Career Concerns: A Human Capital Perspective

Braz Camargo
São Paulo School of Economics—FGV

Elena Pastorino
University of Iowa
UAB—Barcelona GSE

February 15, 2011

Abstract

We introduce human capital accumulation, in the form of learning–by–doing, in a life–cycle model of career concerns and analyze how human capital acquisition affects implicit incentives for performance. We show that standard results from the career concerns literature can be reversed in the presence of human capital acquisition. Namely, implicit incentives need *not* decrease over time and may *decrease* with the degree of uncertainty about an individual's talent. Furthermore, increasing the precision of output measurement *can weaken* rather than strengthen implicit incentives. Overall, our results contribute to shed new light on the ability of markets to discipline moral hazard in the absence of explicit contracts linking pay to performance.

1 Introduction

It has long been recognized that reputational considerations alone can provide incentives for performance. As first discussed by Fama (1980), if an individual's ability is uncertain and his compensation reflects his reputation in the labor market, that is, the market's assessment of his talent, then a desire to influence market beliefs about ability can motivate the individual to exert effort even in the absence of any explicit contractual link between pay and performance. This dynamic incentive mechanism originating from career concerns was first formalized in Holmström (1982, 1999). Besides being central for explaining incentives for performance in firms (Baker, Jensen, and Murphy (1988) and Prendergast (1999)), career concerns are also important for understanding incentives in settings in which rewards for performance are typically informal or constrained, as is the case in the judiciary system (Posner (1993) and Levy (2005)), in the political arena (Persson and Tabellini (2002), Mattozzi and Merlo (2008), and Martinez (2009a)), and in the public sector (Dewatripont, Jewitt, and Tirole (1999a, 1999b)).

An assumption in Holmström's analysis and in the subsequent literature on career concerns is that an individual's labor input does not affect his future productivity. Yet, the quantitative importance of human capital accumulation through experience in the labor market, and its impact on wage and productivity growth, has been widely documented.[1] A prominent explanation for the accumulation of human capital in the labor market is the acquisition of new productive skills on the job while working, the so–called process of learning–by–doing.[2] The possibility for labor input to affect future output gives rise to an

[1]See, for example, Shaw (1989), Wolpin (1992), Keane and Wolpin (1997), Altuğ and Miller (1998), and the review by Rubinstein and Weiss (2007).

[2]An alternative explanation for the acquisition of human capital in the labor market is on–the–job training. The difference between on–the–job training and learning–by–doing is that the latter does not entail a tradeoff between time spent working and time devoted to the accumulation of new skills for individuals who acquire human capital while working. Due to data limitations and identification issues, few studies have attempted to distinguish between the importance of learning–by–doing and on–the–job training in explaining returns to experience. Among them, Heckman and Lochner (2005) study the impact of wage taxes and subsidies on skill formation. They interpret their findings as evidence in favor of human capital accumulation through learning–by–doing. Foster and Rosenzweig (1995) analyze the pattern of adoption and the profitability of high–yield seed varieties in India during the Green Revolution. They document the existence of learning–by–doing and of important variation in the time pattern of its returns.

additional channel through which effort influences reputation, which may affect how career concerns discipline moral hazard. In light of this, in this paper we assume that an individual's effort influences his future productivity and study how this process of human capital acquisition through learning–by–doing affects career concerns incentives.[3] We show, contrary to the case without accumulation of human capital, that career concerns incentives need not decrease with experience in the labor market and may decrease with the degree of uncertainty about an individual's ability. We also show that increasing the precision with which output is measured can have an adverse impact on implicit incentives.

Our starting point is a finite horizon version of the model in Holmström (1999) (HM) in which an individual can be either risk neutral or risk averse. In this benchmark case, an individual's effort increases with the uncertainty about his ability: the higher the uncertainty about an individual's talent, the greater his scope to manipulate the market's assessment of his ability, which in turn strengthens incentives for performance. This is the standard career concerns effect identified by Holmström. Also as in HM, an increase in output noise, by diminishing an individual's ability to influence his reputation, reduces implicit incentives. However, unlike in HM, for a given level of uncertainty about his ability, an individual's incentive to exert effort decreases with his age: since the individual is finitely lived, his gain from influencing the market's perception of his talent decreases over time. In particular, in contrast to HM, the effect of age on incentives implies that an individual's effort is strictly decreasing over time even if idiosyncratic shocks to ability keep uncertainty about ability constant over time.

We then contrast the benchmark case with the case in which an individual's productivity can improve over time through a learning–by–doing component of effort. There are two aspects to the investment in human capital that can potentially alter how age and uncertainty about ability affect career concerns incentives. The first aspect is that the return to learning–by–doing, that is, the rate at which effort increases future productivity, may change with an individual's experience in the labor market. The second aspect is that the return to

[3] Our model does not allow for a labor–leisure choice on the part of workers. Nevertheless, our setting captures an essential feature of the process of learning–by–doing, which is that an individual's labor input, whether interpreted as hours worked or effort on the job, determines how much human capital the individual accumulates in the labor market.

learning–by–doing may depend on an individual's ability. The interpretation of talent as ability to learn is discussed, for instance, in Baker, Gibbs, and Holmström (1994).

Consider first the possibility that the return to learning–by–doing changes with age. As in the benchmark case, an individual's effort, by affecting his current output, influences his reputation, and thus his wage compensation, in all subsequent periods. However, in the presence of human capital acquisition, an individual's effort also affects his future productivity, which implies an additional impact of effort on his reputation, the size of which depends on the return to learning–by–doing. It is not surprising, then, that when the return to learning–by–doing changes with an individual's age, his effort need not decrease monotonically over time even if uncertainty about ability diminishes with experience in the labor market.

Consider now the possibility that the return to learning–by–doing depends on ability. It is still true that higher uncertainty about ability gives an individual greater latitude to manipulate his reputation. However, as the impact of effort on future reputation now depends on ability, and so is uncertain, the return to an individual from exerting effort is risky, and this risk increases with the uncertainty about his ability. Thus, unlike the case without accumulation of human capital, greater uncertainty about ability can weaken rather than strengthen career concerns incentives. The flip side of this is that an increase in output noise, by making performance less informative about ability, reduces the risk from exerting effort on the job, which can have a positive effect on career concerns incentives. We derive conditions, depending on the worker's risk aversion, under which an increase in uncertainty about ability or a reduction in the precision of output have an adverse effect on career concerns incentives. Therefore, in the presence of human capital acquisition (but *not* in its absence), when individuals differ in their ability to accumulate human capital, risk aversion is central to the relationship between the strength of career concerns incentives and both the uncertainty about talent and the precision of output.

In the remainder of this section, we discuss the related literature. In Section 2, we consider the benchmark career concerns case. In Section 3, we examine the case in which accumulation of human capital through learning–by–doing is possible. We divide the analysis in two parts. In Subsection 3.1, we study the case in which the returns to learning–by–doing depend on age. In Subsection 3.2, we study the case in which the returns to learning–by–doing vary in an individual's ability. For simplicity, and to facilitate comparison with HM, we assume risk–neutrality in Section 2 and Subsection 3.1. As we discuss in Subsection 3.2, the results in the previous parts of the paper extend to the case in which workers are risk averse. Section 4 concludes and suggests directions for future research. Appendix A collects omitted proofs, Appendix B collects omitted details, and Appendix C contains supplementary material.

Related Literature Beginning with Holmström's seminal work, a growing literature has examined extensions of the basic careers concerns model. A number of papers have investigated the effect of uncertainty about ability on career concerns incentives in settings that are richer than Holmström's. For instance, Kovrijnykh (2007) shows that when ability is career specific, that is, individuals can leave the market and collect a fixed outside option, effort may decrease when uncertainty about ability increases. Martinez (2009b) obtains the same result in a model of career concerns with job assignment.[4]

Dewatripont, Jewitt, and Tirole (1999a) (DJT) compare different information structures, that is, maps from ability and labor input into observable outcomes, in terms of their impact on the strength of career concerns incentives. They derive conditions under which more precise information (in the Blackwell sense) about an individual's performance has an unambiguous effect (positive or negative) on implicit incentives.[5] The formulation in DJT is quite general and allows for learning–by–doing (see remark in p. 185). However, since they restrict attention to a two–period setting, learning–by–doing does not affect career concerns incentives. Indeed, as we will see in our analysis, accumulation of human capital can only affect implicit incentives if there are three or more periods. In particular, the nature of our result that more precise information about output can have an adverse effect on implicit incentives is quite different from the corresponding result in DJT.

Finally, in our environment, there exists a complementarity between an individual's effort

[4]See also Martinez (2006), who studies career concerns in a framework with a more general wage determination process than in HM.

[5]Other authors have examined how changes in information structure affect career concerns incentives. Jeon (1996) considers team production and analyzes how group composition and output sharing within team members affect career concerns incentives. Ortega (2003) and Blanes–i–Vidal (2007) analyze the impact of delegation on career concerns incentives. Bar–Isaac and Ganuza (2008) explore the importance of different recruitment and training strategies for career concerns incentives.

and his ability when the returns to learning–by–doing depend on ability. As in Dewatripont, Jewitt, and Tirole (1999b), this complementarity can lead to multiple equilibria. In the spirit of the literature on games with strategic complementarities (see Milgrom and Roberts (1990) and Milgrom and Shannon (1994)), our comparative statics results in Subsection 3.2 are set–valued: we consider how changes in uncertainty about ability or noise in output affect the set of equilibrium effort choices. We return to this point in our concluding remarks.

2 Benchmark Case

In this section we study the benchmark environment. The main result is Proposition 2, which describes the effect of experience in the labor market (age), uncertainty about ability, and noise in output on career concerns incentives.

2.1 Setup

We consider a risk neutral worker in a competitive labor market. Time is discrete and begins in period $t = 1$. The worker lives for $T \geq 2$ periods and has discount factor $\delta \in (0, 1]$. The worker's output in period t is

$$y_t = a_t + k_t + \varepsilon_t,$$

where $a_t \geq 0$ is his private choice of effort, k_t is his human capital, and ε_t is a noise term. The noise terms are independently and normally distributed with mean zero and precision $h_\varepsilon < \infty$ (the precision of a random variable is the inverse of its variance).

The worker's human capital in period t depends on his unknown ability $\theta_t \in \mathbb{R}$ and on his past labor inputs (effort choices). In the benchmark case,

$$k_t = k_t(a_1, \ldots, a_{t-1}, \theta_t) = \theta_t,$$

so that the worker's human capital is given by his ability only. We assume that θ_t evolves stochastically over time according to

$$\theta_{t+1} = \theta_t + \eta_t,$$

where the terms η_t are independently and normally distributed with mean zero and precision $h_\eta < \infty$. As we will see, the presence of shocks to ability allows us to separate the effects of age and uncertainty about ability on career concerns incentives. The worker and the market have a common prior belief about θ_1 that is normally distributed with mean m_1 and precision $h_1 < \infty$. We refer to the market's belief about θ_t as the worker's reputation in period t.

Explicit output–contingent contracts are not possible. In particular, the worker's pay in a period cannot be conditioned on his output in that period. Let w_t be the worker's wage in period t. The worker's payoff from a sequence $\{a_t\}_{t=1}^T$ of effort choices and a sequence $\{w_t\}_{t=1}^T$ of wage payments is

$$\sum_{t=1}^{T} \delta^{t-1}[w_t - g(a_t)],$$

where $g(a)$ is the worker's cost from exerting effort a. The function g is twice differentiable, strictly increasing, and strictly convex, with $g'(0) = 0$ and $\lim_{a\to\infty} g'(a) = \infty$.

Let Y_t be the set of output histories in period t, A_t be the set of labor input histories in period t, and $Z_t = Y_t \times A_t$ be the set of worker histories in period t. We denote a typical element of Y_t by $y^t = (y_1, \ldots, y_{t-1})$, a typical element of A_t by $a^t = (a_1, \ldots, a_{t-1})$, and a typical element of Z_t by z^t. A strategy for the worker is a sequence $\sigma = \{\sigma_t\}_{t=1}^T$, with $\sigma_t : Z_t \to \Delta(\mathbb{R}_+)$, such that $\sigma_t(z^t)$ is the worker's (mixed) choice of effort in period t if z^t is his history.[6] We say the strategy σ is uncontingent if σ_t is constant in Z_t for all $t \in \{1, \ldots, T\}$. In other words, the strategy σ is uncontingent if in every period t it prescribes an effort choice that is the same regardless of the worker's history (but may depend on t).

A wage rule is a sequence $\omega = \{\omega_t\}_{t=1}^T$, with $\omega_t : Y_t \to \mathbb{R}$, such that $\omega_t(y^t)$ is the wage the worker receives in period t if his output history is y^t. A strategy σ for the worker is sequentially rational given a wage rule ω if it maximizes the worker's lifetime payoff after every worker history. Since the market is competitive, the worker's wage w_t in period t is his expected output in that period. In other words, if the market expects the worker to follow the strategy σ and the worker's output history in period t is y^t, then $w_t = \mathbb{E}[y_t|\sigma, y^t]$.

[6]In principle, the worker could also condition his behavior on his past wages. The restriction that the worker does not do so is without loss of generality (see Appendix B for a proof). Note, as is common in the career concerns literature, that we implicitly assume that the worker's outside option is low enough that he always prefers to participate in the labor market.

Definition 1. *An equilibrium is a pair (σ^*, ω^*) such that σ^* is sequentially rational given ω^* and $\omega^*_t(y^t) = \mathbb{E}[y_t|\sigma^*, y^t]$ for all $t \geq 1$ and $y^t \in Y^t$. We say the equilibrium (σ^*, ω^*) is uncontingent if σ^* is uncontingent.*

As is standard in the career concerns literature, we consider equilibria in which the worker's strategy is pure. Note that if σ is a pure strategy, then the worker's action choice on the path of play is completely determined by his output history.[7] In what follows, we show that there exists a unique equilibrium, that this equilibrium is uncontingent, and provide a complete characterization of it.

2.2 Equilibrium Characterization

Let σ be the worker's strategy, y^t be his output history in period t, and $a_t(y^t|\sigma)$ be his (on the path of play) effort choice in period t given y^t and σ. A standard argument shows that if the worker's reputation in period t is normally distributed with mean m_t and precision h_t, and his output in period t is y_t, then his reputation in period $t+1$ is normally distributed with mean

$$m_{t+1} = m_{t+1}(y^t, y_t) = \mu_t m_t + (1-\mu_t)[y_t - a_t(y^t|\sigma)] \tag{1}$$

and precision

$$h_{t+1} = \frac{(h_t + h_\varepsilon)h_\eta}{h_t + h_\varepsilon + h_\eta}, \tag{2}$$

where

$$\mu_t = \frac{h_t}{h_t + h_\varepsilon}. \tag{3}$$

Since the prior belief about θ_1 is normally distributed, we then have that regardless of the strategy the worker follows, his reputation after every output history is normally distributed. This also holds for the worker's posterior belief about his ability—which coincides with his reputation on the path of play—after every worker history. Equation (2) implies that the evolution of the precision h_t is deterministic and independent of the worker's behavior.

[7] Let σ be a pure strategy and adopt the convention that $Z_1 = \{\emptyset\}$. The worker's effort in the first period is $\sigma_1(\emptyset) = a_1$. If the worker's output in the first period is y_1, then his effort in the second period is $\sigma_2(a_1, y_1) = a_2(y_1)$. Likewise, if the worker's output in the second period is y_2, then his effort in the third period is $\sigma_3(\emptyset, a_1, y_1, a_2(y_1), y_2) = a_3(y_1, y_2)$, and so on.

Suppose the market expects the worker to follow a strategy $\widehat{\sigma}$ that is uncontingent from period $t+1$ on, with $t \leq T-1$, and let $\widehat{a}_s$ be the worker's conjectured choice of effort in period $s \geq t+1$. The worker's wage in period $s \geq t+1$ is then given by

$$w_s(y_1, \ldots, y_{s-1}) = m_s(y_1 \ldots, y_{s-1}) + \widehat{a}_s.$$

Notice, by (1) and the market's conjecture about the worker's behavior, that

$$\begin{aligned} m_s(y_1, \ldots, y_{s-1}) = m_t(y_1, \ldots, y_{t-1}) \prod_{\tau=t}^{s-1} \mu_\tau + (1-\mu_t) \prod_{\tau=t+1}^{s-1} \mu_\tau [y_t - a_t(y^t|\widehat{\sigma})] \\ + \sum_{q=t+1}^{s-1} (1-\mu_q) \prod_{\tau=q+1}^{s-1} \mu_\tau (y_q - \widehat{a}_q), \end{aligned}$$

for all $s \geq t+1$, where we adopt the convention that $\prod_{\tau=t}^{t-1} \mu_\tau = 1$.

Consider now the worker's choice of effort a in period t when his history is $z^t = (y^t, a^t)$ and he behaves according to $\widehat{\sigma}$ from period $t+1$ on. For this, let $\mathbb{E}[\theta_t|z^t]$ be the mean of the worker's posterior belief about his ability in period t. Since, by assumption, y_q does not depend on a for all $q \geq t+1$ and

$$\mathbb{E}[y_t - a_t(y^t|\widehat{\sigma})|z^t, a_t = a] = \mathbb{E}[\theta_t|z^t] + a - a_t(y^t|\widehat{\sigma}),$$

we then have that the worker's expected wage in period $s \geq t+1$ when he chooses a in period t is

$$\mathbb{E}[w_s|z^t, a_t = a] = (1-\mu_t) \prod_{\tau=t+1}^{s-1} \mu_\tau a + \text{constant}.$$

Thus, the worker's optimal choice of effort in period t is the unique solution a_t^* to the (necessary and sufficient) first–order condition

$$\sum_{s=t+1}^{T} \delta^{s-t} \frac{\partial}{\partial a} \mathbb{E}[w_s|z^t, a_t = a] = (1-\mu_t) \sum_{s=t+1}^{T} \delta^{s-t} \prod_{\tau=t+1}^{s-1} \mu_\tau = g'(a). \tag{4}$$

Since a_t^* is independent of z^t, we can conclude that if the worker's equilibrium behavior from period $t+1$ on is uncontingent, then his equilibrium behavior in period t is uncontingent as well. Given that in any equilibrium the worker's choice of effort in period T is zero, and thus uncontingent, a straightforward induction argument shows that all equilibria are uncontingent. Moreover, since the solution to (4) is independent of the market's conjecture

$\widehat{\sigma}$, it is immediate to see that the pair (σ^*, ω^*), with σ^* such that $\sigma_t^*(z^t) \equiv a_t^*$ and ω^* such that $\omega_t^*(y^t) = \mathbb{E}[y_t|\sigma^*, y^t]$, is the unique equilibrium. Note that (4) also characterizes the worker's choice of effort in period T, in which case it reduces to $g'(a) = 0$. We have thus established the following result.

Proposition 1. *There exists a unique equilibrium, which is uncontingent. The worker's choice of effort in period $t \geq 1$ is the unique solution a_t^* to (4).*

2.3 The Effect of Age, Uncertainty about Ability, and Noise in Output on Incentives

For each period t, equation (4) defines a_t^* as a function of μ_t, and thus as a function of h_t and h_ε. Write $a_t^*(h_t, h_\varepsilon)$ to denote the worker's effort in period t as a function of h_t and h_ε. From HM, the sequence $\{\mu_t\}$ given by (3) is such that

$$\mu_{t+1} = \frac{1}{2 + \varrho - \mu_t}, \tag{5}$$

where $\varrho = h_\varepsilon / h_\eta$. Moreover, regardless of the initial precision h_1 about the worker's ability, $\{\mu_t\}$ converges monotonically to the unique steady–state μ_∞ of (5), that is, μ_t increases monotonically to μ_∞ if $\mu_1 < \mu_\infty$ and μ_t decreases monotonically to μ_∞ if $\mu_1 > \mu_\infty$. Note that $\mu_\infty \in (0, 1)$ for all $\varrho > 0$. Now let

$$b_k(\mu_t) = (1 - \mu_t) \prod_{s=1}^{k} \mu_{t+s}, \tag{6}$$

with the convention that $b_0(\mu_t) = (1 - \mu_t)$. From HM, b_k is decreasing in μ_t for all $k \geq 0$. Observe that (4) reduces to

$$g'(a) = \sum_{s=t+1}^{T} \delta^{s-t} b_{s-t-1}(\mu_t). \tag{7}$$

We can then prove the following result.

Proposition 2. *For each $t \leq T - 1$, $a_t^*(h_t, h_\varepsilon)$ is strictly decreasing in h_t and strictly increasing in h_ε. Moreover, for all $h, h_\varepsilon > 0$, $a_1^*(h, h_\varepsilon) > a_2^*(h, h_\varepsilon) > \cdots > a_T^*(h, h_\varepsilon) \equiv 0$.*

The first part of the proposition follows from the fact that b_k is decreasing in μ_t for all $k \geq 0$, and μ_t is strictly increasing in h_t and strictly decreasing in h_ε. The second part follows immediately from the equilibrium condition (7). We refer to the result that the worker's effort is strictly decreasing in the precision about his ability (except in the last period) as the *precision effect*. The intuition for the precision effect is straightforward: the higher the uncertainty about the worker's ability, the greater his scope to manipulate his reputation, and thus his future compensation, which increases incentives to exert effort. Likewise, a reduction of the noise in output also increases the impact of the worker's effort on his reputation. We refer to the result that for a given precision about his ability, the worker's effort is strictly decreasing over time as the *age effect*. The age effect follows from the fact that the worker is finitely lived, and so his gain from influencing his reputation decreases with his age.

When $\mu_1 = \mu_\infty$, the uncertainty about the worker's ability is constant over time. In this case, the time path of the worker's effort is dictated solely by the age effect, which acts to decrease the worker's labor input as he ages. Suppose now that $\mu_1 < \mu_\infty$. Since $h_1/(h_1 + h_\varepsilon)$ is strictly increasing in h_1, $\mu_1 < \mu_\infty$ implies that the uncertainty about the worker's ability decreases over time, in which case the precision effect works together with the age effect to reduce effort over time. Thus, $\mu_1 \leq \mu_\infty$ implies the worker's labor input decreases strictly over time. The same is true even when $\mu_1 > \mu_\infty$, as long as μ_1 is close enough to μ_∞. Indeed, when $\mu_1 > \mu_\infty$, the uncertainty about the worker's ability increases over time, and so the precision effect and the age effect work in opposite directions. However, the latter dominates the former if μ_1 is close enough to μ_∞. These results stand in contrast to HM, where the worker's effort is constant over time when $\mu_1 = \mu_\infty$ and strictly increasing over time when $\mu_1 > \mu_\infty$. To summarize, we have the following result.[8]

Proposition 3. *There exists $1 > \overline{\mu} > \mu_\infty$ such that if $\mu_1 \leq \overline{\mu}$, then the equilibrium choice of effort is strictly decreasing over time.*

[8]Martinez (2006) considers the same setting that we do in the benchmark case except that he allows w_t to be any increasing function of the mean of the worker's reputation. He assumes that $\mu_1 = \mu_\infty$, though.

3 Learning–by–Doing

We now consider the case in which the worker can accumulate new productive skills over time through a learning–by–doing component of effort. As discussed in the Introduction, there are two aspects of learning–by–doing that can potentially alter how career concerns incentives discipline moral hazard. The first is that the return to learning–by–doing, that is, the rate at which current effort increases future productivity, may change with the worker's age. The second is that the return to learning–by–doing may depend on the worker's ability, and so be uncertain. We consider both features of learning–by–doing in what follows.

3.1 Time–Dependent Returns

Suppose the worker's human capital in period t is

$$k_t = k_t(a_1, \ldots, a_{t-1}, \theta_t) = \theta_t + \sum_{s=1}^{t-1} \gamma_s a_s,$$

where γ_1 to γ_{T-1} are non–negative constants. Hence, the worker's effort increases his future productivity, but the rate at which it does so depends on the worker's age.[9] The definition of an equilibrium in this environment is the same as in the previous section and we again restrict attention to equilibria in which the worker follows a pure strategy.

Let σ be the worker's strategy and y^t and k_t be his output history and human capital in period t, respectively. Assume, as before, that the worker's reputation in period t is normally distributed with mean m_t and precision h_t. The same argument that leads to (1) shows that if the worker's output in period t is y_t, then his reputation in period $t+1$ is normally distributed with mean

$$m_{t+1} = m_{t+1}(y^t, y_t) = \mu_t m_t + (1-\mu_t)[y_t - a_t(y^t|\sigma) - k_t] \tag{8}$$

and precision h_{t+1} given by (2). Since the prior belief about θ_1 is normally distributed, we again have that no matter the worker's strategy, both the worker's reputation after any

[9]The results in this subsection easily extend to the case in which $k_t(a_1, \ldots, a_{t-1}, \theta_t) = \theta_t + \sum_{s=1}^{t-1} h_s(a_s)$, with the functions h_1 to h_{T-1} differentiable, strictly increasing, and (weakly) concave. We can also adapt our analysis to the case in which human capital depreciates, as in $k_t(a_1, \ldots, a_{t-1}, \theta_t) = \theta_t + \sum_{s=1}^{t-1} \lambda^{t-s-1} h_s(a_s)$, where $\lambda \in (0,1)$ is the depreciation rate of human capital, for instance.

output history and the worker's posterior belief about his ability after any worker history are normally distributed.

In order to understand the impact of learning–by–doing on career concerns incentives, consider the case in which the market expects the worker to follow an uncontingent strategy $\widehat{\sigma}$ and let $\widehat{a}_t$ be the worker's conjectured choice of effort in period t. In this case, the worker's wage in period $s \geq 2$ is

$$w_s(y_1, \ldots, y_{s-1}) = m_s(y_1, \ldots, y_{s-1}) + \widehat{a}_s + \sum_{q=1}^{s-1} \gamma_q \widehat{a}_q.$$

Equation (8) and the market's conjecture about the worker's behavior imply that

$$m_s(y_1, \ldots, y_{s-1}) = m_t(y_1, \ldots, y_{t-1}) \prod_{\tau=t}^{s-1} \mu_\tau + \sum_{q=t}^{s-1} (1-\mu_q) \prod_{\tau=q+1}^{s-1} \mu_\tau \left(y_q - \widehat{a}_q - \sum_{r=1}^{q-1} \gamma_r \widehat{a}_r \right)$$

for all s and t with $s \geq t+1$.

First note that the worker's optimal choice of effort in period T is zero no matter the market's conjecture about his behavior. Consider now the worker's choice of effort a in period $t \leq T-1$ when his history is z^t and he behaves according to the uncontingent strategy $\widehat{\sigma}$ from period $t+1$ on. Since

$$\mathbb{E}[y_q - \widehat{a}_q - \sum_{r=1}^{q-1} \gamma_r \widehat{a}_r | z^t, a_t = a] = \begin{cases} \mathbb{E}[\theta_t | z^t] + a - \widehat{a}_t & \text{if} \quad q = t \\ \\ \mathbb{E}[\theta_t | z^t] + \gamma_t (a - \widehat{a}_t) & \text{if} \quad q \geq t+1 \end{cases},$$

we then have that for all $s \geq t+1$,

$$\frac{\partial}{\partial a} \mathbb{E}[w_s | z^t, a_t = a] = \underbrace{(1-\mu_t) \prod_{\tau=t+1}^{s-1} \mu_\tau}_{\text{A}} + \underbrace{\gamma_t \sum_{q=t+1}^{s-1} (1-\mu_q) \prod_{\tau=q+1}^{s-1} \mu_\tau}_{\text{B}}. \tag{9}$$

Equation (9) shows the marginal impact of an increase in effort in period t on the worker's expected wage in all subsequent periods. The worker's effort in period t affects his future compensation in two ways. First, by affecting the worker's output in period t, effort directly influences his reputation, and thus his wage, in all future periods. This is the standard career concerns effect identified by Holmström and it corresponds to the term A in (9). The second effect is due to the learning–by–doing component of effort. Since the worker's choice of effort in period t affects his productivity from period $t+1$ on, effort has an additional impact on his output from period $t+1$ on. Therefore, the worker's choice of effort in period t has a further impact on his reputation from period $t+2$ on. This second effect, which corresponds to the term B in (9) and is only present when $t+2 \leq T$, is proportional to γ_t, the return to learning–by–doing in period t.

From (9) and (6), the worker's optimal choice of effort in period t is the unique solution a_t^* to the (necessary and sufficient) first–order condition

$$\sum_{s=t+1}^{T} \delta^{s-t} \left[b_{s-t-1}(\mu_t) + \gamma_t \sum_{q=t+1}^{s-1} b_{s-q-1}(\mu_q) \right] = g'(a). \tag{10}$$

Note that (10) also characterizes the worker's optimal choice of effort in $t = T$, in which case it reduces to $g'(a) = 0$. Since a_t^* is independent of z^t and the market's conjecture $\widehat{\sigma}$, a straightforward induction argument shows that the pair (σ^*, ω^*), where σ^* is such that $\sigma_t^*(z^t) \equiv a_t^*$ and ω^* is such that $\omega_t^*(y^t) = \mathbb{E}[y_t | \sigma^*, y^t]$, is an uncontingent equilibrium, and the unique such equilibrium. Moreover, the same argument used in Section 2 to show that all equilibria in the benchmark environment are uncontingent shows that the same is true in this case. We have thus established the following result.

Proposition 4. *There exists a unique equilibrium, which is uncontingent. The worker's choice of effort in period $t \geq 1$ is the unique solution a_t^* to* (10)*.*

Since the right side of (5) is increasing in μ_t, an increase in μ_t leads to an increase in μ_q for all $q \geq t+1$. Thus, the left side of (10) is strictly decreasing in μ_t for all $t \leq T-1$. Hence, as in the benchmark case, the precision effect holds. Moreover, also as in the benchmark case, a reduction of the noise in output increases effort in all periods but the last. However, if we keep the precision about the worker's ability constant, the left side of (10) need not diminish with the worker's age. Whether this is the case depends on how the return to learning–by–doing is affected by the worker's age. A necessary condition for the left side of (10) to not diminish with age is that the return to learning–by–doing does not decrease over time. This will be the case, for instance, if the returns from learning–by–doing are initially increasing with the worker's experience. We then have the following result.

Proposition 5. *Suppose that $\mu_1 \leq \mu_\infty$. The worker's effort need not decrease with his age. A necessary condition for this result is that γ_t is not monotonically decreasing in t.*

3.2 Uncertain Returns

Suppose now the worker's human capital in period t is

$$k_t = k_t(a_1, \ldots, a_{t-1}, \theta_t) = (\alpha_0 + a_1 + \cdots + a_{t-1})\theta_t,$$

where α_0 is a positive constant. Hence, the rate at which the worker's effort increases his future productivity is constant, but uncertain, as it depends on the worker's ability. For simplicity, we assume the worker's ability is constant, that is, $\theta_t \equiv \theta_1$ for all $t \geq 2$. We obtain the same results if the worker's ability is subject to idiosyncratic shocks.[10] The definition of an equilibrium is the same as in Section 2 and once more we restrict attention to equilibria in which the worker follows a pure strategy. For tractability, we assume the worker lives for $T = 3$ periods. We discuss this restriction at the end of the subsection.

The analysis that follows shows that the worker must be risk averse for uncertainty in the return to learning–by–doing to matter for his choice of effort. So, from now on, we assume the worker's payoff from a sequence $\{w_t\}_{t=1}^T$ of wage payments and a sequence $\{a_t\}_{t=1}^T$ of effort choices is

$$-\exp\left(-r\left\{\sum_{t=1}^{T}\delta^{t-1}[w_t - g(a_t)]\right\}\right), \tag{11}$$

where $r > 0$ is the worker's coefficient of absolute risk aversion. This specification of preferences implies the worker is indifferent between any pair of wage and effort sequences that have the same present discounted lifetime value, as if the worker had access to perfect capital markets and so could smooth his consumption over time completely. This allows us to focus on the role of uncertainty about ability on career concerns incentives.

It is simple to adapt the analysis in Section 2 and in Subsection 3.1 to the case in which the worker's preferences are given by (11). In both instances, since there is no uncertainty

[10] The purpose of assuming shocks to ability in Section 2 was to allow a clear separation between the effects of experience in the labor market and uncertainty about ability on career concerns incentives. The aim now is to examine how uncertainty about ability affects career concern incentives when the return to learning–by–doing depends on ability. As such, there is no qualitative gain in assuming shocks to ability.

about the impact of effort on output (current or future), the equilibrium characterization is the same as when the worker is risk neutral. In other words, Propositions 1 and 4, and thus Propositions 2, 3, and 5, are still valid when the worker has preferences given by (11).

Let $\xi(a) = (1+a)^{-1}g(a)$. In the analysis that follows we assume that ξ is nondecreasing with $\lim_{a\to\infty}\xi(a) = \infty$. The restriction that ξ is nondecreasing is satisfied if g' is convex.[11] A sufficient condition for $\lim_{a\to\infty}\xi(a) = \infty$ is that $\lim_{a\to\infty} g''(a) = \infty$.

We begin the analysis of the case in which the returns from learning–by–doing depend on the worker's ability with a preliminary discussion. We then proceed to the main analysis. We conclude with a discussion of the assumption that $T = 3$.

3.2.1 Preliminary Discussion

As in Subsection 3.1, the worker's effort affects his future compensation in two ways. First, by affecting his current output, effort influences the worker's reputation in all subsequent periods. Second, by influencing his productivity in all future periods, effort has an additional impact on the worker's reputation two periods into the future and on. Therefore, since $T = 3$, uncertainty about the return to learning–by–doing can only matter for the worker's behavior in period one.

In order to understand the impact of the learning–by–doing component of effort on the worker's behavior in period one, it is useful to first think about the case in which the worker's effort does not affect his current output. In this case, his behavior in the first period is completely determined by the learning–by–doing component of effort. Since output is noisy, given a conjecture about the worker's effort in period one, the greater the uncertainty about his ability, the more a good performance in period two is interpreted as evidence that the worker is of high ability. This is the same channel through which the precision effect works in the benchmark case. However, the impact of effort in period one on the worker's reputation in period three depends on his ability, and thus is uncertain. This uncertainty can dampen the worker's implicit incentives for performance as he is risk averse.

To see the negative impact of uncertainty about ability on career concerns incentives,

[11] Since $g'(0) = 0$, g' convex implies that $g''(a) \geq g'(a)/a$ for all a, from which the desired result follows.

consider as an example the extreme case in which there is no noise in the worker's output, so that the channel through which the precision effect works is not present. For this case to be of interest, we need to assume that $\alpha_0 = 0$, otherwise the worker's output in period one fully reveals his ability, and the worker has no incentive to exert effort (as this effort has no impact on his reputation in period three). Suppose then that

$$y_t = (a_1 + \cdots + a_{t-1})\theta_1,$$

with the convention that $a_1 + \cdots + a_{t-1} = 0$ if $t = 1$. We solve for the equilibrium by backward induction.

First note that the worker's equilibrium choice of effort in period three is zero. Consider now the worker's choice of effort in the second period. Since y_2 does not depend on a_2 by assumption, the worker's effort in period two does not affect his wage in period three, which can only depend on y_1 and y_2. Thus, the workers' equilibrium choice of effort in period two is also zero. Finally, consider the worker's choice of effort in the first period. For this, let the market's conjecture about the worker's effort in this period be $\widehat{a}_1 > 0$, which must be correct in equilibrium. Given that y_1 is independent of θ_1, the worker's reputation in period two is the same as in period one, which implies that $w_2 = \widehat{a}_1 m_1$. Given that there is no noise in output, if the worker produces y_2 in period two, the market's belief in period three is that $\theta_1 = y_2/\widehat{a}_1$. Thus, $w_3 = \widehat{a}_1(y_2/\widehat{a}_1) = y_2$. Thus implies that the worker's choice of effort in period one is the value of a_1 that maximizes

$$U(a_1) = -\mathbb{E}\left\{\exp\left(-r\left[\delta^2 w_3 - g(a_1)\right]\right)\right\} = -\exp\left\{-r\left[\delta^2 a_1 m_1 - g(a_1) - \frac{1}{2} r\delta^4 a_1^2 \sigma_1^2\right]\right\},$$

where $\sigma_1^2 = 1/h_1 > 0$ is the variance of the prior about θ_1.[12] Since g is strictly convex, U is strictly quasi–concave. Therefore, the optimal choice of a_1 (assuming it is interior) is the unique solution to the first–order condition

$$\delta^2 m_1 - ra\delta^4\sigma_1^2 = g'(a). \tag{12}$$

The left side of (12) has a clear interpretation. The term $\delta^2 m_1$ is the worker's marginal

[12] In the derivation of $U(a_1)$ we use the fact that if a random variable x is normally distributed with mean m and variance σ^2, then $\mathbb{E}[\exp(-rx)] = \exp(-rm + r^2\sigma^2/2)$.

benefit (in monetary terms) from effort in period one, while the term

$$MC_{\text{risk}}(a, r, \sigma_1^2, \delta) = ra\delta^4\sigma_1^2$$

is the worker's marginal cost of effort a in period one due to his risk aversion. Since $g'(a)$ is strictly increasing with $g'(0) = 0$ and $\lim_{a\to\infty} g'(a) = \infty$, it is immediate to see that (12) has a positive solution if, and only if, $m_1 > 0$, and that this solution is independent of $\widehat{a}_1$. We have thus established the following: there exists a unique equilibrium in which the worker exerts effort in period one if, and only if, $m_1 > 0$. In this equilibrium, the worker's effort in period one is the solution a_1^* to (12). The fact that $m_1 > 0$ is necessary for the worker to exert effort in the first period is intuitive: the worker is only willing to exert effort if the expected return to doing so is positive. That $m_1 > 0$ is also sufficient for the worker to exert effort in the first period follows from the fact that $MC_{\text{risk}}(0, r, \sigma_1^2, \delta) \equiv 0$.

Let then $m_1 > 0$. Since the left side of (12) is strictly decreasing in σ_1^2, it is immediate to see that a_1^* is strictly decreasing in σ_1^2. The intuition for this result is straightforward. Suppose the market expects the worker to exert effort in period one. The benefit to the worker from exerting effort in this period is that since m_1 is positive, effort increases y_2 on average, leading to a higher expected wage in period three. However, the return to exerting effort in the first period is also proportional to the worker's ability. Hence, the greater σ_1^2, the greater the variance in the return to a given choice of effort. Given that the worker is risk averse, an increase in σ_1^2 reduces his willingness to exert effort.[13]

3.2.2 Main Analysis

As discussed in the first paragraph of Subsection 3.2.1, uncertainty in the return to learning–by–doing only matters for the worker's choice of effort in the first period. Given this, in what follows we assume that

$$y_t = \vartheta_t a_t + k_t + \varepsilon_t,$$

[13]A natural concern is the extent to which the result that higher uncertainty about ability leads to lower effort in period one is driven by the fact that θ_1 can be negative (which implies a negative return to effort). In this regard, note that the result that a_1^* is decreasing in σ_1^2 holds for all $m_1 > 0$. In particular, this result holds when m_1 is large enough for the worker to place a negligible probability on the event that $\theta_1 < 0$.

where $\vartheta_1 = 1$ and $\vartheta_2 = \vartheta_3 = 0$. Thus, effort only affects current output in the first period. For simplicity, we also assume that $\alpha_0 = 1$ and set $\delta = 1$. The assumption that $\vartheta_3 = 0$ is without loss of generality, since the worker's (equilibrium) effort in period three is always zero regardless of ϑ_3. The assumption that $\vartheta_2 = 0$ does not substantively alter our results, but it simplifies the analysis. Indeed, it is immediate to see that effort in period two is always zero (and so uncontingent) when $\vartheta_2 = 0$. Instead, when $\vartheta_2 > 0$ effort in period two cannot be determined independently of the choice of effort in period one. Intuitively, when $\vartheta_2 > 0$ effort in period one affects the worker's productivity in period two, and thus his return to effort in the latter period. Under the assumption that g' is convex, when $\vartheta_2 > 0$ effort in period two is uncontingent and the analysis is similar to the analysis when $\vartheta_2 = 0$. We discuss the case in which $\vartheta_2 > 0$ in the supplementary appendix (Appendix C).

As discussed in the previous paragraph, effort in periods two and three is always zero. Once more, let $\widehat{a}_1$ be the market's conjecture about the worker's effort in period one. Note that we do not require $\widehat{a}_1$ to be greater than zero. Moreover, let $\sigma_2^2 = \sigma_1^2\sigma_\varepsilon^2/(\sigma_1^2 + \sigma_\varepsilon^2)$, where $\sigma_\varepsilon^2 = 1/h_\varepsilon > 0$ is the variance of the noise terms. The worker's wage in period two is $w_2 = (1 + \widehat{a}_1)\mathbb{E}[\theta|y_1]$, where

$$\mathbb{E}[\theta|y_1] = m_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_\varepsilon^2}\left(y_1 - \widehat{a}_1 - m_1\right)$$

by a standard argument. The worker's wage in period three is $w_3 = (1 + \widehat{a}_1)\mathbb{E}[\theta|y_2, y_1]$, where, also by a standard argument, we have that

$$\begin{aligned}
\mathbb{E}[\theta|y_2, y_1] &= \mathbb{E}[\theta|y_1] + \frac{(1+\widehat{a}_1)\sigma_2^2}{(1+\widehat{a}_1)^2\sigma_2^2 + \sigma_\varepsilon^2}\left\{y_2 - (1+\widehat{a}_1)\mathbb{E}[\theta|y_1]\right\} \\
&= \frac{\sigma_1^2}{[1+(1+\widehat{a}_1)^2]\sigma_1^2 + \sigma_\varepsilon^2}y_1 + \frac{(1+\widehat{a}_1)\sigma_1^2}{[1+(1+\widehat{a}_1)^2]\sigma_1^2 + \sigma_\varepsilon^2}y_2 \; + \; \text{constant.}
\end{aligned}$$

The same argument that leads to the first–order condition (12) in the preliminary discussion shows that the worker's optimal choice of effort in period one is the unique solution to the first–order condition

$$\begin{aligned}
&\frac{(1+\widehat{a}_1)\sigma_1^2}{\sigma_1^2 + \sigma_\varepsilon^2} + \frac{(1+\widehat{a}_1)\sigma_1^2}{[1+(1+\widehat{a}_1)^2]\sigma_1^2 + \sigma_\varepsilon^2} + \frac{(1+\widehat{a}_1)^2\sigma_1^2 m_1}{[1+(1+\widehat{a}_1)^2]\sigma_1^2 + \sigma_\varepsilon^2} \\
&\qquad -r\left\{\frac{(1+\widehat{a}_1)^2\sigma_1^2}{[1+(1+\widehat{a}_1)^2]\sigma_1^2 + \sigma_\varepsilon^2}\right\}^2(1+a)\sigma_1^2 + \lambda = g'(a), \qquad (13)
\end{aligned}$$

where $\lambda \geq 0$ and $\lambda = 0$ if the solution is positive; λ is the multiplier associated with the constraint that effort is non–negative. As with (12), the left side of (13) has a clear interpretation. The term

$$MB(\widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, m_1) = \frac{(1+\widehat{a}_1)\sigma_1^2}{\sigma_1^2 + \sigma_\varepsilon^2} + \frac{(1+\widehat{a}_1)\sigma_1^2}{[1+(1+\widehat{a}_1)^2]\sigma_1^2 + \sigma_\varepsilon^2} + \frac{(1+\widehat{a}_1)^2\sigma_1^2 m_1}{[1+(1+\widehat{a}_1)^2]\sigma_1^2 + \sigma_\varepsilon^2} \tag{14}$$

is the worker's marginal benefit from effort in period one, while the term

$$MC_{\text{risk}}(a, \widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) = r \left\{ \frac{(1+\widehat{a}_1)^2\sigma_1^2}{[1+(1+\widehat{a}_1)^2]\sigma_1^2 + \sigma_\varepsilon^2} \right\}^2 (1+a)\sigma_1^2 \tag{15}$$

is the worker's marginal cost of effort a in the same period due to his risk aversion. Straightforward algebra shows that $MB(\widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, m_1)$ is strictly increasing in $\widehat{a}_1$.

Let $\alpha_1 = \alpha_1(\widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)$ denote the solution to (13). The possible values for the worker's effort in period one, a_1^*, are the fixed points of the map $\widehat{a}_1 \mapsto \alpha_1(\widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)$. In other words, the possible values of a_1^* are the solutions to

$$\begin{aligned} H(a, \lambda, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) = g'(a) - \lambda - \frac{(1+a)\sigma_1^2}{(\sigma_1^2 + \sigma_\varepsilon^2)} - \frac{(1+a)\sigma_1^2}{[1+(1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2} \\ - \frac{(1+a)^2\sigma_1^2}{[1+(1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2} \left\{ m_1 - \frac{r(1+a)^3\sigma_1^4}{[1+(1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2} \right\} = 0, \end{aligned} \tag{16}$$

where $\lambda \geq 0$ and $\lambda = 0$ if $a_1^* > 0$. Since

$$H(0, \lambda, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) = H(0, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) - \lambda,$$

it is immediate to see that $a_1^* = 0$ is a solution to (16) if $H(0, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) \geq 0$. Suppose now that $H(0, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) < 0$. Given that H is continuous in a and $\lim_{a\to\infty} \xi(a) = \infty$ implies that $\lim_{a\to\infty} H(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) = \infty$, equation (16) has a positive solution by the intermediate value theorem. Thus, an equilibrium always exists.

Unlike the example in the preliminary discussion, equilibria may not be unique, though. The reason for this is that $MB(\widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)$ is strictly increasing in $\widehat{a}_1$. The complementarity between the worker's marginal benefit of effort and the market's expectation about the worker's effort can lead to multiple equilibria, as it can imply that $\alpha_1(\widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)$ is nondecreasing in $\widehat{a}_1$.[14] When this is the case, the map $\widehat{a}_1 \mapsto \alpha_1(\widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)$ may have multiple fixed points.

[14] Indeed, if $\alpha_1(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)$ is interior, the implicit function theorem implies that

$$\frac{\partial \alpha_1}{\partial \widehat{a}_1}(a) = \frac{\partial MB(\alpha_1(a))/\partial \widehat{a}_1 - \partial MC_{\text{risk}}(\alpha_1(a))/\partial \widehat{a}_1}{g''(\alpha_1(a)) + \partial MC_{\text{risk}}(\alpha_1(a))/\partial a},$$

For each $\chi \in \mathbb{R}^4_{++}$, let

$$A_1(\chi) = \{a \in \mathbb{R}_+ : \exists \lambda \geq 0 \text{ with } H(a, \lambda, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) = 0 \text{ and } \lambda a = 0\}$$

be the set of possible effort choices for the worker in the first period. Since H is continuous in a, $A_1(\chi)$ is closed for all $\chi \in \mathbb{R}^4_{++}$. Moreover, since

$$H(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) \geq g'(a) - (1 + a) - (1 + m_1),$$

we have that for each $\chi \in \mathbb{R}^4_{++}$, any solution to (16) is bounded above by the only solution $\overline{a}_1 = \overline{a}_1(m_1)$ to $g'(a) = 2 + a + m_1$.[15] Hence, for all $\chi \in \mathbb{R}^4_{++}$, $A_1(\chi)$ is compact, and thus has a smallest and a greatest element. Denote the smallest and greatest elements of $A_1(\chi)$ by $a^*_{1,\min} = a^*_{1,\min}(\chi)$ and $a^*_{1,\max} = a^*_{1,\max}(\chi)$, respectively. Standard arguments from general equilibrium theory (see Mas–Colell, Whinston, and Green (1995)) can be used to prove that the set of $\chi \in \mathbb{R}^4_{++}$ for which $A_1(\chi)$ has a finite number of elements is generic (open and full measure). For completeness, we include a proof of this result in Appendix B. We have then established the following result.

Proposition 6. *An equilibrium always exists. Moreover, there exists an open and full measure subset S of $\mathbb{R}^4_{++}$ such that for all $\chi \in S$, the number of equilibria is finite.*

Next, we establish necessary and sufficient conditions on σ_1^2 for any solution to (16) to be positive. For this, let

$$h(\sigma_1^2, \sigma_\varepsilon^2, m_1, r) = 2 + m_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_\varepsilon^2} - \frac{r\sigma_1^4}{2\sigma_1^2 + \sigma_\varepsilon^2}.$$

It is possible to show that if $2 + m_1 > 0$, then for each σ_ε^2 and r there exists a unique and positive cutoff $\overline{\Sigma}_1^2 = \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r)$ such that $h(\sigma_1^2, \sigma_\varepsilon^2, m_1, r)$ is greater than zero if, and only if, $\sigma_1^2 \in (0, \overline{\Sigma}_1^2)$. Moreover, $\overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r)$ is strictly increasing in σ_ε^2, $\lim_{\sigma_\varepsilon^2 \to \infty} \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r) = \infty$, $\lim_{r \to 0} \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r) = \infty$, and $\lim_{\sigma_\varepsilon^2 \to 0} \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r) = 2(3 + m_1)/r$. See Appendix B for the details. These properties of $\overline{\Sigma}_1^2$ will prove useful below. We then have the following result, the proof of which is in Appendix A.

where we omit the dependence of α_1 on σ_1^2, σ_ε^2, m_1, and r for ease of notation. Note that the denominator of $\partial\alpha_1(a)/\partial\widehat{a}_1$, is positive. Since $\partial MB(\alpha_1(a))/\partial\widehat{a}_1 > 0$, we then have that α_1 may be nondecreasing in $\widehat{a}_1$.

[15] The uniqueness of $\overline{a}_1$ follows from the assumption that $\xi(a)$ is nondecreasing.

Proposition 7. *Suppose that $2 + m_1 > 0$. The worker's effort in period one is positive if, and only if, $\sigma_1^2 \in (0, \overline{\Sigma}_1^2)$.*

From now on we assume that $2 + m_1 > 0$. Note that $\sigma_1^2 < \overline{\Sigma}_1^2$ if, and only if,

$$r\sigma_1^2 \cdot \frac{\sigma_1^2}{2\sigma_1^2 + \sigma_\varepsilon^2} < 2 + m_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_\varepsilon^2}.$$

So, in order for a_1^* to be positive, we need either m_1 large relative to $r\sigma_1^2$ or $\sigma_\varepsilon^2/\sigma_1^2$ large. In the first case, the benefit of effort in period one is substantial. In the other case, output is not too informative about ability, which constrains the variation in the worker's period three wage. This makes the marginal cost of effort due to risk aversion small.

Comparative statics analysis in the presence of multiple equilibria may be problematic. Nevertheless, in what follows we are able to establish comparative statics results that apply to the entire set of period one effort choices. We begin with some definitions taken from the monotone comparative statics literature. Let S_1 and S_2 be two subsets of the real line. We say that S_1 is *smaller* than S_2, and write $S_1 \leq S_2$, if for all $y_1 \in S_1$ there exists $y_2 \in S_2$ such that $y_1 \leq y_2$, and for all $y_2 \in S_2$ there exists $y_1 \in S_1$ such that $y_2 \geq y_1$. We say that S_1 is *strictly smaller* than S_2, and write $S_1 < S_2$, if S_1 is smaller than S_2, there exists $y_2 \in S_2$ such that $y_2 > y_1$ for all $y_1 \in S_1$, and there exists $y_1 \in S_1$ such that $y_1 < y_2$ for all $y_2 \in S_2$. Now let $\Gamma : I \rightrightarrows \mathbb{R}$, where $I \subseteq \mathbb{R}$ is an interval, be a correspondence. We say that Γ is increasing if $\Gamma(x_1) \leq \Gamma(x_2)$ for all $x_1 < x_2$ in I, and that Γ is strictly increasing if $\Gamma(x_1) < \Gamma(x_2)$ for all $x_1 < x_2$ in I. Decreasing and strictly decreasing correspondences are defined similarly. Suppose Γ is compact–valued and for each $x \in I$, let $\gamma_{\min}(x)$ and $\gamma_{\max}(x)$ be the smallest and greatest elements of $\Gamma(x)$, respectively. A necessary and sufficient condition for Γ to be strictly increasing is that both $\gamma_{\min}$ and $\gamma_{\max}$ are strictly increasing in x. Likewise, Γ is strictly decreasing if, and only if, both $\gamma_{\min}$ and $\gamma_{\max}$ are strictly decreasing in x.

Define $\mathcal{A}_1 : \mathbb{R}_{++}^4 \rightrightarrows \mathbb{R}_+$ to be such that $\mathcal{A}_1(\chi) = A_1(\chi)$. By construction, $\mathcal{A}_1$ is the correspondence that maps the set of parameters of the model into the set of equilibrium effort choices in period one. From above, $\mathcal{A}_1$ is strictly increasing (decreasing) in σ_1^2 if both $a_{1,\min}^*$ and $a_{1,\max}^*$ are strictly increasing (decreasing) in σ_1^2. Since $\lim_{a \to \infty} H(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) = \infty$ and $A_1(\chi)$ is bounded above by $\overline{a}_1 = \overline{a}_1(m_1)$, we then have that $\mathcal{A}_1$ is strictly increasing (decreasing) in σ_1^2 over some interval I if $\sigma_1^2 \in I$ implies that $\sigma_1^2 < \overline{\Sigma}_1^2$ and $H(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)$

is strictly decreasing (increasing) in σ_1^2 for all $a \leq \overline{a}_1$. Indeed,

$$H(0,0,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) = -\frac{\sigma_1^2}{2\sigma_1^2+\sigma_\varepsilon^2}h(\sigma_1^2,\sigma_\varepsilon^2,m_1,r),$$

and so $\sigma_1^2 < \overline{\Sigma}_1^2$ implies that $H(0,0,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) < 0$. In this case, $a^*_{1,\min} > 0$ and $a^*_{1,\max} > 0$ are, respectively, the smallest and greatest solutions of $H(a,0,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) = 0$. Hence, if an increase in σ_1^2 shifts H up (down) when $a \in [0,\overline{a}_1]$, then this increase reduces both $a^*_{1,\min}$ and $a^*_{1,\max}$. Note that a change in σ_1^2 does not change $\overline{a}_1$. Analogous comparative statics results hold when $H(a,0,\sigma_1^2,\sigma_\varepsilon^2,m_1,r)$ is strictly monotonic in σ_ε^2 or r.

The first comparative statics results we derive are about the dependence of $\mathcal{A}_1$ on the uncertainty about the worker's ability. There are two opposing effects when one increases σ_1^2. The first effect, which is the force behind the precision effect, is that an increase in the uncertainty about the worker's ability increases his return to exerting effort (in period one). Indeed, from (14), the worker's marginal benefit of effort is strictly increasing in σ_1^2. The second effect is that an increase in the uncertainty about the worker's ability increases the variance in the return to learning–by–doing, which dampens incentives for effort. In fact, from (15), the worker's marginal cost of effort due to risk aversion is also increasing in σ_1^2. When r is small, the positive effect dominates the negative one, and we are as in the benchmark case (except that now there may be multiple equilibria). More precisely, we have the following result.

Proposition 8. *Fix σ_ε^2 and m_1. For all $\Sigma_1^2 > 0$, there exists $\overline{r} > 0$ such that if $r \leq \overline{r}$, then $\mathcal{A}_1$ is strictly increasing in σ_1^2 for all $\sigma_1^2 \in (0,\Sigma_1^2)$.*

The sketch of the proof of Proposition 8 is as follows. The omitted details are in Appendix A. We know from above that both $MB(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2,m_1)$ and $MC_{\text{risk}}(a,\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2,r)$ are strictly increasing in σ_1^2. However, for given Σ_1^2, $MB(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2,m_1)$ increases in σ_1^2 at a faster rate than $MC_{\text{risk}}(a,\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2,r)$ if the worker's degree of risk aversion is small enough, as long as σ_1^2 is smaller than Σ_1^2 and a is smaller than $\overline{a}_1$. Since

$$H(a,0,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) = g'(a) - MB(a,\sigma_1^2,\sigma_\varepsilon^2,m_1) + MC_{\text{risk}}(a,a,\sigma_1^2,\sigma_\varepsilon^2,r),$$

we can then conclude that for each Σ_1^2, there exists $\overline{r} > 0$ such that $H(a,0,\sigma_1^2,\sigma_\varepsilon^2,m_1,r)$ is strictly increasing in $\sigma_1^2 \in (0, \Sigma_1^2]$ for all $a \leq \overline{a}_1$ if $r \leq \overline{r}$. Since $\lim_{r\to 0} \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r) = \infty$, we can choose $\overline{r}$ to be such that $\Sigma_1^2 < \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, \overline{r})$, so that $\Sigma_1^2 < \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r)$ for all $r \leq \overline{r}$.

The next comparative statics result considers the effect of an increase in σ_1^2 in two different cases: when σ_1^2 is small and when σ_1^2 is close to $\overline{\Sigma}_1^2$.

Proposition 9. *Fix σ_ε^2, m_1, and r. There exist $0 < \Sigma_{10}^2 < \Sigma_{11}^2 < \overline{\Sigma}_1^2$ such that $\mathcal{A}_1$ is strictly increasing in σ_1^2 if $\sigma_1^2 \in (0, \Sigma_{10}^2)$, and a_1^* is unique and strictly decreasing if $\sigma_1^2 \in (\Sigma_{11}^2, \overline{\Sigma}_1^2)$.*

Here we just present a sketch of the proof of Proposition 9. The omitted details are in Appendix A. Fix σ_ε^2, m_1, and r. First observe that while $MB(\widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, m_1)$ is proportional to σ_1^2, $MC_{\text{risk}}(a, \widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, r)$ is proportional to σ_1^6. Hence, when σ_1^2 is small, the rate at which $MB(\widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, m_1)$ increases in σ_1^2 is greater than the rate at which $MC_{\text{risk}}(a, \widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, r)$ increases in σ_1^2. As in the proof of Proposition 8, this implies that $\mathcal{A}_1$ is strictly increasing in σ_1^2 when σ_1^2 is small. From Proposition 7, the worker's choice of effort in period one is uniquely defined when $\sigma_1^2 \geq \overline{\Sigma}_1^2$. It is intuitive then that this choice of effort is also uniquely defined when $\sigma_1^2 < \overline{\Sigma}_1^2$, as long as σ_1^2 is close enough to $\overline{\Sigma}_1^2$. Moreover, since H is continuous, it must be that as σ_1^2 increases to $\overline{\Sigma}_1^2$, the worker's effort in period one converges to the effort he exerts when $\sigma_1^2 = \overline{\Sigma}_1^2$, which is zero (uniqueness is key for this). Since the worker's effort in period one is positive when $\sigma_1^2 < \overline{\Sigma}_1^2$, it must be the case that this effort eventually becomes strictly decreasing in σ_1^2 as σ_1^2 increases to $\overline{\Sigma}_1^2$.

Proposition 9 is a local result. It tells us how the set of possible period one effort choices for the worker responds to an increase in σ_1^2 when σ_1^2 is in a neighborhood of either zero or $\overline{\Sigma}_1^2$. The next result we obtain shows that as noise in output reduces, the lower bound on σ_1^2 above which a_1^* is unique and strictly decreasing in σ_1^2 converges to zero. For this, let $\overline{\Sigma}_1^2(0, m_1, r) = \lim_{\sigma_\varepsilon^2 \to 0} \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r) = 2(3 + m_1)/r$.

Proposition 10. *For each m_1, r, and $\Sigma_1^2 \in (0, \overline{\Sigma}_1^2(0, m_1, r))$, there exists $\Sigma_\varepsilon^2 > 0$ such that if $\sigma_\varepsilon^2 < \Sigma_\varepsilon^2$, then a_1^* is unique and strictly decreasing in σ_1^2 for all $\sigma_1^2 \in (\Sigma_1^2, \overline{\Sigma}_1^2)$.*

Recall that $\overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r)$ is strictly increasing in σ_ε^2. Therefore, $\Sigma_1^2 < \overline{\Sigma}_1^2(0, m_1, r)$ implies that $\Sigma_1^2 < \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r)$ for all σ_ε^2. The proof of Proposition 10 is in Appendix A. The intuition is as follows. As discussed above, an increase in σ_1^2 has both a positive and a negative effect on the worker's incentive for effort (in the first period). The positive effect comes from the precision effect, which depends on the worker's output being noisy. In particular, the rate at which the worker's return to exerting effort increases in σ_1^2 diminishes as σ_ε^2 gets smaller. Indeed, when the ratio $\sigma_\varepsilon^2/\sigma_1^2$ is small, the worker's performance is very informative about his ability, and so an increase in σ_1^2 has a small impact on the return from effort. On the other hand, the rate at which the marginal cost of effort due to risk aversion is increasing in σ_1^2 increases as σ_ε^2 gets smaller. Formally, $\partial^2 MC_{\text{risk}}/\partial\sigma_\varepsilon^2\partial\sigma_1^2 < 0$. Thus, the negative effect of an increase in σ_1^2 dominates the positive effect when σ_ε^2 is small, which is the opposite of what happens when r is small.

We now discuss the effect of changes in σ_ε^2 on career concerns incentives. As in the benchmark case, an increase in σ_ε^2, by making output less informative about the worker's ability, reduces the worker's return to exerting effort. Indeed, $MB(\widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, m_1)$ is strictly decreasing in σ_ε^2. However, differently from the benchmark case, if output becomes less informative about ability, this reduces the variation in the worker's wage in period three. Thus, an increase in σ_ε^2 reduces the uncertainty in the return to learning–by–doing, which is good for incentives. In fact, $MC_{\text{risk}}(a, \widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, r)$ is also strictly decreasing in σ_ε^2. Consequently, an increase in the noise in output, as an increase in the uncertainty about the worker's ability, has an ambiguous effect on incentives for effort. Since $\overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r)$ is strictly increasing in σ_ε^2 and $\lim_{\sigma_\varepsilon^2\to\infty} \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r) = \infty$, an immediate consequence of Proposition 7 is that an increase in σ_ε^2 can increase effort when uncertainty about ability is high enough. In particular, if $\sigma_1^2 > \overline{\Sigma}_1^2(0, m_1, r)$, implicit incentives are *not* maximal when the noise in output is minimized. Since $\lim_{r\to\infty} \overline{\Sigma}_1^2(0, m_1, r) = 0$, the following result holds.

Proposition 11. *For each m_1 and σ_1^2, there exists $\overline{r} > 0$ such that if $r \geq \overline{r}$, then implicit incentives are not maximal when output noise is minimized.*

The intuition for the above result is simple. Since $MC_{\text{risk}}(a, \widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, r)$ is proportional to r, the rate at which it decreases in σ_ε^2 is proportional to r. Hence, the greater the worker's risk aversion, the stronger is the positive effect of an increase in output noise.

3.2.3 Longer Time Horizons

The main challenge in extending the analysis in this subsection to the case where the worker lives for $T > 3$ periods is to obtain an analytical characterization of equilibria. This is already apparent when $T = 4$. Indeed, with a few straightforward modifications, the analysis in 3.2.2. provides a complete characterization of behavior from periods two, three, and four when

$$y_t = \vartheta_t a_t + k_t + \varepsilon_t,$$

with $\vartheta_1 = \vartheta_2 = 1$ and $\vartheta_3 = \vartheta_4 = 0$. Since the worker's choice of effort in period two depends on the mean of his belief about his ability, equilibrium behavior in the second period is no longer uncontingent. This prevents us from obtaining a closed–form characterization of equilibrium behavior in period one.

Nevertheless, the same forces at play when $T = 3$ are present when $T \geq 4$. As before, there are two channels through which the worker's effort affects his reputation: (i) by affecting current output, effort influences the worker's reputation in all subsequent periods; (ii) by affecting productivity from the next period on, effort has an additional impact on the worker's reputation in all periods after the next. The first channel—the standard channel through which career concerns work—implies that an increase in uncertainty about ability (or a decrease in output noise) increases the return to exerting effort. However, the second channel implies that an increase in uncertainty about ability (or a decrease in output noise) increases the risk associated with exerting effort, which has a negative impact on implicit incentives. This suggests that the comparative statics results obtained when $T = 3$ survive in the more general case when $T \geq 4$.

4 Conclusion

In this paper we extend the standard career concerns model to allow for the possibility that an individual's human capital can change over time through a learning–by–doing component of effort. We show that learning–by–doing has substantive implications for the ability of career concerns incentives to discipline moral hazard. Namely, well–known results on the effect of experience in the labor market, uncertainty about ability, and noise in output on the strength of implicit incentives can be reversed once learning–by–doing is present. In addition, in the presence of learning–by–doing, risk aversion is central to the relationship between the power of career concerns incentives and the level of uncertainty in the environment.

Our analysis has abstracted from the possibility of explicit contracting. In related work (Camargo and Pastorino (2010)), we analyze optimal contracting in the presence of career concerns and show that human capital acquisition through learning–by–doing offers a natural and empirically plausible explanation for the conflicting evidence on the sensitivity of pay to performance for executives at different stages of their working life.

Appendix A: Omitted Proofs

Proof of Proposition 7

Fix σ_ε^2, m_1, and r. First note that $H(0, \lambda, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) = 0$ if, and only if,

$$\lambda = -\frac{\sigma_1^2}{2\sigma_1^2 + \sigma_\varepsilon^2} h(\sigma_1^2, \sigma_\varepsilon^2, m_1, r),$$

and that $\lambda \geq 0$ if, and only if, $\sigma_1^2 \geq \overline{\Sigma}_1^2$. Hence, $a_1^* = 0$ is a solution to (16) only when $\sigma_1^2 \geq \overline{\Sigma}_1^2$. Since (16) has a solution for all $\sigma_1^2 > 0$, it must then be that the worker's effort in period one is positive if $\sigma_1^2 \in (0, \overline{\Sigma}_1^2)$. We now show that (16) has no positive solution when $\sigma_1^2 \geq \overline{\Sigma}_1^2$. For this, let $\mathcal{H}(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) = (1 + a)^{-1} H(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)$. By construction, (16) has a positive solution if, and only if, the equation $\mathcal{H}(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) = 0$ has a positive solution. Since $\xi(a)$ is nondecreasing, straightforward algebra shows that

$$\frac{\partial \mathcal{H}}{\partial a}(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) > \frac{4r\sigma_1^6(1+a)^4(\sigma_1^2 + \sigma_\varepsilon^2)}{\{[1 + (1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2\}^3} - \frac{\sigma_1^2 m_1}{[1 + (1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2}.$$

Now observe that $\sigma_1^2 \geq \overline{\Sigma}_1^2$ implies that

$$\begin{aligned} 4r\sigma_1^6(1+a)^4(\sigma_1^2 + \sigma_\varepsilon^2) &\geq 4m_1\sigma_1^2(1+a)^4(\sigma_1^2 + \sigma_\varepsilon^2)(2\sigma_1^2 + \sigma_\varepsilon^2) \\ &\geq 2m_1\sigma_1^2\{[1 + (1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2\}^2. \end{aligned}$$

Hence, $\mathcal{H}$ is strictly increasing in a for all $\sigma_1^2 \geq \overline{\Sigma}_1^2$. The desired result follows from the fact that $\mathcal{H}(0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) \geq 0$ when $\sigma_1^2 \geq \overline{\Sigma}_1^2$.

Proof of Proposition 8

Fix Σ_1^2, σ_ε^2, and m_1. Now note that

$$\begin{aligned}
\frac{\partial H}{\partial \sigma_1^2}(a,0,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) &= -\frac{(1+a)\sigma_\varepsilon^2}{(\sigma_1^2+\sigma_\varepsilon^2)^2} - \frac{(1+a)\sigma_\varepsilon^2[1+(1+a)m_1]}{\{[1+(1+a)^2]\sigma_1^2+\sigma_\varepsilon^2\}^2} \\
&\quad + \frac{r(1+a)^5\sigma_1^4}{\{[1+(1+a)^2]\sigma_1^2+\sigma_\varepsilon^2\}^2}\{[1+(1+a)^2]\sigma_1^2+3\sigma_\varepsilon^2\} \\
&\leq -\frac{(1+m_1)\sigma_\varepsilon^2}{(2\Sigma_1^2+\sigma_\varepsilon^2)^2} + \frac{3r(1+\overline{a}_1)^5\Sigma_1^4}{\{[1+(1+\overline{a}_1)^2]\Sigma_1^2+\sigma_\varepsilon^2\}^2},
\end{aligned}$$

and so there exists $\overline{r}>0$ such that $r \leq \overline{r}$ implies that $\partial H(a,0,\sigma_1^2,\sigma_\varepsilon^2,m_1,r)/\partial\sigma_1^2 < 0$ for all $(\sigma_1^2, a) \in (0,\Sigma_1^2] \times [0,\overline{a}_1]$. The desired result follows from the argument in the main text.

Proof of Proposition 9

Fix m_1, r, and σ_ε^2. Note, from the proof of Proposition 8, that

$$\frac{\partial H}{\partial \sigma_1^2}(a,0,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) \leq -\frac{(1+m_1)\sigma_\varepsilon^2}{(2\sigma_1^2+\sigma_\varepsilon^2)^2} + \frac{3r(1+\overline{a}_1)^5\sigma_1^4}{\{[1+(1+\overline{a}_1)^2]\sigma_1^2+\sigma_\varepsilon^2\}^2}.$$

Thus, there exists $\Sigma_{10}^2 \in (0,\overline{\Sigma}_1^2)$ such that $\partial H(a,0,\sigma_1^2,\sigma_\varepsilon^2,m_1,r)/\partial\sigma_1^2 < 0$ for all $\sigma_1^2 \in (0,\Sigma_{10}^2)$ and all $a \in [0,\overline{a}_1]$, which implies that $\mathcal{A}_1$ is strictly increasing in σ_1^2 when $\sigma_1^2 \in (0,\Sigma_{10}^2)$. Now note, from the proof of Proposition 7, that

$$\frac{\partial \mathcal{H}}{\partial a}(a,\overline{\Sigma}_1^2,\sigma_\varepsilon^2,m_1,r) > \frac{\overline{\Sigma}_1^2 m_1}{[1+(1+a)^2]\overline{\Sigma}_1^2+\sigma_\varepsilon^2} \geq \frac{\overline{\Sigma}_1^2 m_1}{[1+(1+\overline{a}_1)^2]\overline{\Sigma}_1^2+\sigma_\varepsilon^2} = \rho > 0$$

for all $a \in [0,\overline{a}_1]$. Given that $\partial\mathcal{H}/\partial a$ is uniformly continuous in a and σ_1^2 when $(a,\sigma_1^2) \in [0,\overline{a}_1] \times [\Sigma_{10}^2,\overline{\Sigma}_1^2]$, we have that there exists $\Sigma_{10}^2 \leq \Sigma_{11}^2 < \overline{\Sigma}_1^2$ such that

$$\frac{\partial\mathcal{H}}{\partial a}(a,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) \geq \frac{\rho}{2}$$

for all $\sigma_1^2 > \Sigma_{11}^2$ and all $a \in [0,\overline{a}_1]$. Moreover, since $\mathcal{H}$ is also uniformly continuous in a and σ_1^2 when $(a,\sigma_1^2) \in [0,\overline{a}_1] \times [\Sigma_{10}^2,\overline{\Sigma}_1^2]$, we have, increasing Σ_{11}^2 if necessary, that

$$\mathcal{H}(a,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) > \mathcal{H}(a,\overline{\Sigma}_1^2,\sigma_\varepsilon^2,m_1,r) - \rho/4 \geq \mathcal{H}(0,\overline{\Sigma}_1^2,\sigma_\varepsilon^2,m_1,r) - \rho/4 = -\rho/4$$

for all $\sigma_1^2 > \Sigma_{11}^2$ and all $a \in [0,\overline{a}_1]$. Finally, given that

$$\frac{\partial H}{\partial a}(a,0,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) = \mathcal{H}(a,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) + \frac{\partial\mathcal{H}}{\partial a}(a,\sigma_1^2,\sigma_\varepsilon^2,m_1,r),$$

we can then conclude that $H(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)$ is strictly increasing in a when $a \in [0, \overline{a}_1]$ as long as $\sigma_1^2 > \Sigma_{11}^2$, which implies that a_1^* is unique when $\sigma_1^2 \in (\Sigma_{11}^2, \overline{\Sigma}_1^2)$. Since H is continuous and $a = 0$ is the only solution to $H(a, 0, \overline{\Sigma}_1^2, \sigma_\varepsilon^2, m_1, r) = 0$, a straightforward argument shows that a_1^* converges to zero as σ_1^2 increases to $\overline{\Sigma}_1^2$. The desired result follows from the fact that a_1^* is positive for all $\sigma_1^2 \in (0, \overline{\Sigma}_1^2)$.

Proof of Proposition 10

Fix m_1 and r, and let $\Sigma_1^2 \in (0, \overline{\Sigma}_1^2(0, m_1, r))$. Since $\xi(a)$ is nondecreasing in a,

$$\frac{\partial \mathcal{H}}{\partial a}(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) \geq \frac{2(1+a)\sigma_1^4[1 + (1+a)m_1]}{\{[1 + (1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2\}^2} - \frac{\sigma_1^2 m_1}{[1 + (1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2}.$$

Hence, $\sigma_\varepsilon^2 < \sigma_1^2/m_1$ implies that

$$\frac{\partial \mathcal{H}}{\partial a}(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) \geq \frac{(1+a)\sigma_1^4}{\{[1 + (1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2\}^2} = R(a, \sigma_1^2, \sigma_\varepsilon^2).$$

Now observe that for each $(a, \sigma_1^2) \in [0, \overline{a}_1] \times [\Sigma_1^2, \overline{\Sigma}_1^2(0, m_1, r)]$, $R(a, \sigma_1^2, \sigma_\varepsilon^2)$ increases to

$$R_\infty(a, \sigma_1^2) = \frac{2(1+a)}{1 + (1+a)^2} \geq \frac{2(1+\overline{a})}{1 + (1+\overline{a})^2} > 0$$

as σ_ε^2 converges to zero. Dini's Theorem then implies that $R(a, \sigma_1^2, \sigma_\varepsilon^2)$ converges uniformly to $R_\infty(a, \sigma_1^2)$ as σ_ε^2 converges to zero when $(a, \sigma_1^2) \in [0, \overline{a}_1] \times [\Sigma_1^2, \overline{\Sigma}_1^2(0, m_1, r)]$. Therefore, there exists $\Sigma_\varepsilon^2 > 0$ such that $\sigma_\varepsilon^2 < \Sigma_\varepsilon^2$ implies that $\mathcal{H}$ is strictly increasing in a for all $a \in [0, \overline{a}_1]$ when $\sigma_1^2 \in [\Sigma_1^2, \overline{\Sigma}_1^2(0, m_1, r)]$. Consequently, $\sigma_\varepsilon^2 < \Sigma_\varepsilon^2$ implies that the equation $\mathcal{H}(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) = 0$ has a unique and positive solution when $\sigma_1^2 \in (\Sigma_1^2, \overline{\Sigma}_1^2(0, m_1, r))$, so that a_1^* is unique when $\sigma_1^2 \in (\Sigma_1^2, \overline{\Sigma}_1^2(0, m_1, r))$.

To finish the proof, note that

$$\begin{aligned}\frac{\partial \mathcal{H}}{\partial \sigma_1^2}(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) \quad \geq \quad & \frac{r(1+a)^4\sigma_1^4}{\{[1 + (1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2\}^2} - \frac{\sigma_\varepsilon^2}{(\sigma_1^2 + \sigma_\varepsilon^2)^2} \\ & - \frac{\sigma_\varepsilon^2}{\{[1 + (1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2\}^2} - \frac{(1+a)\sigma_\varepsilon^2 m_1}{\{[1 + (1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2\}^2}. \end{aligned} \tag{17}$$

It is easy to see that the last three terms on the right side of (17) converge pointwise to zero as σ_ε^2 converges to zero and that this convergence becomes monotonic when $\sigma_\varepsilon^2 \leq \sigma_1^2$. Since the first term on the right side of (17) converges pointwise and monotonically to

$$\widetilde{R}(a) = \frac{r(1+a)^4}{\{1 + (1+a)^2\}^2} \geq \frac{r}{2} > 0$$

as σ_ε^2 converges to zero, Dini's theorem then implies that $\partial\mathcal{H}/\partial\sigma_1^2$ converges uniformly to $\widetilde{R}$ as σ_ε^2 converges to zero when $(a,\sigma_1^2) \in [0,\overline{a}_1] \times [\Sigma_1^2, \overline{\Sigma}_1^2(0,m_1,r)]$. Hence, reducing Σ_ε^2 if necessary, we have that a_1^* is strictly decreasing in σ_1^2 when $\sigma_1^2 \in (\Sigma_1^2, \overline{\Sigma}_1^2(0,m_1,r))$.

References

[1] Altuğ, S., and R.A. Miller (1998): "The Effect of Work Experience on Female Wages and Labour Supply", *Review of Economic Studies* 65(1), 45–85.

[2] Baker, G., R. Gibbons and K.J. Murphy (1994): "Subjective Performance Measures in Optimal Incentive Contracts", *Quarterly Journal of Economics* 109(4), 1125–1156.

[3] Bar–Isaac, H., and J. Ganuza (2008): "Recruitment, Training and Career Concerns", *Journal of Economics & Management Strategy* 17(4), 839–864.

[4] Baker, G., M.C. Jensen, and K.J. Murphy (1988): "Compensation and Incentives: Practice vs. Theory" *Journal of Finance* 63(3), 593–616.

[5] Blanes–i–Vidal, J. (2007): "Delegation of Decision Rights and the Winner's Curse", *Economics Letters* 94(2), 163–169.

[6] Camargo, B., and E. Pastorino (2010): "Optimal Incentive Contracts in the Presence of Career Concerns and Human Capital Acquisition", mimeo.

[7] Dewatripont, M., I. Jewitt, and J. Tirole (1999a): "The Economics of Career Concerns, Part I: Comparing Information Structures", *Review of Economic Studies* 66(1), 183–198.

[8] Dewatripont, M., I. Jewitt, and J. Tirole (1999b): "The Economics of Career Concerns, Part I: Application to Missions and Accountability of Government Agencies", *Review of Economic Studies* 66(1), 199–217.

[9] Echenique, F., and I. Komunjer (2009): "Testing Models with Multiple Equilibria by Quantile Methods", *Econometrica* 77(4), 1281–1297.

[10] Fama, E.F. (1980): "Agency Problems and the Theory of the Firm", *Journal of Political Economy* 88(2), 288–307.

[11] Foster, A.D., and M.R. Rosenzweig (1995): "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture", *Journal of Political Economy* 103(6), 1176–1209.

[12] Guillemin, V., and A. Pollack (1974): *Differential Topology*, Prentice–Hall, Englewood Cliffs, New Jersey.

[13] Heckman, J.J., and L. Lochner (2005): "Distinguishing Between On–the–Job Training and Learning–by–Doing", mimeo.

[14] Holmström, B. (1982): "Managerial Incentive Problems: A Dynamic Perspective", in *Essays in Economics and Management in Honor of Lars Wahlbeck*, Swedish School of Economics, Helsinki.

[15] Holmström, B. (1999): "Managerial Incentive Problems: A Dynamic Perspective", *Review of Economic Studies* 66(1), 169–182.

[16] Jeon, S. (1996): "Moral Hazard and Reputational Concerns in Teams: Implications for Organizational Choice", *International Journal of Industrial Organization* 14, 297–315.

[17] Keane, M.P., and K.I. Wolpin (1997): "The Career Decisions of Young Men", *Journal of Political Economy* 105(3), 473–522.

[18] Kovrijnykh, A. (2007): "Career Uncertainty and Dynamic Incentives", University of Chicago, mimeo.

[19] Levy, G. (2005): "Careerist Judges and the Appeals Process", *RAND Journal of Economics* 36(2), 275–297.

[20] Martinez, L. (2006): "Reputation and Career Concerns", Federal Reserve Bank of Richmond, Working Paper WP 06–01.

[21] Martinez, L. (2009a): "A Theory of Political Cycles", *Journal of Economic Theory*, 144(3), 1166–1186.

[22] Martinez, L. (2009b): “Reputation, Career concerns, and Job Assignments”, *B.E. Journal of Theoretical Economics* 9(1), Article 15.

[23] Mas–Colell, A., M.D. Whinston, and J.R. Green (1995): *Microeconomic Theory*, Oxford University Press, New York, New York.

[24] Mattozzi, A., and A. Merlo (2008): “Political Careers or Career Politicians?”, *Journal of Public Economics*, 92 (3-4), 597–608.

[25] Milgrom, P., and J. Roberts (1990): “Rationalizability, Learning, and Equilibrium Selection in Games with Strategic Complementarities”, *Econometrica*, 58(6), 1255–1277.

[26] Milgrom, P., and C. Shannon (1994): “Monotone Comparative Statics”, *Econometrica*, 62(1), 157–180.

[27] Ortega, J. (2003): “Power in the Firm and Managerial Career Concerns”, *Journal of Economics & Management Strategy* 12(1), 1–29.

[28] Persson, T., and G. Tabellini (2002): *Political Economics: Explaining Economic Policy*, MIT Press, Cambridge, Massachussets.

[29] Posner, R.A. (1993): “What Do Judges and Justices Maximize? (The Same Thing Everybody Else Does)”, *Supreme Court Economic Review* 3, 1–44.

[30] Prendergast, C., (1999): “The Provision of Incentives in Firms”, *Journal of Economic Literature* 37(1), 7–63.

[31] Rubinstein, Y., and Y. Weiss (2007): “Post Schooling Wage Growth: Investment, Search and Learning”, in *Handbook of the Economics of Education*, Volume I, E.A. Hanushek and F. Welch Eds., North–Holland, the Netherlands.

[32] Rudin, W. (1991): *Functional Analysis*, Second Edition, McGraw–Hill, New York, New York.

[33] Shaw, K.L. (1989): “Life–Cycle Labor Supply with Human Capital Accumulation”, *International Economic Review* 30(2), 431–456.

[34] Wolpin, K.I. (1992): "The Determinants of Black–White Differences in Early Employment Careers: Search, Layoffs, Quits, and Endogenous Wage Growth", *Journal of Political Economy* 100(3), 535–560.

Appendix B: Omitted Details (Not for Publication)

Proof of Claim in Footnote 6

Suppose the worker conditions his behavior on his past wages. Let W_t, with typical element $w^t = (w_1, \ldots, w_{t-1})$, be the set of wage histories for the worker in period t, and let $\widetilde{Z}_t = Z_t \times W_t$, with typical element $\widetilde{z}^t = (z^t, w^t)$, be the set of period–$t$ histories for the worker. A wage rule is still a sequence $\omega = \{\omega_t\}_{t=1}^T$, with $\omega_t : Y_t \to \mathbb{R}$. Now, however, a strategy for the worker is a sequence $\widetilde{\sigma} = \{\widetilde{\sigma}_t\}_{t=1}^T$, with $\widetilde{\sigma}_t : \widetilde{Z}_t \to \Delta(\mathbb{R}_+)$. The definition of an equilibrium is the same as in the main text. Suppose then that $(\widetilde{\sigma}^*, \omega^*)$ is an equilibrium and let $w^t(y_1, \ldots, y_{t-1}) = (w_1, \ldots, w_{t-1})$ be such that $w_s = \omega_s^*(y_1, \ldots, y_{s-1})$ for all $s \in \{1, \ldots, t\}$. Now define $\sigma^* = \{\sigma_t^*\}_{t=1}^T$, with $\sigma_t^* : Z_t \to \mathbb{R}_+$, to be such that if $z^t = (y^t, a^t)$, then $\sigma_t^*(z^t) = \widetilde{\sigma}_t^*(z^t, w^t(y^t))$. It is immediate to see that (σ^*, ω^*) is an equilibrium according to the definition in the main text, and its outcome coincides with that of $(\widetilde{\sigma}^*, \omega^*)$.

Proof of Proposition 6

First note that $A_1(\chi)$ has a finite number of elements if, and only if,

$$A_1^{++}(\chi) = \{a \in \mathbb{R}_{++} : H(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) = 0\}$$

has a finite number of elements. Let $\widetilde{H} : \mathbb{R}_{++} \times \mathbb{R}_{++}^4 \to \mathbb{R}$ be given by $\widetilde{H}(a, \chi) = H(a, 0, \chi)$ and for each $\chi \in \mathbb{R}_{++}^4$, define $\widetilde{H}_\chi : \mathbb{R}_{++} \to \mathbb{R}$ to be such that $\widetilde{H}_\chi(a) = \widetilde{H}(a, \chi)$. Since

$$\frac{\partial \widetilde{H}}{\partial m_1}(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) = -\frac{(1+a)^2\sigma_1^2}{[1+(1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2} < 0$$

for all $(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) \in \mathbb{R}_{++} \times \mathbb{R}_{++}^4$, we have that zero is a regular value of $\widetilde{H}$. By the transversality theorem (see Guillemin and Pollack (1974)), the set Ξ of $\chi \in \mathbb{R}_{++}^4$ for which zero is a regular value of $\widetilde{H}_\chi$ has full measure. It is immediate to see that Ξ is open as well. To finish, note, by the implicit function theorem, that if $\chi \in \Xi$, then the elements of $A_1^{++}(\chi)$ are locally isolated. Since $A_1^{++}(\chi)$ is closed, and thus compact, for all $\chi \in \mathbb{R}_{++}^4$, we can then conclude that $A_1^{++}(\chi)$ has a finite number of elements if $\chi \in \Xi$.

4.1 Properties of $\overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r)$

First note that $h(0, \sigma_\varepsilon^2, m_1, r) = 2 + m_1 > 0$ and $\lim_{\sigma_1^2 \to \infty} h(\sigma_1^2, \sigma_\varepsilon^2, m_1, r) = -\infty$. Hence, for each σ_ε^2, m_1, and r, the equation $h(\sigma_1^2, \sigma_\varepsilon^2, m_1, r) = 0$ has a solution by the intermediate value theorem. Since

$$\begin{aligned}\frac{\partial h}{\partial \sigma_1^2}(\sigma_1^2, \sigma_\varepsilon^2, m_1, r) &= \frac{\sigma_\varepsilon^2}{(\sigma_1^2 + \sigma_\varepsilon^2)^2} - \frac{2r(\sigma_1^4 + \sigma_1^2\sigma_\varepsilon^2)}{(2\sigma_1^2 + \sigma_\varepsilon^2)^2} \\ &< \frac{1}{\sigma_1^2 + \sigma_\varepsilon^2} - \frac{2r\sigma_1^4}{2\sigma_1^2 + \sigma_\varepsilon^2} = \frac{1}{2\sigma_1^2 + \sigma_\varepsilon^2}\left(1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_\varepsilon^2} - \frac{2r\sigma_1^4}{2\sigma_1^2 + \sigma_\varepsilon^2}\right),\end{aligned}$$

we then have that $\partial h(\sigma_1^2, \sigma_\varepsilon^2, m_1, r)/\partial \sigma_1^2 < 0$ when $h(\sigma_1^2, \sigma_\varepsilon^2, m_1, r) = 0$. Thus, for all σ_ε^2, m_1, and r, the solution to the equation $h(\sigma_1^2, \sigma_\varepsilon^2, m_1, r) = 0$ is unique. Denote this solution by $\overline{\Sigma}_1^2 = \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r)$. Note that $\overline{\Sigma}_1^2 > 0$. By construction, $h(\sigma_1^2, \sigma_\varepsilon^2, m_1, r)$ is positive if, and only if, $\sigma_1^2 \in (0, \overline{\Sigma}_1^2)$.

Now note that $\partial h(\overline{\Sigma}_1^2, \sigma_\varepsilon^2, m_1, r)/\partial \sigma_1^2 < 0$. Moreover,

$$\frac{\partial h}{\partial \sigma_\varepsilon^2}(\sigma_1^2, \sigma_\varepsilon^2, m_1, r) = -\frac{\sigma_1^2}{(\sigma_1^2 + \sigma_\varepsilon^2)^2} + \frac{r\sigma_1^4}{(2\sigma_1^2 + \sigma_\varepsilon^2)^2} > \frac{1}{2\sigma_1^2 + \sigma_\varepsilon^2}\left(\frac{r\sigma_1^4}{2\sigma_1^2 + \sigma_\varepsilon^2} - 1 - \frac{\sigma_1^2}{\sigma_1^2 + \sigma_\varepsilon^2}\right),$$

and so $\partial h(\overline{\Sigma}_1^2, \sigma_\varepsilon^2, m_1, r)/\partial \sigma_\varepsilon^2 > 0$. Hence, by the implicit function theorem, we have that $\overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r)$ is strictly increasing in σ_ε^2. Next observe that

$$\overline{\Sigma}_1^4 = \frac{2\overline{\Sigma}_1^2 + \sigma_\varepsilon^2}{r}\left(2 + m_1 + \frac{\overline{\Sigma}_1^2}{\overline{\Sigma}_1^2 + \sigma_\varepsilon^2}\right) > \frac{\overline{\Sigma}_1^2}{r} + \frac{\sigma_\varepsilon^2(2 + m_1)}{r},$$

which implies that

$$\overline{\Sigma}_1^2 > \frac{1}{2r} + \left\{\frac{1}{4r^2} + \frac{\sigma_\varepsilon^2(2 + m_1)}{r}\right\}^{1/2}. \tag{18}$$

In particular, $\lim_{r \to 0} \overline{\Sigma}_1^2(\sigma_\varepsilon, m_1, r) = \lim_{\sigma_\varepsilon^2 \to \infty} \overline{\Sigma}_1^2(\sigma_\varepsilon, m_1, r) = \infty$. To finish, suppose, by contradiction, that $\lim_{\sigma_\varepsilon^2 \to 0} \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r) \neq 2(3 + m_1)/r$. This implies that there exists a sequence $\{\sigma_{\varepsilon,n}^2\}$ converging to zero such that if $\overline{\Sigma}_{1,n}^2 = \overline{\Sigma}_1^2(\sigma_{\varepsilon,n}^2, m_1, r)$, then $\{\overline{\Sigma}_{1,n}^2\}$ does not converge to $2(3 + m_1)/r$. Without loss of generality, assume that $\{\overline{\Sigma}_{1,n}^2\}$ is convergent and denote its limit by $\overline{\Sigma}_{1,\infty}^2$. Note that $\overline{\Sigma}_{1,\infty}^2 \geq 1/2r > 0$ by (18). Since $2(3 + m_1)/r$ is the only solution to $h(\sigma_1^2, 0, m_1, r) = 0$ and $\overline{\Sigma}_{1,\infty}^2 > 0$ implies that $h(\cdot, \cdot, m_1, r)$ is jointly continuous in $(\sigma_1^2, \sigma_\varepsilon^2)$ at $(\overline{\Sigma}_{1,\infty}^2, 0)$, we then have that $\lim_n h(\overline{\Sigma}_{1,n}^2, \sigma_{\varepsilon,n}^2, m_1, r) = h(\overline{\Sigma}_{1,\infty}^2, 0, m_1, r) \neq 0$. However, $h(\overline{\Sigma}_{1,n}^2, \sigma_{\varepsilon,n}^2, m_1, r) = 0$ for all n, and so $\lim_n h(\overline{\Sigma}_{1,n}^2, \sigma_{\varepsilon,n}^2, m_1, r) = 0$, a contradiction.

Appendix C: Supplementary Material (Not for Publication)

Here we consider the case of uncertain returns to learning–by–doing when $\vartheta_2 > 0$. In what follows we assume that g' is convex and, as in the main text, that $\lim_{a\to\infty} \xi(a) = \infty$. Recall that $\xi(a) = (1+a)^{-1}g(a)$. We also assume, again as in the main text, that $2+m_1 > 0$.

We know the worker's effort in period three is always zero. In order to determine effort in period two, let $\widehat{a}_1$ and $\widehat{a}_2$ be the market's conjectures about the worker's effort in periods one and two, respectively. Note that $\widehat{a}_2$ can depend on the worker's output in period one. We omit the dependence of $\widehat{a}_2$ on y_1 for ease of notation. Moreover, as we will see, in equilibrium the worker's effort in period two is uncontingent. The worker's wage in period three is

$$w_3 = (1+\widehat{a}_1+\widehat{a}_2)\mathbb{E}[\theta|y_1,y_2] = (1+\widehat{a}_1+\widehat{a}_2)\left\{m_2 + \frac{(1+\widehat{a}_1)\sigma_2^2}{(1+\widehat{a}_1)^2\sigma_2^2+\sigma_\varepsilon^2}\left[y_2-(1+\widehat{a}_1)m_2\right]\right\},$$

where $\sigma_2^2 = \sigma_1^2\sigma_\varepsilon^2/(\sigma_1^2+\sigma_\varepsilon^2)$ and

$$m_2 = \mathbb{E}[\theta|y_1] = m_1 + \frac{\sigma_1^2}{\sigma_1^2+\sigma_\varepsilon^2}\left(y_1-\widehat{a}_1-m_1\right).$$

Since $y_2 = a_2 + k_2 + \varepsilon_2$, conditional on the worker's effort and output in period one, the variance of the worker's wage in period three does not depend on his effort in period two. From this (and the fact that $g'(0) = 0$), it follows that the worker's optimal effort in period two is the unique and positive solution to

$$g'(a) = \frac{(1+\widehat{a}_1+\widehat{a}_2)(1+\widehat{a}_1)\sigma_2^2}{(1+\widehat{a}_1)^2\sigma_2^2+\sigma_\varepsilon^2} = \frac{(1+\widehat{a}_1+\widehat{a}_2)(1+\widehat{a}_1)\sigma_1^2}{\left[1+(1+\widehat{a}_1)^2\right]\sigma_1^2+\sigma_\varepsilon^2}. \tag{19}$$

From (19), the possible values for the worker's effort in period two are the solutions to

$$J_2(a,\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2) = g'(a) - \frac{(1+\widehat{a}_1+a)(1+\widehat{a}_1)\sigma_1^2}{\left[1+(1+\widehat{a}_1)^2\right]\sigma_1^2+\sigma_\varepsilon^2} = 0. \tag{20}$$

Since g' is convex, it follows that (20) has a unique and positive solution. Moreover, the solution to (20) does not depend on y_1, that is, it is uncontingent. Equation (20) defines the worker's effort in period two implicitly as a function of $\widehat{a}_1$, σ_1^2, and σ_ε^2. Denote this function by α_2. Note that

$$J_2(a,\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2) \geq g'(a) - 1 - a.$$

Therefore, regardless of $\widehat{a}_1$, σ_1^2, and σ_ε^2, $\alpha_2(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2)$ is bounded above by $\overline{a}_2$, where $\overline{a}_2 > 0$ is the unique solution to $g'(a) = 1+a$. Hence, α_2 is uniformly bounded. Since $g''(a) \geq g'(a)/a$

for all $a > 0$, we have that

$$\frac{\partial J_2}{\partial a}(\alpha_2(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2),\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2) = g''(\alpha_2(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2)) - \frac{(1+\widehat{a}_1)\sigma_1^2}{[1+(1+\widehat{a}_1)^2]\,\sigma_1^2+\sigma_\varepsilon^2}$$
$$\geq \frac{1+\widehat{a}_1+\alpha_2(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2)}{\alpha_2(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2)}\frac{(1+\widehat{a}_1)\sigma_1^2}{\{[1+(1+\widehat{a}_1)^2]\,\sigma_1^2+\sigma_\varepsilon^2\}} - \frac{(1+\widehat{a}_1)\sigma_1^2}{[1+(1+\widehat{a}_1)^2]\,\sigma_1^2+\sigma_\varepsilon^2} > 0$$

if $\widehat{a}_1 \geq 0$. Thus, α is also differentiable in $\widehat{a}_1$, σ_1^2, and σ_ε^2 by the Implicit Function Theorem.[16]

It is immediate to see from (20) that α_2 is strictly increasing in σ_1^2 and strictly decreasing in σ_ε^2. The effect of an increase in $\widehat{a}_1$ on α_2 is ambiguous, though. Indeed,

$$\frac{\partial\alpha_2}{\partial\widehat{a}_1}(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2) = \frac{1}{g''(\alpha_2(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2)) - \dfrac{(1+\widehat{a}_1)\sigma_1^2}{[1+(1+\widehat{a}_1)^2]\sigma_1^2+\sigma_\varepsilon^2}}\frac{\sigma_1^2}{[1+(1+\widehat{a}_1)^2]\sigma_1^2+\sigma_\varepsilon^2}$$
$$\times\left\{2(1+\widehat{a}_1)+\alpha_2(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2) - \frac{2[1+\widehat{a}_1+\alpha_2(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2)](1+\widehat{a}_1)^2\sigma_1^2}{[1+(1+\widehat{a}_1)^2]\sigma_1^2+\sigma_\varepsilon^2}\right\} \quad (21)$$

Therefore, $\partial\alpha_2(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2)/\partial\widehat{a}_1 > 0$ when $\widehat{a}_1$ is small. However, since α_2 is uniformly bounded, $\partial\alpha_2(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2)/\partial\widehat{a}_1 < 0$ when $\widehat{a}_1$ is large. The intuition for the non monotonicity of α_2 in $\widehat{a}_1$ is as follows. The worker's wage in any period is proportional to both his expected human capital and his reputation. Thus, an increase in $\widehat{a}_1$, which implies an increase in the worker's expected human capital in period three, increases the worker's return from manipulating his reputation in period three. On the other hand, the effect of the worker's output in period two on his reputation in period three decreases as $\widehat{a}_1$ increases, making it more costly for the worker to influence the market's beliefs in the third period. When $\widehat{a}_1$ is small, the first effect dominates the second. When $\widehat{a}_1$ is large, the second effect prevails.

Suppose now the worker's effort in period two is $\alpha_2(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2)$. An argument similar to the one used in the main text shows that the worker's optimal choice of effort in the first period is the solution to

$$\frac{(1+\widehat{a}_1)\sigma_1^2}{\sigma_1^2+\sigma_\varepsilon^2} + \frac{[1+\widehat{a}_1+\alpha_2(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2)]\sigma_1^2}{[1+(1+\widehat{a}_1)^2]\sigma_1^2+\sigma_\varepsilon^2} + \frac{(1+\widehat{a}_1)[1+\widehat{a}_1+\alpha_2(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2)]\sigma_1^2 m_1}{[1+(1+\widehat{a}_1)^2]\sigma_1^2+\sigma_\varepsilon^2}$$
$$-r\left\{\frac{(1+\widehat{a}_1)[1+\widehat{a}_1+\alpha_2(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2)]\sigma_1^2}{[1+(1+\widehat{a}_1)^2]\sigma_1^2+\sigma_\varepsilon^2}\right\}^2[1+a+\alpha_2(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2)]\sigma_1^2+\lambda = g'(a),$$

[16] We are implicitly assuming that for each σ_1^2 and σ_ε^2, $\alpha_2(\cdot,\sigma_1^2,\sigma_\varepsilon^2)$ is defined in a neighborhood of zero. This is not a problem since equation (20) has a unique and positive solution when $\widehat{a}_1$ is negative as long as $\widehat{a}_1$ is close to zero.

where $\lambda \geq 0$ and $\lambda = 0$ if the solution is positive. The term

$$\begin{aligned} MB(\widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, m_1) &= \frac{(1+\widehat{a}_1)\sigma_1^2}{\sigma_1^2+\sigma_\varepsilon^2} + \frac{[1+\widehat{a}_1+\alpha_2(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2)]\sigma_1^2}{[1+(1+\widehat{a}_1)^2]\sigma_1^2+\sigma_\varepsilon^2} \\ &\quad + \frac{(1+\widehat{a}_1)[1+\widehat{a}_1+\alpha_2(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2)]\sigma_1^2 m_1}{[1+(1+\widehat{a}_1)^2]\sigma_1^2+\sigma_\varepsilon^2} \end{aligned}$$

is the worker's marginal benefit from effort in period one, while the term

$$MC_{\text{risk}}(a, \widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, r) = -r \left\{ \frac{(1+\widehat{a}_1)[1+\widehat{a}_1+\alpha_2(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2)]\sigma_1^2}{[1+(1+\widehat{a}_1)^2]\sigma_1^2+\sigma_\varepsilon^2} \right\}^2 [1+a+\alpha_2(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2)]\sigma_1^2$$

is the worker's marginal cost of effort a in period one due to his risk aversion. Thus, the possible values for the worker's choice of effort in period one, a_1^*, are the non negative solutions to

$$\begin{aligned} J_1(a, \lambda, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) &= g'(a) - \lambda - \frac{(1+a)\sigma_1^2}{\sigma_1^2+\sigma_\varepsilon^2} - \frac{[1+a+\alpha_2(a,\sigma_1^2,\sigma_\varepsilon^2)]\sigma_1^2}{[1+(1+a)^2]\sigma_1^2+\sigma_\varepsilon^2} \\ &\quad - \frac{(1+a)[1+a+\alpha_2(a,\sigma_1^2,\sigma_\varepsilon^2)]\sigma_1^2}{[1+(1+a)^2]\sigma_1^2+\sigma_\varepsilon^2} \left\{ m_1 - \frac{r(1+a)[1+a+\alpha_2(a,\sigma_1^2,\sigma_\varepsilon^2)]^2\sigma_1^4}{[1+(1+a)^2]\sigma_1^2+\sigma_\varepsilon^2} \right\} = 0, \end{aligned} \quad (22)$$

where $\lambda \geq 0$ and $\lambda = 0$ if a_1^* is positive. By construction, the worker's choice of effort in period two is $\alpha_2(a_1^*, \sigma_1^2, \sigma_\varepsilon^2)$, which is unique given a_1^*.

As in the case where $\vartheta_2 = 0$, non–uniqueness of the worker's choice of effort in period one is a possibility. Since

$$J_1(a, \lambda, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) \geq g'(a) - (1+a) - (1+a+\overline{a}_2)(1+m_1),$$

the possible values for a_1^* are bounded above by $\overline{a}_1 = \overline{a}_1(m_1)$, where $\overline{a}_1$ is the unique solution to $g'(a) = (1+a) + (1+a+\overline{a}_2)(1+m_1)$. The same argument as in the main text proves that for each $\chi = (\sigma_1^2, \sigma_\varepsilon^2, m_1, r) \in \mathbb{R}_{++}^4$, the set

$$A_1(\chi) = \{a \in \mathbb{R}_+ : \exists \lambda \geq 0 \text{ with } J_1(a, \lambda, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) = 0 \text{ and } \lambda a = 0\}$$

of equilibrium effort choices in the first period is compact. In particular, for each $\chi \in \mathbb{R}_{++}^4$ there exist a smallest and a greatest a_1^*, which we denote by $a_{1,\min}^*$ and $a_{1,\max}^*$, respectively. We first establish that (22) has always a solution and that for generic χ, the number of solutions to (22) is finite.

Proposition 12. *An equilibrium always exists. The set of $\chi \in \mathbb{R}_{++}^4$ for which the number of equilibria is finite is open and has full measure.*

Proof: Suppose first that $J_1(0,0,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) \geq 0$. Given that

$$J_1(0,\lambda,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) = J_1(0,0,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) - \lambda,$$

it is immediate to see that $a_1^* = 0$ is a solution to (22) in this case. Suppose now that $J_1(0,0,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) < 0$. Notice that α_2 uniformly bounded and $\lim_{a\to\infty}\xi(a) = \infty$ imply that $\lim_{a\to\infty} J_1(a,0,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) = \infty$. Since J_1 is continuous in a (since α_2 is continuous in a), the intermediate value theorem implies that (22) has a positive solution in this case. Thus, an equilibrium always exists. The same argument as in the proof of Proposition 6 in the main text establishes the second part of the claim. □

We now identify necessary and sufficient conditions for the worker's choice of effort in period one to be positive. For this, let

$$j_1(\sigma_1^2,\sigma_\varepsilon^2,m_1,r) = 1 + \frac{\sigma_1^2}{\sigma_1^2+\sigma_\varepsilon^2} + [1+\alpha_2(0,\sigma_1^2,\sigma_\varepsilon^2)](1+m_1) - \frac{r\sigma_1^4}{2\sigma_1^2+\sigma_\varepsilon^2}[1+\alpha_2(0,\sigma_1^2,\sigma_\varepsilon^2)]^3.$$

Note that α_2 uniformly bounded implies that $\lim_{\sigma_1^2\to\infty} j_1(\sigma_1^2,\sigma_\varepsilon^2,m_1,r) = -\infty$. Given that $j_1(0,\sigma_\varepsilon^2,m_1,r) = 2+m_1 > 0$ and j_1 is continuous in σ_1^2 (as α_2 is continuous in σ_1^2), the equation $j_1(\sigma_1^2,\sigma_\varepsilon^2,m_1,r) = 0$ has a solution, and its solutions are positive. Now observe that

$$\begin{aligned}
\frac{\partial j_1}{\partial \sigma_1^2}(\sigma_1^2,\sigma_\varepsilon^2,m_1,r) = {} & \frac{\sigma_\varepsilon^2}{(\sigma_1^2+\sigma_\varepsilon^2)^2} + (1+m_1)\frac{\partial \alpha_2}{\partial \sigma_1^2}(0,\sigma_1^2,\sigma_\varepsilon^2) - \frac{2r\sigma_1^2(\sigma_1^2+\sigma_\varepsilon^2)}{(2\sigma_1^2+\sigma_\varepsilon^2)^2}[1+\alpha_2(0,\sigma_1^2,\sigma_\varepsilon^2)]^3 \\
& - \frac{3r\sigma_1^4}{2\sigma_1^2+\sigma_\varepsilon^2}[1+\alpha_2(0,\sigma_1^2,\sigma_\varepsilon^2)]^2\frac{\partial \alpha_2}{\partial \sigma_1^2}(0,\sigma_1^2,\sigma_\varepsilon^2) \\
< {} & \frac{1}{\sigma_1^2}\underbrace{\left\{\frac{\sigma_1^2}{\sigma_1^2+\sigma_\varepsilon^2} - \frac{r\sigma_1^4}{2\sigma_1^2+\sigma_\varepsilon^2}[1+\alpha_2(0,\sigma_1^2,\sigma_\varepsilon^2)]^3\right\}}_{A} \\
& + \frac{\partial \alpha_2}{\partial \sigma_1^2}(0,\sigma_1^2,\sigma_\varepsilon^2)\underbrace{\left\{(1+m_1) - \frac{3r\sigma_1^4}{2\sigma_1^2+\sigma_\varepsilon^2}[1+\alpha_2(0,\sigma_1^2,\sigma_\varepsilon^2)]^2\right\}}_{B}
\end{aligned}$$

Since $1+(1+m_1)[1+\alpha_2(0,\sigma_1^2,\sigma_\varepsilon^2)] > 0$, the term A is negative when $j_1(\sigma_1^2,\sigma_\varepsilon^2,m_1,r) = 0$. Likewise, $1+(\sigma_1^2+\sigma_\varepsilon^2)^{-1}\sigma_1^2 > 0$ implies that B is also negative when $j_1(\sigma_1^2,\sigma_\varepsilon^2,m_1,r) = 0$. Given that $\partial\alpha_2(0,\sigma_1^2,\sigma_\varepsilon^2)/\partial\sigma_1^2 > 0$ by (20), we then have that $\partial j_1(\sigma_1^2,\sigma_\varepsilon^2,m_1,r)/\partial\sigma_1^2 < 0$ when $j_1(\sigma_1^2,\sigma_\varepsilon^2,m_1,r) = 0$. Thus, there exists a unique $\overline{\Sigma}_1^2 = \overline{\Sigma}_1^2(\sigma_\varepsilon^2,m_1,r) > 0$ such that $j_1(\sigma_1^2,\sigma_\varepsilon^2,m_1,r) > 0$ if, and only if, $\sigma_1^2 \in (0,\overline{\Sigma}_1^2)$. We have the following result, which is useful in the analysis that follows.

Lemma 1. *$\overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r)$ is strictly increasing in σ_ε^2. Moreover, $\lim_{\sigma_\varepsilon^2\to\infty} \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r) = \infty$, $\lim_{r\to 0} \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r) = \infty$, $\lim_{r\to\infty} r\overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r) = \infty$, and*

$$\lim_{\sigma_\varepsilon^2\to 0} \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r) = \frac{1}{r[1+a_2^0]^3}\left\{2 + [1+a_2^0](1+m_1)\right\},$$

where a_2^0 is the unique solution to $g'(a) = (1+a)/2$.

Proof: Note that

$$\begin{aligned}
\frac{\partial j_1}{\partial \sigma_\varepsilon^2}(\sigma_1^2, \sigma_\varepsilon^2, m_1, r) &= -\frac{\sigma_1^2}{(\sigma_1^2+\sigma_\varepsilon^2)^2} + (1+m_1)\frac{\partial \alpha_2}{\partial \sigma_\varepsilon^2}(0, \sigma_1^2, \sigma_\varepsilon^2) + \frac{r\sigma_1^4\left[1+\alpha_2(0,\sigma_1^2,\sigma_\varepsilon^2)\right]^3}{(2\sigma_1^2+\sigma_\varepsilon^2)^2} \\
&\quad -3\left[1+\alpha_2(0,\sigma_1^2,\sigma_\varepsilon^2)\right]^2 \frac{r\sigma_1^4}{2\sigma_1^2+\sigma_\varepsilon^2}\frac{\partial \alpha_2}{\partial \sigma_\varepsilon^2}(0,\sigma_1^2,\sigma_\varepsilon^2) \\
&> \frac{1}{2\sigma_1^2+\sigma_\varepsilon^2}\underbrace{\left\{\frac{r\sigma_1^4\left[1+\alpha_2(0,\sigma_1^2,\sigma_\varepsilon^2)\right]^3}{2\sigma_1^2+\sigma_\varepsilon^2} - 1 - \frac{\sigma_1^2}{\sigma_1^2+\sigma_\varepsilon^2}\right\}}_{C} \\
&\quad + \frac{\partial \alpha_2}{\partial \sigma_\varepsilon^2}(0,\sigma_1^2,\sigma_\varepsilon^2)\underbrace{\left\{1+m_1 - \frac{3r\sigma_1^4\left[1+\alpha_2(0,\sigma_1^2,\sigma_\varepsilon^2)\right]^2}{2\sigma_1^2+\sigma_\varepsilon^2}\right\}}_{D}.
\end{aligned}$$

Since both C and D are negative when $\sigma_1^2 = \overline{\Sigma}_1^2$, and α_2 is strictly decreasing in σ_ε^2, we have that $\partial j_1(\sigma_1^2, \sigma_\varepsilon^2, m_1, r)/\partial \sigma_\varepsilon^2 > 0$ when $\sigma_1^2 = \overline{\Sigma}_1^2$. Hence, by the Implicit Function Theorem, $\overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r)$ is strictly increasing in σ_ε^2.

Now note that

$$\overline{\Sigma}_1^4 = \frac{2\overline{\Sigma}_1^2+\sigma_\varepsilon^2}{r[1+\alpha_2(0,\overline{\Sigma}_1^2,\sigma_\varepsilon^2)]^3}\left\{1 + \frac{\overline{\Sigma}_1^2}{\overline{\Sigma}_1^2+\sigma_\varepsilon^2} + [1+\alpha_2(0,\overline{\Sigma}_1^2,\sigma_\varepsilon^2)](1+m_1)\right\}, \tag{23}$$

and so

$$\overline{\Sigma}_1^2 > \frac{\sigma_\varepsilon^2\left\{1+[1+\alpha_2(0,\overline{\Sigma}_1^2,\sigma_\varepsilon^2)](1+m_1)\right\}}{r[1+\alpha_2(0,\overline{\Sigma}_1^2,\sigma_\varepsilon^2)]^3}.$$

Since α_2 is uniformly bounded by $\overline{a}_2$, $\lim_{\sigma_\varepsilon^2\to\infty} \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r) = \lim_{r\to 0} \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r) = \infty$. Next, note that (23) implies that

$$\overline{\Sigma}_1^4 - \frac{2(1+m_1)\overline{\Sigma}_1^2}{r[1+\alpha_2(0,\overline{\Sigma}_1^2,\sigma_\varepsilon^2)]^2} - \frac{\sigma_\varepsilon^2}{r[1+\alpha_2(0,\overline{\Sigma}_1^2,\sigma_\varepsilon^2)]^3} > 0,$$

from which we obtain that

$$\overline{\Sigma}_1^2 > \frac{1+m_1}{r[1+\alpha_2(0,\overline{\Sigma}_1^2,\sigma_\varepsilon^2)]^2} + \left\{\frac{(1+m_1)^2}{r^2[1+\alpha_2(0,\overline{\Sigma}_1^2,\sigma_\varepsilon^2)]^4} + \frac{\sigma_\varepsilon^2}{r[1+\alpha_2(0,\overline{\Sigma}_1^2,\sigma_\varepsilon^2)]^3}\right\}^{1/2}.$$

Therefore,

$$r\overline{\Sigma}_1^2 > \left\{ \frac{4r\sigma_\varepsilon^2}{[1+\alpha_2(0,\overline{\Sigma}_1^2,\sigma_\varepsilon^2)]^3} \right\}^{1/2},$$

and so, once more since α_2 is uniformly bounded by $\overline{a}_2$, $\lim_{r\to\infty} r\overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r) = \infty$.

To finish, note from (20) that $\alpha_2(0,\sigma_1^2,0) \equiv a_2^0$, and so

$$\frac{1}{r[1+a_2^0]^3}\left\{2+[1+a_2^0](1+m_1)\right\}$$

is the only solution to $j_1(\sigma_1^2, 0, m_1, r) = 0$. Given $\Sigma_1^2 > 0$ implies that $j_1(\cdot,\cdot,m_1,r)$ is jointly continuous in $(\sigma_1^2,\sigma_\varepsilon^2)$ at $(\Sigma_1^2, 0)$, the same argument used in Appendix B to prove that $\lim_{\sigma_\varepsilon^2\to 0}\overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r) = 2(3+m_1)/r$ when $\vartheta_2 = 0$ can be used to show that

$$\lim_{\sigma_\varepsilon^2\to 0}\overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r) = \frac{1}{r[1+a_2^0]^3}\left\{2+[1+a_2^0](1+m_1)\right\},$$

which completes the proof. □

Proposition 13 below shows that the worker's effort in period one can only be zero if $\sigma_1^2 \geq \overline{\Sigma}_1^2$, and that this effort is zero if σ_1^2 is large enough (given σ_ε^2, m_1, and r). Moreover, for each m_1 and r, there exists a lower bound on σ_ε^2 above which $\sigma_1^2 \geq \overline{\Sigma}_1^2$ implies that the worker does not exert effort in the first period. Likewise, for each m_1 and σ_ε^2, there exists a lower bound on r above which $\sigma_1^2 \geq \overline{\Sigma}_1^2$ implies that the worker's effort in period one is zero.

Proposition 13. *The worker's effort in the first period is positive if $\sigma_1^2 \in (0, \overline{\Sigma}_1^2)$ and there exists $\widetilde{\Sigma}_1^2 = \widetilde{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r) \geq \overline{\Sigma}_1^2$ such that the worker's effort in period one is zero if $\sigma_1^2 \geq \widetilde{\Sigma}_1^2$. Moreover, for each m_1 and r, there exists $\Sigma_\varepsilon^2 \geq 0$ such that $\widetilde{\Sigma}_1^2 = \overline{\Sigma}_1^2$ if $\sigma_\varepsilon^2 > \Sigma_\varepsilon^2$, and for each m_1 and σ_ε^2, there exists $\overline{r} \geq 0$ such that $\widetilde{\Sigma}_1^2 = \overline{\Sigma}_1^2$ if $r > \overline{r}$.*

Proof: First note that $J_1(0,\lambda,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) = 0$ if, and only if,

$$\lambda = -\frac{\sigma^2}{2\sigma_1^2+\sigma_\varepsilon^2} j_1(\sigma_1^2,\sigma_\varepsilon^2,m_1,r),$$

and that $\lambda \geq 0$ if, and only if, $\sigma_1^2 \geq \overline{\Sigma}_1^2$. Hence, $a_1^* = 0$ is a solution to (22) only when $\sigma_1^2 \geq \overline{\Sigma}_1^2$. Since, by Proposition 12, a solution to (22) always exists, we then have that the solutions to (22) are positive when $\sigma_1^2 \in (0, \overline{\Sigma}_1^2)$.

We now show that there exists $\widetilde{\Sigma}_1^2 = \widetilde{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r) \geq \overline{\Sigma}_1^2$ such that $a_1^* = 0$ is the only solution to (22) when $\sigma_1^2 \geq \widetilde{\Sigma}_1^2$. For this, let $\mathcal{J}_1(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) = (1+a)^{-1} J_1(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)$. Note that (22) has no positive solution if $\mathcal{J}_1(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) > 0$ for all $a > 0$. Hence, we are done with this part of the argument if we show that there exists $\widetilde{\Sigma}_1^2 \geq \overline{\Sigma}_1^2$ such that $\mathcal{J}_1(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) > 0$ for all $a > 0$ when $\sigma_1^2 \geq \widetilde{\Sigma}_1^2$. For ease of notation, we omit the dependence of α_2 on σ_1^2 and σ_ε^2 in the remainder of the proof.

To start, let $\widetilde{a}_1$ be the value of a such that $(1+a)^2\sigma_1^2 = \sigma_1^2 + \sigma_\varepsilon^2$. We claim that $\sigma_1^2 \geq \overline{\Sigma}_1^2$ implies that $\mathcal{J}_1(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) > 0$ for all $a \in (0, \widetilde{a}_1)$. First observe from (21) that

$$1 + \frac{\partial \alpha_2}{\partial \widehat{a}_1}(a) \geq g''(\alpha_2(a)) \left\{ g''(\alpha_2(a)) - \frac{(1+a)\sigma_1^2}{[1+(1+a)^2]\,\sigma_1^2 + \sigma_\varepsilon^2} \right\}^{-1}$$

if $a \in (0, \widetilde{a}_1)$. Since $g''(a) \geq g'(a)/a$ for all $a > 0$, we also have that

$$g''(\alpha_2(a)) > \frac{(1+a)^2\sigma_1^2}{\alpha_2(a)\left\{[1+(1+a)^2]\,\sigma_1^2 + \sigma_\varepsilon^2\right\}} + \frac{(1+a)\sigma_1^2}{[1+(1+a)^2]\,\sigma_1^2 + \sigma_\varepsilon^2} \tag{24}$$

for all $a > 0$. Therefore,

$$1 + \frac{\partial \alpha_2}{\partial \widehat{a}_1}(a) \geq \frac{1 + a + \alpha_2(a)}{1+a}$$

if $a \in (0, \widetilde{a}_1)$. In particular, $a \in (0, \widetilde{a}_1)$ implies that $(1+a)^{-1}[1+a+\alpha_2(a)]^2$ is strictly increasing in a and

$$2(1+a)[1+a+\alpha_2(a)]^2 \left[1 + \frac{\partial \alpha_2}{\partial \widehat{a}_1}(a)\right] \geq 2[1+a+\alpha_2(a)]^3.$$

Now note (omitting the algebra) that

$$\begin{aligned}
\frac{\partial \mathcal{J}_1}{\partial a}(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) = \xi'(a) &+ \frac{[1+a+\alpha_2(a)]\sigma_1^2}{(1+a)^2\{[1+(1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2\}} + \frac{2[1+a+\alpha_2(a)]\sigma_1^4}{\{[1+(1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2\}^2} \\
&+ \frac{2(1+a)[1+a+\alpha_2(a)]\sigma_1^4 m_1}{[1+(1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2} - \left[1 + \frac{\partial \alpha_2(a)}{\partial \widehat{a}_1}\right] \frac{A\sigma_1^2}{[1+(1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2} \\
&+ \frac{Br\sigma_1^6}{\{[1+(1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2\}^2} - \frac{4r(1+a)^2[1+a+\alpha_2(a)]^3\sigma_1^8}{\{[1+(1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2\}^3},
\end{aligned}$$

where

$$A = \frac{1}{1+a} + m_1 + \frac{r\sigma_1^4(1+a)[1+a+\alpha_2(a)]^2}{[1+(1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2}$$

and

$$B = [1+a+\alpha_2(a)]^3 + 2(1+a)[1+a+\alpha_2(a)]^2 \left[1 + \frac{\partial \alpha_2}{\partial \widehat{a}_1}(a)\right].$$

From the previous paragraph, we have that

$$\begin{aligned} A &< 1+m_1 - \frac{r(1+a)^2\sigma_1^4}{[1+(1+a)^2]\sigma_1^2+\sigma_\varepsilon^2}\frac{[1+a+\alpha_2(a)]^2}{1+a} < 1+m_1 - \frac{[1+\alpha_2(0)]^2 r\sigma_1^4}{2\sigma_1^2+\sigma_\varepsilon^2} \\ &= \frac{1}{1+\alpha_2(0)}\left[j_1(\sigma_1^2,\sigma_\varepsilon^2,m_1,r) - 1 - \frac{\sigma_1^2}{\sigma_1^2+\sigma_\varepsilon^2}\right] < 0 \end{aligned}$$

if $a \in (0,\widetilde{a}_1)$ and $\sigma_1^2 \geq \overline{\Sigma}_1^2$, and that $B \geq 3[1+a+\alpha_2(a)]^3$ if $a \in (0,\widetilde{a}_1)$. Given that $a \in (0,\widetilde{a}_1)$ also implies that

$$\frac{4r(1+a)^2[1+a+\alpha_2(a)]^3\sigma_1^8}{\{[1+(1+a)^2]\sigma_1^2+\sigma_\varepsilon^2\}^3} \leq \frac{2r[1+a+\alpha_2(a)]^3\sigma_1^6}{\{[1+(1+a)^2]\sigma_1^2+\sigma_\varepsilon^2\}^2},$$

we can conclude that $\partial\mathcal{J}_1(a,\sigma_1^2,\sigma_\varepsilon^2,m_1,r)/\partial a > 0$ for all $a \in (0,\widetilde{a}_1)$ when $\sigma_1^2 \geq \overline{\Sigma}_1^2$. Since $\mathcal{J}_1(0,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) \geq 0$ if $\sigma_1^2 \geq \overline{\Sigma}_1^2$, we then have that $\mathcal{J}_1(a,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) > 0$ for all $a \in (0,\widetilde{a}_1)$ when $\sigma_1^2 \geq \overline{\Sigma}_1^2$, as claimed.

We now establish that for each σ_ε^2, m_1, and r, there exists $\widetilde{\Sigma}_1^2 \geq \overline{\Sigma}_1^2$ with the property that $\mathcal{J}_1(a,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) > 0$ for all $a \geq \widetilde{a}_1$ if $\sigma_1^2 \geq \widetilde{\Sigma}_1^2$. For this, let

$$\begin{aligned} G(a,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) &= \xi(a) - \frac{\sigma_1^2}{\sigma_1^2+\sigma_\varepsilon^2} - \frac{[1+a+\alpha_2(a)]\sigma_1^2}{(1+a)\{[1+(1+a)^2]\sigma_1^2+\sigma_\varepsilon^2\}} \\ &\quad - \frac{[1+a+\alpha_2(a)]\sigma_1^2 m_1}{[1+(1+a)^2]\sigma_1^2+\sigma_\varepsilon^2} + \frac{r(1+a)^4\sigma_1^6}{\{[1+(1+a)^2]\sigma_1^2+\sigma_\varepsilon^2\}^2}. \end{aligned}$$

By construction, $G(a,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) \leq \mathcal{J}_1(a,\sigma_1^2,\sigma_\varepsilon^2,m_1,r)$ for all $a \geq 0$. We claim that G is strictly increasing in a if $a \geq \widetilde{a}_1$. Since $(1+a)\sigma_1^2/\{[1+(1+a)^2]\sigma_1^2+\sigma_\varepsilon^2\}$ is strictly decreasing in a when $a \geq \widetilde{a}_1$, and

$$\frac{[1+a+\alpha_2(a)]\sigma_1^2 m_1}{[1+(1+a)^2]\sigma_1^2+\sigma_\varepsilon^2} = \frac{1+a}{[1+(1+a)^2]\sigma_1^2+\sigma_\varepsilon^2}\frac{1+a+\alpha_2(a)}{1+a},$$

the desired result holds as long as $\zeta(a) = (1+a)^{-1}[1+a+\alpha_2(a)]$ is decreasing in a when $a \geq \widetilde{a}_1$. Now observe from (21) and (24) that

$$\frac{\partial\alpha_2}{\partial\widehat{a}_1}(a) \leq \frac{\alpha_2(a)}{(1+a)^2}\left\{2(1+a)+\alpha_2(a) - \frac{2(1+a)^2[1+a+\alpha_2(a)]\sigma_1^2}{[1+(1+a)^2]\sigma_1^2+\sigma_\varepsilon^2}\right\},$$

and so $\partial\alpha_2(a)/\partial\widehat{a}_1 \leq (1+a)^{-1}\alpha_2(a)$ if $a \geq \widetilde{a}_1$. Given that ζ is decreasing in a if, and only if, $(1+a)\partial\alpha_2(a)/\partial\widehat{a}_1 \leq \alpha_2$, G is indeed strictly increasing in a when $a \geq \widetilde{a}_1$. Consequently, $\mathcal{J}_1(a,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) > 0$ for all $a \geq \widetilde{a}_1$ as long as $G(\widetilde{a}_1,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) \geq 0$. Since $\lim_{\sigma_1^2\to\infty} G(\widetilde{a}_1,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) = \infty$, there exists $\widetilde{\Sigma}_1^2 \in [\overline{\Sigma}_1^2,\infty)$ such that $G(\widetilde{a}_1,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) \geq 0$

for all $\sigma_1^2 \geq \widetilde{\Sigma}_1^2$. By construction, $\sigma_1^2 \geq \widetilde{\Sigma}_1^2$ implies that $\mathcal{J}_1(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) > 0$ for all $a \geq \widetilde{a}_1$, the desired result.

To finish the proof, we show that for each m_1 and r, there exists $\Sigma_\varepsilon^2 \geq 0$ such that $G(\widetilde{a}_1, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) \geq 0$ for all $\sigma_1^2 \geq \overline{\Sigma}_1^2$ when $\sigma_\varepsilon^2 > \Sigma_\varepsilon^2$, and that for each m_1 and σ_ε^2, there exists $\overline{r} \geq 0$ such that $G(\widetilde{a}_1, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) \geq 0$ for all $\sigma_1^2 \geq \overline{\Sigma}_1^2$ when $r > \overline{r}$. First note, since α_2 is uniformly bounded by $\overline{a}_2$, that $\sigma_1^2 \geq \overline{\Sigma}_1^2$ implies that

$$\begin{aligned} G(\widetilde{a}_1, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) &= \xi(\widetilde{a}_1) - \frac{\sigma_1^2}{\sigma_1^2 + \sigma_\varepsilon^2} - \frac{g'(\alpha_2(\widetilde{a}_1, \sigma_1^2, \sigma_\varepsilon^2))}{1 + \widetilde{a}_1}\left[\frac{1}{1+\widetilde{a}_1} + m_1\right] + \frac{r\sigma_1^2}{4} \\ &> \xi\left(\sqrt{1 + \frac{\sigma_\varepsilon^2}{\overline{\Sigma}_1^2}} - 1\right) - 1 - (1+m_1)g'(\overline{a}_2) + \frac{r\overline{\Sigma}_1^2}{4} = \mathcal{G}(\sigma_\varepsilon^2, m_1, r) \end{aligned}$$

Now let $\Sigma_\varepsilon^2(m_1, r) = \inf\{\sigma_\varepsilon^2 > 0 : \mathcal{G}(\sigma_\varepsilon^2, m_1, r) \geq 0\}$. Since $\lim_{\sigma_\varepsilon^2 \to \infty} \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r) = \infty$ by Lemma 1, we have that $\lim_{\sigma_\varepsilon^2 \to \infty} \mathcal{G}(\sigma_\varepsilon^2, m_1, r) = \infty$, which implies that $\Sigma_\varepsilon^2 < \infty$. Likewise, let $\overline{r}(m_1, \sigma_\varepsilon^2) = \inf\{r > 0 : \mathcal{G}(\sigma_\varepsilon^2, m_1, r) \geq 0\}$. Since $\lim_{r \to \infty} r\overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r) = \infty$ by Lemma 1, we also have that $\lim_{r \to \infty} \mathcal{G}(\sigma_\varepsilon^2, m_1, r) = \infty$, which implies that $\overline{r} < \infty$ as well. By construction, $G(\widetilde{a}_1, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) \geq 0$ for all $\sigma_1^2 \geq \overline{\Sigma}_1^2$ when either $\sigma_\varepsilon^2 > \Sigma_\varepsilon^2$ or $r > \overline{r}$, in which case $\widetilde{\Sigma}_1^2 = \overline{\Sigma}_1^2$. This completes the proof. □

As in Proposition 7 in the main text, Proposition 13 shows that for each σ_ε^2, m_1, and r, there exists a cutoff $\overline{\Sigma}_1^2 = \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r)$ such that the worker's effort in period one is positive if $\sigma_1^2 \in (0, \overline{\Sigma}_1^2)$. Unlike Proposition 7, $\sigma_1^2 \geq \overline{\Sigma}_1^2$ does not necessarily imply that the worker exerts no effort in the first period: for each σ_ε^2, m_1, and r, effort is necessarily zero only if $\sigma_1^2 \geq \widetilde{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r)$, where $\widetilde{\Sigma}_1^2$ can be greater than $\overline{\Sigma}_1^2$, and the cutoffs coincide if either the noise in output is high enough or the worker is sufficiently risk averse. The next result shows that there exist cost functions g for which $\widetilde{\Sigma}_1^2 = \overline{\Sigma}_1^2$ regardless σ_ε^2, m_1, and r.

Corollary 1. *For a non–empty set of cost functions g, effort in the first period is positive if $\sigma_1^2 \in (0, \overline{\Sigma}_1^2)$ and zero otherwise.*

Proof: As in the proof of Proposition 13, we omit the dependence of α_2 on σ_1^2 and σ_ε^2 for ease of notation. We divide the argument in several steps. First we show that $g'(1/5) > 6/5$ implies that

$$\alpha_2(a) \leq (1 + a)/5 \tag{25}$$

for all $a \geq 0$. Indeed, by (20), condition (25) is satisfied if

$$g'\left(\frac{1+a}{5}\right) \geq \frac{6}{5}\frac{(1+a)^2\sigma_1^2}{[1+(1+a)^2]\sigma_1^2+\sigma_\varepsilon^2}.$$

The desired result holds since the right side of the above inequality is bounded above by 6/5 and g' is nondecreasing. In what follows we assume that $g'(1/5) > 6/5$.

Let $\widetilde{\widetilde{a}}_1$ be the value of a such that $(1+a)^2\sigma_1^2 = 2(\sigma_1^2+\sigma_\varepsilon^2)$. Now we show (Step 2) that α_2 is strictly increasing in $\widehat{a}_1$ if $\widehat{a}_1 \in (0, \widetilde{\widetilde{a}}_1)$. Recall from (21) that $\partial\alpha_2(a)/\partial\widehat{a}_1 > 0$ if, and only if,

$$\begin{aligned}
&2(1+a)\left\{1-\frac{(1+a)^2\sigma_1^2}{[1+(1+a)^2]\sigma_1^2+\sigma_\varepsilon^2}\right\}+\alpha_2(a,\sigma_1^2,\sigma_\varepsilon^2)\left\{1-\frac{2(1+a)^2\sigma_1^2}{[1+(1+a)^2]\sigma_1^2+\sigma_\varepsilon^2}\right\}\\
&\propto 2(1+a)(\sigma_1^2+\sigma_\varepsilon^2)+\alpha_2(a)\left[\sigma_1^2+\sigma_\varepsilon^2-(1+a)^2\sigma_1^2\right] > 0.
\end{aligned}$$

It is immediate to see from above that $\partial\alpha_2(a)/\partial\widehat{a}_1 > 0$ if $\widehat{a}_1 \in (0, \widetilde{a}_1]$. So, assume that $a \in [\widetilde{a}_1, \widetilde{\widetilde{a}}_1)$, in which case $\partial\alpha_2(a)/\partial\widehat{a}_1 > 0$ if, and only if,

$$\alpha_2(a) \leq \frac{2(1+a)(\sigma_1^2+\sigma_\varepsilon^2)}{(1+a)^2\sigma_1^2-(\sigma_1^2+\sigma_\varepsilon^2)}.$$

Since $\alpha_2(a) \leq (1+a)$ by (25), a sufficient condition for the last inequality is that $(1+a)\sigma_1^2 \leq 3(\sigma_1^2+\sigma_\varepsilon^2)$, which is satisfied by assumption.

The next step (Step 3) consists in showing that $\mathcal{J}_1(a,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) > 0$ for all $a \in (0, \widetilde{\widetilde{a}}_1)$ if $\sigma_1^2 \geq \overline{\Sigma}_1^2$. First note that $(1+a)^{-1}[1+a+\alpha_2(a)]^2$ is strictly increasing in a when $a \in (0, \widetilde{\widetilde{a}}_1)$ if, and only if,

$$2(1+a)\left[1+\frac{\partial\alpha_2}{\partial\widehat{a}_1}(a)\right]-[1+a+\alpha_2(a)] > 0$$

for all $a \in (0, \widetilde{\widetilde{a}}_1)$, which holds by (25) and Step 2. Now observe that $a \in (0, \widetilde{\widetilde{a}}_1)$ implies that

$$\frac{4r(1+a)^2[1+a+\alpha_2(a)]^3\sigma_1^8}{\{[1+(1+a)^2]\sigma_1^2+\sigma_\varepsilon^2\}^3} \leq \frac{8}{3}\frac{r[1+a+\alpha_2(a)]^3\sigma_1^6}{\{[1+(1+a)^2]\sigma_1^2+\sigma_\varepsilon^2\}^2}$$

and that (25) and Step 2 imply that

$$[1+a+\alpha_2(a)]^3+2(1+a)[1+a+\alpha_2(a)]^2\left[1+\frac{\partial\alpha_2}{\partial\widehat{a}_1}(a)\right] \geq \frac{8}{3}[1+a+\alpha_2(a)]^3$$

when $a \in (0, \widetilde{\widetilde{a}}_1)$. Hence, by the proof of Proposition 13, $\partial\mathcal{J}_1(a,\sigma_1^2,\sigma_\varepsilon^2,m_1,r)/\partial a > 0$ for all $a \in (0, \widetilde{\widetilde{a}}_1)$ when $\sigma_1^2 \geq \overline{\Sigma}_1^2$, which implies the desired result.

Let $G(a,\sigma_1^2,\sigma_\varepsilon^2,m_1,r)$ be the same as in the proof of Proposition 13. From Step 3 and the proof of Proposition 13, we have that $\mathcal{J}_1(a,\sigma_1^2,\sigma_\varepsilon^2,m_1,r)>0$ for all $a>0$ when $\sigma_1^2\geq\overline{\Sigma}_1^2$ as long as

$$G(\widetilde{\widetilde{a}}_1,\sigma_1^2,\sigma_\varepsilon^2,m_1,r)=\xi(\widetilde{\widetilde{a}}_1)-\frac{\sigma_1^2}{\sigma_1^2+\sigma_\varepsilon^2}-\frac{g'(\alpha_2(\widetilde{\widetilde{a}}_1,\sigma_1^2,\sigma_\varepsilon^2))}{1+\widetilde{\widetilde{a}}_1}\left[\frac{1}{1+\widetilde{\widetilde{a}}_1}+m_1\right]+\frac{4r\sigma_1^2}{9}\geq 0$$

for all $\sigma_1^2\geq\overline{\Sigma}_1^2$. Observe that (25) implies that

$$g'(\alpha_2(a,\sigma_1^2,\sigma_\varepsilon^2))\leq\frac{6}{5}\frac{(1+a)^2\sigma_1^2}{[1+(1+a)^2]\sigma_1^2+\sigma_\varepsilon^2},$$

and so $g'(\alpha_2(\widetilde{\widetilde{a}}_1,\sigma_1^2,\sigma_\varepsilon^2))\leq 4/5$. Moreover, $1+\widetilde{\widetilde{a}}_1\geq\sqrt{2}$. Thus,

$$G(\widetilde{\widetilde{a}}_1,\sigma_1^2,\sigma_\varepsilon^2,m_1,r)\geq\xi(\sqrt{2})-1-\frac{4}{5\sqrt{2}}(1+m_1)+\frac{4r\sigma_1^2}{9}.$$

From the definition of $\overline{\Sigma}_1^2$, we have that

$$\frac{4r\overline{\Sigma}_1^2}{9}\geq\frac{8(1+m_1)}{9[1+\alpha_2(0,\overline{\Sigma}_1^2,\sigma_\varepsilon^2)]^2}\geq\frac{50(1+m_1)}{81},$$

where the second inequality follows from (25). Therefore,

$$G(\widetilde{\widetilde{a}}_1,\sigma_1^2,\sigma_\varepsilon^2,m_1,r)\geq\xi(\sqrt{2})-1-\underbrace{\left[\frac{50}{81}-\frac{4}{5\sqrt{2}}\right]}_{>0}(1+m_1),$$

and so $G(\widetilde{\widetilde{a}}_1,\sigma_1^2,\sigma_\varepsilon^2,m_1,r)\geq 0$ for all $\sigma_1^2\geq\overline{\Sigma}_1^2$ as long as $\xi(\sqrt{2})\geq 1$. We can then conclude that if g is such that $g'(1/5)>6/5$ and $g'(\sqrt{2})>1+\sqrt{2}$, then $\sigma_1^2\geq\overline{\Sigma}_1^2$ implies that the worker's effort in period one is zero. $\square$

It is straightforward to show that the set of cost functions for which Corollary 1 holds is open, so that this set is robust to small perturbations.[17]

As in the main text, define $\mathcal{A}_1:\mathbb{R}^4_{++}\to\mathbb{R}_+$ to be such that $\mathcal{A}_1(\chi)=A_1(\chi)$. The first comparative statics result we establish is the counterpart of Proposition 8 in the main text.

[17] One introduces a topology on the set of cost functions as follows. Let $\mathcal{C}$ be the set of functions $g:\mathbb{R}_+\to\mathbb{R}$ that are twice differentiable, strictly increasing, and strictly convex, with $g(0)=g'(0)=0$. Now let $\mathcal{I}$ be the set of closed intervals of the form $[0,n]$, where $n\in\mathbb{N}$. For each $I\in\mathcal{I}$, define $p_I:\mathcal{C}\to\mathbb{R}_+$ to be such that $p_I(g)=\sup_{x\in I}|g(x)|+\sup_{x\in I}|g'(x)|+\sup_{x\in I}|g''(x)|$. The set $\mathcal{P}=\{p_I\}_{I\in\mathcal{I}}$ is a separating family of semi–norms. The topology on $\mathcal{C}$ is the topology induced by $\mathcal{P}$, see Rudin (1991). Since convergence in the topology induced by $\mathcal{P}$ implies uniform convergence of cost functions and their derivatives in the compact subsets of $\mathbb{R}_+$, it is easy to see that if g is a cost function such that $g'(1/5)>6/5$ and $g'(\sqrt{2})>1+\sqrt{2}$, then there exists a neighborhood $\mathcal{N}$ of g such that if $g_0\in\mathcal{N}$, then $g_0'(1/5)>6/5$ and $g_0'(\sqrt{2})>1+\sqrt{2}$.

Proposition 14. *Fix σ_ε^2 and m_1. For all $\Sigma_1^2 > 0$, there exists $\overline{r} > 0$ such that if $r \leq \overline{r}$, then $\mathcal{A}_1$ is strictly increasing in σ_1^2 when $\sigma_1^2 \in (0, \Sigma_1^2)$.*

Proof: TO BE COMPLETED.

□

The next comparative statics result is the counterpart of Proposition 9 in the main text.

Proposition 15. *Fix σ_ε^2, m_1, and r. There exist $0 < \Sigma_{10}^2 < \Sigma_{11}^2 < \widetilde{\Sigma}_1^2$ such that $\mathcal{A}_1$ is strictly increasing in σ_1^2 when $\sigma_1^2 \in (0, \Sigma_{10}^2)$, and a_1^* is unique and strictly decreasing in σ_1^2 when $\sigma_1^2 \in (\Sigma_{11}^2, \widetilde{\Sigma}_1^2)$.*

Proof: TO BE COMPLETED. □

The third comparative statics result is the counterpart of Proposition 10 in the main text. Differently from Proposition 10, the result below does not include a statement about the uniqueness of effort in the first period. Let $\overline{\Sigma}_1^2(0, m_1, r) = \lim_{\sigma_\varepsilon^2 \to 0} \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r)$. Then,

$$\overline{\Sigma}_1^2(0, m_1, r) = \frac{1}{r[1 + a_2^0]^3} \left\{ 2 + [1 + a_2^0](1 + m_1) \right\}.$$

Since $\overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r)$ is strictly increasing in σ_ε^2, we have that $\Sigma_1^2 < \overline{\Sigma}_1^2(0, m_1, r)$ implies that $\Sigma_1^2 < \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r)$ for all σ_ε^2. Recall that a necessary condition for $\mathcal{A}_1$ to be strictly decreasing in σ_1^2 is that $a_{1,\min}^*$ is greater than zero.

Proposition 16. *For each m_1, r, and $\Sigma_1^2 \in (0, \overline{\Sigma}_1^2(0, m_1, r))$, there exists $\Sigma_\varepsilon^2 > 0$ such that if $\sigma_\varepsilon^2 < \Sigma_\varepsilon^2$, then $\mathcal{A}_1$ is strictly decreasing in σ_1^2 when $\sigma_1^2 \in (\Sigma_1^2, \overline{\Sigma}_1^2)$.*

Proof: TO BE COMPLETED. □

We know from Proposition 13 that for each m_1 and r there exists $\Sigma_\varepsilon^2 > 0$ such that $\overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r) = \widetilde{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r)$ when $\sigma_\varepsilon^2 > \Sigma_\varepsilon^2$. Since $\lim_{\sigma_\varepsilon \to \infty} \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r) = \infty$, we then have that an increase in σ_ε^2 can (and eventually does) lead to higher effort in the first period if $\sigma_1^2 > \widetilde{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r)$. Thus, as in the main text, an increase in output noise can have a positive impact on career concerns incentives. However, the counterpart of Proposition 11 in the main text does not always hold. We need to assume that the cost function g is such that $\overline{\Sigma}_1^2 = \widetilde{\Sigma}_1^2$ regardless of σ_ε^2, m_1, and r. In this case, $\sigma_1^2 > \overline{\Sigma}_1^2(0, m_1, r)$ implies that implicit incentives are not maximal when the noise in output is minimized. Since $\lim_{r\to\infty}\overline{\Sigma}_1^2(0,m_1,r)=0$, we then have the following result.

Proposition 17. *For a non–empty set of cost functions we have that for each m_1 and σ_1^2, there exists $\overline{r}>0$ such that if $r\geq\overline{r}$, then implicit incentives are not maximized when the noise in output is minimized.*