026	0.191*	0.172	0.434***	0.487***	0.740***
	(0.107)	(0.096)	(0.097)	(0.091)	(0.100)	(0.104)	(0.114)	(0.119)	(0.133)	(0.173)
Second Set	0.895***	0.695***	0.618***	0.623***	0.585***	0.479***	0.477***	0.503***	0.546***	0.486***
	(0.079)	(0.071)	(0.071)	(0.076)	(0.080)	(0.089)	(0.100)	(0.105)	(0.114)	(0.136)
ln(*Round*)	0.562***	0.580***	0.501***	0.447***	0.382***	0.358***	0.277***	0.207**	0.181**	0.268***
	(0.058)	(0.057)	(0.058)	(0.059)	(0.063)	(0.068)	(0.077)	(0.084)	(0.088)	(0.100)
Constant	4.231***	4.379***	4.379***	4.393***	4.508***	4.539***	4.606***	4.467***	4.075***	2.760***
	(0.114)	(0.108)	(0.112)	(0.118)	(0.125)	(0.138)	(0.156)	(0.164)	(0.181)	(0.200)
Obs (Clusters)	2320	2037	1822	1642	1478	1330	1219	1112	998	886
Clusters	290	290	290	290	289	289	289	288	286	279
R^2	0.000	0.188	0.144	0.128	0.101	0.077	0.060	0.066	0.063	0.060

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6: Period by period comparison of effort.

been at its most attractive while the DC profile should have been at its most attractive at the other end. For a subject to exhibit consistent preferences they should therefore exhibit 1 of 3 choice patterns: 1. All EW. 2. All DC 3. Beginning with a choice of EW for the pair in which EW is its most attractive and then switching to choosing DC at a later profile and choosing the DC profile at any subsequent choices. Our subjects exhibited reasonable consistency in this task with 89% of them being consistent in the EW Fixed treatment and 82% being consistent in the DC Fixed treatment. Due to the fact that it is only this subset of the subjects for whom we can claim that the mechanism might have been successful in identifying a preference, we will use only those subjects who demonstrated consistent choice behavior to provide the support for our final result.

Result 4 *Contrary to the theoretical prediction, a substantial fraction of subjects demonstrated a preference for the EW profile that is not explainable by risk aversion.*

The main question of interest for the consistent subjects is to determine if their choice behavior evidences any systematic preferences which diverge from predicted preferences. As discussed in reference to Prediction 7, relatively risk neutral individuals should always choose the DC profile while even very risk averse subjects should only choose the EW profile once. If we look at the number of times subjects chose EW profiles we find that around 40% of the subjects choose the EW profile 2 or more times which is clearly dominated behavior and not explainable by standard notion of risk preferences.

Having observed the over choosing of EW profiles there are two important questions. The first is whether this behavior is driven by individual characteristics/ demographics or by experiences during the first eight rounds. Table 7 contains regressions for each treatment of the number of times a subject chose EW profiles regressed on the relevant variables. This includes the risk aversion measure, the demographic variables used earlier and three different variables constructed based on their experience during the first eight rounds. These experience variables are the difference between the earnings, job tenure and effort exerted in EW periods and the same variable in the DC periods. We observe that these variables are all generally insignificant indicating the preferences may be due to innate preferences rather than just driven by recent

	EW Fixed			DC Fixed		
	(1)	(2)	(3)	(4)	(5)	(6)
RA	-0.071	-0.043	-0.044	0.134**	0.154**	0.151**
	(0.064)	(0.069)	(0.070)	(0.067)	(0.070)	(0.070)
Earnings Diff	0.0005		0.0003	-0.0003		-0.0004
	(0.0006)		(0.0006)	(0.0005)		(0.0006)
Tenure Diff	-0.023		-0.0007	-0.038		-0.026
	(0.068)		(0.074)	(0.065)		(0.068)
Effort Diff	0.038		0.048	0.090		0.114
	(0.088)		(0.094)	(0.069)		(0.077)
Male		0.242	0.229		-0.133	-0.105
		(0.206)	(0.210)		(0.208)	(0.213)
White		-0.042	-0.057		-0.434**	-0.428**
		(0.220)	(0.226)		(0.210)	(0.210)
Asian		-0.666	-0.628		-1.500**	-1.535**
		(0.495)	(0.503)		(0.738)	(0.736)
GPA		0.002	0.002		-0.320*	-0.320*
		(0.158)	(0.159)		(0.169)	(0.169)
SAT		-0.0003	-0.0001		-0.0002	-0.0004
		(0.0009)	(0.0009)		(0.0009)	(0.0009)
Constant	1.331***	1.567	1.400	0.792***	2.496**	2.622**
	(0.237)	(1.037)	(1.081)	(0.243)	(1.125)	(1.122)
Observations	136	120	120	113	99	99

Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 7: Regressions on the determinants of the number of times subjects chose EW profiles.

experience. We do find that the risk aversion variable is significant in one of the treatments and of the theoretically appropriate sign since risk aversion should make an individual more willing to choose the EW profiles. This is a weak result though and as explained, risk aversion can not rationalize choices of the EW profile more than once. One of the important aspects of this self-selection element was to determine if we find any systematic differences in preferences for wage profiles according to the demographics of the subjects. While we do find some significant coefficients on the White, Asian and GPA variables for the DC Fixed treatment, those variables are not significant for the EW Fixed treatment indicating a lack of robustness of the effects.

Given that we observe so many choices that differ from the theoretical prediction it is important to understand what subjects were foregoing in expected earnings to make

# EW	EW Fixed			DC Fixed		
	Wages	% Max	#	Wages	% Max	#
1	65.8	8%	33	71.2	12%	34
2	133.6	15%	43	153.1	25%	34
3	193.9	20%	9	228.8	37%	8
4	285	42%	5	341.1	55%	3

Table 8: Earnings foregone based on wage profile choices.

those choices. Table 8 shows results related to the total wages foregone in expectation from each number of times of choosing the EW profile rather than the more lucrative DC profile. The losses are presented in terms of total ECUs lost from that choice as well as the percentage of potential earnings lost by making the inferior choice.[8] Also included is the number of individuals making each choice. While choosing the EW profile a single time results in lower earnings, the loss in expected wages is small at around 10% and so it is unsurprising that many might make this small mistake. Choosing the EW profile at the next opportunity involves an expected decrease in earnings of around 20% with losses for continuing to choose the EW reaching 50% for the small number who choose the EW at all opportunities. Those choosing the EW profile two or more times are giving up substantial earnings to do so which indicates potentially strong preferences for the EW profile.

5 Conclusion

In this paper, we compare two forms of compensation that have received much attention in the theoretical literature – efficiency wages and deferred compensation. Empirical tests of their predictions and, in particular, comparisons of relative effectiveness in providing incentives are of interest but have been difficult to accomplish with field data. In our experimental study, we observed effort choices of individuals faced with multiple different wage profiles in the context of an otherwise identical

[8]Specifically each entry in table 8 shows the earnings difference between the EW and DC profiles in the 1st, 2nd, 3rd and 4th choice opportunities. The losses are not calculated as cumulative losses including the previous choices.

life-cycle labor supply problem. We also elicited individual preferences between wage profiles.

On effort provision, we compared EW with DC, and tested empirical predictions of the dynamic effort provision pattern. The theoretical predictions on effort are relatively well supported. Effort is generally over-provided, but over-provision is least for the steeper wage profiles. Our theoretical prediction was that DC should be attractive from the perspective of employers. In our parametrization, increasing the extent of back-loading of compensation (i.e. slope) leads to higher expected effort and lower expected payments. Empirically, we found an effort/earnings edge for DC profiles although the magnitude of the effect was small. This is likely because of the relative over-provision of effort under EW. Interestingly, we confirm also a rather subtle prediction of the theory – effort in early periods is higher in EW but effort in later periods is higher with DC profiles. In other words, DC profiles can attenuate the decline in effort at later stages of the life-cycle, even though the overall advantage over EW is not as much as predicted.

We studied worker choice between wage profiles and found a marked preference for EW wage profiles. The preference was much larger than could be explained by risk-aversion alone. Recall that preference among profiles was elicited in two treatments where (1) EW was made less attractive, and (2) DC was made more attractive. The consistent bias in favor of EW thus amounted to leaving money on the table and, as our calculations indicate, the amounts are not negligible. This has important implications for employers. Employee preference for EW may allow employers to attract them at a lower cost when compared to DC. Combined with the fact of effort over-provision with EW, this may create a distinct advantage for employers from using EW profiles.

We have explored some behavioral mechanisms that might account for the observed deviations from theory. We first considered the possibility of discounting, and find that this does explain an excessive preference for efficiency wages over deferred compensation. However, it leads to effort levels that are lower than those predicted by the risk-neutral no-discounting model of the paper. So the over-provision of effort found here would be even greater relative to such a prediction. We considered loss-aversion in the context of cumulative prospect theory, Tversky and Kahneman (1992).

The difficulty in the application of the model is that, in our particular case, there is no obvious choice of reference point. We tried a number of possibilities involving both fixed and variable reference points. For reference values that lie in between most payoffs and the outside wage we do indeed find that there is an over-provision of effort early (its magnitude declining over time). Additionally, in some cases the extent of the over-provision is greater for the flatter wage profiles. This is indeed consistent with what we observe in our experiment. For other reference point specifications, we find that prospect theory predicts efforts lower than the standard prediction. Considering that this ability of prospect theory to explain the behavior is dependent on a specific choice of a reference point and that there is no clearly applicable theory to use in determining the appropriate reference point for this situation, we hesitate to make a strong claim regarding the ability of prospect theory to explain the results.

It would be interesting to determine whether these result can be established with field data. More precisely, in an environment where providing effort incentives is the only issue, is it the case that EW is more often used (found to be effective). Clearly, many employers choose DC profiles. Assuming our results are accurate, the question is why? One possibility is that profession or employer specific human capital may lead to more experienced employees having higher productivity. Another possibility is that this could be a consequence of a high degree of employee impatience and discounting of future rewards. A third explanation relates to alternative forms of sorting in the presence of worker heterogeneity. We have considered heterogeneity in risk preferences. However, in the presence of differences in ability (e.g. cost of effort, or productivity differences) wages can work as sorting devices (Guasch and Weiss (1980), Guasch and Weiss (1981)). Also, workers' intention to quit may be private information, and DC may screen out workers who intend to quit soon (or who might be attracted away by other employers in periods of high demand). This idea was first developed by Salop and Steven (1976). Related to this is the idea that if workers know themselves to be a better match for a job, they would be more likely to accept a DC wage. Further experimental work evaluating choice in these alternative environments would be desirable, and could contribute to the goal of determining precisely what features of the environment are associated with what types of compensation schemes. An important contribution from the current study though is to point out that it is to one of these alternative considerations that one should look in justifying the use of DC profiles.

References

Akerlof, G. A. (1982). Labor contracts as partial gift exchange. *The Quarterly Journal of Economics 97*(4), 543–569.

Akerlof, G. A. and L. F. Katz (1989). Workers' trust funds and the logic of wage profiles. *The Quarterly Journal of Economics 104*(3), 525–536.

Binswanger, H. P. (1980). Attitudes toward risk: Experimental measurement in rural india. *American Journal of Agricultural Economics 62*, 395–407.

Binswanger, H. P. (1981). Attitudes toward risk: Theoretical implications of an experiment in rural india. *Economic Journal 91*, 867–890.

Eckel, C. C. and P. J. Grossman (2006). Forecasting Risk Attitudes: An Experimental Study of Actual and Forecast Risk Attitudes. Working Paper UT Dallas.

Fehr, E., G. Kirchsteiger, and A. Riedl (1993). Does Fairness Prevent Market Clearing? An Experimental Investigation. *Quarterly Journal of Economics 108*, 437–459.

Fischbacher, U. (2007). z-Tree: Zurich Toolbox For Readymade Economic Experiments. *Experimental Economics 10*(2), 171–178.

Frchette, G. (2009). Laboratory experiments: Professionals versus students. NYU Working Paper.

Gibbons, R. and L. F. Katz (1992). Does Unmeasured Ability Explain Inter-Industry Wage Differentials? *Review of Economic Studies 59*, 515–535.

Gibbons, R. and K. Murphy (1992). Optimal incentive contracts in the presence of career concerns: Theory and evidence. *Journal of Political Economy 100*(3), 468–505.

Greiner, B. (2004). An online recruitment system for experimental economics. In K. Kremer and V. Macho (Eds.), *Forschung und wissenschaftliches Rechnen 2003*, Volume 63, pp. 79–93. Gttingen: Ges. fr Wiss. Datenverarbeitung.

Guasch, J. L. and A. Weiss (1980). Wages as sorting mechanisms in competitive markets with asymmetric information: A theory of testing. *Review of Economic Studies 947*(4), 653–664.

Guasch, J. L. and A. Weiss (1981). Self-selection in the labor market. *American Economic Review 71*(3), 275–284.

Huck, S., A. J. Seltzer, and B. Wallace (2011). Deferred compensation in multi-period labor contracts: An experimental test of lazear's model. forthcoming in the American Economic Review.

Johansson, P. and M. Palme (1996). Do Economic Incentives Affect Work Absence? Empirical Evidence Using Swedish Micro Data. *Journal of Public Economics 59*, 195–218.

Lazear, E. P. (1981). Agency, Earnings Profiles, Productivity, and Hours Restrictions. *American Economic Review 71*(4), 606–620.

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

Lazear, E. P. (2000b). The Future of Personnel Economics. *Economic Journal 110*, F611–F639.

Loewenstein, G. and N. Sicherman (1991). Do workers prefer increasing wage profiles? *Journal of Labor Economics 9*(1), 67–84.

Murphy, K. M. and R. H. Topel (1990). Efficiency Wages Reconsidered: Theory and Evidence. In Y. Weiss and G. Fishelson (Eds.), *Advances in the Theory and Measurement of Unemployment*, pp. 104–240. MacMillan.

Peter, C. and K. Chauvin (1991). An Interplant Test of the Efficiency Wage Hypothesis. *Quarterly Journal of Economics 106*, 769–787.

Prendergast, C. (1999). The provision of incentives in firms. *Journal of Economic Literature 37*(1), 7–63.

Rigdon, M. L. (2002). Efficiency Wages in an Experimental Labor Market. *Proceedings of the National Academy of Sciences 99*(20), 13348–13351.

Ritter, J. A. and L. J. Taylor (1994). Workers as creditors: Performance bonds and efficiency wages. *The American Economic Review 84*(3), 694–704.

Salop, J. and S. Steven (1976). Self-selection and turnover in the labor market. *The Quarterly Journal of Economics 90*(4), 619–627.

Shapiro, C. and J. E. Stiglitz (1984). Equilibrium as a worker discipline device. *The American Economic Review 74*(3), 433–444.

Tversky, A. and D. Kahneman (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty 5*, 297–323.