Career Concerns: A Human Capital Perspective[*]

Braz Camargo[†]

São Paulo School of Economics—FGV

Elena Pastorino

UAB—Barcelona GSE

University of Iowa

May 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 accumulation. 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.

JEL Codes: D82, D83, D86, J24, J41

Keywords: Career Concerns, Reputation, Human Capital

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 compensation reflects reputation in the labor market, that is, the market's assessment of an individual's talent, then a desire to influence market beliefs about ability can motivate the individual to exert effort, even in the absence of explicit contractual links between pay and performance. This dynamic incentive mechanism originating from career concerns was first formalized in Holmström (1982, 1999). Career concerns are known to be central to explaining incentives for performance in firms (Baker, Jensen, and Murphy (1988) and Prendergast (1999)). They are also important for understanding behavior 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)), the political arena (Persson and Tabellini (2002), Mattozzi and Merlo (2008), and Martinez (2009a)), and the public sector (Dewatripont, Jewitt, and Tirole (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 importance of human capital accumulation through experience in the labor market for wage and productivity growth has been widely documented.[1] A prominent explanation for the mechanism through which individuals acquire human capital in the labor market is the accumulation of new productive skills while working, the so–called process of learning–by–doing.[2]

---

[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. An important 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. Due to data limitations and identification issues, few studies have attempted to distinguish empirically between the importance of learning–by–doing and on–the–job training for explaining returns to firm tenure or 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. The presence of learning–by–doing has also been investigated by Foster and Rosenzweig (1995), who analyze the pattern of adoption and the profitability of high–yield seed varieties in India during the Green Revolution. They provide evidence of learning–by–doing and of important variation in the time pattern of its return. More recently, Nagypál (2007) documents the presence of learning–by–doing based on an equilibrium model of worker turnover in the labor market that embeds learning about ability and learning–by–doing.

Naturally, the possibility for an individual's labor input to affect future output gives rise to an additional channel through which effort influences reputation in the labor market, which can have implications for how career concerns discipline moral hazard.

To explore the impact of human capital acquisition on career concerns, in this paper we allow an individual's effort to influence his future productivity and examine how this process of learning–by–doing affects implicit incentives for performance.[3] We show that several insights from the standard career concerns model can be reversed in the presence of human capital accumulation. Specifically, contrary to the case without accumulation of human capital, 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. Moreover, increasing the precision with which output is measured can have an adverse impact on implicit incentives. This last result stands in sharp contrast to the intuition from moral hazard models that better inference about effort improves incentives for performance.

The starting point of our analysis is a finite horizon version of the career concerns model in Holmström (1999) (HM henceforth) in which an individual can be either risk neutral or risk averse. In this benchmark case, an individual's effort is increasing in 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 implicit incentives for performance. This is the standard career concerns effect identified by Holmström. Also as in HM, an increase in output noise, by reducing an individual's ability to influence his reputation, weakens implicit incentives. However, for a given level of uncertainty about ability, unlike in HM, 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 when idiosyncratic shocks to ability keep uncertainty about ability constant.

We then contrast our 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

---

[3]Our model does not feature a labor–leisure choice on the part of individuals. Nevertheless, it 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 he accumulates.

features of 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 learning–by–doing may depend on an individual's ability; the interpretation of talent as ability to learn has been discussed, among others, by 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 compensation, in all subsequent periods. However, in the presence of human capital acquisition, an individual's effort also affects his future productivity, which entails an additional impact of effort on reputation, the magnitude of which depends on the return to learning–by–doing. It is not surprising, therefore, that when the return to learning–by–doing varies with an individual's age, effort need not decrease 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. Higher uncertainty about ability still affords an individual greater latitude to manipulate his reputation, and thus can improve implicit incentives for performance. 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 in the uncertainty about talent. Hence, unlike in the case without accumulation of human capital, greater uncertainty about ability can weaken rather than strengthen career concerns incentives when individuals are risk averse. The counterpart of this result is that an increase in output noise, by making performance less informative about talent, reduces the variability of future reputation and thus the risk of exerting effort, which can have a positive effect on career concerns incentives.

In the paper we derive conditions, which depend on an individual's degree of risk aversion, under which an increase in either uncertainty about ability or the precision of output measurement have an adverse impact on career concerns incentives. Note that an individual's effort and ability become complementary when the return to learning–by–doing depends on ability. As in Dewatripont, Jewitt, and Tirole (1999b), this complementarity can lead to multiple equilibria. We take an agnostic view towards equilibrium selection and derive

4

set–valued comparative statics results: we consider how the set of equilibrium effort choices responds to changes in uncertainty about ability and noise in output.[4] Overall, our findings highlight that 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 career concerns incentives and both the uncertainty about talent and the precision of output measurement.

The rest of the paper is organized as follows. In the remainder of this section, we discuss the related literature. In Section 2, we consider the benchmark career concerns case, while in Sections 3 and 4, we examine the case in which individuals can accumulate human capital through a learning–by–doing component of effort. We study the case in which the return to learning–by–doing depends on age in Section 3, and the case in which the return to learning–by–doing varies with an individual's ability in Section 4. For simplicity, we assume risk neutrality in Sections 2 and 3. As we discuss in Section 4, the results in Sections 2 and 3 extend without any modifications to the case in which workers are risk averse. We conclude in Section 5 with a discussion of our results. Appendix A collects omitted proofs, Appendix B contains supplementary material, and Appendix C collects omitted details.

Related Literature. Beginning with Holmström's seminal work, a growing literature has explored extensions of the basic careers concerns model. For instance, Kovrijnykh (2007) examines career concerns when ability is career specific, that is, when individuals can leave their current career to collect a fixed outside option. Kőszegi and Li (2008) study career concerns when agents are also heterogenous in their responsiveness to incentives. Martinez (2009b) analyzes a model of career concerns with job assignment. Unlike these papers, we consider an environment in which effort has an intertemporal effect on output and examine the implications of this generalization for career concerns incentives.

Other authors have examined how changes in information structure affect career concerns incentives. Dewatripont, Jewitt, and Tirole (1999a) (DJT henceforth) compare different information structures—in their setting, maps from ability and labor input into observable

---

[4]We apply monotone comparative statics techniques to obtain our comparative statics results; see Milgrom and Roberts (1990) and Milgrom and Shannon (1994).

outcomes—in terms of their impact on the strength of career concerns incentives.[5] 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.[6] The formulation in DJT is quite general and allows for learning–by–doing (see remark in p. 185 of their paper). However, since they restrict attention to a two–period setting, learning–by–doing does not influence career concerns incentives. Indeed, as it will become apparent from our analysis, accumulation of human capital can only affect implicit incentives if an individual participates in the labor market for three or more periods. Hence, the nature of our result that more precise information about output can have an adverse impact on career concerns is quite different from the corresponding result in DJT.[7]

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

---

[5]See Cabrales, Gossner, and Serrano (2010) for recent work on general criteria for the comparison of information structures.

[6]Other papers that address the same question in different environments are Jeon (1996), Ortega (2003), Blanes–i–Vidal (2007), and Bar–Isaac and Ganuza (2008). Jeon (1996) considers team production and analyzes how group composition and output sharing within team members influence implicit incentives. Ortega (2003) and Blanes–i–Vidal (2007) analyze the impact of delegation on career concerns. Bar–Isaac and Ganuza (2008) explore the importance of different recruitment and training strategies for career concerns.

[7]In particular, the result in DJT that less noise in output can have a negative impact on implicit incentives cannot arise in the scalar additive–normal framework, which is the environment that we consider.

$h_\varepsilon < \infty$. 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.[8] We say the strategy σ is uncontingent if σ_t is constant in Z_t for all $t \in \{1, \ldots, T\}$.

---

[8]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 C 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.

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 $w_t = \mathbb{E}[y_t|\sigma, y^t]$.

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.[9] 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

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

---

[9]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.

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.

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

$$m_s(y_1, \ldots, y_{s-1}) = m_t(y_1, \ldots, y_{t-1}) \prod_{\tau=t}^{s-1} \mu_\tau + (1 - \mu_t) \prod_{\tau=t+1}^{s-1} \mu_\tau [y_t - a_t(y^t|\widehat{\sigma})]$$

$$+ \sum_{q=t+1}^{s-1} (1 - \mu_q) \prod_{\tau=q+1}^{s-1} \mu_\tau (y_q - \widehat{a}_q),$$

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, and Risk 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

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

We can then prove the following result.

10

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 increases with the uncertainty 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 in 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 level of uncertainty 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.[10]

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.*

---

[10]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 with Time–Dependent Return

In this section and the next, we analyze 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 alter how career concerns discipline moral hazard. The first is that the return to learning–by–doing can change with the worker's age. The second is that the return to learning–by–doing can depend on the worker's ability. We consider the first possibility in this section, and the second possibility in the next section.

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

---

[11]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.

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

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

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

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

Equation (9) shows the marginal impact of an increase in effort in period t on the worker's expected wage in period $s \geq t+1$. 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

13

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 $s \geq t+2$ (so we need $t+1 \leq T$), 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 $t = T$. 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 establishes 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, the precision effect holds. Moreover, a reduction in the noise in output increases effort in all periods but the last one. However, the left side of (10), and thus the worker's effort, need not diminish with the worker's age as the uncertainty about his ability decreases over time. Whether this happens depends on how the return to learning–by–doing is affected by age. A necessary condition for the worker's effort to not diminish with age is that the return to learning–by–doing does not decrease over time. This is typically the case, for instance, at low levels of labor market 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.*

14

4 Learning–by–Doing with Uncertain Return

We now study the case in which the return to learning–by–doing depends on the worker's ability, and so is uncertain. Suppose 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 $\alpha_0 \geq 0$ (the worker's human capital in period one is uncertain if $\alpha_0 > 0$). 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.[12] 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 this section.

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 is common in the literature (see, for instance, Holmström and Milgrom (1987) and Gibbons and Murphy (1992)). It implies that the worker is indifferent between any pair of wage and effort sequences that have the same discounted lifetime value, as if the worker had access to perfect capital markets and so could smooth his consumption over time completely. This property of preferences allows for a transparent analysis of the role of uncertainty about ability for career concerns incentives when the worker is risk averse.

It is simple to adapt the analysis in Sections 2 and 3 to the case in which the worker's preferences are given by (11). In both instances, since there is no uncertainty about the

---

[12]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. Since the aim of this section is to examine how uncertainty about ability affects career concerns incentives when the return to learning–by–doing depends on ability, there is no qualitative gain in assuming shocks to ability.

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).

We begin with a preliminary discussion. We then proceed to the main analysis and conclude with a discussion of the assumption that $T = 3$. In what follows we assume that g' is convex and that $\xi(a) = (1 + a)^{-1} g'(a)$ is such that $\lim_{a \to \infty} \xi(a) = \infty$. Note that ξ is nondecreasing when g' is convex.

4.1 Preliminary Discussion

As in Section 3, 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, given that $T = 3$, uncertainty in 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, that is,

$$y_t = k_t + \varepsilon_t.$$

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 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 when he is risk averse.

To see the negative impact of uncertainty about ability on career concerns incentives, 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). Then, $y_1 = 0$ and

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

for $t \geq 1$. 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, $w_3 = \widehat{a}_1(y_2/\widehat{a}_1) = y_2$. This 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; recall that if x is normally distributed with mean m and variance σ^2, then $\mathbb{E}[\exp(-rx)] = \exp(-rm + r^2\sigma^2/2)$. 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 benefit (in monetary terms) from effort in period one, while the term

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

17

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 results: (i) there exists a unique equilibrium in which the worker exerts effort in period one if, and only if, $m_1 > 0$; (ii) the worker's effort in period one in this unique equilibrium 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 as follows. 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.

4.2 Main Analysis

We know from the preliminary discussion that uncertainty in the return to learning–by–doing only matters for the choice of effort in the first period. Given this, we assume that

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

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 zero regardless of ϑ_3. The assumption that $\vartheta_2 = 0$ does not alter the substance of our results, but it simplifies the analysis to some extent, as it implies that the worker's effort in period two is zero. Instead, when $\vartheta_2 > 0$, effort in period two is not zero, and it cannot be determined independently of the choice of effort in period one. Intuitively, when $\vartheta_2 > 0$, effort in period

18

one affects the worker's productivity in period two, and thus his return to effort in this latter period. This implies a richer dynamics for the worker's choice of effort and it makes for a more subtle analysis. We examine the case in which $\vartheta_2 > 0$ in Appendix B.

We begin with the equilibrium characterization and then proceed to the comparative statics analysis. In what follows we restrict attention to the case in which $m_1 > 0$, so that the expected return to effort in period one is positive.

4.2.1 Equilibrium Characterization

We know from above that effort in periods two and three is always zero. Let $\widehat{a}_1$ be the market's conjecture about the worker's effort in period one.[13] Moreover, let σ_2^2 be such that $\sigma_2^2 = \sigma_1^2 \sigma_\varepsilon^2 / (\sigma_1^2 + \sigma_\varepsilon^2)$, 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 $w_3 = (1 + \widehat{a}_1)\mathbb{E}[\theta|y_1, y_2]$, where, also by a standard argument, we have that

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

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

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

---

[13]Note that, unlike in the example in Subsection 4.1, we do not require $\widehat{a}_1 > 0$. The reason for this is that in the example, the market expects $y_2 = 0$ if $\widehat{a}_1 = 0$, in which case market beliefs are not well–defined when $y_2 \neq 0$. The latter, however, happens with probability one if $a_1 > 0$. Thus, in the example, in order to verify whether zero effort in period one is an optimal response to the conjecture that $\widehat{a}_1 = 0$, one needs to use some refinement to compute beliefs off the path of play. This problem does not appear in the main analysis since the noise in output guarantees that every output level is possible on the path of play.

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. The term

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

on the left side of (13) is the worker's marginal benefit from effort in period one, whereas the term

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

is the worker's marginal cost of effort a in the same period due to his risk aversion. Straight-forward 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

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

$$- \frac{(1+a)^2\sigma_1^2}{[1+(1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2} \left\{ m_1 - \frac{r(1+a)^3\sigma_1^4}{[1+(1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2} \right\} = 0, \quad (14)$$

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 (14) 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 (14) has a positive solution by the intermediate value theorem. Thus, an equilibrium always exists.

Since $MB(\widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, m_1)$ is strictly increasing in $\widehat{a}_1$, the equilibrium may not be unique, though. Indeed, the fact that the worker's marginal benefit of effort depends positively on the market's expectation about his effort implies that $\alpha_1(\widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)$ can be nondecreasing in $\widehat{a}_1$, in which 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]

---

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

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

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 $\widehat{a}_1$.

We finish the equilibrium characterization by establishing necessary and sufficient conditions on σ_1^2 for a solution to (14) to be positive. For this, let

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

In Lemma 1 in Appendix A we show that for each σ_ε^2, m_1, 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, $\overline{\Sigma}_1^2$ is strictly increasing in σ_ε^2, and such that $\lim_{\sigma_\varepsilon^2 \to \infty} \overline{\Sigma}_1^2 = \infty$, $\lim_{r \to 0} \overline{\Sigma}_1^2 = \infty$, and $\lim_{\sigma_\varepsilon^2 \to 0} \overline{\Sigma}_1^2 = 2(3 + m_1)/r$; we omit the dependence of $\overline{\Sigma}_1^2$ on σ_ε^2, m_1, and r when convenient. The properties of $\overline{\Sigma}_1^2$ will prove useful below. We then have the following result, the proof of which is in Appendix A.

Proposition 6. *The worker's effort in period one is positive if, and only if, $\sigma_1^2 \in (0, \overline{\Sigma}_1^2)$.*

A consequence of Proposition 6 is that holding everything else constant, implicit incentives are muted when uncertainty about ability is sufficiently high. 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 to be 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 second 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.

4.2.2 Comparative Statics

Comparative statics analysis in the presence of multiple equilibria can be problematic. In what follows we take an agnostic view towards equilibrium selection and establish comparative statics results that apply to the entire set of period one effort choices. These results not only illustrate the dependence of effort on the parameters of interest, namely the level of uncertainty about ability and the noise in output measurement, but could also be useful to test for properties of the human capital accumulation function. We return to this point in our concluding remarks.

We begin with some definitions borrowed from the literature on monotone comparative statics (see Milgrom and Shannon (1994)). 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 $y_1 \in S_1$ such that $y_1 < y_2$ for all $y_2 \in S_2$, and there exists $y_2 \in S_2$ such that $y_2 > y_1$ for all $y_1 \in S_1$. 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.

For each $\chi = (\sigma_1^2, \sigma_\varepsilon^2, m_1, r) \in \mathbb{R}_{++}^4$, 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 period one.[15] Now let $\mathcal{A}_1 : \mathbb{R}_{++}^4 \rightrightarrows \mathbb{R}_+$ be such that $\mathcal{A}_1(\chi) = A_1(\chi)$. By construction, $\mathcal{A}_1$ is the correspondence that takes the set of parameters of the model into the set of equilibrium effort choices in the first period. We first consider the dependence of $\mathcal{A}_1$ on σ_1^2. Then, we consider the dependence of $\mathcal{A}_1$ on σ_ε^2.

Uncertainty and Career Concerns. As in the benchmark case, an increase in uncertainty about ability increases the worker's return to exerting effort, which is the force behind the precision effect. Indeed, the worker's marginal benefit of effort is strictly increasing in σ_1^2. However, an increase in uncertainty about ability also increases the variance in the return to learning–by–doing, which weakens incentives for effort. In fact, the worker's marginal cost of effort due to risk aversion is increasing in σ_1^2 as well. Thus, unlike the benchmark case, the effect of an increase in σ_1^2 is ambiguous.

Since $MC_{\text{risk}}(a, \widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, r)$ is proportional to r, the positive effect on implicit incentives of an increase in σ_1^2 dominates the negative effect when the worker's risk aversion is small. Thus, the effect on implicit incentives of an increase in σ_1^2 is the same as in the benchmark case when r is small. Formally, we have the following result; see Appendix A for a proof.

---

[15]Standard arguments from general equilibrium theory (see Mas–Colell, Whinston, and Green (1995)) can be used to prove that there exists an open and full measure subset of $\mathbb{R}_{++}^4$ such that $A_1(\chi)$ is finite for all χ in this subset. We include a proof of this result in Appendix C for completeness.

Proposition 7. *Fix σ_ε^2 and m_1. For all $\Sigma_1^2 > 0$, there exists $\bar{r} > 0$ such that if $r \leq \bar{r}$, then $\mathcal{A}_1$ is strictly increasing in σ_1^2 if $\sigma_1^2 \in (0, \Sigma_1^2)$.*

Now observe that when σ_1^2 is small, $MB(\widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, m_1)$ is roughly proportional to σ_1^2, while $MC_{\text{risk}}(a, \widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, r)$ is roughly proportional to σ_1^6.[16] Hence, when σ_1^2 is small, the rate at which the first term increases in σ_1^2 is greater than the rate at which the second term increases in σ_1^2. Thus, an increase in uncertainty about ability has a positive impact on implicit incentives when σ_1^2 is small. We now argue that the opposite happens when σ_1^2 is large. From Proposition 6, the worker's choice of effort in period one is uniquely defined when $\sigma_1^2 \geq \overline{\Sigma}_1^2$. Hence, this choice of effort should also be unique when $\sigma_1^2 < \overline{\Sigma}_1^2$, as long as σ_1^2 is close enough to $\overline{\Sigma}_1^2$. Given that H is continuous, it must then be the case 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. Since the worker's effort in period one is positive when $\sigma_1^2 < \overline{\Sigma}_1^2$, we can then conclude that this effort eventually becomes strictly decreasing in σ_1^2 as σ_1^2 increases to $\overline{\Sigma}_1^2$. The following result formalizes this discussion; see the Appendix A for a proof.

Proposition 8. *Fix σ_ε^2, m_1, and r. There exist $\Sigma_{10}^2, \Sigma_{11}^2 \in (0, \overline{\Sigma}_1^2)$, with $\Sigma_{10}^2 < \Sigma_{11}^2$, such that $\mathcal{A}_1$ is strictly increasing in σ_1^2 if $\sigma_1^2 \in (0, \Sigma_{10}^2)$ and single–valued and strictly decreasing in σ_1^2 if $\sigma_1^2 \in (\Sigma_{11}^2, \overline{\Sigma}_1^2)$.*

Proposition 8 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 decreases, the lower bound on σ_1^2 above which a^* is unique and strictly decreasing in σ_1^2 converges to zero. For this, recall that $\lim_{\sigma_\varepsilon^2 \to 0} \overline{\Sigma}_1^2 = 2(3 + m_1)/r$ and $\overline{\Sigma}_1^2$ is strictly increasing in σ_ε^2, so that $\Sigma_1^2 < 2(3 + m_1)/r$ implies that $\Sigma_1^2 < \overline{\Sigma}_1^2$ regardless of σ_ε^2.

Proposition 9. *Fix m_1, r, and $\Sigma_1^2 \in (0, 2(3 + m_1)/r)$. There exists $\Sigma_\varepsilon^2 > 0$ such that if $\sigma_\varepsilon^2 \in (0, \Sigma_\varepsilon^2)$, then $\mathcal{A}_1$ is single–valued and strictly decreasing in σ_1^2 for all $\sigma_1^2 \in (\Sigma_1^2, \overline{\Sigma}_1^2)$.*

The proof of Proposition 9 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

---

[16]Precisely, $MB(\widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, m_1) = O(\sigma_1^2)$ and $MC_{\text{risk}}(a, \widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, r) = O(\sigma_1^6)$ when σ_1^2 converges to zero.

for effort. 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 to effort. On the other hand, the rate at which the marginal cost of effort due to risk aversion increases in σ_1^2 becomes larger as σ_ε^2 gets smaller. Indeed, $\partial^2 MC_{\mathrm{risk}}/\partial\sigma_\varepsilon^2\partial\sigma_1^2 < 0$. Thus, the negative effect of an increase in σ_1^2 dominates the positive effect when σ_ε^2 is small.

Risk and Career Concerns. As in the case of a change in σ_1^2, a change in σ_ε^2 also has opposite effects on the marginal benefit of effort and the marginal cost of effort due to risk aversion. Indeed, an increase in σ_ε^2 makes output less informative about the worker's ability, which reduces his return to exerting effort: $MB(\widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, m_1)$ is strictly decreasing in σ_ε^2. On the other hand, 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, in the same way as a decrease in σ_1^2, reduces the uncertainty in the return to learning–by–doing, which is good for incentives: $MC_{\mathrm{risk}}(a, \widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2, r)$ is strictly decreasing in σ_ε^2. Consequently, an increase in the noise on output also has an ambiguous effect on career concerns incentives.

Since a decrease in σ_ε^2 acts on incentives for effort in the same way as an increase in σ_1^2, the same considerations made in the above discussion about the effect of changes in σ_1^2 on $\mathcal{A}_1$ apply. In particular, the positive effect of an increase in output noise is small compared to the negative effect when the worker's risk aversion is small. The same is true when uncertainty about ability is small. The following result formalizes this discussion; see Appendix A for a proof.

Proposition 10. *Fix σ_1^2 and m_1. For all $\Sigma_\varepsilon^2 > 0$, there exists $\underline{r} > 0$ such that if $r \in (0, \underline{r})$, then $\mathcal{A}_1$ is strictly decreasing in σ_ε^2 if $\sigma_\varepsilon^2 \in (0, \Sigma_\varepsilon^2)$. Now fix m_1 and r. There exists $\Sigma_1^2 > 0$ with the property that if $\sigma_1^2 \in (0, \Sigma_1^2)$, then there exists $\Sigma_\varepsilon^2 > 0$ such that $\mathcal{A}_1$ is strictly decreasing in σ_ε^2 when $\sigma_\varepsilon^2 \in (0, \Sigma_\varepsilon^2)$.*

Now observe that since $\overline{\Sigma}_1^2$ is strictly increasing in σ_ε^2, with $\lim_{\sigma_\varepsilon^2 \to \infty} \overline{\Sigma}_1^2 = \infty$, an immediate consequence of Proposition 6 is that an increase in σ_ε^2 can, and eventually does, increase

effort when uncertainty about ability is high enough for effort to be initially zero. In partic-
ular, implicit incentives are increasing in σ_ε^2 when output noise is small enough; recall that
if $\sigma_1^2 > 2(3 + m_1)/r$, then $\sigma_1^2 > \overline{\Sigma}_1^2$ when σ_ε^2 is small enough. Since $\lim_{r \to \infty} 2(3 + m_1)/r = 0$,
we then have the following result; see Appendix A for a proof.

Proposition 11. *An increase in σ_ε^2 can, and eventually does, increase effort when uncer-
tainty about ability is sufficiently high. Moreover, for each m_1 and σ_1^2, there exists $\overline{r} > 0$ with
the property that if $r > \overline{r}$, then there exists $\Sigma_\varepsilon^2 > 0$ such that $\mathcal{A}_1$ is single–valued, increasing,
and non–constant when $\sigma_\varepsilon^2 \in (0, \Sigma_\varepsilon^2)$.*

4.3 Longer Time Horizons

The main challenge in extending the analysis in this subsection to the case in which the
worker lives for $T > 3$ periods is to obtain an analytical characterization of equilibria. This
difficulty already emerges when $T = 4$. Indeed, with a few straightforward modifications,
the analysis conducted in Subsection 4.2.1 provides a complete characterization of behavior
in all periods but the first when

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

with $\vartheta_1 = \vartheta_2 = 1$ and $\vartheta_3 = \vartheta_4 = 0$. However, 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 effort 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 in a period affects his reputation:
(i) by affecting current output, effort influences the worker's reputation in all subsequent
periods; (ii) by affecting productivity from next period on, effort has an additional impact
on the worker's reputation in all periods after the next one. 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

25

noise) increases the risk associated with exerting effort, which has a negative impact on implicit incentives. These observations suggest that the comparative statics results obtained when $T = 3$ survive in the more general case when $T \geq 4$.

5   Concluding Remarks

In this paper we extend the standard career concerns model to allow for the possibility that an individual's human capital can improve over time through a learning–by–doing component of effort. We show that this form of human capital acquisition has substantive implications for the ability of career concerns incentives to discipline moral hazard. Namely, standard results about 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, when individuals are heterogeneous in their ability to accumulate human capital, risk aversion becomes central to the relationship between the power of career concerns incentives and the degree of uncertainty and risk in the environment.

When the return to learning–by–doing depends on an individual's ability, a complementarity arises between ability and effort, which can lead to multiple equilibria. Echenique and Komunjer (2009) derive testable implications of the complementarity between explanatory and dependent variables for a large class of economic models.[17] In our setting, their approach could be useful to test for some characteristics of the human capital accumulation process, such as uncertainty in the returns to learning–by–doing.[18] The exploration of the empirical content of a version of the career concerns model that allows for human capital acquisition and complementarity between ability and effort is the focus of current research.

A basic tenet of moral hazard models is the difficulty of gauging an individual's labor input from the observation of his output. A standard result in the literature is that incentive problems worsen as output measurement becomes less precise. In particular, contractual

---

[17]Specifically, they show that in the presence of complementarity, monotone comparative statics arguments allow the derivation of testable restrictions on the conditional quantiles of the dependent variable.

[18]Intuitively, since complementarity between ability and effort only arises in our setting when the return to human capital acquisition depends on ability, testing for complementarity would amount to testing whether the return to human capital acquisition is uncertain, conditional on learning–by–doing being present.

arrangements become lower–powered as noise in output increases. Our comparative statics result that implicit incentives can be *stronger* in situations in which performance measures are *less accurate* implies that greater risk does not necessarily lead to inferior incentives. So, the absence of explicit contracts may just be an indication that incentive conflicts are not too severe. Thus, our result on the impact of noise on career concerns can help explain both the mixed evidence on the importance of noise for incentives (see Gibbs, Merchant, van der Stede, and Vargus (2009) and the references therein) and the fact that explicit performance measures are seldom used in employment contracts (see Prendergast and Topel (1993)).[19] Moreover, the result that the relationship between the power of implicit incentives and the noise in output measurement can be either positive or negative suggests that estimates of either no relation or a positive relation between risk and explicit incentives do not necessarily invalidate the implications of moral hazard theory (see Prendergast (1999, 2002a, 2002b)).[20]

Our analysis abstracts from the possibility of explicit contracting. In related work (Camargo and Pastorino (2010)), we examine optimal contracting in the presence of career concerns and analyze the relationship between explicit and implicit incentives for performance when individuals can accumulate human capital. In that paper, we 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 (see Murphy (1999) and Prendergast (1999) for a review of the debate on performance–pay elasticities).

---

[19]For instance, one reason why CEO compensation is more closely tied to performance than the compensation of middle managers could be that performance measures for middle managers are noisier (and their risk aversion possibly higher, as risk aversion typically decreases in wealth), which would imply that implicit incentives for middle managers can be stronger than for CEOs. See Milkovich and Newman (2008) for a discussion of the fact that dimensions like team–work, joint divisional–level output indicators, and aspects of performance harder to verify, like a manager's attitude towards clients, frequently provide the basis for the subjective assessment of middle managers' contribution to firm value rather than explicit output measures. Besides, as cited in Kőszegi and Li (2008), according to the Employee Compensation Survey of the BLS (2004), the use of work hours, attendance, and volunteering for difficult assignments to determine pay at large and medium sized enterprises has increased from 17% of employees in 1983 to 42% in 1997. This fact suggests that when firms have the ability to choose how to monitor performance, in some circumstances they prefer to motivate workers by using noisier measures of effort than output indicators.

[20]The fact that contract choice is not exogenous could be a further source of difficulty in interpreting the available evidence. For instance, see Paarsch and Shearer (1999) for an illustration of the challenge in recovering the sign of the correlation between measures of productivity and incentives in the presence of unmeasured characteristics affecting the choice of contract.

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] Baker, G., M. Gibbs and B. Hölmstrom (1994): "The Internal Economics of the Firm: Evidence from Personnel Data", *Quarterly Journal of Economics* 109(4), 881–919.

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

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

[7] Cabrales, A., O. Gossner, and R. Serrano (2010): "Entropy and the Value of Information for Investors", mimeo.

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

[9] 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.

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

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

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

[13] 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.

[14] Gibbons, R., and K. J. Murphy (1992): "Optimal Incentive Contracts in the Presence of Careers Concerns: Theory and Evidence", *Journal of Political Economy* 100(3), 468–505.

[15] Gibbs, M. J., K. A. Merchant, W. A. van der Stede, and M. E. Vargus (2009): "Performance Measure Properties and Incentive System Design", *Industrial Relations* 48(2), 237–264.

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

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

[18] 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.

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

[20] Holmström, B., and P. Milgrom (1987): "Aggregation and Linearity in the Provision of Intertemporal Incentives", *Econometrica* 55(2), 303–328.

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

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

[23] Kőszegi, B., and W. Li (2008): "Drive and Talent", *Journal of the European Economic Association* 6(1), 210–236.

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

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

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

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

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

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

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

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

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

[33] Milkovich, G. T., and J. M. Newman (2008): *Compensation*, Ninth Edition, McGraw–Hill, New York, New York.

[34] Murphy, K. J. (1999): "Executive Compensation", in *Handbook of Labor Economics*, Volume III, Part II, O.C. Ashenfelter and D. Card Eds., North–Holland, the Netherlands.

[35] Nagypál, E. (2007): "Learning by Doing vs. Learning about Match Quality: Can We Tell Them Apart?", *Review of Economic Studies* 74(2), 537–566.

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

[37] Paarsch, H. J., and B. S. Shearer (1999): "The Response of Worker Effort to Piece Rates: Evidence from the British Columbia Tree–Planting Industry", *Journal of Human Resources* 34(4), 643-667.

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

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

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

[41] Prendergast, C. (2002a): "Uncertainty and Incentives", *Journal of Labor Economics* 20(2), S115–S137.

[42] Prendergast, C. (2002b): "The Tenuous Tradeoff between Incentives and Risk", *Journal of Political Economy* 110(5), 1071-102.

[43] Prendergast, C., and R. Topel (1993): "Discretion and Bias in Performance Evaluation", *European Economic Review* 37(2-3), 355–365.

[44] 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.

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

[46] 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 A: Omitted Proofs

Lemma 1 and Proof

Lemma 1. *For each σ_ε^2, m_1, and r there exists a unique $\overline{\Sigma}_1^2 = \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r) > 0$ such that $h(\sigma_1^2, \sigma_\varepsilon^2, m_1, r)$ is greater than zero if, and only if, $\sigma_1^2 \in (0, \overline{\Sigma}_1^2)$. The cutoff $\overline{\Sigma}_1^2$ is strictly increasing in σ_ε^2, with $\lim_{\sigma_\varepsilon^2 \to \infty} \overline{\Sigma}_1^2 = \infty$, $\lim_{r \to 0} \overline{\Sigma}_1^2 = \infty$, and $\lim_{\sigma_\varepsilon^2 \to 0} \overline{\Sigma}_1^2 = 2(3 + m_1)/r$.*

Proof: Note that $h(0, \sigma_\varepsilon^2, m_1, r) = 2 + m_1 > 0$ and $\lim_{\sigma_1^2 \to \infty} h(\sigma_1^2, \sigma_\varepsilon^2, m_1, r) = -\infty$; recall that $m_1 > 0$ by assumption. 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

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

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

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 $\overline{\Sigma}_1^2 = \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r)$. Observe that $\overline{\Sigma}_1^2 > 0$. By construction, $h(\sigma_1^2, \sigma_\varepsilon^2, m_1, r)$ is positive if, and only if, $\sigma_1^2 \in (0, \overline{\Sigma}_1^2)$.

It it immediate to see that $\partial h(\overline{\Sigma}_1^2, \sigma_\varepsilon^2, m_1, r)/\partial \sigma_1^2 < 0$. Moreover,

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

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

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

from which we obtain that

$$\overline{\Sigma}_1^2 > \frac{1}{2r} + \left\{\frac{1}{4r^2} + \frac{\sigma_\varepsilon^2(2 + m_1)}{r}\right\}^{1/2}. \tag{15}$$

In particular, $\lim_{r \to 0} \overline{\Sigma}_1^2 = \lim_{\sigma_\varepsilon^2 \to \infty} \overline{\Sigma}_1^2 = \infty$. To finish the characterization of $\overline{\Sigma}_1^2$, suppose, by contradiction, that $\lim_{\sigma_\varepsilon^2 \to 0} \overline{\Sigma}_1^2 \neq 2(3+m_1)/r$. 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 with limit $\overline{\Sigma}_{1,\infty}^2$. Note that $\overline{\Sigma}_{1,\infty}^2 \geq 1/2r > 0$ by (15). Since the only solution to $h(\sigma_1^2, 0, m_1, r) = 0$ is $2(3+m_1)/r$ and $\overline{\Sigma}_{1,\infty}^2 > 0$ implies that $h(\cdot, \cdot, m_1, r)$ is jointly continuous at $(\sigma_1^2, \sigma_\varepsilon^2) = (\overline{\Sigma}_{1,\infty}^2, 0)$, 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. $\qquad\square$

Proof of Proposition 6

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 (14) only when $\sigma_1^2 \geq \overline{\Sigma}_1^2$. Since (14) 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 (14) 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, (14) 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 $\sigma_1^2 \geq \overline{\Sigma}_1^2$. $\qquad\square$

Proof of Proposition 7

We need the following result in order to establish Proposition 7.

Lemma 2. *Suppose there exist $\Sigma_{10}^2, \Sigma_{11}^2 \in (0, \overline{\Sigma}_1^2)$, with $\Sigma_{10}^2 < \Sigma_{11}^2$, such that for all $a \leq \overline{a}_1$, $H(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)$ is strictly decreasing (increasing) in σ_1^2 when $\sigma_1^2 \in (\Sigma_{10}^2, \Sigma_{11}^2)$. Then $\mathcal{A}_1$ is strictly increasing (decreasing) in σ_1^2 when $\sigma_1^2 \in (\Sigma_{10}^2, \Sigma_{11}^2)$.*

Proof: 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 $\chi \in \mathbb{R}_{++}^4$, the set $A_1(\chi)$ is bounded above by the only solution $\overline{a}_1 = \overline{a}_1(m_1)$ to $g'(a) = 2 + a + m_1$; the uniqueness of $\overline{a}_1$ follows from the assumption that $\xi(a)$ is nondecreasing. Since H is continuous in a, the set $A_1(\chi)$ is closed as well. Hence, $A_1(\chi)$ is compact for all $\chi \in \mathbb{R}_{++}^4$. 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. It is immediate to see that $\mathcal{A}_1$ is strictly increasing (decreasing) in σ_1^2 if, and only if, $a_{1,\min}^*$ and $a_{1,\max}^*$ are strictly increasing (decreasing) in σ_1^2.

Now observe that

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

for all $\sigma_1^2 \in (0, \overline{\Sigma}_1^2)$. Thus, $\sigma_1^2 \in (0, \overline{\Sigma}_1^2)$ implies that $a_{1,\min}^*$ and $a_{1,\max}^*$ are the smallest and greatest solutions of $H(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) = 0$, respectively, and $a_{1,\min}^*$ is positive. Suppose then that there exist $\Sigma_{10}^2, \Sigma_{11}^2 \in (0, \overline{\Sigma}_1^2)$, with $\Sigma_{10}^2 < \Sigma_{11}^2$, such that for all $a \leq \overline{a}_1$, $H(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)$ is strictly decreasing (increasing) in σ_1^2 when $\sigma_1^2 \in (\Sigma_{10}^2, \Sigma_{11}^2)$. Given that $\lim_{a\to\infty} H(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) = \infty$, it is easy to see that $a_{1,\min}^*$ and $a_{1,\max}^*$ are strictly increasing (decreasing) in σ_1^2 when $\sigma_1^2 \in (\Sigma_{10}^2, \Sigma_{11}^2)$; note that a change in σ_1^2 does not affect $\overline{a}_1$. This establishes the desired result. $\square$

We can now prove Proposition 7. Fix Σ_1^2, σ_ε^2, and m_1, and note that if $a \leq \overline{a}_1$, then

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

So, there exists $\overline{r} > 0$ such that $\partial H(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)/\partial\sigma_1^2 < 0$ for all $(\sigma_1^2, a) \in (0, \Sigma_1^2) \times [0, \overline{a}_1]$

34

if $r \leq \bar{r}$. Since $\lim_{r \to 0} \overline{\Sigma}_1^2 = \infty$, we can choose $\bar{r}$ to be such that $\Sigma_1^2 < \overline{\Sigma}_1^2(\sigma_\varepsilon^2, m_1, r)$ for all $r \leq \bar{r}$. The desired result now follows from Lemma 2. $\qquad\square$

Proof of Proposition 8

Fix m_1, r, and σ_ε^2. Note, from the proof of Proposition 7, that if $a \leq \bar{a}_1$, then

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

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 $a \in [0, \bar{a}_1]$, which implies (from Lemma 2 above) 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 6, that

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

for all $a \in [0, \bar{a}_1]$. Given that $\partial \mathcal{H}/\partial a$ is uniformly continuous in a and σ_1^2 when $(a, \sigma_1^2) \in [0, \bar{a}_1] \times [\Sigma_{10}^2, \overline{\Sigma}_1^2]$, 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 $a \in [0, \bar{a}_1]$. Moreover, since $\mathcal{H}$ is also uniformly continuous in a and σ_1^2 when $(a, \sigma_1^2) \in [0, \bar{a}_1] \times [\Sigma_{10}^2, \overline{\Sigma}_1^2]$, 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 $a \in [0, \bar{a}_1]$. Finally, given that

$$\frac{\partial H}{\partial a}(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) = \mathcal{H}(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) + (1 + a)\frac{\partial \mathcal{H}}{\partial a}(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r),$$

we can then conclude that $H(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)$ is strictly increasing in a when $a \in [0, \bar{a}_1]$ 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 $\sigma_1^2 \in (0, \overline{\Sigma}_1^2)$. $\qquad\square$

Proof of Proposition 9

Fix m_1 and r, and let $\Sigma_1^2 \in (0, 2(3+m_1)/r)$. 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 $(a, \sigma_1^2) \in [0, \bar{a}_1] \times [\Sigma_1^2, 2(3+m_1)/r]$, $R(a, \sigma_1^2, \sigma_\varepsilon^2)$ increases to

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

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 $(a, \sigma_1^2) \in [0, \bar{a}_1] \times [\Sigma_1^2, 2(3+m_1)/r]$. 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 $a \in [0, \bar{a}_1]$ when $\sigma_1^2 \in [\Sigma_1^2, 2(3+m_1)/r]$. 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 $\sigma_1^2 \in (\Sigma_1^2, 2(3+m_1)/r)$, so that a_1^* is unique when $\sigma_1^2 \in (\Sigma_1^2, 2(3+m_1)/r)$.

To finish the proof, note that

$$\frac{\partial \mathcal{H}}{\partial \sigma_1^2}(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) \geq \frac{r(1+a)^4 \sigma_1^4}{\{[1 + (1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2\}^2} - \frac{\sigma_\varepsilon^2}{(\sigma_1^2 + \sigma_\varepsilon^2)^2}$$
$$- \frac{\sigma_\varepsilon^2}{\{[1 + (1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2\}^2} - \frac{(1+a)\sigma_\varepsilon^2 m_1}{\{[1 + (1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2\}^2}. \quad (16)$$

It is easy to see that the last three terms on the right side of (16) 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 (16) 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 $(a, \sigma_1^2) \in [0, \bar{a}_1] \times [\Sigma_1^2, 2(3+m_1)/r]$. Hence, reducing Σ_ε^2 if necessary, we have that a_1^* is strictly decreasing in σ_1^2 when $\sigma_1^2 \in (\Sigma_1^2, 2(3+m_1)/r)$. $\square$

Proof of Proposition 10

We prove the second part of Proposition 10; the proof of the first part is very similar. Fix m_1 and r, and let $\Sigma_1^2 < \overline{\Sigma}_1^2(0, m_1, r)$. Then, $\sigma_1^2 \in (0, \Sigma_1^2)$ implies that a_1^* is positive regardless of σ_ε^2. Now observe that

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

$$\geq \frac{1}{(\sigma_1^2 + \sigma_\varepsilon^2)^2} + \frac{m_1}{(2\sigma_1^2 + \sigma_\varepsilon^2)^2} - \frac{2r(1+\overline{a}_1)^4\sigma_1^4}{\{[1+(1+\overline{a}_1)^2]\sigma_1^2 + \sigma_\varepsilon^2\}^3}.$$

Reducing Σ_1^2 if necessary, we have that $\sigma_1^2 \in (0, \Sigma_1^2)$ implies that

$$\frac{1 + m_1/4}{\sigma_1^4} - \frac{2r(1+\overline{a}_1)^6}{[1+(1+\overline{a}_1)^2]^3\sigma_1^2} > 0.$$

Hence, for each $\sigma_1^2 \in (0, \Sigma_1^2)$, there exists $\Sigma_\varepsilon^2 > 0$ such that $\partial H(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)/\partial \sigma_\varepsilon^2 > 0$ for all $(a, \sigma_\varepsilon^2) \in [0, \overline{a}_1] \times (0, \Sigma_\varepsilon^2)$. The desired result now follows from the same argument used in the proof of Proposition 7; it is straightforward to adapt Lemma 2 to cover the case where the comparative statics is with respect to σ_ε^2. $\qquad\square$

Proof of Proposition 11

Fix m_1 and σ_1^2 and let $\overline{r} > 0$ be the unique value of r such that $2(3 + m_1)/r = \sigma_1^2$. Now let $r > \overline{r}$. Since $\lim_{\sigma_\varepsilon^2 \to 0} \overline{\Sigma}_1^2 = 2(3 + m_1)/r$ and $\overline{\Sigma}_1^2$ is strictly increasing in σ_ε^2, there exists $\Sigma_{1,\varepsilon}^2 > 0$ such that a_1^* is positive if, and only if, $\sigma_\varepsilon^2 > \Sigma_{1,\varepsilon}^2$. The same argument as in the proof of Proposition 8 shows that there exists $\Sigma_\varepsilon^2 > \Sigma_{1,\varepsilon}^2$ such that a_1^* is unique if $\sigma_1^2 \in (\Sigma_{1,\varepsilon}^2, \Sigma_\varepsilon^2)$. Given that a_1^* is greater than zero when $\sigma_\varepsilon^2 \in (\Sigma_{1,\varepsilon}^2, \Sigma_\varepsilon^2)$, we then have, reducing Σ_ε^2 if necessary, that a_1^* must be strictly increasing in σ_ε^2 when $\sigma_\varepsilon^2 \in (\Sigma_{1,\varepsilon}^2, \Sigma_\varepsilon^2)$. This establishes the desired result. $\qquad\square$

Appendix B: Supplementary Material

Here we consider the case in which the return to learning–by–doing is uncertain and $\vartheta_2 > 0$. As in the main text, we assume that g' is convex and 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 $m_1 > 0$, so that the expected return to effort in period one is positive. We begin with the equilibrium characterization and then obtain our comparative statics results.

Equilibrium Characterization

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

$$g'(a) = \frac{(1+\widehat{a}_1 + \widehat{a}_2)(1+\widehat{a}_1)\sigma_2^2}{(1+\widehat{a}_1)^2\sigma_2^2 + \sigma_\varepsilon^2} = \frac{(1+\widehat{a}_1 + \widehat{a}_2)(1+\widehat{a}_1)\sigma_1^2}{\left[1 + (1+\widehat{a}_1)^2\right]\sigma_1^2 + \sigma_\varepsilon^2}. \tag{17}$$

From (17), the possible values for the worker's effort in period two are the solutions to

$$J_2(a, \widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2) = g'(a) - \frac{(1+\widehat{a}_1 + a)(1+\widehat{a}_1)\sigma_1^2}{\left[1 + (1+\widehat{a}_1)^2\right]\sigma_1^2 + \sigma_\varepsilon^2} = 0. \tag{18}$$

Since g' is convex, it follows that (18) has a unique and positive solution. Moreover, the solution to (18) does not depend on y_1, that is, is uncontingent. Equation (18) defines the

38

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

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

Hence, 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 $g'(a) = 1 + a$. In other words, α_2 is uniformly bounded. Since $g''(a) \geq g'(a)/a$ for all $a > 0$, we have from (17) that

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

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

Thus, α is differentiable in $\widehat{a}_1$, σ_1^2, and σ_ε^2 by the Implicit Function Theorem.[21]

It is immediate to see from (18) 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,

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

$$\times \left\{ 2(1 + \widehat{a}_1) + \alpha_2(\widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2) - \frac{2[1 + \widehat{a}_1 + \alpha_2(\widehat{a}_1, \sigma_1^2, \sigma_\varepsilon^2)](1 + \widehat{a}_1)^2\sigma_1^2}{[1 + (1 + \widehat{a}_1)^2]\sigma_1^2 + \sigma_\varepsilon^2} \right\}. \quad (19)$$

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 period three reputation. 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 market 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

---

[21]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 (18) has a unique and positive solution when $\widehat{a}_1$ is in a neighborhood of zero.

period is the solution to

$$\frac{(1+\widehat{a}_1)\sigma_1^2}{\sigma_1^2+\sigma_\varepsilon^2} + \frac{[1+\widehat{a}_1+\alpha_2(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2)]\sigma_1^2}{[1+(1+\widehat{a}_1)^2]\sigma_1^2+\sigma_\varepsilon^2} + \frac{(1+\widehat{a}_1)[1+\widehat{a}_1+\alpha_2(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2)]\sigma_1^2 m_1}{[1+(1+\widehat{a}_1)^2]\sigma_1^2+\sigma_\varepsilon^2}$$

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

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

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

$$+\frac{(1+\widehat{a}_1)[1+\widehat{a}_1+\alpha_2(\widehat{a}_1,\sigma_1^2,\sigma_\varepsilon^2)]\sigma_1^2 m_1}{[1+(1+\widehat{a}_1)^2]\sigma_1^2+\sigma_\varepsilon^2}$$

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

$$J_1(a,\lambda,\sigma_1^2,\sigma_\varepsilon^2,m_1,r) = g'(a) - \lambda - \frac{(1+a)\sigma_1^2}{\sigma_1^2+\sigma_\varepsilon^2} - \frac{[1+a+\alpha_2(a,\sigma_1^2,\sigma_\varepsilon^2)]\sigma_1^2}{[1+(1+a)^2]\sigma_1^2+\sigma_\varepsilon^2}$$

$$-\frac{(1+a)[1+a+\alpha_2(a,\sigma_1^2,\sigma_\varepsilon^2)]\sigma_1^2}{[1+(1+a)^2]\sigma_1^2+\sigma_\varepsilon^2}\left\{m_1 - \frac{r(1+a)[1+a+\alpha_2(a,\sigma_1^2,\sigma_\varepsilon^2)]^2\sigma_1^4}{[1+(1+a)^2]\sigma_1^2+\sigma_\varepsilon^2}\right\} = 0, \quad (20)$$

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^*.

Proposition 12. *An equilibrium always exists.*

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 (20) 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 (as α_2 is continuous in a), the intermediate value theorem implies that (20) has a positive solution. $\square$

As in the case where $\vartheta_2 = 0$, the worker's choice of effort in period one need not be unique. 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 (18), 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 3. *The cutoff $\overline{\Sigma}_1^2$ is strictly increasing in σ_ε^2, with $\lim_{\sigma_\varepsilon^2 \to \infty} \overline{\Sigma}_1^2 = \infty$. Moreover, $\lim_{r \to 0} \overline{\Sigma}_1^2 = \infty$, $\lim_{r \to \infty} r\overline{\Sigma}_1^2 = \infty$, and*

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

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

Proof: Note that

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

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, $\overline{\Sigma}_1^2$ is strictly increasing in σ_ε^2.

Now note that

$$
\overline{\Sigma}_1^4 = \frac{2\overline{\Sigma}_1^2 + \sigma_\varepsilon^2}{r[1 + \alpha_2(0, \overline{\Sigma}_1^2, \sigma_\varepsilon^2)]^3}\left\{ 1 + \frac{\overline{\Sigma}_1^2}{\overline{\Sigma}_1^2 + \sigma_\varepsilon^2} + [1 + \alpha_2(0, \overline{\Sigma}_1^2, \sigma_\varepsilon^2)](1 + m_1) \right\}, \tag{21}
$$

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, we then have that $\lim_{\sigma_\varepsilon^2 \to \infty} \overline{\Sigma}_1^2 = \lim_{r \to 0} \overline{\Sigma}_1^2 = \infty$. Next, note that (21) 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, $\lim_{r \to \infty} r\overline{\Sigma}_1^2 = \infty$.

42

To finish, note from (18) 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 that $\Sigma_1^2 > 0$ implies that $j_1(\cdot, \cdot, m_1, r)$ is continuous at $(\sigma_1^2, \sigma_\varepsilon^2) = (\Sigma_1^2, 0)$, the same argument used in Appendix B to prove that $\lim_{\sigma_\varepsilon^2 \to 0} \overline{\Sigma}_1^2 = 2(3 + m_1)/r$ when $\vartheta_2 = 0$ can be used to show that

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

which completes the proof. $\qquad\qquad\square$

The next result, Proposition 13, 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 necessarily 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 $\sigma_1^2 \in (0, \overline{\Sigma}_1^2)$. 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,

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

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 (20) only when $\sigma_1^2 \geq \overline{\Sigma}_1^2$. Since, by Proposition 12, a solution to (20) always exists, we then have that the solutions to (20) 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 (20) 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 (20) 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

43

$\mathcal{J}_1(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) > 0$ 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 (19) 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

$$g''(\alpha_2(a)) > \frac{(1 + a)^2 \sigma_1^2}{\alpha_2(a)\left\{[1 + (1 + a)^2]\sigma_1^2 + \sigma_\varepsilon^2\right\}} + \frac{(1 + a)\sigma_1^2}{[1 + (1 + a)^2]\sigma_1^2 + \sigma_\varepsilon^2} \tag{22}$$

for all $a > 0$. Therefore,

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

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

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

where

$$A = \frac{1}{1 + a} + m_1 + \frac{r\sigma_1^4(1 + a)[1 + a + \alpha_2(a)]^2}{[1 + (1 + a)^2]\sigma_1^2 + \sigma_\varepsilon^2} \quad \text{and} \quad B = 1 + a + \alpha_2(a) + 2(1 + a)\left[1 + \frac{\partial \alpha_2}{\partial \widehat{a}_1}(a)\right].$$

From the previous paragraph, we have that: (i)

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

if $a \in (0, \widetilde{a}_1)$ and $\sigma_1^2 \geq \overline{\Sigma}_1^2$; and (ii) $B \geq 3[1 + a + \alpha_2(a)]$ 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 then 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

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

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 (19) and (22) 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, $(1+a)\partial \alpha_2(a)/\partial \widehat{a}_1 \leq \alpha_2$, the function 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

45

exists $\bar{r} \geq 0$ such that $G(\tilde{a}_1, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) \geq 0$ for all $\sigma_1^2 \geq \overline{\Sigma}_1^2$ when $r > \bar{r}$. First note, since α_2 is uniformly bounded by $\bar{a}_2$, that $\sigma_1^2 \geq \overline{\Sigma}_1^2$ implies that

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

Now let $\Sigma_\varepsilon^2 = \Sigma_\varepsilon^2(m_1, r) = \inf\{\sigma_\varepsilon^2 > 0 : \mathcal{G}(\sigma_\varepsilon^2, m_1, r) \geq 0\}$. Since $\lim_{\sigma_\varepsilon^2 \to \infty} \overline{\Sigma}_1^2 = \infty$ by Lemma 3, 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 $\bar{r} = \bar{r}(m_1, \sigma_\varepsilon^2) = \inf\{r > 0 : \mathcal{G}(\sigma_\varepsilon^2, m_1, r) \geq 0\}$. Since $\lim_{r \to \infty} r\overline{\Sigma}_1^2 = \infty$ by Lemma 3, we also have that $\lim_{r \to \infty} \mathcal{G}(\sigma_\varepsilon^2, m_1, r) = \infty$, which implies that $\bar{r} < \infty$ as well. By construction, $G(\tilde{a}_1, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) \geq 0$ for all $\sigma_1^2 \geq \overline{\Sigma}_1^2$ when either $\sigma_\varepsilon^2 > \Sigma_\varepsilon^2$ or $r > \bar{r}$, in which case $\widetilde{\Sigma}_1^2 = \overline{\Sigma}_1^2$. This completes the proof. $\qquad\square$

As in Proposition 6 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 $\sigma_1^2 \in (0, \overline{\Sigma}_1^2)$. Moreover, for each σ_ε^2, m_1, and r, 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 $\sigma_1^2 \geq \widetilde{\Sigma}_1^2$, where $\widetilde{\Sigma}_1^2 = \overline{\Sigma}_1^2$ 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* of 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: In what follows, we omit the dependence of α_2 on σ_1^2 and σ_ε^2 when convenient. We divide the argument in several steps. First we show (Step 1) that $g'(1/5) > 6/5$ implies that

$$\alpha_2(a) \leq (1 + a)/5 \tag{23}$$

for all $a \geq 0$. Indeed, by (18), condition (23) 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 $\widetilde{\widehat{a}}_1$ 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 $\widehat{a}_1 \in (0, \widetilde{\widehat{a}}_1)$. Recall from (19) that $\partial\alpha_2(a)/\partial\widehat{a}_1 > 0$ if, and only if,

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

We know from the proof of Proposition 13 that $\partial\alpha_2(a)/\partial\widehat{a}_1 > 0$ if $\widehat{a}_1 \in (0, \widetilde{a}_1]$. So, assume that $a \in [\widetilde{a}_1, \widetilde{\widehat{a}}_1)$, 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 (23), a sufficient condition for the last inequality is that $(1+a)^2\sigma_1^2 \leq 3(\sigma_1^2 + \sigma_\varepsilon^2)$, 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 $a \in (0, \widetilde{\widehat{a}}_1)$ 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 $a \in (0, \widetilde{\widehat{a}}_1)$ 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 $a \in (0, \widetilde{\widehat{a}}_1)$, which holds by (23) and Step 2. Now observe that $a \in (0, \widetilde{\widehat{a}}_1)$ 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 (23) 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 $a \in (0, \widetilde{\widehat{a}}_1)$. 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 $a \in (0, \widetilde{\widehat{a}}_1)$ 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

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

for all $\sigma_1^2 \geq \overline{\Sigma}_1^2$. Observe that (23) implies that

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

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

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

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 (23). Therefore,

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

and so $G(\widetilde{a}_1, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) \geq 0$ 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 immediate to see from the proof of Corollary 1 that the set of cost functions for which effort in period one is positive if, and only if, $\sigma_1^2 \in (0, \overline{\Sigma}_1^2)$ is robust to small perturbations.

Comparative Statics

For each $\chi = (\sigma_1^2, \sigma_\varepsilon^2, m_1, r) \in \mathbb{R}_{++}^4$, let

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

and define $\mathcal{A}_1 : \mathbb{R}_{++}^4 \rightrightarrows \mathbb{R}_+$ to be such that $\mathcal{A}_1(\chi) = A_1(\chi)$. As in the main text, $\mathcal{A}_1$ is the correspondence that maps the set of parameters of the model into the set of possible period one effort choices by the worker. Since

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

for each $\chi \in \mathbb{R}_{++}^4$, the elements of $\mathcal{A}_1(\chi)$ are bounded above by $\bar{a}_1 = \bar{a}_1(m_1)$, where $\bar{a}_1$ is the unique solution to $g'(a) = (1+a) + (1+a+\bar{a}_2)(1+m_1)$.

We begin our comparative statics analysis considering the effect of changes in uncertainty about ability on implicit incentives. Then we study the effect of changes in output noise on career concerns. The first result we obtain is the analogue of Proposition 7 in the main text.

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

Proof: Fix σ_ε^2 and m_1 and let $\Sigma_1^2 > 0$. Since $\lim_{r \to 0} \overline{\Sigma}_1^2 = \infty$, we can assume that $\Sigma_1^2 < \overline{\Sigma}_1^2$ without loss of generality. Since α_2 is strictly increasing in σ_1^2, we have that

$$
\frac{\partial J_1}{\partial \sigma_1^2}(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) \leq -\frac{(1+a)\sigma_\varepsilon^2}{(\sigma_1^2 + \sigma_\varepsilon^2)^2} - \frac{[1 + a + \alpha_2(a, \sigma_1^2, \sigma_\varepsilon^2)]\{1 + (1+a)m_1\}}{\{[1 + (1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2\}^2}
$$
$$
+ \frac{3[1 + a + \alpha_2(a, \sigma_1^2, \sigma_\varepsilon^2)]^2(1+a)^2 r\sigma_1^4}{\{[1 + (1+a)^2]\sigma_1^2 + \sigma_\varepsilon^2\}^2}\left[1 + a + \alpha_2(a, \sigma_1^2, \sigma_\varepsilon^2) + \sigma_1^2\frac{\partial\alpha_2}{\partial\sigma_1^2}(a, \sigma_1^2, \sigma_\varepsilon^2)\right].
$$

Now observe, from (18) and the fact that $g''(a) \geq g'(a)/a$ for all $a > 0$, that

$$
\frac{\partial\alpha_2}{\partial\sigma_1^2}(a, \sigma_1^2, \sigma_\varepsilon^2) \leq \frac{\alpha_2(a, \sigma_1^2, \sigma_\varepsilon^2)\sigma_\varepsilon^2}{\sigma_1^2\{[1 + (1+a)]^2\sigma_1^2 + \sigma_\varepsilon^2\}}.
$$

Hence, since α_2 is uniformly bounded by $\bar{a}_2$, we have that

$$
\frac{\partial J_1}{\partial \sigma_1^2}(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) \leq -\frac{\sigma_\varepsilon^2(1 + m_1)}{(2\Sigma_1^2 + \sigma_\varepsilon^2)^2} + \frac{3[1 + \bar{a}_1 + \bar{a}_2]^2(1 + \bar{a}_1)^4 r\Sigma_1^4}{\{[1 + (1+\bar{a}_1)^2]\Sigma_1^2 + \sigma_\varepsilon^2\}^2}(1 + \bar{a}_1 + 2\bar{a}_2)
$$

if $a \leq \bar{a}_1$. Therefore, there exists $\bar{r} > 0$ such that if $r \leq \bar{r}$, then $\partial J_1(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)/\partial\sigma_1^2 < 0$ for all $(\sigma_1^2, a) \in (0, \Sigma_1^2) \times [0, \bar{a}_1]$, which implies the desired result.[22] $\qquad\square$

The next comparative statics result we establish is the analogue of Proposition 8 in the main text. Note that the result that a_1^* is unique and strictly decreasing in σ_1^2 if $\sigma_1^2 < \overline{\Sigma}_1^2$ and σ_1^2 is close enough to $\overline{\Sigma}_1^2$ is only valid if either the noise in output is high enough or the worker is sufficiently risk averse.

Proposition 15. *Fix σ_ε^2, m_1, and r. There exists $\Sigma_{10}^2 \in (0, \widetilde{\Sigma}_1^2)$ such that $\mathcal{A}_1$ is strictly increasing in σ_1^2 when $\sigma_1^2 \in (0, \Sigma_{10}^2)$. Moreover, there exists $\Sigma_\varepsilon^2 \geq 0$ ($\bar{r} \geq 0$) with the property that if $\sigma_\varepsilon^2 > \Sigma_\varepsilon^2$ ($r > \bar{r}$), then there exists $\Sigma_{11}^2 \in (\Sigma_{10}^2, \overline{\Sigma}_1^2)$ such that $\mathcal{A}_1$ is single–valued and strictly decreasing in σ_1^2 when $\sigma_1^2 \in (\Sigma_{11}^2, \overline{\Sigma}_1^2)$.*

---

[22]It is straightforward to adapt the proof of Lemma 2 in Appendix A to the case under consideration (with $J_1(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)$ in place of $H(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)$.

Proof: Fix σ_ε^2, m_1, and r. We know from the proof of the last result that

$$\frac{\partial J_1}{\partial \sigma_1^2}(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) \le -\frac{\sigma_\varepsilon^2(1 + m_1)}{(2\sigma_1^2 + \sigma_\varepsilon^2)^2} + \frac{3[1 + \bar{a}_1 + \bar{a}_2]^2(1 + \bar{a}_1)^4 r \sigma_1^4}{\{[1 + (1 + \bar{a}_1)^2]\sigma_1^2 + \sigma_\varepsilon^2\}^2}(1 + \bar{a}_1 + 2\bar{a}_2)$$

for all $a \le \bar{a}_1$. Thus, there exists $\Sigma_{10}^2 \in (0, \overline{\Sigma}_1^2)$ such that $\partial J_1(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)/\partial \sigma_1^2 < 0$ for all $\sigma_1^2 \in (0, \Sigma_{10}^2)$ and all $a \in [0, \bar{a}_1]$, from which we obtain that $\mathcal{A}_1$ is strictly increasing in σ_1^2 when $\sigma_1^2 \in (0, \Sigma_{10}^2)$.

We know from Proposition 13 that there exists $\Sigma_\varepsilon^2 > 0$ such that if $\sigma_\varepsilon^2 > \Sigma_\varepsilon^2$, then a_1^* is positive if, and only if, $\sigma_1^2 \in (0, \overline{\Sigma}_1^2)$. Suppose then that $\sigma_\varepsilon^2 > \Sigma_\varepsilon^2$. Note, from the proof of Proposition 13, that $\mathcal{J}_1(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) > 0$ and

$$\frac{\partial \mathcal{J}_1}{\partial a}(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) \ge \frac{[1 + a + \alpha_2(a, \sigma_1^2, \sigma_\varepsilon^2)]\sigma_1^2}{(1 + a)^2\{[1 + (1 + a)^2]\sigma_1^2 + \sigma_\varepsilon^2\}}$$

for all $a \in (0, \tilde{a}_1)$ when $\sigma_1^2 \ge \overline{\Sigma}_1^2$; recall that $\mathcal{J}(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) = (1+a)^{-1}J_1(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)$ and $\tilde{a}_1$ is the only value of a such that $(1 + a)^2\sigma_1^2 = \sigma_1^2 + \sigma_\varepsilon^2$. Hence,

$$\frac{\partial J_1}{\partial a}(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) = \mathcal{J}_1(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) + (1 + a)\frac{\partial \mathcal{J}_1}{\partial a}(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)$$

$$\ge \frac{\overline{\Sigma}_1^2}{[1 + (1 + \tilde{a}_1)^2]\overline{\Sigma}_1^2 + \sigma_\varepsilon^2} > 0$$

for all $a \in (0, \tilde{a}_1)$ when $\sigma_1^2 \ge \overline{\Sigma}_1^2$. Consequently, since $\partial J_1/\partial a$ is continuous in σ_1^2, there exists $\underline{\Sigma}_1^2 \in (0, \overline{\Sigma}_1^2)$ such that $\partial J_1(a, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)/\partial a > 0$ for all $a \in (0, \tilde{a}_1)$ when $\sigma_1^2 \in (\underline{\Sigma}_1^2, \overline{\Sigma}_1^2)$. We show below that this implies that there exists $\Sigma_{11}^2 \in (\underline{\Sigma}_1^2, \overline{\Sigma}_1^2)$ such that a_1^* is unique when $\sigma_1^2 \in (\Sigma_{11}^2, \overline{\Sigma}_1^2)$.

From the proof of Proposition 13, when $\sigma_1^2 \in (0, \overline{\Sigma}_1^2)$ the possible values of a_1^* are the solutions to the equation

$$J_1(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) = 0, \tag{24}$$

which are all positive. Now observe that there exists $\Sigma_{11}^2 \in [\underline{\Sigma}_1^2, \overline{\Sigma}_1^2)$ such that any solution to (24) must belong to the interval $(0, \kappa)$, where $\kappa = (1 + \sigma_\varepsilon^2/\overline{\Sigma}_1^2)^{1/2} - 1$; note that $\tilde{a}_1 > \kappa$ for all $\sigma_1^2 < \overline{\Sigma}_1^2$. Suppose not. Then there exist sequences $\{\sigma_{1,n}^2\}$ and $\{a_{1,n}\}$ such that $\sigma_{1,n}^2 \uparrow \overline{\Sigma}_1^2$, $a_{1,n} \ge \kappa$ and $J_1(a_{1,n}, 0, \sigma_{1,n}^2, \sigma_\varepsilon^2, m_1, r) = 0$ for all $n \in \mathbb{N}$. Since $a_{1,n} \le \bar{a}_1$ for all $n \in \mathbb{N}$, the sequence $\{a_{1,n}\}$ has a convergent subsequence. Assume, without loss of generality, that $\{a_{1,n}\}$ itself is convergent, and denote its limit by $a_{1,\infty}$; note that $a_{1,\infty} \ge \kappa$.

Since $J_1(a_{1,n}, 0, \sigma_{1,n}^2, \sigma_\varepsilon^2, m_1, r) \to J_1(a_{1,\infty}, 0, \overline{\Sigma}_1^2, \sigma_\varepsilon^2, m_1, r)$ and, by Proposition 13, $a = 0$ is the only solution to (24) when $\sigma_1^2 = \overline{\Sigma}_1^2$, we can then conclude that $a_{1,\infty} = 0$, a contradiction. The desired result follows from the fact that if $\sigma_1^2 \in (\Sigma_{11}^2, \overline{\Sigma}_1^2)$, then J_1 is strictly increasing in a when $a \in (0, \kappa)$.

The same argument as in the proof of Proposition 8 shows, increasing Σ_{11}^2 if necessary, that a_1^* is strictly decreasing in σ_1^2 when $\sigma_1^2 \in (\Sigma_{11}^2, \overline{\Sigma}_1^2)$. Exactly the same argument as above shows that there exists $\overline{r} \geq 0$ such that if $r > \overline{r}$, then there exists $\Sigma_{11}^2 \in (\Sigma_{10}^2, \overline{\Sigma}_1^2)$ such that a_1^* is unique and strictly decreasing in σ_1^2 when $\sigma_1^2 \in (\Sigma_{11}^2, \overline{\Sigma}_1^2)$. $\qquad\square$

The third comparative statics result we establish is the analogue of Proposition 9 in the main text. Let

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

Since $\overline{\Sigma}_1^2$ is strictly increasing in σ_ε^2 with $\lim_{\sigma_\varepsilon^2 \to 0} \overline{\Sigma}_1^2 = \alpha$, we have that $\Sigma_1^2 < \alpha$ implies that $\Sigma_1^2 < \overline{\Sigma}_1^2$ regardless of σ_ε^2.

Proposition 16. *For each m_1, r, and $\Sigma_1^2 \in (0, \sigma^2)$, 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: Fix m_1, r, and $\Sigma_1^2 \in (0, \sigma^2)$. Since α_2 is strictly increasing in σ_1^2, uniformly bounded by $\overline{a}_2$, and

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

we have that

$$\begin{aligned}
\frac{\partial J_1}{\partial \sigma_1^2}(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) &\geq -\frac{(1 + a)\sigma_\varepsilon^2}{(\sigma_1^2 + \sigma_\varepsilon^2)^2} - \frac{[1 + a + 2\alpha_2(a, \sigma_1^2, \sigma_\varepsilon^2)][1 + (1 + a)m_1]\sigma_\varepsilon^2}{\{[1 + (1 + a)^2]\sigma_1^2 + \sigma_\varepsilon^2\}^2} \\
&\quad + \frac{r(1 + a)^2[1 + a + \alpha_2(a, \sigma_1^2, \sigma_\varepsilon^2)]^3\sigma_1^4}{\{[1 + (1 + a)^2]\sigma_1^2 + \sigma_\varepsilon^2\}^2} \\
&\geq -\frac{(1 + a)\sigma_\varepsilon^2}{(\sigma_1^2 + \sigma_\varepsilon^2)^2} - \frac{[1 + a + 2\overline{a}_2][1 + (1 + a)m_1]\sigma_\varepsilon^2}{\{[1 + (1 + a)^2]\sigma_1^2 + \sigma_\varepsilon^2\}^2} \\
&\quad + \frac{r(1 + a)^5\sigma_1^4}{\{[1 + (1 + a)^2]\sigma_1^2 + \sigma_\varepsilon^2\}^2}.
\end{aligned}$$

Now observe that the right side of the last inequality is strictly increasing in σ_ε^2 if $\sigma_\varepsilon^2 < \sigma_1^2$. The same argument as in the proof of Proposition 9 then shows that there exists $\Sigma_\varepsilon^2 > 0$

51

such that $\partial J_1(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)/\partial\sigma_1^2 > 0$ for all $(a, \sigma_1^2) \in [0, \overline{a}_1] \times (\Sigma_1^2, \overline{\Sigma}_1^2)$ when $\sigma_\varepsilon^2 \in (0, \Sigma_\varepsilon^2)$, which implies the desired result. □

We now consider the effect of changes in output noise on career concerns incentives. The first result is the analogue of Proposition 10 in the main text.

Proposition 17. *Fix σ_1^2 and m_1. For all $\Sigma_\varepsilon^2 > 0$, there exists $\underline{r} > 0$ such that if $r \in (0, \underline{r})$, then $\mathcal{A}_1$ is strictly decreasing in σ_ε^2 if $\sigma_\varepsilon^2 \in (0, \Sigma_\varepsilon^2)$. Now fix m_1 and r. There exists $\Sigma_1^2 > 0$ with the property that if $\sigma_1^2 \in (0, \Sigma_1^2)$, then there exists $\Sigma_\varepsilon^2 > 0$ such that $\mathcal{A}_1$ is strictly decreasing in σ_ε^2 when $\sigma_\varepsilon^2 \in (0, \Sigma_\varepsilon^2)$.*

Proof: We only prove the second part of the proposition; the proof of the first part is very similar. Fix m_1 and r, and let $\Sigma_1^2 < \alpha$. Thus, $\sigma_1^2 \in (0, \Sigma_1^2)$ implies that effort in period one is positive regardless of σ_ε^2. Straightforward algebra shows that

$$\frac{1}{\sigma_1^2}\frac{\partial J_1}{\partial\sigma_\varepsilon^2}(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r) \geq \frac{1+m_1}{(2\sigma_1^2 + \sigma_\varepsilon^2)^2} - \frac{2r(1+\overline{a}_1)^6(1+\overline{a}_1+\overline{a}_2)^3\sigma_1^4}{\{[1+(1+\overline{a}_1)^2]\sigma_1^2 + \sigma_\varepsilon^2\}^3}$$
$$-\frac{\partial\alpha_2}{\partial\sigma_\varepsilon^2}(a, \sigma_1^2, \sigma_\varepsilon^2)\left\{\frac{1+m_1}{[1+(1+\overline{a}_1)^2]\sigma_1^2 + \sigma_\varepsilon^2} - \frac{3r(1+\overline{a}_1)^4(1+\overline{a}_1+\overline{a}_2)^2\sigma_1^4}{\{[1+(1+\overline{a}_1)^2]\sigma_1^2 + \sigma_\varepsilon^2\}^2}\right\};$$

recall that α_2 is decreasing in σ_ε^2. Reducing Σ_1^2 if necessary, we have that

$$\min\left\{\frac{1+m_1}{4\sigma_1^4} - \frac{2r(1+\overline{a}_1)^6(1+\overline{a}_1+\overline{a}_2)^3}{[1+(1+\overline{a}_1)^2]^3\sigma_1^2}, \frac{1+m_1}{[1+(1+\overline{a}_1)^2]\sigma_1^2} - \frac{3r(1+\overline{a}_1)^4(1+\overline{a}_1+\overline{a}_2)^2}{[1+(1+\overline{a}_1)^2]^2}\right\}$$

is positive for all $\sigma_1^2 \in (0, \Sigma_1^2)$. Hence, for each $\sigma_1^2 \in (0, \Sigma_1^2)$, there exists $\Sigma_\varepsilon^2 > 0$ such that $\partial J_1(a, 0, \sigma_1^2, \sigma_\varepsilon^2, m_1, r)/\partial\sigma_\varepsilon^2 > 0$ for all $(a, \sigma_\varepsilon^2) \in [0, \overline{a}_1] \times (0, \Sigma_\varepsilon^2)$, which implies the desired result (see comment at the end of the proof of Proposition 10). □

We know from Proposition 13 that for each m_1 and r there exists $\Sigma_\varepsilon^2 > 0$ such that $\overline{\Sigma}_1^2 = \widetilde{\Sigma}_1^2$ when $\sigma_\varepsilon^2 > \Sigma_\varepsilon^2$. Since $\lim_{\sigma_\varepsilon \to \infty} \overline{\Sigma}_1^2 = \infty$, we then have that an increase in σ_ε^2 can (and eventually does) lead to higher effort in the first period if $\sigma_1^2 > \widetilde{\Sigma}_1^2$. Thus, as in the main text, an increase in output noise can have a positive impact on career concerns incentives. However, the analogue 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, $\sigma_1^2 > \alpha$ implies implicit incentives are increasing in σ_ε^2 when output noise is small enough (as $\sigma_1^2 > \alpha$ implies that $\sigma_1^2 < \overline{\Sigma}_1^2$ for σ_ε^2 small). Since $\lim_{r \to \infty} \alpha = 0$, we then have the following result; its proof is identical to the proof of Proposition 11.

Proposition 18. *An increase in σ_ε^2 can, and eventually does, increase effort when uncertainty about ability is sufficiently high. Moreover, for a non–empty set of cost functions, we have that for each m_1 and σ_1^2, there exists $\bar{r} > 0$ with the property that if $r > \bar{r}$, then there exists $\Sigma_\varepsilon^2 > 0$ such that $\mathcal{A}_1$ is single–valued, increasing, and non–constant when $\sigma_\varepsilon^2 \in (0, \Sigma_\varepsilon^2)$.*

Appendix C: Omitted Details (Not for Publication)

Proof of Claim in Footnote 8

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 $\widetilde{Z}_t = Z_t \times W_t$, 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 $\sigma_t^*(z^t) = \widetilde{\sigma}_t^*(z^t, w^t(y^t))$ when $z^t = (y^t, a^t)$. 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 Statement in Footnote 15

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 $\chi \in \Xi$.