Risk or Loss Aversion?

Evidence from Personnel Records

Emil P. Iantchev*

Syracuse University

July 8, 2011

Abstract

I use data on individual pay and output from a private firm to estimate a principal-agent model with moral hazard in which the agent's preferences are allowed to feature deviations–reference dependence and loss aversion–from the standard case of risk aversion. Estimating the preference parameters via GMM shows that a specification where loss aversion equals 2.11 and the reference point is determined by the prevalent contract fits the data better than any of the other nested specifications. Counterfactual estimates suggest that ignoring loss aversion if present would have increased the firm's compensation costs by an average of about 28 percent.

*Department of Economics, 110 Eggers Hall, Syracuse, NY 13244; e-mail: iantchev@syr.edu.

Acknowledgements to be added.

1 Introduction

In designing an optimal mechanism, a rational principal will take into account the attitudes of the participating agents. Since in most instances the optimal way to provide incentives depends on the preferences of the parties involved, the incentive contract chosen by the principal contains valuable information about the agents' preferences. Given a sufficiently rich dataset on individual characteristics and contracts, an economist could infer which of several preference specifications conforms most closely with the empirical evidence.

This paper uses a dataset on individual pay and productivity from the Safelite autoglass company, the same dataset analyzed by Lazear (2000), to estimate the production and preference parameters of the classic principal-agent model with moral hazard. This is a useful exercise since the incomplete risk-sharing implied by moral hazard allows us to distinguish, based on contractual outcomes, among different specifications of attitudes towards risk. The main conceptual departure of the paper lies in the specification used to model the agent's preferences. In particular, I consider a piecewise quadratic functional form which nests approximations to various specifications of reference dependent preferences with or without loss aversion and diminishing sensitivity as well as an approximation to the standard case of CARA utility.

Using two-stage GMM, I obtain a point estimate for the coefficient of loss aversion of 2.11, which is comparable to estimates derived from experiments, typically ranging between 2 and 2.5. I also estimate two alternative specifications–no loss aversion and CARA. The value for the objective function in each of the restricted models is much higher than the objective value of the unrestricted model, which suggests that the unrestricted model fits the data better than either of the alternatives.

In addition, I assess the goodness-of-fit for the different preference specifications by comparing the simulated and actual contracts for a random sample which was not used in estimating the model parameters. In terms of the root mean squared error (RMSE) criterion, the unrestricted model outperforms the model restricted to the CARA approximation by about 18 percent. Furthermore, ignoring loss aversion and diminishing sensitivity if present would have increased the firm's compensation costs by an average of about 28 percent–$650 per person-month–which implies that loss aversion and diminishing sensitivity could have played a significant role in shaping the optimal incentive contract adopted by Safelite.

I use the Safelite dataset mainly for two reasons. First, it contains information on individual output under two incentive regimes that elicit different effort levels. Hence the production process can be estimated under both high and low effort, which helps in identifying the production parameters of the model. In addition, the incentive contract chosen by Safelite is used commonly in practice, which suggests that any insight about preferences should be generally applicable.

I first formalize the environment in which the dataset was generated. Specifically, I derive the properties of the optimal deviation to a high powered incentive contract by a principal that operates in a competitive market where the prevalent contract is low powered. Under loss aversion, diminishing sensitivity, and the assumption that the agent's reference point is the expected value under the prevalent contract in the market[1], the optimal contract has several properties that distinguish it from the insurance-incentives trade-off characteristic of pure risk aversion. In particular, the optimal contract postulates that if output is below a critical value, the principal should impose on the agent the biggest possible punishment, a transfer of $w^{\min}$, with probability one. When output is above this critical value but below an upper threshold, the agent receives a constant amount equal to his market wage (outside option). And if output exceeds the upper threshold, there is a region with increasing additional rewards. Hence contrary to the consumption-smoothing result of the model with a risk averse agent, incentives under loss aversion and diminishing sensitivity are provided globally rather than locally.

The constant transfer in the middle part of the output distribution is due to reference dependence and loss aversion. Loss aversion makes the agent very risk averse around the

[1] This rule is motivated by Rayo and Becker (2007). We use it instead of the rule by Koszegi and Rabin (2006) for empirical reasons. See section 3.3. for the details.

reference point so full insurance is locally optimal. The agent is risk averse for gains relative to the reference transfer, hence incentives are partially provided by the region with increasing rewards. On the other hand, the agent is locally risk-loving for values below the reference point making a corner solution, the biggest possible punishment, optimal.

On the surface, this optimal contract fits well with evidence from the personnel literature in general and Safelite in particular. Specifically, the company switched its workforce to an incentive contract which featured a piece rate coupled with a minimum hourly guarantee. That is, if the weekly piece rate pay for a worker came below the guarantee, he would receive the guaranteed amount. The guarantee region was economically significant with about 68 percent of person-months featuring some guaranteed pay. In addition, more than 93 percent of employees were given a guarantee exactly equal to their previous wage, hence outside option, which is also consistent with the theoretical model.

However, the flat region in the observed incentive contract could well have been due to a non-monotonicity in the production process. Moreover, there was no explicit provision for punishment in the case of poor performance. To address the first issue, I estimate a production process which is flexible enough to incorporate violations of the Monotone Likelihood Ratio Condition (MLRC). As it turns out, the data is consistent with MLRC being satisfied, which rules out a non-monotone output process as the reason for the existence of the guarantee region.

As far as the punishment region is concerned, the model is consistent with the evidence if we interpret punishment as the forfeiture of an implicit performance bond in the event of separation. In particular, the estimates from a logistic hazard model suggest that workers with contemporaneous output in the left tail of the output distribution were significantly more likely to separate from the firm during the incentive regime than workers with current output in the middle and right parts of the distribution. I thus interpret termination as the ultimate punishment that the principal can impose on each agent and treat the amount, $w^{\min}$, forfeited by the agent as something to be estimated and which in principle can depend on the agent's observable characteristics.

Section 2 provides a summary of the related literature. Section 3 contains the theoretical part of the paper. Section 4 describes the dataset, while section 5 presents the empirical strategy and estimation results. Section 6 contains the goodness-of-fit simulations and section 7 concludes.

2 Related Literature

The paper is related to three different strands of economic analysis. First, it complements the existing literature that tries to recover preference parameters using regular market participants. For the case of risk attitudes, prominent examples are Cicchetti and Dubin (1994), who look at individual decisions to insure against telephone wire failure, Chetty (2006), who infers risk preferences from labor supply choices, and Cohen and Einav (2007), who estimate the entire distribution of absolute risk aversion underlaying deductible choices for automobile insurance in Israel. Moreover, our estimate of loss aversion is derived using market data. Most of the evidence on the coefficient of loss aversion is either from lab experiments such as Tversky and Kahneman (1992), or field experiments such as the TV show "Deal or No Deal?" used by Post et al. (2007). In this sense, our paper can shed some light on how comparable the estimates from the experimental and market based approaches are.

This paper also contributes to the literature on the applications of prospect theory preferences. Camerer (2001) identifies ten market phenomena that are inconsistent with the classic formulation of expected utility theory but can be explained by some feature of prospect theory, typically loss aversion or diminishing sensitivity. The most prominent applications are in the area of finance–Benartzi and Thaler (1995), Barberis, Huang, and Santos (2001)–where it is argued that reference dependence and loss aversion can successfully account for the equity premium puzzle. But prospect theory preferences have also been used to explain behavior in a variety of other situations. Examples include inter-temporal choice–Bowman, Minehart, and Rabin (1999)–labor supply–Camerer et al. (1997), Goette, Fehr, and Huffman (2004)–and seller behavior in the residential real estate market–Genesove and Mayer (2001).

Initially, papers in this area aimed at explaining peculiar forms of behavior that violated predictions of the standard theory. Recently, however, works such as Barberis, Huang, and Santos (2001), Mas (2006), and Heidhues and Koszegi (2005) have been investigating more systematically the general implications of prospect theory preferences. Our paper contributes to this trend by incorporating prospect theory preferences in the classic principal-agent model with moral hazard. The conclusions of the model are then tested with available data and their merit evaluated on the basis of how well they explain that data.

In this sense, our paper also contributes to the empirical literature on the structural estimation of principal-agent models with moral hazard. As noted by Chiappori and Salanie (2003), this literature is still very much in its infancy. However, as a result of the complementary increases in the availability of micro datasets and computational power, the structural estimation of such models has become more feasible. Recent examples of this approach can be found in Flinn (1997), Ferrall and Shearer (1999), Paarsch and Shearer (2000), Vera-Hernandez (2003), and Bandiera, Barankay, and Rasul (2005) and (forthcoming). The last two papers are especially relevant since they use the detailed nature of personnel records to test hypotheses about social preferences, a task which this paper also tries to make progress on.

3 Theoretical Preliminaries

In this section I provide a description of the environment in which the principal-agent interaction takes place. Consider a *competitive labor market* where principals and agents trade labor services (actions) for monetary remuneration (earnings) via *spot contracts.* A contract in this environment is a recommended action for the agent and a payment schedule which may depend on the output produced by the agent. There are many principals and many agents and a principal may sign contracts with more than one agent but not vice versa. There are *no frictions and no specificity to any particular relationship,* so long term contracts are ruled out[2].

Upon signing a contract, an agent can *post a performance bond up to an exogenous limit* $w^{\min}$ that may depend on his characteristics. After the contract is signed, the agent has to chose an action from the set $\{H, L\}$. The action L denotes the level of effort that can be observed without cost by the principal. This includes whether the agent comes to work on time, slacks around the shop, etc. Beyond this minimum effort level, whether any "heart", H, is put into the job is the agent's private information.

After the agent has chosen an action, his output Y gets realized. Individual output is a random variable that can take on one of S possible values with $Y_1 < Y_2 < ... < Y_S$. I assume that the agent can freely dispose of output if he finds that to be advantageous. Furthermore, conditional on free disposal, I assume that individual output is *observable* but *may or may not be contractible.* As is standard in the incentives literature, I also assume that the agent's action influences the probability with which each outcome occurs. In particular, let $p_{sH} = \Pr(Y = Y_s|H)$ and $p_{sL} = \Pr(Y = Y_s|L)$ where presumably $\{p_{sH}\}$ and $\{p_{sL}\}$ are such that $E[Y|H] > E[Y|L]$. Note that I have implicitly assumed that no team production takes place. In the case of Safelite, however, this is a good approximation to reality.

To summarize, within each period we have the following timing of events: (i) contracts are signed and any bond is posted if necessary; (ii) action is chosen by the agent; (iii) output uncertainty is realized; (iv) ex post transfer takes place as specified by the contract and the period ends.

[2] These assumptions appear to be plausible for the industry under investigation. In particular, there seems to be a strong seasonality to the business. Hence most contracts are short term and it is not uncommon for workers to be dismissed in periods of low output demand.

3.1 Preferences

The principals in this market are risk neutral and their objective is to maximize expected profit. The agents on the other hand maximize expected utility. Unlike the principals, the agents are not risk neutral. In particular, from the agents' point of view, in each period there are two carriers of utility–earnings and effort. Utility after the output uncertainty of the current period has been realized is given by the function

$$v(b - r) + D(a = L)e(r_a), \tag{1}$$

where b is the realized current money payoff to the agent, D is an indicator function that takes on 1 if the action chosen by the agent was L and takes on 0 otherwise, $e(r_a) > 0$ is the current disutility of high effort (H relative to L), and (r, r_a) are reference values to be described shortly.

For the empirical section only, I will assume that

$$v(x) = \left\{ \begin{array}{ll} x + \gamma_g \cdot x^2 & \text{for } x \geq 0 \\ \Lambda x + \gamma_l \cdot x^2 & \text{for } x < 0 \end{array} \right\}, \tag{2}$$

where $\Lambda \geq 1$ is a coefficient of loss aversion while γ_g and γ_l represent the curvature of the function above and below the reference point.

The above function is a quadratic approximation to several specifications of reference-dependent preferences as well as the standard case of CARA utility. For instance, the value function in Kahneman and Tversky (1978) corresponds to the restrictions $r = b_{-1}$ (last period payoff), $\Lambda > 1$, $\gamma_g \leq 0$, and $\gamma_l \geq 0$. Koszegi and Rabin's version with deterministic reference point in their 2006 paper implies the same restrictions except that now $r = E_C[b]$ where C is the contract the agent expects to face. Similarly, the specification derived by Rayo and Becker (2007) corresponds to $r = E_C[b]$, $\Lambda = 1$, $\gamma_g < 0$, and $\gamma_l > 0$.

As far as the case of CARA utility, note that for any concave $\widetilde{v}$,

$$\widetilde{v}(b) \approx \widetilde{v}(r) + \widetilde{v}'(r)(b-r) + \frac{1}{2}\widetilde{v}''(r)(b-r)^2. \tag{3}$$

Now normalize $\widetilde{v}(r) = 0$ and let $v(b) = \widetilde{v}(b)/\widetilde{v}'(r)$. Then clearly, if $\widetilde{v}(b)$ is of the CARA type, $v(b)$ will be isomorphic to its quadratic approximation around r with $\Lambda = 1$ and $\gamma_g = \gamma_l = \widetilde{v}''(r)/2\widetilde{v}'(r)$. Hence the case of constant absolute risk aversion can be defined as the pair of restrictions $\Lambda = 1$ and $\gamma_g = \gamma_l < 0$.

3.2 Competitive equilibrium with non-contractible output

As a benchmark, I will first describe the equilibrium contract under the following assumption:

A1. Individual output is NOT verifiable/ NOT contractible.

The equilibrium concept I use is Bertrand-Rotschild-Stigliz equilibrium since the market for windshield installers is competitive. Formally, I define competitive equilibrium in the following way.

Definition: *A contract $C = (b, a)$ is a competitive equilibrium if it's robust to the introduction of new contracts. That is, there is no other contract which is at least as good to the agents and makes positive expected profit.*

It is immediately obvious that any equilibrium must feature *zero* expected profit. If that is not the case, then there exists another contract which is feasible, makes positive profit, and strictly increases the agent's expected utility. But then the current contract cannot be a competitive equilibrium. Hence if $C^* = (b^*, a^*)$ is an equilibrium, I must have $b^* = E[Y|a^*]$.

In equilibrium, no extra effort can then be induced from the agent. At the effort choosing stage, the agent decides between $v(b^*-r)+e(r_a)$ and $v(b^*-r)$. Clearly, given that $e(r_a) > 0$, the agent will never choose to exert extra effort[3]. Given that the equilibrium transfer is

[3] Note that this conclusion critically depends on A1 as well as the assumptions of no frictions or relationship specificity. In particular, if there is a strictly positive future surplus from continuing the current relationship

constant and equilibrium profits are zero, we can conclude that $b^* = E[Y|L]$.

3.3 Optimal deviation when output becomes contractible for one principal

Suppose now that assumption A1 is no longer satisfied for one particular principal even though it holds for all other principals in the market. Since this is a deviation from equilibrium by just one firm, the rest of the principals behave in the usual manner.

I will initially take (r, r_a) as given and then endogenize it. Let $\underline{V}$ be the agent's (ex-ante) reservation utility in equilibrium with non-contractible output. I know that in any such equilibrium $a^* = L$ and $b^* = E[Y|L]$, which implies that the agent's reservation value is given by

$$\underline{V} = \sum_{s=1}^{S} p_{sL} v(b^* - r) + e = e, \tag{4}$$

where I have suppressed the dependence of the disutility of effort on r_a. I want to assume that the production technology is such that if an agent who is currently employed were to own the technology, it will be individually rational to exert high effort. That is,

$$E[v(Y - r)|H] - E[v(Y - r)|L] > e. \tag{5}$$

The firm that deviates can now provide incentives using piece rate contracts. Hence b CAN be state-contingent. Let $b_s \geq w^{\min}$ be any state contingent ex-post transfer. The deviating principal would like to implement effort H since it is the more productive action. Moreover, the principal would like to structure the incentive contract in such a way that ex ante, the agents are just indifferent between working for him with high effort and their outside

compared with the outside option, incentives can be provided through a long term contract that features an implicit piece rate–see MacLeod and Malcomson (1989) option $\underline{V}$. This is because for a given effort level, providing a higher level of expected utility increases expected expenditures. Whether it's possible to drive expected utility completely down to $\underline{V}$ depends on the value of $w^{\min}$[4]. I will assume that the limit ($w^{\min}$) is lax enough so that the participation constraint binds with equality. The profit maximizing spot deviation contract can then be described by

$$\pi(\underline{V}) = \max_{\{b_s\}} \sum_{s=1}^{S} p_{sH}\{Y_s - b_s\}$$

$$\text{s.t.} \quad b_s \geq w^{\min},$$

$$(IC): \quad \sum_{s=1}^{S}(p_{sH} - p_{sL})v(b_s - r) \geq e \qquad [\mu]$$

$$(PC): \quad \sum_{s=1}^{S} p_{sH}v(b_s - r) = \underline{V} \qquad [\lambda]$$

$$v(b_s - r) \geq v(b_y - r) \qquad \text{for all } s \text{ and } y < s,$$

where b_s can in general be stochastic as long as any possible outcome of a lottery is bigger than $w^{\min}$. I can then prove the following proposition for the case of diminishing sensitivity and strong loss aversion.

Proposition 1: *Suppose sabotage is not possible. For $v : [\underline{x}, \infty] \mapsto \mathbb{R}$ exhibiting diminishing sensitivity–$v''_{x>0}(x) < 0$ and $v''_{x<0}(x) \geq 0$–and strong loss aversion–$v'_{x\uparrow 0}(x)/v'_{x\downarrow 0}(x) > 1$ and $v'(x) > 1$ for $x \in [\underline{x}, 0]$–the optimal deviation contract is given by:*

1. Punishment region. If the realized state of the world s is such that the likelihood ratio $1 - p_{sL}/p_{sH} < (\underline{x}/v(\underline{x}) - \lambda)/\mu$, with μ and λ as the multipliers on the incentive and

[4] See Laffont and Martimort (2002) section 4.3.

participation constraints, the agent receives the minimum allowable transfer $\underline{x}$ *with probability one.*

2. Inactivity (draw) region. For states where the likelihood ratio is within the open interval from $(\underline{x}/v(\underline{x}) - \lambda)/\mu$ *to* $(v'_{x\downarrow 0}(x) - \lambda)/\mu$*, the agent receives a constant transfer of* r *with probability one.*

3. Incentive region. For states where $1 - p_{sL}/p_{sH} > (v'_{x\downarrow 0}(x) - \lambda)/\mu$*, the agent receives a transfer of* $b_s = r + x_s$ *with* x_s *such that* $1 - p_{sL}/p_{sH} = [1/v'(x_s) - \lambda]/\mu$*.*

Proof: See Appendix A.

A few comments about the optimal contract will prove useful once we move to the empirical section. The properties of the contract clearly depend on the way the reference point is determined. If we do not have a well-formed theory of how the values of (r, r_a) are determined, we will not be able to put any sharp restrictions on behavior. One possibility is the *personal equilibrium* concept of Koszegi and Rabin (2006) in which the principal is always able to influence the agents' reference points. In that case, the optimal contract is the solution to a fixed point problem. That is, $r^* = E[b^*|(r^*, r_a^*), e(r_a^*)]$ and similarly for r_a^*, where $\{b_s^*\}_{s=1}^S$ is the solution to the principal's profit maximization problem given (r^*, r_a^*). This specification suggests that the agent's guarantee is set *above* his previous wage rate since the agent must be compensated for exerting the extra effort.

Another possibility is that the agent's reference point is the expected value under the prevalent contract in the market, which is consistent with Rayo and Becker (2007). In that scenario, no single principal can influence the agents' preferences by unilaterally deviating from the prevalent contract in the market. For instance, if every principal in the market offers a suboptimal contract that provides utility U, when a principal switches to a contract that provides utility $U + \varepsilon$, the preferences of all agents in the market, including the agents employed by the deviating principal, will not change. This is because agents form their reference points based on the prevalent contract, while each principal is infinitesimal relative to the entire market. There needs to be a positive mass of principals deviating simultaneously to the same contract in order for the agents' reference points (contractual expectations) to be affected. In that case, we will have $r^* = E[b^*|L]$ and $r_a^* = L$, where $b^* = E[Y|L]$ is the prevalent (equilibrium) contract in the market. In this specification, the deviating principal will set the agent's guarantee at *exactly* his previous wage rate.

How the reference point is determined is hence an empirical matter. When Safelite switched its workers to incentive pay, over 93 percent of the employees were provided with an hourly guarantee that was *precisely* equal to their previous wage (outside option). This is exactly the prediction generated in an environment where the reference point is determined by the prevalent contract in the market. Given that our goal is to find the model that best interprets the data, I use the reference point rule derived by Rayo and Becker rather than Koszegi and Rabin.

The punishment region can be interpreted as states where the current principal-agent relationship is involuntarily terminated. In particular, imagine the following situation. At the beginning of a relationship, the agent posts a performance bond in the amount of $w^{\min}$. As long as performance is satisfactory (outside of the punishment region) the agent receives a transfer greater than or equal to r and the performance bond is rolled over or reimbursed if a voluntary separation occurs. However, if performance falls into the punishment region, the bond is forfeited and the agent has to post a new bond next period, which is akin to starting a fresh relationship, even if it is with the same principal. In addition, note that not every contract must necessarily contain a punishment region. Whether punishment exists or not will depend on the production process being considered. If, for instance, the production process is such that $1 - p_{1L}/p_{1H} > (\underline{x}/v(\underline{x}) - \lambda)/\mu$, the optimal contract will consist only of the draw and incentive regions. Whether a punishment region exists is thus a matter which depends on the particular production technology in question and which must ultimately be settled empirically.

The optimal contract in the case of CARA preferences features the well known local trade-off between insurance and incentives. The payoff to the agent varies smoothly with the value of the likelihood ratio. In the case where the production technology satisfies the monotone likelihood ratio condition (MLRC) or sabotage is impossible, the optimal contract will not feature a flat region above the minimum transfer $w^{\min}$. If MLRC is not satisfied, however, and the agent can destroy output, the optimal contract will call for a constant payoff to the agent over the non-monotone region of the output distribution. It is therefore crucial that loss aversion is not imposed a priori by requiring the MLRC to be satisfied. Such flexibility can be achieved by assuming an output process which is general enough to allow for the violation of MLRC. As it turns out, the estimation in Section 5.2. suggests that the MLRC is satisfied for the production process generating the Safelite data.

4 Dataset Description

I now attempt to formally test the theory developed thus far by analyzing a panel dataset on individual pay and productivity courtesy of the Safelite Glass Corporation. The company is located in Columbus, Ohio and is the largest installer of automobile glass in the US. Historically, Safelite had been paying its workers an hourly wage that did not depend explicitly on the number of units installed. During the period 1994-1995, workers were gradually switched from an hourly wage to a performance pay plan (PPP). The incentive plan consisted of a constant piece rate per unit installed coupled with an hourly wage guarantee. If their weekly piece rate pay came below the guarantee, workers were paid the guaranteed amount.

The contractual change and the resulting increase in productivity and profits are extensively analyzed in a paper by Lazear (2000). Lazear (2000, p. 1350) says that the new managers "changed the compensation scheme because they felt productivity was below where it should have been." The change was then clearly done for incentive reasons and can be analyzed as a move from a low powered to a high powered incentive regime.

At the time of the change, the rest of the companies in the industry were using the hourly wage system previously used by Safelite. Hence the change to PPP fits reasonably well with the off-equilibrium deviation analyzed in the previous section. What fundamentals laid behind the change is a very interesting question in itself. The computerized information system owned by Safelite no doubt played a role. Without it, keeping verifiable records of individual output would have been very expensive. In addition, technological improvements in capital may have increased the marginal product of effort, especially if capital and effort are complements. The principal will then want workers to exert more effort, claiming that productivity is not where it should be. Finally, there was a change in management prior to the switch, which suggests that the adoption of incentive pay might have been a profit opportunity identified by the new management.

Whatever the reason behind the change to the high powered incentive regime, the switch itself is where the dataset becomes especially valuable for our purposes. The environment in the company fits well with the classical model of moral hazard. First, Safelite operates in a competitive industry. In addition, individual output is measurable and quality concerns are easy to address. Multi-tasking and efficiency wage considerations are hence not likely to play an important role. And since the switch to the incentive plan was phased in over a nineteen month period, there is reliable data on individual output and compensation under both the low and high powered incentive regimes, enabling us to estimate the production process under both high and low effort.

The dataset used here is exactly the same as the dataset used by Lazear. Since I would like to examine the goodness-of-fit for various preference specifications, I split the workers in two random samples. One of the samples is used for estimating the production and preference parameters. Once the estimation is done, we can compute the optimal contract that these estimates imply for each of the data points in the simulation sample. In this way, we would be able to compare the contracts implied by different preference specification with the actual contracts implemented by the company. In addition, we will be able to rank the overall goodness-of-fit for different specifications using a criterion such as root mean squared

: Table 1: Descriptive statistics–mean [sdev]–for selected variables.

	Full Sample		Est. Sample	Sim. Sample
	Fixed Wage	Incentives		
Units-a-day	2.72 [1.42]	3.30 [1.58]	3.02 [1.56]	3.00 [1.50]
Tenure (years)	4.11 [4.36]	3.97 [3.99]	4.01 [4.46]	3.97 [3.85]
Hourly wage or guarantee	11.73 [3.27]	11.40 [2.52]	11.56 [3.00]	11.51 [2.79]
Piecerate in incentive regime[i]	– [–]	41.65 [119.13]	43.41 [125.51]	39.99 [112.78]
Attrition probability[ii]	0.05 [0.05]	0.04 [0.04]	0.05 [0.05]	0.05 [0.04]
Fraction of person-months in incentive regime	0.50		0.49	0.51
Number of person-months[iii]	12,570	14,073	13,802	14,219
Number of persons	2,547		1,274	1,273

Notes: i. imputed from monthly incentive pay and output; ii. estimates from a logistic hazard model; iii. 1378 transition months are ommitted from the fixed wage and incentive subsamples.

error (RMSE). The simulation stage will thus allow us to see which preference specification comes most closely to rationalizing the observed contracts.

Table 1 presents summary statistics for some of the key variables for the full sample by incentive regime and for each of the two subsamples. The basic measure of output is units per day–the monthly average of the number of windshields that an individual installs per eight hours of labor during a given month. As evident from the table, average daily output increased significantly after the company switched to the performance pay plan. Lazear argues that about half of the increase is due to incentives–workers exert more effort, while the other half is the result of selection–with performance pay, the firm attracts and is able to retain a better pool of workers.

Another interesting fact is that both the guarantee and the piece rate varied significantly across workers. This is consistent with the hypothesis that the principal tailors the incentive contract to each agent based on the agent's observable characteristics. Moreover, when workers were switched to PPP pay, they were given an hourly guarantee equal to their previous hourly wage. In particular, over 93 percent of associates were given an hourly guarantee exactly equal to their previous hourly wage. This finding is consistent with the hypothesis that the agents' reference points continued to be at their competitive wages since the switch by Safelite was a deviation by just one principal and thus did not influence the agents' preferences.

5 Estimation

In order to bring the theoretical model closer to the data, I introduce a few modifications prior to estimation. Specifically, I allow agents to be heterogeneous along a vector of observable characteristics Z_t. These characteristics may influence the production process, the agent's disutility of effort, and the constraint on the minimum allowable transfer.

The disutility of effort is assumed to be stochastic and can vary both across agents and periods even after controlling for observable characteristics. Specifically, assume that $e_{it} = e(Z_{it}; \xi) + u_{it}$, where u_{it} is iid with $E[u_{it}|Z_{it}] = 0$ and $\min e(Z_{it}; \xi) + \min u_{it} > 0$. u_{it} is realized at the very beginning of tenure period t and is observed by both the principal and the agent but not by the econometrician. In addition, I assume that there is no heterogeneity in the parameters Λ, γ_g, and γ_l that govern attitudes towards earnings.

The two assumptions mentioned above are made for tractability as relaxing either of them will substantially complicate the analysis both theoretically and empirically. Although neither assumption is likely to hold completely in reality, they are reasonable as a first approximation, especially in light of the incentive contracts implemented by the firm. In particular, asymmetric information about heterogeneous preferences would introduce adverse selection on top of the moral hazard. We would then expect to observe a menu of contracts offered by the firm for the workers to self-select into. However, no such menu was observed in Safelite's case. Even though contracts were individually tailored, they all had the same basic structure–piece rate plus a minimum guarantee. Observed preference heterogeneity on the other hand is consistent with the individually tailored contracts and can be accommodated by making the preference parameters functions of observable characteristics. This is a fairly straightforward generalization of the current model but it significantly increases the number of preference parameters that need to be estimated in the GMM step. Given the computational cost, I do not pursue this extension here, mainly because the goodness-of-fit ordering of the preference models is not likely to be drastically affected[5].

The timing of the interactions is the same as before except that the principal and each agent can now decide to separate after e_{it} is realized in the beginning of each period. Let $\overline{\pi}$ be the maximum one-period expected profit that a principal deviating to an incentive contract can obtain, where the maximization is done over the support of (Z_t, u_t) subject to the agent receiving his reservation utility in expectation. Moreover since t influences Z only through its effect on labor market experience (lme), I will assume that if the relationship with worker i is terminated after τ periods, the principal can costlessly find a worker j such that $lme_{j,t=0} = lme_{i,t=\tau+1}$, so that $Z_{j,0} = Z_{i,\tau+1}$. In that case, the principal will employ only agents with characteristics (Z_t, u_t) such that the corresponding expected profit is $\pi_t = \overline{\pi}$. Note however that the distribution of people in terms of (Z, u) employed by the deviating principal need not be degenerate. This is because production, limits to liability, and the disutility of effort are allowed to depend on Z. Hence there can be multiple values of (Z, u) which yield the same maximum profit level $\overline{\pi}$.

[5] For instance, allowing Λ to have a non-degenerate distribution will add greater flexibility to the reference-dependent model and hence improve its goodness-of-fit relative to the CARA ($\Lambda = 1$) approximation.

5.1 Attrition

Recall that the model is consistent with attrition on realized output if punishment is interpreted as termination of the current relationship. In addition, separations may occur as a result of the shock to the disutility of effort. Finally, for some measurement reason, workers may drop out of the dataset even if they continue to be employed by the firm. In order to account for these possibilities, I first estimate attrition rates $\widehat{h}_{it}$ and construct inverse probability weights, $1/(1-\widehat{h}_{it})$, which are subsequently used in estimating separations as well as the production and preference parameters.

Since it's best to impose as few a priori restrictions on the data as possible, I estimate a general attrition process. In particular, let h_{it} be the probability that individual i drops out of the dataset after tenure period t. I will model attrition with a discrete hazard model. Specifically, consider

$$h_{it} = prob(T_i = t | T_i \geq t, X_{it}), \tag{6}$$

where T_i is a discrete random variable representing the tenure at which the observation spell for worker i ends. The vector of explanatory variables X_{it} consists of: W_{it}–dummy variables for calendar month and year; individual characteristics Z_{it}–job category, a quadratic polynomial in tenure, current output, and base wage in the first period the worker is observed in the data; and the regime variable $effort$, which equals one in the incentive regime and zero otherwise, as well as interactions of $effort$ with the tenure variables and output.

Jenkins (1995) shows that when the data is properly organized, the likelihood function for such hazard models is equivalent to the likelihood of a pooled discrete response model, such as probit or logit. The properties of the distribution function depend on the distribution of u_{t+1}, the shock to the disutility of effort. Given that we have not imposed any distributional assumption on u_{t+1}, I tried both probit and logit to make sure the results are not sensitive to the specification. The results for the attrition logit are given in Appendix B. I ended up using logit instead of probit in estimating the attrition model because logit gives better goodness of fit in terms of pseudo R-squared.

5.2 Production

Assume that monthly output is given by,

$$Y_{it} \;=\; A_{it} \cdot D_{it}, \tag{7}$$

where Y_{it} is the total number of units installed, D_{it} is the working days (8 hours of labor) of individual i during data-period t, and A_{it} is a random variable denoting average daily productivity.

Furthermore, assume that $\ln A_{it}$ is $N(\mu_{it}, \sigma_{it}^2)$ with

$$\mu_{it} = \beta_e^\mu \cdot effort_{it} + Z_{it}\beta_z^\mu + W_{it}\beta_w^\mu \tag{8}$$

and

$$\ln \sigma_{it}^2 = \beta_e^\sigma \cdot effort_{it} + Z_{it}\beta_z^\sigma + W_{it}\beta_w^\sigma \tag{9}$$

where $effort_{it} \in \{0, 1\}$ denotes the effort level, 0 for L and 1 for H, and Z_{it} are the observable characteristics, which include job category, quadratic in years of tenure and months under the incentive regime, and the initial base wage/guarantee for each worker. I also include dummy variables for month and year, W_{it}. I use the initial base wage/guarantee as a measure of individual ability since by definition, it is the market's best estimate for each worker's average productivity. In addition, using wage as a measure of ability is consistent with the equilibrium selection model developed by Flinn (1997).

To estimate the parameters of the output distribution, we can then use the equation

$$\ln uad_{it} = \beta_e^{\mu} \cdot effort_{it} + Z_{it}\beta_z^{\mu} + W_{it}\beta_w^{\mu} + v_{it}, \tag{10}$$

where $\ln uad_{it}$ is the actual units-per-day for individual i during period t. If the disturbances v_{it} are not correlated with the explanatory variables, that is if

$E[v_{it} \cdot X_{it}] = 0$ for any $X_{it} \in \{effort_{it}, Z_{it}, W_{it}\}$,

the parameters β^{μ} can be consistently estimated by pooled OLS using the estimated inverse survivor rates as weights. Once this is done and we have the residuals $\widehat{v}_{it}$, we can then estimate β^{σ} by running pooled OLS on

$$\ln \widehat{u}_{it}^2 = \beta_e^{\sigma} \cdot effort_{it} + Z_{it}\beta_z^{\sigma} + W_{it}\beta_w^{\sigma} + k_{it}, \tag{11}$$

using the weighted sample, provided again that

$E[k_{it} \cdot X_{it}] = 0$ for any $X_{it} \in \{effort_{it}, Z_{it}, W_{it}\}$.

It is not obvious, however, whether the above orthogonality conditions are satisfied. On theoretical grounds, the conditions make sense. In an industry where turnover is high and output is frequently observed, any systematic productivity difference among agents should be reflected in their market wages. Table 2 provides results from estimating equations (10) and (11) when the individual controls Z_{it} are excluded and included. The purpose of this comparison is twofold. First, to see whether individual characteristics account for a significant portion of the variation in productivity. This turns out to be true as including the controls Z_{it} increases the adjusted R-squared from 0.03 to 0.29 for the mean and from 0.01 to 0.14 for the variance.

The second purpose of the comparison is to see whether the estimates of β_e^{μ} and β_e^{σ} are sensitive to omitting relevant individual characteristics. Given that it is the values of β_e^{μ}

and β_e^σ that determine the properties of the likelihood ratio distribution, if the estimates are overly sensitive to individual controls, one would be worried that pooled OLS estimation may result in omitted variable bias. However, as evident from Table 2, the hypothesis that the estimated coefficients are equal in each of the two specifications cannot be rejected at any conventional level of significance. Finally, the coefficient on effort is close in magnitude to the value of 0.197 that Lazear (2000) reports in his Table 3 using the entire sample and individual fixed effects, which further suggests that the omitted variable bias if present is not too strong.

Alternatively, one can consider a model where unobserved heterogeneity is explicitly incorporated. First differences can then be used to eliminate the unobserved effect. The parameters can thus be estimated via pooled OLS on the differenced sample. The problem with this approach is that if the goal is to calculate the likelihood ratios, the unobserved individual effects need to be estimated from the time series for each individual. This is problematic since as evident from Table 1, we have only about 11 observations per individual on average. As a result, the production process and all subsequent parameters will be subject to small sample bias. Not to mention that estimating individual fixed effects will introduce a massive amount of nuisance parameters, which may significantly influence the precision with which the rest of the model can be estimated. Thus, given the theoretical appeal of the orthogonality conditions and the relatively smaller observational requirements that pooled OLS imposes on the data, it seems reasonable to estimate the production process via pooled OLS.

Table 2 also provides summaries for the distributions of μ and σ. In particular, note the signs of the estimates for β_e^μ and β_e^σ. As expected, being on incentive pay leads to an increase in average productivity. A more interesting finding is that the variance of log average product actually declines as a result of moving from fixed wages to performance pay. Although this effect is somewhat more unexpected, it is by no means unreasonable. Intuitively, as the stakes involved in making an error increase, agents would be expected to

: Table 2: Estimates of production parameters: pooled OLS (top part) coefficients and descriptive statistics for the mean and sdev of log daily productivity (bottom part).

Control	Dependent Variable			
	ln(uad)	ln(uad)	$\ln(u^2)$	$\ln(u^2)$
Incentive Regime Dummy	0.2396 [0.0183]	0.2814 [0.0234]	−0.1818 [0.0603]	−0.3194 [0.0654]
Individual Characteristics†	YES	NO	YES	NO
Month & Year Dummies	YES	YES	YES	YES
R-squared	0.29	0.03	0.14	0.01
N	13,802	13,802	13,802	13,802

	Estimated Parameters (Incentive Regime)			
	High Effort		Low Effort	
	mu_{it}	$sigma_{it}$	mu_{it}	$sigma_{it}$
Mean	1.0410	0.2411	0.8014	0.2641
SDev	0.3672	0.1484	0.3672	0.1625
N	6,828	6,828	6,828	6,828

Notes: †Individual characteristics include job category, quadratic in tenure, and initial wage or hourly guarantee.

be more careful and allow fewer errors in the production process.

These estimates have important implications for the shape of the optimal incentive contract. The combination of an increase in average productivity with a decrease in variance suggests that the likelihood ratio is increasing with output except at very high output levels. The presence of the guarantee region thus cannot be due to non-monotonicity in the production process. If anything, the estimated production parameters suggest that there should be a cap on performance pay rather than an hourly guarantee for medium and low output levels.

Once we have estimated $(\beta^{\mu}, \beta^{\sigma})$, we can construct a discrete approximation to the distribution of $\ln A_{it}$ for each it over some common support $[\underline{y}, \overline{y}]$. If we have S total states, we can divide the support into $S-1$ cells with nodes $\{y_1, y_2, ..., y_S\}$ and compute

$$\widehat{p}_{se}(Z_{it}) = \widehat{F}_{it}(y_s) - \widehat{F}_{it}(y_{s-1}) \tag{12}$$

for $s = 1, ..., S-1$ where $\widehat{F}_{it}$ is the cdf of the distribution $N(\widehat{\mu}_{it}, \widehat{\sigma}^2_{it})$ and $\widehat{p}_{Se}(Z_{it}) = 1 - \widehat{F}_{it}(y_{S-1})$.

In estimating the preference parameters (Section 5.4.), I use the interval $[-0.75, 2.25]$ as the support for $\ln A$ and discretize it into 999 cells, so $S = 1000$. In addition, since the model does not have anything to say about the quantity of hours worked, in all subsequent calculations I normalize the total hours of work per month to 176, or 22 labor days.

When translated into actual units installed, the bounds on the productivity support correspond to 10.4 and 208.7 units per month of labor. This range covers ± 2 standard deviations around μ_{it} in both incentive regimes for over 90 percent (91.4% to be exact) of the person-months in the estimation sample. In addition, since the difference between the upper and lower bounds is roughly 200, a natural lower bound on the number of states is 200 since an increment in that case roughly corresponds to an additional unit installed during the month.

5.3 Separations and Punishment Region

Before we estimate the preference parameters, we need to determine whether there is any relation between contemporaneous output and the probability of separation. To do that, I again estimate a discrete hazard model with

$$s_{it} = \Pr(T_i = t | T_i \geq t, X_{it}), \tag{13}$$

where s_{it} is the probability that a separation occurs after tenure period t and X_{it} are observable characteristics which include the current output realization. Note that the theoretical model predicts a non-linear relationship between current output and the likelihood of separation for performance reasons. To account for this, I create 10 categorical variables which split the pooled distribution of $\ln uad$ into equal parts. Remember that we are interested in separations for performance reasons in the incentive regime, and in particular, whether the introduction of performance pay had any non-linear influence on the relationship between separations and current output. The variables of interest thus consist of the interaction between each of the output categories and the dummy variable for the incentive regime. I label these variables as $catK\_ppp$ where K is an integer between 1 and 10.

Table 3 presents the estimates from two logistic hazard models. The first regression includes all ten categories as dummy variables. It turns out that $cat2\_ppp = 1$ predicts failure perfectly, so the 34 observations where $cat2\_ppp = 1$ are dropped from the first regression. Nevertheless, during the incentive regime there is a clear non-linear relationship between current output and the probability of leaving the firm at the end of that particular period. Since the coefficients on $cat3\_ppp$ through $cat6\_ppp$ are not significantly different from each other, I estimate a second regression where I pool together all observations from categories 1 through 6. The results are shown in the right side of table 3. This is the model which I use to infer the likelihood of falling into the punishment region for each observation in the dataset.

To estimate the punishment region, let $\alpha_{it} \in \{0, 1\}$ be an index variable which is equal to 0 if a separation for incentive reasons has taken place after tenure period t and equal to 1 otherwise. We can then express the survival probability for a given relationship as

$$\begin{aligned} 1 - s_{it} &= \Pr(\alpha_{it} = 1 | \ln uad\_ppp) \cdot \\ [1 - \Pr(T_i &= t | T_i \geq t, \widetilde{X}_{it}, \alpha_{it} = 1)], \end{aligned} \tag{14}$$

where $\widetilde{X}_{it}$ includes all variables in X_{it} except the interactions between the categorical

: Table 3: The influence of current output on the probability of separation under incentive pay, estimates from discrete logistic hazard models.

Control		Dependent Variable			Control
		Separ. Prob.	Separ. Prob.		
cat1_PPP		1.3213	1.4785	*	cat1_6_PPP
		[0.8742]	[0.4840]		
cat3_PPP	*	2.5542	0.5530		cat7_PPP
		[0.8017]	[0.4370]		
cat4_PPP	*	2.3777	0.1819		cat8_PPP
		[0.6963]	[0.4428]		
cat5_PPP	*	1.6234	-0.0713		cat9_10_PPP
		[0.6789]	[0.5388]		
cat6_PPP	*	1.3527			
		[0.5052]			
cat7_PPP		0.5941			
		[0.4383]			
cat8_PPP		0.2251			
		[0.4447]			
cat9_PPP		-0.0195			
		[0.5417]			
cat10_PPP		-0.1123			
		[1.1432]			
Pseudo R-sq		0.0662	0.0651		Pseudo R-sq
N		13,161	13,195		N

Notes: Each model includes month and year dummies, job category, tenure, individual wage, and months on incentive pay as controls.
* = coefficient significant at 5% level.

output variables and the incentive dummy. The first term in the factorization is the probability of separation conditional on the current output realization, while the second term is the survival probability conditional on characteristics and the event that no incentive punishment has been implemented. Using the above relationship, we can calculate $\widehat{\alpha}_{it} = \Pr(\alpha_{it} = 1|\ln uad\_ppp)$ since we have estimates for s_{it} and $\Pr(T_i = t|T_i \geq t, \widetilde{X}_{it}, \alpha_{it} = 1)$ from the logistic hazard model.

If α_s denotes the probability of non-punishment in state s, we know that

$$\widehat{\alpha}_{it} = \sum_s \widehat{p}_{sH}(Z_{it})\alpha_s. \tag{15}$$

Now the theoretical model implies that α_s is either 0 or 1 depending on the value of the likelihood ratio. Since only the coefficient on $cat1\_6\_ppp$ has a significant influence on separations, I assume that $\alpha_s = 1$ for states such that $cat1\_6\_ppp = 0$. Hence to determine the number of states covered by $cat1\_6\_ppp$ for which $\alpha_s = 0$, I use the estimated output distribution. In particular, I set $\alpha_s = 0$ for states where $\widehat{F}_{it}(s) \leq 1 - \widehat{\alpha}_{it}$, where again $\widehat{F}_{it}$ is the cdf of $N(\widehat{\mu}_{it}, \widehat{\sigma}^2_{it})$. Otherwise, I set $\alpha_s = 1$.

The mean of the estimated separation probability for incentive reasons is 0.072. This number implies that on average an incentive termination per relationship occurs every 13.8 months. This may seem like an unusually low duration. However, the severity of the punishment will depend not only on the probability but also on the amount forfeited. In particular, if the agent forfeits only a small amount as a result of a separation, the associated expected value can be quite reasonable even if the probability of the event is relatively high.

5.4 Preference Parameters

5.4.1 GMM Estimation

A principal that maximizes profit will offer to agent i for tenure period t the contract $C_{it} = (\{\alpha_s, b_s\}, H)$ that makes the agent's participation and incentive constraints bind. The two constraints imposed by the model are:

$$\begin{aligned} u_{it} = & \\ & \sum_{s=1}^{S} (p_{sH} - p_{sL})\{\alpha_s(bonus_s + \gamma_g \cdot bonus_s^2) + \\ & (1-\alpha_s)[\Lambda(w^{\min}(Z_{it}, \psi) - guar_{it}) \\ & + \gamma_l \cdot (w^{\min}(Z_{it}, \psi) - guar_{it})^2)]\} - e(Z_{it}; \xi) \end{aligned} \tag{16}$$

and

$$
\begin{aligned}
u_{it} \;=\; & \\
& \sum_{s=1}^{S} p_{sH}\{\alpha_s(bonus_s + \gamma_g \cdot bonus_s^2) + \\
& (1-\alpha_s)[\Lambda(w^{\min}(Z_{it},\psi) - guar_{it}) + \\
& \gamma_l \cdot (w^{\min}(Z_{it},\psi) - guar_{it})^2)]\} - e(Z_{it};\xi), \qquad (17)
\end{aligned}
$$

where $bonus_s = \max\{pay_s - guar_{it}, 0\}$ is the piece rate bonus while $w^{\min}(Z_{it}, \psi)$ is the minimum allowable ex-post transfer. That is, $w^{\min}(Z_{it}, \psi) - guar_{it}$ is the amount that the agent will forfeit if the relationship is terminated for incentive reasons.

Our goal is to estimate the parameters $(\Lambda, \gamma_g, \gamma_l, \xi, \psi)$ and perform various tests on them. In order to do that, I use data on the actual contract $C_{it} = (\{bonus_s\}, guar_{it}, H)$ and observable characteristics Z_{it}–job category, tenure, and initial wage–as well as the estimates for $\{p_{sH}, p_{sL}, \alpha_s\}$ derived from the production and separation models. Provided that the shocks, u_{it}, at the true values of the parameters are orthogonal to Z_{it}, the preference parameters can be consistently estimated by two-stage GMM on the system of equations. I again weigh each observation by $1/(1 - \widehat{h}_{it})$ in order to correct for any bias induced by attrition on $(\varepsilon_{it}, u_{it})$. This produces consistent estimates since inverse probability weighting works for general M-estimation, including GMM.

Table 4 presents two-stage GMM estimates of the unrestricted model as well as two restricted models–diminishing sensitivity without loss aversion ($\gamma_g \leq 0$, $\gamma_l \geq 0$, and $\Lambda = 1$) and CARA utility ($\Lambda = 1$ and $\gamma_g = \gamma_l$). The estimates are calculated using the system of equations consisting of the incentive and participation constraints. For each person-month the likelihood ratio distribution is approximated using 1,000 states. Roughly, this means that an increment in the state space is equivalent to 0.20 additional units of output per month. Sensitivity with respect to the number of states is explored in the next subsection.

The point estimate of Λ from the unrestricted model is 2.11, which is surprisingly close

: Table 4: GMM estimates of preference parameters, unrestricted and two restricted models.

	Lam	G_g	G_l	$E(w_{it}-guar_{it})$	$E(e_{it})$	Obj. Value
1st Stage	3.04	-0.000019	-0.005351	-109.73	210.99	8.23784×10^{11}
2nd Stage Unrestricted (M1)	2.11	-0.000022	0.000000	-241.43	230.23	340,556
CARA (M2) (Lam = 1, $G_g = G_l$)	1.00	-0.000022	-0.000022	-398.72	220.05	396,868
No Loss Aversion (M3) (Lam = 1, $G_g < 0$, $G_l > 0$)	1.00	-0.000022	0.000000	-423.23	237.97	494,666

to estimates, typically ranging between 2 and 2.5, obtained from experiments. Tversky and Kahneman (1992), for instance, obtain $\Lambda = 2.25$ using laboratory experiments, while Post et al. (2007) report point estimates ranging from 2.38 in the Netherlands to 4.53 in the USA using choices from "Deal or No Deal?". Although there is no strong justification for why the estimates are so similar, especially given the different approaches and circumstances used to obtain them, it is encouraging that the orders of magnitude are comparable.

The estimates for the CARA approximation suggest a coefficient of absolute risk aversion around 0.000044. This is also roughly consistent with other evidence. For instance, Jullien and Salanie (2000) estimate a slightly negative coefficient of absolute risk aversion (corresponding to $\gamma_g = 0.03$ in our specification), while Cohen and Einav (2007) conclude that most consumers in their sample are close to risk neutral ($\gamma_g = 0$ in our case) with respect to a symmetric \$100 lottery.

Conceptually, it is not surprising that the estimated CARA approximation is close to risk neutrality, especially given Rabin's calibration theorem (2000) and the nontrivial amount of risk involved in the contract. In particular, any significantly higher degree of risk aversion will help in matching the flat region of the contract. However, it will have problems fitting both the increasing part of each contract as well as the region where separations occur for incentive reasons. In fact, to reasonably fit the observed contracts, it appears than one would need a utility function with highly variable curvature, which is essentially what loss aversion and diminishing sensitivity provide.

Table 4 also suggests that, based on the values of the objective functions, the unrestricted model fits the data in the estimation sample much better than either of the two alternatives. We have to be cautious, however, in deriving any general conclusions from this observation. In particular, the GMM estimation involves a lot of nuisance (production and punishment) parameters that influence the asymptotic distribution of the preference estimates. Hence the standard quasi-likelihood ratio test cannot be used to reject either of the two restricted models. Moreover, due to the number of parameters involved and the fact that the GMM equations are complicated non-linear functions of the nuisance parameters, calculating the correct asymptotic covariance matrix for the estimates is quite challenging. On top of that, bootstrapping is not a feasible option either since an iteration of the GMM procedure can take, depending on the optimization method involved, anywhere from several hours to several days to complete.

It thus seems that the only feasible way in which we can address the variability underlaying the estimation procedure is by exploring the sensitivity of the preference estimates to the uncertainty generated by estimating other features of the model. I do that below but more importantly, in the next section I also investigate how well the distribution of simulated optimal contracts using $M1$ and $M2$ fits to the actual data of a simulation sample. In some sense this method is preferable to calculating standard errors since it is a stronger test of the model and is also closer to what we ultimately care about–overall model fit.

5.4.2 Sensitivity of Estimates

To address any concerns associated with the uncertainty underlaying the preference estimates, I undertake two steps. First I explore the sensitivity of the preference estimates with respect to the number of states S used to approximate the output space. In addition, I investigate how the preference parameters respond to changes in the values of the nuisance parameters. In particular, I estimate two additional specifications, where the value for each nuisance parameter is set to its point estimate *plus* or *minus* two times its associated standard error.

The top two rows of table 5 show the second stage estimates from the best fitting model (in terms of objective value) for two discretizations of the state space corresponding to $S = 200$ and $S = 1,000$ (benchmark) respectively. As already mentioned, the case of $S = 200$ is of natural interest since a state increment there roughly corresponds to an additional unit of output. Furthermore, for computational reasons, in the next section I will test the preference specifications by simulating the optimal incentive pay at 200 rather than 1,000 points even though I will still be using the benchmark parameter estimates. Thus it is important to understand how comparable the two discretizations are.

As can be seen from the table, the point estimates vary somewhat between the two discretizations. Nevertheless, the variation is in a direction we would expect. First, given a fixed number of parameters, it is more difficult to fit a nonlinear function at 1,000 points than at 200 points. Thus as the grid becomes finer, the minimum value of the objective function increases. On the other hand, as the grid becomes coarser, the approximate contract deviates more and more from the nonlinear actual contract. Thus one would expect the estimated coefficient of loss aversion to decrease as the approximation worsens. Even with 200 states, however, the approximate contracts seem sufficiently nonlinear to result in a point estimate for Λ above 1.5, while the curvature parameters do not seem to be influenced at all by the size of the grid. Loss aversion and diminishing sensitivity thus seem to survive variations in the discretization of the state space.

The bottom two rows of table 5 present point estimates for the preferences parameters in the two situations where each nuisance (production and punishment) parameter is set to a value equal to its point estimate plus or minus 2 standard deviations. These specifications are quite extreme and should result in preference estimates that differ substantially from the

: Table 5. Sensitivity of preference estimates with respect to the number of output states used and values of nuisance (production and punishment) parameters.

	Lam	G_g	G_l	$E(w_{it}-guar_{it})$	$E(e_{it})$	Obj. Value
M1 with S=200	1.60	−0.000022	0.000000	−315.74	187.89	233,563
M1 with S=1,000 (benchmark)	2.11	−0.000022	0.000000	−241.43	230.23	340,556
Benchmark with nuissance pars. = point ests. + 2 SDs	1.06	−0.000004	0.000000	−192.77	287.79	169.92
= point ests. − 2 SDs	2.35	−0.000061	0.000000	−98.04	4.52	16,460

benchmark. They do allow us, however, to quantify the influence of estimation uncertainty in the production and punishment stages on the GMM estimates of the preference parameters.

The point estimates from the bottom two sensitivity estimations also go in the direction we would intuitively expect. For instance, when the value of each nuisance parameter is increased by two standard deviations, the likelihood ratio in the production process increases and so does the probability of termination due to incentive reasons. In combination, these two effects mean that a given agent will be in the flat portion of his contract with a lower probability than under the benchmark specification. Naturally, this leads to a lower estimate for the coefficient of loss aversion with the new value of 1.06 barely above zero. In addition, since this new shape is further away from full insurance than the benchmark, one would need a higher disutility of effort in order to justify the deviation. Hence the higher estimate of 287.89 for the average disutility of effort relative to the benchmark estimate of 230.23.

Finally, the estimates when two standard deviations are subtracted from the value of each nuisance parameter are the mirror image, relative to the benchmark, of the estimates when two standard deviations are added. The effect however is not symmetric. This is because compared with subtracting two standard deviations, adding two standard deviations has a much stronger effect on the probability of termination due to incentive reasons–it reduces the average value of $\widehat{\alpha}_{it}$ from 0.93 to 0.13, while adding two standard deviations increases

$\widehat{\alpha}_{it}$ on average from 0.93 to above 0.99. Hence the biggest source of parameter uncertainty appears to come from the possibility that the degree of punishment for incentive reasons may be underestimated. This, however, is unlikely to be a problem. In particular, recall that we observe separations rather than punishments. Hence if anything, the degree of punishment for incentive reasons may be overestimated. The parameter bounds shown in the bottom row of table 5 should thus be much closer to the true parameter values that those in the row above it. This is encouraging as even with a trivial probability of punishment (less than 0.01), the model produces estimates for the preference parameters that are quite reasonable given our priors.

6 Discussion on Modelling and Estimation Procedures

In this part I try to address some of the issues that can potentially be raised against the modelling and estimation decisions made in the previous sections.

All theoretical and empirical results rely on the assumption that the firm is maximizing profits both before and after the switch, which is not realistic.

It will be hard to analyze optimal contracts if the firm's objective is not well defined and profit maximization seems the most reasonable objective. However, for identification purposes, it is not required than the firm was maximizing profits prior to the change in the incentive regime. Maybe it was the case that the firm (like the entire market) was behaving sub-optimally by paying fixed wages at the going market rate even though it had the opportunity to contract on output. All conclusions we have derived will go through in this alternative situation as long as there was some optimizing reason behind the contractual change. Lazear reports that profits increased significantly after the switch to the new regime, which makes the case for profit maximization easier to swallow. It is true however that if the firm chose the new incentive contracts randomly, they would not contain any useful information about workers' preferences and the analysis in this paper would be erroneous.

The coefficient of loss aversion is identified only from the punishment decision, which is not directly observable. Hence it is not clear that what is estimated is at all loss aversion.

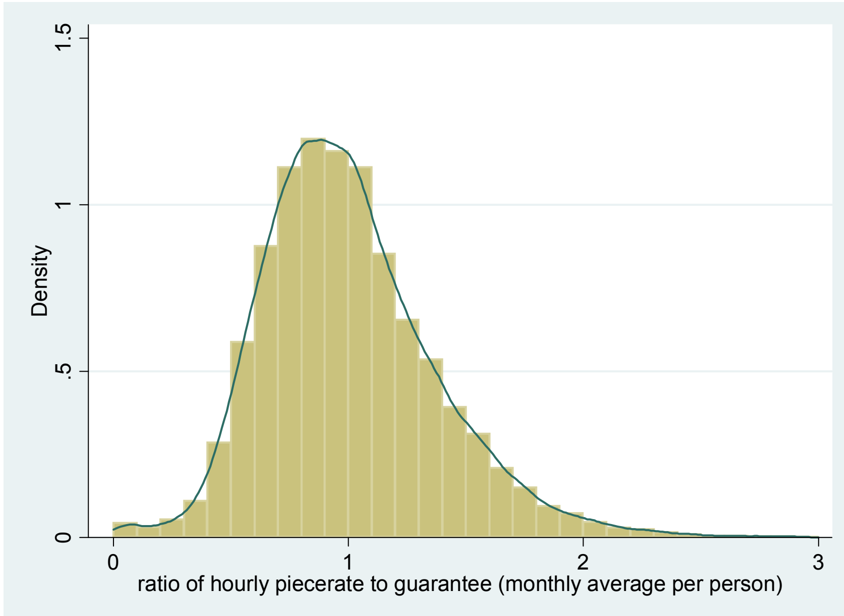
Whether any punishment will be present depends on the production and preference parameters[6]. In fact, observing no punishment implies that the coefficient of loss aversion is above a particular value. Hence even then one can estimate a lower bound on the coefficient of loss aversion. Refer for instance to the last row of table 5. There, the average punishment probability is reduced to about 0.01, which increases the estimate of the coefficient of loss aversion to 2.35. Given that the punishment probabilities in this specification are set to two standard deviations below their point estimates, even if the punishment decision were to be completely removed, the corresponding lower bound for Λ will not change drastically. If anything, allowing for punishment works against the goal of finding loss aversion even if it is present in the data.

The model assumes that effort does not vary over time within a month. It is much more likely that effort is contingent on the realization of the individual productivity shock for each month.

Consider the following alternative specification. Each worker gets a productivity shock at the beginning of every month that will last for the entire month. He observes it sooner rather than later and makes his effort decision conditional on this shock. So then if he has gotten a good shock, he knows he is productive and will be in the piece-rate region, so he works hard. On the other hand, if he has received a bad shock, he knows that he cannot reach the incentive region and works just hard enough to not get punished. This is essentially the model explored in the original paper by Lazear.

Suppose we plot the distribution of (hourly pure piece-rate pay)/(hourly guarantee). Now, if workers behave according to the above model, we should observe bunching and gaps for an interval below 1 in this distribution, or at least something like bimodality. However, figure 1 suggests nothing of this sort. The distribution is uni-modal with the mode around

[6] See the discussion below the proposition in section 3.3.



: Figure 1: *Distribution of piece-rate/guarantee ($ per hour). Bi-modality will suggest effort is chosen contingent on productivity realization.*

0.8. So it doesn't seem that people were deciding period by period whether to be in the incentive region or to stay in the guarantee region and shirk. Rather, it looks more like they were working harder after the switch than during the flat wage regime and sometimes they made it into the incentive region and sometimes they did not.

Other explanations for the chosen contract are much more likely.

I will discuss only the most prominent alternatives as in principle there are infinitely many other possibilities:

(i) The flat region will be observed in case of a *limited liability constraint*, and that's more likely than loss aversion.

Although this observation is correct, it is not at all obvious why the guarantee should be so high. Why not set the guarantee at the minimum wage? And why do most other firms that use such contracts typically set the guarantee well above the minimum wage?

(ii) The contract features are more likely due to concerns over workers' *preference for horizontal equity.*

If that were the case, why did both the guarantee and piece rate vary significantly across workers? Concerns for horizontal equity will compress the distribution of hourly guarantees. However, there is no evidence for this in the data.

(iii) The contract features are more likely to be due to concerns over *multi-tasking.*

One of the reasons this particular dataset is useful is because multi-tasking concerns are easy to address since the quality of the installation is easy to verify. In particular, any worker who defectively installed a windshield had to reinstall it on his own time. However, he was not charged for any of the wasted materials and his time was covered by the guarantee. Hence it seems that the firms was trying to insure workers against the variability of both the quantity and quality of output produced.

(iv) *The fact that most workers were given a guarantee equal to their previous wage has nothing to do with reference dependence. It was probably used to ensure worker cooperation in the transition to the new regime.*

If that were the case, one would expect the guarantee to disappear after some time. However, that did not happen and the firm is using the same type of contract to this day.

7 Testing: Goodness-of-fit Simulations

In this section I assess the goodness of fit of the unrestricted model, the approximate CARA model, and a combination of the two in determining the shape of the observed incentive contracts. In particular, I perform three simulations using the data (simulation sample) that was not used in estimating the preference parameters. For each of the three models and each of the 7,245 remaining observations, the simulation computes the optimal (cost-minimizing) contract implied by the preference parameters and subject to the underlaying incentive, participation, and monotonicity constraints.

In solving the optimal contract for each person-month, I use an approximation for the state space with $S = 200$, or about a unit of output per month as the increment. This rougher approximation is required for computational reasons. In particular, each optimal contract is the solution to a non-linear constrained optimization problem where the number of choice variables equals the number of states. Given that we have to solve such a problem a total of 21,735 times, approximations with significantly more than 200 states are not computationally feasible.

In some sense, however, the roughness of the discretization works in favor of the CARA approximation. This is because the nonlinearities associated with both the actual and CARA contracts become more pronounced as the grid becomes finer[7]. This will tend to increase the squared error associated with the CARA model in the flat region of the contract, while the unrestricted model will not really be affected. Hence, the goodness-of-fit results presented below can be considered a lower bound on the difference in the fit of the three models.

The results of the three simulations are summarized in Table 6. The first column shows the root mean squared error associated with the unrestricted model $M1$. The second model, $M1'$, uses the same parameters as $M1$ except that it imposes the ad hoc restrictions $\Lambda = 1$ and $\gamma_l = \gamma_g = -0.000022$. In this sense, the difference in the goodness-of-fit of these two simulations gives us an idea of the particular influence loss aversion and diminishing sensitivity play in determining the optimal provision of incentives, holding everything else fixed. Finally, the third column gives the goodness-of-fit associated with the CARA model $M2$.

The unrestricted model outperforms both alternatives. In particular, $M1'$ and $M2$ fit the actual data about 22 percent worse than the unrestricted model. In terms of dollars, this is an average difference of \$650 dollars per person-month–a significant amount given that the mean pay in the simulation sample is around \$2,300 per labor month (176 hours of work per month). Moreover, comparing the RMSEs for $M1'$ and $M2$ suggests that the superior fit of $M1$ is almost entirely due to the loss aversion and diminishing sensitivity parameters. In

[7] To visualize this process, start with the case where say $S = 3$.

: Table 6: Assessing goodness-of-fit with the simulation sample. Root mean squared errors for three models–unrestricted, CARA approximation, and a combination of the two.

	Unrestricted (M1)	(M1'): M1 with Lam = 1, $G_g = G_l = -0.000022$	CARA (M2)
RMSE	3,048	3,707	3,699
% worse than M1	0	22	21
N	7,245	7,245	7,245

particular, if the parameters for either $w^{\min}$ or effort played a significant role for the difference between $M1$ and $M2$, the RMSE associated with $M1'$ would have been much closer to the RMSE of the unrestricted model. Since the fits of $M1'$ and $M2$ are virtually identical, we can conclude that adding loss aversion and diminishing sensitivity significantly improves our ability to fit the observed incentive contracts.

These numbers can also be interpreted in the following ways. First, if the true population parameters are those given by $M2$, the firm would have been losing on average around \$650 per person-month by designing a contract that takes loss aversion and diminishing sensitivity into account. Alternatively, if the true population parameters are those of the unrestricted model, the firm would have incurred additional labor costs of about \$650 per person-month on average if they had designed an incentive mechanism that assumes no loss aversion. Regardless of which interpretation is used, the number just seems too large to be ignored by a competitive firm, even if the firm is not maximizing profits exactly.

8 Conclusions

Incomplete market models can help us to distinguish among different specifications of attitudes towards risk. This is because market incompleteness leads to imperfect sharing of risk with the resulting sharing rule typically being sensitive to assumptions about risk preferences. In this paper I analyzed a particular instance of this general idea. I used micro data on individual pay and output from a private company to structurally estimate a principal-agent model with moral hazard in which the agent's utility function nests approximations to various specifications of reference dependence, diminishing sensitivity, and loss aversion, as well as the standard case of CARA utility. The empirical results show that the model with loss aversion and diminishing sensitivity fits the data better than the model where preferences approximate the CARA type. This is because loss aversion and diminishing sensitivity provide significantly greater flexibility in the degree of local risk aversion. This flexibility allows us to match better the nonlinear features of observed incentive contracts.

Appendix

A Proof of Proposition 1

Let $\underline{x} = w^{\min}$. Note that the principal's optimization problem can be represented by

$$\pi(\underline{V}) = \max_{\{\alpha_s, b_s\}} \sum_{s=1}^{S} p_{sH}\{Y_s - (1-\alpha_s)w^{\min} - \alpha_s b_s\}$$

$$\begin{aligned} \text{s.t.} \quad 0 &\leq \alpha_s \leq 1 \\ b_s &\geq r, \end{aligned}$$

$$(IC): \quad \sum_{s=1}^{S}(p_{sH} - p_{sL})[\alpha_s v(b_s - r) + (1-\alpha_s)v(w^{\min} - r)] \geq e \qquad [\mu]$$

$$(PC): \quad \sum_{s=1}^{S} p_{sH}[\alpha_s v(b_s - r) + (1-\alpha_s)v(w^{\min} - r)] = \underline{V} \qquad [\lambda].$$

The above contract uses the set of lotteries between $w^{\min}$ and r to convexify the utility possibilities set. I can restrict attention to the frontier of the utility possibilities set because any point not on the frontier can be dominated either by its certainty equivalent or a lottery involving $w^{\min}$ and r. In particular, for any interior point, there exists a degenerate lottery or a lottery between $w^{\min}$ and r that gives the same expected utility at a lower expected expenditures. But then the original contract cannot be a solution to the principal's optimization problem.

This problem is a concave programming problem over the utility levels $\{v_s\}_{s=1}^{S}$ where

$$v_s = \alpha_s v(b_s - r) + (1 - \alpha_s) v(w^{\min} - r).$$

In particular, all constraints are convex while the cost of providing utility v_s, $c(v_s) = \alpha_s(b_s - w^{\min})$ is convex in v_s since it is linear in v_s for $v_s < 0$ and convex for $v_s > 0$. Hence the objective function is concave. The first order conditions are then necessary and sufficient for an optimum. They are given by:

$$[\alpha_s < 1] : \quad \lambda + \mu(1 - \frac{p_{sL}}{p_{sH}}) = \frac{w^{\min} - r}{v(w^{\min} - r)} - \xi_s,$$

where ξ_s is the multiplier on the constraint $\alpha_s \geq 0$. Note that for states where the left hand side is strictly smaller than $(w^{\min} - r)/v(w^{\min} - r)$, I must have $\xi_s > 0$, which implies $\alpha_s = 0$. That is, the agent is given the biggest possible punishment with probability *one*.

Now obviously for states where the left hand side of the condition is strictly greater than $(w^{\min} - r)/v(w^{\min} - r)$, I must have $\alpha_s = 1$. Now for $b_s > r$, the first order condition for b_s is given by:

$$[b_s > r] : \qquad \lambda + \mu(1 - \frac{p_{sL}}{p_{sH}}) = \frac{1}{v'(b_s - r)}.$$

: Table A2: Results for the logistic hazard model used to control for attrition.

Dep. Variable Attrition={0,1}	Coef.	Std. Err.
jan	0.1881	0.3020
feb	0.6476	0.2843
mar	0.8355	0.2794
apr	0.3198	0.2921
may	0.9887	0.2739
jun	1.2602	0.2677
jul	0.4000	0.2859
aug	0.9429	0.2787
sep	0.9278	0.2767
oct	0.1471	0.3131
nov	0.4909	0.2931
yr95	-0.1744	0.1497
job4015	0.8090	0.1368
job4016	2.1530	0.3288
job4017	1.7670	0.2077
job4025	0.1102	0.1820
t	-0.0094	0.0020
t_sq	0.0000	0.0000
t_ppp	-0.0402	0.0526
t_ppp_sq	0.0022	0.0039
ppp_half	0.0931	0.1840
ppp_one	0.4563	0.2121
w_0	0.0161	0.0186
uad	-0.3658	0.0488
uad_ppp	-0.0197	0.0635
_cons	-3.2929	0.3462
N = 13,802	Ps R-sq = 0.0829	

Clearly, the smallest value that the right hand side of this equation can take is given by $1/v'_{x\downarrow 0}(x)$, since $v'(x)$ is decreasing for $x > 0$. Hence for states where the left hand side is smaller than $1/v'_{x\downarrow 0}(x)$, I must have $b_s = r$. But with strong loss aversion, $v'(w^{\min} - r) > v'_{x\downarrow 0}(x)$. On the other hand, convexity for values below the reference point implies that $v(w^{\min} - r)/(w^{\min} - r) \geq v'(w^{\min} - r) > v'_{x\downarrow 0}(x)$. Hence for states where the left hand side of each first order condition is between $(w^{\min} - r)/v(w^{\min} - r)$ and $1/v'_{x\downarrow 0}(x)$, the optimal contract features $\alpha_s = 1$ and $b_s = r$. *Q.E.D.*

B Attrition Logit

References

[1] Bandiera, O., I. Barankay, and I. Rasul. (2005). Social preferences and the response to incetives: evidence from personnel data. *Quarterly Journal of Economics* 120: 917-962.

[2] Bandiera, O., I. Barankay, and I. Rasul. (forthcoming). Social connections and incentives in the workplace: evidence from personnel data. *Econometrica*

[3] Barberis, N., M. Huang, and T. Santos. (2001). Prospect theory and asset prices. *Quarterly Journal of Economics* 116: 1-53.

[4] Benartzi, S., and R. Thaler. (1995). Myopic loss aversion and the equity premium puzzle. *Quarterly Journal of Economics* 110: 73-92.

[5] Bowman, D., D. Minehart, and M. Rabin. (1999). Loss aversion in a consumption-savings model. *Journal of Economic Behavior and Organization* 38: 155-178.

[6] Camerer, C. (2001). Prospect theory in the wild: Evidence from the field. In *Choices, Values, and Frames*, D. Kahneman and A. Tversky, eds. Cambridge: Cambridge University Press.

[7] Camerer, C., L. Babcock, G. Lowenstein, and R. Thaler. (1997). Labor supply of New York city cab drivers: One day at a time. *Quarterly Journal of Economics* 111: 408-441.

[8] Chetty, R. (2006). A new method of estimating risk aversion. *American Economic Review* 96: 1821-1834.

[9] Chiappori, P-A., and B. Salanie. (2003). Testing contract theory: a survey of some recent work. In *Advances in Economics and Econometrics,* M. Dewatripont, L. P. Hansen, and S. J. Turnovsky, eds. Cambridge: Cambridge University Press.

[10] Cicchetti, C., and J. Dubin. (1994). A microeconometric analysis of risk aversion and the decision to self-insure. *Journal of Political Economy* 102: 169-186.

[11] Cohen, A., and L. Einav. (2007). Estimating risk preferences from deductible choice. *American Economic Review* 97: 745-788.

[12] Flinn, C. (1997). Equilibrium wage and dismissal processes. *Journal of Business and Economic Statistics* 15: 221-236.

[13] Ferrall, C., and B. Shearer. (1999). Incentives and transactions costs within the firm: estimating an agency model using payroll records. *Review of Economic Studie*s 66: 309–338.

[14] Genesove, D., and C. Mayer. (2001). Loss aversion and seller behavior: Evidence from the housing market. *Quarterly Journal of Economics* 116: 1233-1260.

[15] Goette, L., E. Fehr, and D. Huffman. (2004). Loss aversion and labor supply. *Journal of the European Economic Association* 2: 216-228.

[16] Heidhues, P. and B. Koszegi. (2008). Competition and price variation when consumers are loss averse. American Economics Review 98: 1245-1268.

[17] Jenkins, S. (1995). Easy estimation methods for discrete-time duration models. *Oxford Bulletin of Economics and Statistics* 57: 129-138.

[18] Jullien, B., and B. Salanie. (2000). Estimating preferences under risk: The case of racetrack bettors. *Journal of Political Economy* 108: 503–530.

[19] Kahneman, D., and A. Tversky. (1979). Prospect theory: An analysis of decision under risk. *Econometrica* 47: 263-292.

[20] Koszegi, B., and M. Rabin. (2006). A model of reference-dependent preferences. *Quarterly Journal of Economics* 121: 1133-1165.

[21] Laffont, J-J., and D. Martimort. (2002). *The Theory of Incentives: The Principal-agent Model*. Princeton: Princeton University Press.

[22] Lazear, E. (2000). Performance pay and productivity. *American Economic Review* 90: 1346-1361.

[23] MacLeod, W.B., and J. Malcomson. (1989). Implicit contracts, incentive compatibility, and involuntary unemployment. *Econometrica* 57: 447-480.

[24] Mas, A. (2006). Pay, reference points, and police performance. *Quarterly Journal of Economics* 121: 783-821.

[25] Paarsch, H., and B. Shearer. (2000). Piece rates, fixed wages, and incentive effects: statistical evidence from payroll records. *International Economic Review* 41: 59–92.

[26] Post, T., M. Van den Assem, G. Baltussen, and R. Thaler. (2007). Deal or no deal? Decision making under risk in a large-payoff game show. *American Economic Review* 98: 38-72.

[27] Rabin, M. (2000). Risk aversion and expected-utility theory: A calibration theorem. *Econometrica* 68: 1281-1292.

[28] Rayo, L., and G. Becker. (2007). Evolutionary efficiency and happiness. *Journal of Political Economy* 115: 302-337.

[29] Tversky, A., and D. Kahneman. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5: 297–323.

[30] Vera-Hernandez, M. (2003). Structural estimation of a principal-agent model: moral hazard in medical insurance. *RAND Journal of Economics* 34: 670-693.