Ambition and Talent*

Botond Kőszegi, UC Berkeley Wei Li, MIT

May 24, 2002

Abstract

We develop a career concerns model in which agents differ in taste for income in addition to ability, and derive basic implications of this framework. We argue that the model captures important aspects of ambition. Since ambitious agents are expected to work harder–and therefore be paid more–than unambitious ones, everyone might be induced to work hard to prove that they are ambitious. On the other hand, proving one's ambition can be detrimental, because past outputs will be taken by the principal to reflect lower ability. Thus, "ambition-proving incentives" are likely to increase effort early in the career and decrease it later. Over a long horizon, ambition-proving incentives have a tendency to bootstrap themselves, and, if this effect is strong enough, to create significant incentives with little else motivating the agent. Finally, we discuss in detail two consequences of our framework for organizational design. To maximize effort, the principal wants to cater incentives to the best-performing employees, and wants to observe a measure of the agent's effort (say, his hours) early, but not late, in the career.

*We thank George Akerlof, Abhijit Banerjee, Mathias Dewatripont, Glenn Ellison, Bengt Holmström, Michael Jansson, Jon Levin, Markus Möbius, Antonio Rangel, Emmanuel Saez, Chris Shannon, Jean Tirole, Muhamet Yildiz, and seminar participants at MIT, UC Berkeley, and University of Chicago GSB for great comments.

1 Introduction

In most existing career concerns models and many models of signaling and screening, the domain of uncertainty that drives the analysis is an agent's inherent "ability" or "talent." Since talented agents are more productive, perceived talent is rewarded in the market, and agents have an incentive to convince the principal of their inherent abilities.

While discussions of talent certainly feature prominently in evaluations at real life organizations, they are usually accompanied by assessments about an equally important attribute: ambition, or an agent's commitment to his career (Kanter 1977, Landers, Rebitzer, and Taylor 1996).[1] In particular, not only is the principal interested in whether an employee is talented, she also wants to know whether he is ambitious. As evidenced by statements like "he is talented, but not very ambitious," or vice versa, observations about these two qualities often go hand in hand. Similarly, both qualities ultimately affect an agent's productivity. Hence it is somewhat odd that economic theory has focused so exclusively on ability.

In this paper we attempt to shed light on the behavior of agents and the structure of organizations when employers make inferences about the ambition of their employees, and show how these inferences are connected to observations about ability. We believe that a number of possible formalizations of ambition can lead to similar effects, but in our model we focus on just one interpretation: ambition is identified with the marginal utility of income of the agent. In other words, more ambitious people derive more happiness than their less ambitious counterparts from any given level of success. Formally, we assume that–besides differing in their ability–agents also differ in their marginal utility of income m, a measure which they know but the principal does not. We examine incentives in an otherwise standard career concerns framework: an agent's output depends on his ability, the unobservable level of effort and noise. The principal pays a competitive but fixed wage to the agent, which is equal to the output she expects him to deliver. Thus, if the principal perceives the agent to be either more talented or more hard-working, she will offer him a higher wage.

Since more talented agents produce more on average, higher output is taken as a sign of ability

[1] Throughout this paper, the principal is assumed to be female and the agent is assumed to be male.

by the principal. As is well known since Fama (1980) and Holmström (1999), this creates an incentive for the agent to increase his level of effort in an attempt to "convince" the principal that he is talented. But unlike in Holmström's model and its subsequent extensions, agents respond differently to these career concerns incentives due to different levels of ambition. Therefore, the agent's output is a signal of both his ability and his marginal utility of income. Just as he wants to manipulate the signal the principal receives about his ability, the agent wants to manipulate the principal's impressions about his ambition, a consideration we label the *ambition-proving incentive*.

Ambition-proving incentives are composed of two basic forces. First, since more ambitious agents are expected to work harder, a high output can be used to signal one's ambition, which in turn translates into expectations about harder work in the future. This *forward attribution* increases effort. Second, if the principal becomes convinced that the agent is ambitious, given the levels of past output she must downgrade her opinion of the agent's talent–since he must have worked hard in the past, his output tends to reflect more toil than talent. This *backward attribution* decreases the agent's effort. The relative strength of these effects plays a major role in determining the agent's effort choices in different periods of his career and in different environments.

To capture the range of results that are implied by ambition-proving incentives, we gradually increase the scope and duration of the informational asymmetry between the principal and the agent. For a major part of our paper, we focus on variants of a three-period model. Also, in many of our models, m denotes the agent's marginal utility of income in the last period. First, in a benchmark model, the principal learns m before setting the wage in period 3. Therefore, ambition-proving incentives enter only in the agent's attempt to influence the wage in period 2. We show that more ambitious agents work harder in both periods 1 and 2, and that in this case ambition-proving incentives unambiguously increase effort relative to a standard career concerns model. Next, if m remains unknown to the principal in period 3, effort in period 2 decreases–since all types of agents exert zero effort in period 3, only the backward attribution operates at that point.

In general, the backward attribution becomes relatively more important as the agent's career progresses. Therefore, under reasonable conditions, heterogeneity in ambition increases incentives early in the career but decreases them later. As we demonstrate in an extension of the three-period model, this also implies that the agent is likely to work hardest on average right after he has figured out how ambitious he is. In other words, ambition-proving incentives have the (some say unfortunate) implication that people have to work hardest in their careers exactly when they are also learning about crucial aspects of their personal lives. Finally, we analyze the case when ambition differs across agents throughout their career and show that these conclusions can be reversed, but only if the career is very short.

Next, we consider an infinite-horizon model and assume that heterogeneity in ambition extends to all periods, with evolving marginal utility of income and ability for each individual. In this case, ambition-proving incentives have a self-reinforcing feature. Given any amount of persistent heterogeneity in ability (or essentially any monetary incentive), ambitious agents respond more strongly to it than unambitious ones. But, in addition to the standard career concerns incentives, this now provides ambition-proving incentives as well, to which ambitious agents also respond more strongly, creating further ambition-proving incentives. Thus, the ambition-proving incentive bootstraps itself and becomes stronger and stronger. If this effect is strong enough–for which we provide a necessary and sufficient condition–even a trivial amount of heterogeneity in ability leads to non-trivial levels of effort in steady state.

We devote considerable attention to our model's implications for organizational design. For a sufficiently long career, a straightforward consequence of our basic model is that the principal wants to increase the degree to which ambitious agents differ from unambitious ones. If ambitious agents work much harder in equilibrium, all agents will want to prove that they are ambitious, improving everyone's incentives. Somewhat more subtlely, the employer wants to create the right conditions for signaling one's ambition. One way to do so is to observe a measure of the agent's effort but not of his talent, for example, through observing how many hours he stays in his office. In stark contrast to standard career concerns models–where observation of hours is either useless or makes it harder to prove ability–observing the agent's hours early in the career unambiguously increases incentives. The intuition is that the agent's increased opportunity to "single out" and prove his ambition outweighs the loss in opportunity to prove his talent. When the principal downgrades her view of the agent's talent, she also concludes that he must have been less lucky, dampening the backward attribution. She does not, however, make inferences about future luck, so the agent benefits fully from the forward attribution. In contrast, the firm does not want to observe the agent's hours late in the career; in fact, we provide a reason for the firm to *forget* early observation of hours, because this dampens the backward attribution operating at the end of the career.

From a theoretical point of view, our model takes the logical next step implied but not taken by most career concerns models. In career concerns models, agents take actions to boost their perceived productivity. Instead, agents in our model signal their responsiveness to career concerns, in essence trying to show that they consider their careers very important. These "concerns about career concerns" are potentially important in any career concerns application. To our knowledge, they surface only in Levin (2001), although he does not explicitly frame his model in these terms.[2]

Our work is also related to Aron (1987) and Landers, Rebitzer, and Taylor (1996). Like our model, both of these papers feature heterogeneity in employee preferences (though not in ability). But in contrast to our focus on signaling, they focus on screening. For example, partners might make long hours a prerequisite for promotion of associates, to select those who will work hard as partners (Landers, Rebitzer, and Taylor 1996). We believe that both signaling and screening are important for the organizational design problem. In many cases, including consulting and investment banking, our signaling model seems to fit reality better, and it seems relevant even for law firms. If law firms can require long hours for screening purposes, why could they not similarly require long hours from partners? Rather, hours might constitute a variable on which firms do not explicitly want to condition incentives, but use them consciously as part of an implicit incentive structure. Indeed, as Landers, Rebitzer, and Taylor (1996) document, hours requirements are usually not explicit in law firms, and even when they are, the expectation is for associates to work much more.

The paper is organized as follows. In section 2, we set up our basic three-period model and derive some of its implications. Section 3 considers an infinite-horizon model and shows our bootstrapping result. Section 4 tackles the organizational design questions. Section 5 discusses extensions of our

[2] In Levin's model, collective reputation is essentially the strength of career concerns that operate within a group. By joining a given group, agents signal the importance which they will attach to their careers.

model to the case of evolving ambition and other organizational design issues, and identifies a force inherent to our model to generate multiple equilibria. Section 6 concludes.

2 A Basic Model of Ambition-Proving Incentives

This section introduces our first model of ambition-proving incentives, the notion that agents might not only want to prove that they are talented, but also that they are hard-working. In addition to heterogeneity in talent, as in standard career concerns models, agents also differ in their levels of commitment to work.

2.1 Setup

There are three periods, labeled 1, 2, and 3. A risk-neutral principal employs an agent, who produces output q_t in period t. The output q_t is composed of three additive terms in each period: the agent's time-invariant ability a, his effort level e_t and a noise term ϵ_t distributed normally with mean zero and variance σ_ϵ^2. The principal's and the agent's priors over the ability parameter a are also normal, with mean zero and variance σ_a^2. The agent's utility cost of exerting effort e is $c(e) = \frac{1}{2}ke^2$, which is additively separable from the utility from consumption.[3]

In the basic model, risk neutral agents are assumed to differ in their marginal utility of income in period 3, denoted by m. This variable is distributed normally in the population with mean $\mu_m >> 0$ and variance σ_m^2. The agent's utility from consumption is $u(w_t) = w_t$ for $t = 1, 2$ and $u(w_3) = mw_3$. The agent knows his taste for income, and the principal does not. Importantly, while in many career concerns models talent can be assumed to be symmetric information, heterogeneity in ambition only makes a difference when it is private information. a, m, and the noise terms in the output are all independent.

The assumption that m only represents marginal utility in period 3–instead of all periods–is a simplification that makes our model more transparent. It might also be empirically relevant

[3] It seems to us that the basic logic in our models would still apply with a more general cost function, however, we do not want the agents' incentives to depend on the convexity or the smoothness of a more general cost function. Moreover, a very kinked cost function may lead to non-existence of equilibria.

if agents mostly differ in their valuation of some late-career reward, like becoming a CEO. The implications of changing this assumption are taken up in section 2.4.

As in other career concerns models, we assume that the principal is either prevented from using an explicit incentive contract, or does not wish to do so; she is restricted to using fixed-wage contracts.[4] The labor market is perfectly competitive; therefore, in each period, the principal offers the agent a wage w_t equal to his expected product conditional on his *past* performances. There is no discounting or saving. Denote the history of past output at time t as h^{t-1}; then, the agent maximizes:

$$\sum_{t=1}^{3} E[u(w_t(h^{t-1})) - c(e_t(h^{t-1}))] \tag{1}$$

In many of our models, we will be interested in the changes in incentives as we vary the importance the agent attaches to different periods of his career. For example, period 2 (mid-career) may be the crucial part of the agent's working life. To capture this, we can assign a weight $\omega > 1$, which is common knowledge, to the wage of the important period. Since this would be an obvious extension and adds notation, we analyze the model without the extra weights.

The agent's level of ambition is assumed to be drawn from a normal distribution mostly for technical reasons. With the principal updating about multiple attributes of the agent, a normality assumption on the distribution of ambition makes this model more tractable. A normal distribution does have the unattractive property that it assigns a positive probability to negative marginal utility. This is not a crucial part of our model. A lower bound of zero would make it significantly more complicated to analyze the model, but would not change the logic of its mechanism.[5] One natural alternative model is to assume that people have discrete levels of ambitions. Such an assumption, however, does not make updating less cumbersome: it can lead to highly nonlinear incentives because different types of agents' effort may concentrate around different peaks.

[4] Standard reasons for not using explicit incentive contracts are the unverifiability of output, multitasking (the concern that the agent neglects some tasks if he is rewarded for others), and sabotage of other workers' output (if workers are competing for a pool of bonus money). See Holmström and Milgrom (1991) and Lazear (1989).

[5] For m near zero, the truncation will make updating considerably more complex and will change the equilibrium. In particular–due to the truncation–agents with a small marginal utility of income are less likely to be confused with others of different ambition, so their behavior is less affected by signaling about m. As m increases, the problem will resemble one without truncation more closely. Since $\mu_m >> 0$, the probability of having negative marginal utility is quite small. Thus our results should hold approximately for most values of m.

We look for the rational expectations equilibria of this signaling game. The equilibrium is defined by each type of agent choosing his effort level optimally given the principal's anticipated inferences, and the principal updating about the agent's type in a Bayesian way, given the expectations about the agent's behavior. We focus our attention on (pure-strategy) linear equilibria, in which the effort level is a linear function of m in each period: $e_t = \underline{e}_t + \alpha_t m$.[6] Linear equilibria seem like the most natural candidates to consider, because payoffs increase linearly with m and the cost function is quadratic.[7],[8]

Although we are ultimately interested in effort levels, many of our results will be phrased in terms of the α_t's. Due to the straightforward relationship between α_t and e_t, α_t reveals much or all about the average effort level. In the current model, $e_1 = \frac{1}{k}\frac{\partial w_2}{\partial q_1} + \alpha_1 m$ and $e_2 = \alpha_2 m$. In the model where m differs throughout the agent's life (sections 2.4 and 3), $e_t = \alpha_t m$.[9]

2.2 What Does This Capture About Ambition?

We interpret ambition as a general term for the "importance" people attach to their careers. This could have numerous possible facets, many of which are compatible with our models and lead to similar results. We take a unified approach throughout the paper and assume that ambition is captured by the marginal utility agents attach to their future wages. Alternatively, more ambitious agents may enjoy work more than their less ambitious counterparts, which can be captured in their disutility of effort. There are at least two places where heterogeneity in disutility of effort

[6] Since m (and potentially α_t) can be negative, effort levels can be negative. We think of it as destroying output. However, just as allowing m to be negative is not important for the intuition of most of our results, allowing e_t to be negative is not crucial, either.

[7] In our model, however, the agent's behavior endogenously changes the noise with which the variables a and m are observed. More precisely, how much agents respond to their level of ambition determines the accuracy with which ambition and talent can be inferred. But the responsiveness itself depends on the inference. Thus, we cannot rule out the possibility that some non-linear strategy changes the noise structure in a way that ends up justifying itself. We find this less interesting and do not consider it in this paper.

[8] Besides the linearity in m, this restriction also ensures that the agent's strategy does not depend on how smart he thinks he is. Even if a is symmetric information at the beginning, it becomes asymmetric information once the parties observe q_1: since the agent knows e_1, he has a better signal about ability. Allowing strategies to depend on beliefs about a would complicate the model considerably.

[9] Whether the equilibrium effort level contains a non-zero constant $\underline{e}_t$ depends on whether marginal utility is symmetric information in any of the periods. The constant captures the agent's response to incentives coming from future period(s) in which marginal utility is known.

might creep in: in the constant k and in the cost-minimizing level of effort.[10] In the current paper, heterogeneity in m and heterogeneity in k would lead to similar results and we do not explicitly consider this possibility.[11] If the heterogeneity is in the cost-minimizing level of effort, the effort level of different agents just differs by a constant no matter what incentives they are facing. Since this is a time-invariant shift in productivity, from a technical point of view it is identical to an increase in the variance of a. In the language of our model, it would imply that the forward attribution always outweighs the backward attribution. A third possibility is that ambitious people plan to stay in their current careers for a long time, whereas unambitious ones intend to switch jobs or quit the labor market altogether. This assumption would generate a slightly different model from ours, because the mere fact that the agent shows up for work provides information about his type. We briefly discuss what difference this would make in section 2.3.

Of course, ambition also has aspects whose analysis requires a completely different model from ours. For example, people might have an intrinsic valuation for being considered competent, completely independently of the compensation this implies; furthermore, this intrinsic valuation need not conform to what the principal finds valuable.

2.3 Ambition-Proving Incentives Over the Career

As a benchmark, we consider the case when m becomes known to the principal before wages are set in period 3. This is the simplest possible model that still generates novel effects. Also, it corresponds more closely to the interpretation of ambition as the agent's privately known length of career.[12]

To start with, $e_3 = 0$ for all types of agents, as they have no further use of reputation. Since

[10] That is, the agent's utility function might be $c(e) = \frac{1}{2}k(e - e^0)^2$, with heterogeneity in e^0.

[11] To be more precise, a model in which the heterogeneity is in k is isomorphic to one in which the heterogeneity is in m, and where this heterogeneity extends to all periods. Instead of writing the agent's utility function as $\sum_{t=1}^{3} w_t - \frac{1}{2}kc(e_t)$, we can write it as $\sum_{t=1}^{3} \frac{1}{k}w_t - \frac{1}{2}c(e_t)$, transforming the model into one with heterogeneity in m.

[12] To see this, imagine that there are only two types of agents, where the more ambitious type Y works all three periods and the less ambitious type N retires after the second period. Therefore, if an agent shows up for work at $t = 3$, the principal knows his ambition for sure, that is, she knows that he is type Y. Thus the agent's ambition m is fully revealed the day he might stop showing up for work. Of course, the revelation of this information is likely to be more gradual, at least when the length of career can take on a continuum of values. But considering this extreme case does highlight the difference between the two models of ambition. The intuition of our results still applies if m is revealed more gradually (though faster than observing it from output alone).

m is observed in period 3, the principal can deduce the effort levels e_1, e_2 that are expected of the agent in equilibrium. Therefore, just as in a standard career concerns model, she can extract two signals $a + \epsilon_1$ and $a + \epsilon_2$ from observing the outputs in the first two periods. Thus, we have

$$w_3 = E[a|q_1, q_2] = \frac{\sigma_a^2}{2\sigma_a^2 + \sigma_\epsilon^2}(q_1 - \alpha_1 m - \underline{e}_1) + \frac{\sigma_a^2}{2\sigma_a^2 + \sigma_\epsilon^2}(q_2 - \alpha_2 m - \underline{e}_2). \tag{2}$$

From the agent's point of view, q_t is distributed normally with mean $\mu_a + e_t$, where μ_a is the mean of his beliefs about ability (given past output). Thus, an increase in e_t just shifts the distribution of q_t to the right, increasing expected wages in period 3. Since agents with a higher level of ambition care more about this wage, we must have $\alpha_1, \alpha_2 > 0$. Consequently, the wage paid to the agent in the second period is an increasing function of the principal's mean beliefs about ambition m–given the level of ability, an agent who has to worry about the future works harder and thus produces more.

Although an easy extension of Holmström's (1999) career concerns model, this result already has some important implications. For example, if women are more likely to abandon their careers later in life, then–holding constant their qualifications–they will be paid less by their employers than men. (This is also true holding constant q_1, but to show that requires our analysis below.) This wage discrimination depends on the employee's attachment to the *labor market*, not the firm itself as in many previous models.[13] Even if *all* workers leave the firm with probability one after one period, agents' career concerns connect their incentives in the current period with their future plans.

Now we study the principal's updating problem after period 1. Upon observing $q_1 = a + \underline{e}_1 + \alpha_1 m + \epsilon_1$, the principal will offer the agent the wage $w_2 = E[a + \alpha_2 m|q_1]$. This can be rewritten as

$$E\left[E[a|m, q_1] + \alpha_2 m|q_1\right] = E\left[\frac{\sigma_a^2}{\sigma_a^2 + \sigma_\epsilon^2}(q_1 - \alpha_1 m - \underline{e}_1) + \alpha_2 m \middle| q_1\right] \tag{3}$$

[13] Existing explanations of discriminatory wage practices rely on some sort of turnover costs, which are argued to be higher for women since they are more likely to leave their firm. For starters, turnover can lead to direct hiring or other costs. It can also have an indirect effect on a firm's labor costs, because the firm loses its investment into the employee's human capital (Kuhn 1993), the firm has to resort to costly monitoring (Goldin 1986), or a higher efficiency wage is necessary to ensure that workers are not shirking (Bulow and Summers 1986). In our model, women would be expected to receive a different wage from men even when they are not more likely to leave the firm, when hiring costs are zero, or when there is no expropriable investment in human capital on the part of the firm.

where the outermost expectation is with respect to m. Therefore,

$$w_2 = \underbrace{\frac{\sigma_a^2}{\sigma_a^2+\sigma_\epsilon^2}q_1}_{\text{career concerns}} + \underbrace{\left(\alpha_2 - \frac{\sigma_a^2}{\sigma_a^2+\sigma_\epsilon^2}\alpha_1\right)E[m|q_1]}_{\text{ambition-proving incentives}} - \frac{\sigma_a^2}{\sigma_a^2+\sigma_\epsilon^2}\underline{e}_1. \tag{4}$$

The last term in expression 4 is a constant outside the agent's control, so his incentives are only affected by the first two terms. Notice that the first of these terms depends only on q_1, and is equal to the term we would have in a standard career concerns model. The second term, which we have labeled the *ambition-proving incentives*, depends on the principal's inferences about m. Thus, this term reflects the agent's incentives to change the principal's beliefs about his ambition, *holding output constant*. Incidentally, the sign of ambition-proving incentives also tells us whether heterogeneity in ambition increases or decreases incentives relative to a standard career concerns model. Of the two multiplicative parts that enter ambition-proving incentives, first consider $E[m|q_1]$. As we have noted above, more ambitious agents work harder, giving $\alpha_1 > 0$. Therefore, when output is higher, the principal will attribute a higher degree of ambition to the agent. By increasing effort, the agent is thus more likely to "convince" the principal that he is ambitious. As a result, the fact that more ambitious agents work harder in light of period 3 allows all types to work hard and boost their wage in period 2, about which they care equally.

Next, consider $\alpha_2 - \frac{\sigma_a^2}{\sigma_a^2+\sigma_\epsilon^2}\alpha_1$. There are two effects. On the one hand, if the principal thinks the agent cares more about period 3, she expects him to exert a higher effort in period 2. This *forward attribution*, the α_2 part in the agent's ambition-proving incentives, increases the agent's wages in period 2. On the other hand, there is a completely different, opposing effect. If the principal believes that the agent is likely to be an ambitious type, she also thinks that the agent must have worked hard in the first period. Thus, given the level of output, the principal downgrades her beliefs about the agent's ability. This *backward attribution* decreases the agent's wages in period 2.[14] Which of the above two effects is stronger is determined by the sign of $\alpha_2 - \frac{\sigma_a^2}{\sigma_a^2+\sigma_\epsilon^2}\alpha_1$.

[14] At this point, the distinction between the forward and backward attributions might seem somewhat artificial. In the current model, it might be interpreted as a fancy way of stating the simple fact that the agent cannot fully signal his ambition, because the need to signal his ability interferes in the process. However, the two effects will prove very useful for developing intuition for our results, as they are affected differently by different environments.

From a type m agent's first-order conditions, and using expression 2,

$$\begin{aligned} ke_2 &= m\frac{\sigma_a^2}{2\sigma_a^2+\sigma_\epsilon^2}, \text{ and} \\ ke_1 &= \frac{\partial}{\partial e_1}E[w_2(q_1)|e_1] + m\frac{\sigma_a^2}{2\sigma_a^2+\sigma_\epsilon^2}. \end{aligned} \tag{5}$$

From these equations, it is easy to see that

$$\alpha^* \equiv \alpha_1 = \alpha_2 = \frac{1}{k}\cdot\frac{\sigma_a^2}{2\sigma_a^2+\sigma_\epsilon^2}. \tag{6}$$

Since $\alpha_1 = \alpha_2$ in equilibrium, the penalty due to the backward attribution is only a fraction of the marginal benefit from the forward attribution. Intuitively, when the principal updates her beliefs about the agent's past level of effort, she also updates her beliefs about the agent's luck (ϵ_1)–since the agent's output now looks smaller, the principal will think that he was less lucky. This dampens the punishment due to the backward attribution. Since there is no updating about future luck, the reward for the forward attribution is not dampened. Therefore, ambition-improving incentives increase effort in the first period. Precisely, the responsiveness of w_2 to q_1 is

$$\frac{\partial w_2}{\partial q_1} = \frac{\sigma_a^2 + \alpha^{*2}\sigma_m^2}{\sigma_a^2+\alpha^{*2}\sigma_m^2+\sigma_\epsilon^2}, \tag{7}$$

which follows immediately from updating normals.

Note two immediate properties of ambition-proving incentives. Clearly, $\alpha^* \to 0$ as $\sigma_\epsilon \to \infty$. As in the case when there is no heterogeneity in ambition, if output is very noisily observed, no non-trivial level of effort can be sustained. Since able and less able agents become indistinguishable, there is no reason to exert effort to prove one's talent. Then, ambition-proving incentives are also eliminated. Similarly, if there is little heterogeneity in ability, ambition-proving incentives are close to zero, even if there is substantial heterogeneity in ambition and output is accurately observed. This highlights an important general property of ambition-proving incentives: they rely on other incentives on which they can "piggy-back." Unlike in standard career-concerns models, however, ambition-proving incentives are also eliminated when observation of output is very accurate. As expression 4 demonstrates, the backward and forward attributions offset each other, and the ambition-proving incentive is close to zero. This is true even though an increase in production signals a higher ambition to the principal. But in as much as the principal attributes an increase in output to ambition, she also discounts past production, a discount that is attributed entirely to the agent's ability due to the lack of noise. Thus a belief of higher ambition will not lead the principal to attribute a higher *productivity* to the agent.

When the principal does not find out the agent's ambition in period 3, the analysis is more complicated because she has to infer the agents' level of ambition from past outputs. The following theorem describes the agent's effort choices in equilibrium:

Theorem 1 *Suppose the principal never learns m. Then, in the unique linear equilibrium, $\alpha_1 = \alpha_2 \equiv \alpha^{**} > 0$, and α^{**} satisfies*

$$k\alpha^{**} = \frac{\sigma_a^2}{\sigma_\epsilon^2 + 2(\sigma_a^2 + \alpha^{**2}\sigma_m^2)}. \tag{8}$$

Furthermore, the right-hand side of this equation is the derivative of the agent's period 3 wage with respect to q_1 and q_2.

Proof: See the Appendix. □

To determine how the principal updates about the agent's ability and level of ambition, the proof of Theorem 1 takes advantage of the updating rule for multivariate normals. The random variables a, $a + \alpha_1 m + \epsilon_1$, and $a + \alpha_2 m + \epsilon_2$ are multivariate normal, with a variance-covariance matrix that is easy to write down given our assumptions. Although the problem is symmetric, it now requires some work to prove that $\alpha_1 = \alpha_2$. If it were the case that, say, an increase in α_1 increased the marginal period 3 payoff to increasing output in period 1, we might get an asymmetric equilibrium. However, exactly the opposite is the case. When α_1 increases, the principal will attribute more of an increase in q_1 to the agent's ambition rather than ability, decreasing the responsiveness of w_3 to output in period 1.

As we can see from equation 8, the heterogeneity in ambition decreases the marginal period 3 benefit of increasing output in the first two periods. The basic intuition is simple. When agents differ in their ambition (and that makes them exert different levels of effort), the principal cannot be sure whether a high output is due to ambition or inherent ability. Since ambition does not matter in period 3, heterogeneity in m thus acts as noise that diminishes incentives by making it harder to try to "prove" one's ability. This is the backward attribution we identified above: if the principal becomes convinced that the agent is ambitious, she also becomes convinced that he works hard, making her more pessimistic about ability.

But backward attribution is more serious than mere noise in the observation of output, because inferences about ambition made from output in a given period affect the interpretation of other outputs as well.[15] Imagine an employee, well into his career, who has some way of convincing his employer that he is ambitious. If he does so, his employer will conclude that he must have worked hard all these years. Given his performance, the principal thus downgrades her opinion of his talent. And since the agent's efforts cannot change past output at this point, he is discouraged from working hard.[16] This general intuition implies that heterogeneity in ambition undermines incentives late in the career.

There is an opposite force acting at the beginning of the career, however. Given α^{**}, the determination of the agent's period 2 wage is similar to our benchmark case. In particular,

$$\frac{\partial w_2}{\partial q_1} = \frac{\sigma_a^2 + \alpha^{**2}\sigma_m^2}{\sigma_a^2 + \alpha^{**2}\sigma_m^2 + \sigma_\epsilon^2}. \tag{9}$$

Though the principal cannot perfectly infer whether an increase in output is due to the agent's high level of effort or inherent ability, this does not weaken incentives in this case, because the principal simply *does not care*. If the agent worked hard in period 1, she will also work hard in period 2. Thus, while in period 3 heterogeneity in ambition acts as noise that decreases incentives, in period 2 it becomes variation in productivity that increases it.

It is easy to prove that both of these effects–the increase in $\frac{\partial w_2}{\partial q_1}$ and the decrease in $\frac{\partial w_3}{\partial q_1}$ and $\frac{\partial w_3}{\partial q_2}$–are stronger when σ_m^2 is higher. From equation 8, an increase in σ_m^2 leads to a decrease in α^{**} and therefore an increase in $\alpha^{**2}\sigma_m^2$. And $\frac{\partial w_2}{\partial q_1}$ is increasing in $\alpha^{**2}\sigma_m^2$, while $\frac{\partial w_3}{\partial q_1}$ and $\frac{\partial w_3}{\partial q_2}$ decrease in it. More generally, under the reasonable assumption that the agent's career is not too short (period 2 is important enough), our model indicates that ambition-proving incentives

[15] That is why coefficient on the term $\alpha^{**}\sigma_m^2$ in expression 8 is 2 instead of 1.

[16] The effect going the other way, that inferences about m from earlier output affect the interpretation of later outputs, is similar to the ratchet effect. If the agent increases output in period 1, the principal concludes that he must be more ambitious, thus expecting him to work harder in period 2. If the agent does not deliver, beliefs about his ability decrease. We call both of these effects the backward attribution, because ultimately both derive from period 3 wage setting.

increase effort early in the career, and decrease it later. At the beginning, the forward attribution dominates, and as sunk past performances accumulate, forward attribution becomes less and less important relative to backward attribution.[17]

It is worth comparing this model with the previous one, in which m is revealed before the wage is set in period 3. Clearly, $\alpha^* > \alpha^{**}$, which also means that period 2 incentives are stronger when m is eventually observed. Observing m eliminates the backward attribution that would otherwise operate in period 3. Instead of using output to infer the agent's ambition, the principal instead just observes it directly. Therefore, an increase in output just increases the principal's perceived ability of the agent, without being accompanied by a more "cynical" view of past performances.[18]

Two more comparative statics results are noteworthy. First, just as in a standard career concerns model, the level of effort decreases over time. From exerting effort in period 1, the agent derives an extra benefit of $\frac{\partial w_2}{\partial q_1} = \frac{\sigma_a^2 + \alpha^{**2}\sigma_m^2}{\sigma_a^2 + \alpha^{**2}\sigma_m^2 + \sigma_\epsilon^2}$ in expectation. Therefore, the decrease in effort over time is more pronounced as the heterogeneity in ambition σ_m^2 increases. Empirically, it might be quite difficult to disentangle this effect from the decrease in effort over time that is predicted by many career concerns models. However, our model does make the qualitative point that differences in the level of ambition can make the effect much more serious.

Second, note that α^{**} (and α^*) decreases with k. Not surprisingly, when agents are responsive to incentives (k is small), the difference in behavior between the ambitious and the unambitious is greater. A similar result holds when the agent's productivity is higher. This implies that more of the heterogeneity in observed output is attributed to ambition rather than ability. Consequently, for populations of agents who are sensitive to incentives, a large part of their motivation derives

[17] Besides the relative importance of periods 2 and 3, the effect of ambition heterogeneity on period 1 effort also depends on the noise (σ_ϵ^2). If the noise in the observation of output is very small (σ_ϵ^2 is close to zero), heterogeneity in ambition decreases first-period incentives, while if the noise is large, it increases them. When observation of output is very accurate, w_2 is already as responsive to q_1 as it could possibly be, so heterogeneity in ambition will not help very much. On the other hand, it will decrease the still significant responsiveness of w_3 to q_1. When observation is quite noisy, ambition helps by adding another dimension of heterogeneity that induces agents to prove themselves. It seems that in most applications, output is not very accurately observed, so this qualification is unlikely to reverse the conclusion that ambition heterogeneity increases effort early in the career.

[18] As backward attribution in period 3 is diminished, there is a kind of "multiplier effect" that further increases incentives in period 1. Since ambitious agents now work harder relative to unambitious agents, it becomes desirable for *everybody* to prove that they are ambitious. This kind of multiplier effect is an important property of ambition-proving incentives, and is discussed in section 3.

from trying to prove their ambition, and the part coming from signaling about ability is negligible. This might explain why ambition-proving seems to be more important in industries or occupations where effort levels tend to be high in general due to high stakes or responsiveness to incentives. Also, our model predicts that the decrease in effort over time is steeper in these industries.

But our result on decreasing effort over the career requires at least one qualification. In the above model, we have assumed that the agent knows his ambition from the beginning of his career. In reality, people's ambitions may become known later, perhaps due to changes in needs or circumstances. For example, the agent might get married and find that his wife has a particularly expensive taste in sports cars, increasing his marginal utility of income. To account for this possibility, in section 5, we study an extension in which the agent learns m only in the second period.

To close our discussion of the basic model, we comment on the generality of our key effects. Many models featuring implicit or explicit monetary incentives would generate the forward attribution. As long as ambitious agents exert more effort, the increased output that results from it cannot be completely rewarded in an explicit contract, and there is competition for more productive workers, agents want to prove their ambition to better their position in the market. Thus incentives due to signaling about ambition are much more general than our career-concerns framework would suggest. The backward attribution is limited to a framework where the principal is also making judgments about the agent's ability. Thus, in models with some incentives but without inferences about ability, ambition-proving incentives will always tend to increase effort. We choose the career concerns framework in order to study the interaction between inferences about ability and ambition.

2.4 When Ambition Differs Through One's Career

In our first model, we have assumed that the marginal utility of income m that summarizes ambition differs only in the last period. This assumption makes the model more analytically tractable, because it introduces a kind of symmetry between periods 1 and 2. Although an agent's overall *level* of effort differs between the two periods–in fact, this is a central part of our results–the *difference* between agents of various levels of ambition does not change ($\alpha_1 = \alpha_2$).

One, perhaps more realistic, alternative would be to assume that agents' marginal utility of income differs in all periods, and is the same over time for each type of agent. Extending the model in this way introduces several new effects, which we now discuss.

When m differs throughout the agent's career, the equilibrium efforts are $e_t = \alpha_t m$. The equilibrium conditions for α_1 and α_2 become

$$\begin{aligned} k\alpha_1 &= \frac{\sigma_a^2 + \alpha_1\alpha_2\sigma_m^2}{\sigma_a^2 + \alpha_1^2\sigma_m^2 + \sigma_\epsilon^2} + \frac{\sigma_a^2\left(\alpha_2(\alpha_2-\alpha_1)\sigma_m^2 + \sigma_\epsilon^2\right)}{\sigma_\epsilon^2(2\sigma_a^2+\sigma_\epsilon^2) + \sigma_m^2\left((\alpha_1^2+\alpha_2^2)\sigma_\epsilon^2 + (\alpha_1-\alpha_2)^2\sigma_a^2\right)} \\ k\alpha_2 &= \frac{\sigma_a^2\left(\alpha_1(\alpha_1-\alpha_2)\sigma_m^2 + \sigma_\epsilon^2\right)}{\sigma_\epsilon^2(2\sigma_a^2+\sigma_\epsilon^2) + \sigma_m^2\left((\alpha_1^2+\alpha_2^2)\sigma_\epsilon^2 + (\alpha_1-\alpha_2)^2\sigma_a^2\right)} \end{aligned} \quad (10)$$

The only difference relative to our basic model is the first term in the expression for α_1, which is the derivative $\frac{\partial w_2}{\partial q_1}$. Since more ambitious agents respond more strongly to this incentive, the slope of period 1 effort with respect to ambition is affected.

The properties of equilibrium are summarized in the following theorem:

Theorem 2 *Suppose heterogeneity in m extends to all three periods. Then*

1. *In any equilibrium, $\alpha_2 > 0$, and $\alpha_1 \neq 0$.*
2. *An equilibrium with $\alpha_1, \alpha_2 > 0$ exists.*
3. *In any equilibrium in which α_1 and α_2 are positive, $\alpha_1 > \alpha_2$.*

Proof: See the Appendix. □

In this model, it is not in general true that $\alpha_1 \geq 0$. If the principal expects the agent to produce less when he is ambitious, he might have an incentive to *destroy* output to prove that he is ambitious. And since ambitious agents care more about the wage in period 2, they destroy output more. This effect sounds unreasonable at first, but note that it is limited to a framework in which the agent can exert costly effort to destroy output.[19] It cannot happen if there is a lower bound of zero on the effort level, and it can only happen at the beginning of the agent's career.

The fact that $\alpha_2 < \alpha_1$ in the (positive) equilibrium of this model introduces caveats to our previous discussion. Namely, it constitutes a force that acts against two of our earlier claims: that

[19] And in that framework, it is not *that* unreasonable. For example, consider a soccer player looking to move to a different club. He might well be paid more if he scores more goals against his future club.

heterogeneity in ambition increases effort at the beginning of the career, and decreases it later. As can be seen from expressions 10, an asymmetry between α_1 and α_2 introduces new terms into the numerators for $\frac{\partial w_3}{\partial q_1}$ and $\frac{\partial w_3}{\partial q_2}$; the new term in $\frac{\partial w_3}{\partial q_1}$ is negative, while the one in $\frac{\partial w_3}{\partial q_2}$ is positive. In addition, $\frac{\partial w_2}{\partial q_1}$ features $\alpha_1 \alpha_2 \sigma_m^2$ in the numerator instead of $\alpha_1^2 \sigma_m^2$ as in the previous model, further tending to decrease incentives in period 1.

Intuitively, when α_2 is smaller, the forward attribution is weaker relative to the backward attribution, weakening incentives in period 1. If the agent is expected to slack off in the next period, there is less of a point in proving ambition, since this will be rewarded less generously. At the same time, the backward attribution still operates. The effect increasing incentives in period 2 is more subtle. Consider for a moment $\alpha_1 > 0$ and $\alpha_2 = 0$. Then, any increase in q_2 is attributed to ability, not ambition. Given q_1, this *decreases* the principal's beliefs about the agent's ambition, since the same output now seems to have been achieved with less effort. A similar effect survives when α_2 is positive, but much smaller than α_1. But when the principal's impression about ambition decreases, this leads to a further increase in perceived ability, as effort in period 2 is perceived to be smaller.[20] Finally, a mirror image of this effect decreases incentives in period 1. With a smaller α_2, more inference is made about ambition from q_1 (relative to q_2), so as long as $\alpha_2 > 0$, the backward attribution operating through the period 2 effort level is exacerbated.

These effects can be strong enough to reverse our conclusion that heterogeneity in ambition increases effort early in the career, and decreases it later. However, they are end-effects, because to be strong they rely on α_2 being considerably smaller than α_1. As there are fewer periods to care about on the agent's horizon, the difference between ambitious and unambitious agents decreases as time goes by. In a three-period model, this decrease is drastic, so α_2 can be much smaller than α_1. In a long-horizon setting, a one-period decrease in the horizon should have a small effect on overall incentives. Therefore, the decrease in α_t should be slower at first, and then steeper only as the agent nears the end of the horizon. This intuition seems to imply that, intriguingly, the predictions of a long-horizon model with heterogeneity across ambition extending to all periods are more similar to those of a three-period model in which m differs only at the end than to those of

[20] Of course, in reality the principal does all the updating at the same time. The above is merely a heuristic argument that helps to understand why the responsiveness of the wage to period 2 output can remain high.

a three-period model in which m differs in all periods. In fact, though we have not been able to solve the long-horizon model in general, we have been able to verify that for a sufficiently small σ_m^2, heterogeneity in ambition increases incentives at the beginning of the career, and decreases them at later stages. Thus, the effects derived from the three-period model in which m differs in all periods should be limited to settings where the length of career is very short.

3 Ambition-Proving Over the Long Term

In the previous section, we have provided basic insights about ambition-proving incentives that can be identified from a short-horizon model with heterogeneity in marginal utility of income (m). We then made the claim that these insights extend to a long-horizon model where m differs across agents in all periods. However, a long horizon introduces an important new mechanism, the bootstrapping of ambition-proving incentives, that is non-existent in our three-period model.

Consider a variant of the model presented in section 2. Instead of three periods, we now assume that the horizon is infinite. There is a constant discount factor δ. Reasonably for this longer horizon, we assume that agents' marginal utility of income differs in every period. In addition, marginal utility of income, as well as ability, evolve over time. Denote marginal utility and ability in period t by m_t and a_t respectively. We assume that m_t and a_t evolve according to

$$\begin{aligned} m_{t+1} &= m_t + \nu_t,\ \nu_t \sim N(0, \sigma_\nu^2) \\ a_{t+1} &= a_t + \eta_t,\ \eta_t \sim N(0, \sigma_\eta^2), \end{aligned}$$

where the errors are all independent.

We are still looking for linear rational expectations equilibria, that is, equilibria in which $e_t = \alpha_t m_t$.[21] This problem turns out to be very difficult to solve in general, so we further restrict our attention: we look for the steady-state level of α_t (and thus the steady-state level of average effort).

If α_t is a constant α, the principal does not need to keep track of a_t and m_t separately–she only cares about the linear combination $a_t + \alpha m_t$. As will be clear below, a sufficient condition for

[21] As we have mentioned above, since the marginal utility of income now differs in all periods, there is no constant in the effort choice rule.

α_t to be constant is that the variance of $a_t + \alpha m_t$ is constant. Lemma 1 in the Appendix shows that this is also necessary. Therefore, we look for pairs of parameters compatible with the steady state: a steady state responsiveness to ambition α and a steady state variance of $a_t + \alpha m_t$, which we denote by $\sigma^2_{a+\alpha m}$.

We derive two equations that need to hold for α and $\sigma^2_{a+\alpha m}$. First, by observing $q_t = a_t + \alpha m_t + \epsilon_t$ in period t, the principal makes an inference about $a_t + \alpha m_t$, making her beliefs more precise. The variance of her posterior is $\frac{\sigma^2_\epsilon \sigma^2_{a+\alpha m}}{\sigma^2_\epsilon + \sigma^2_{a+\alpha m}}$. At the same time, a_t and m_t change, increasing the variance by $\sigma^2_\eta + \alpha^2 \sigma^2_\nu$. In order for the the principal's belief to have variance $\sigma^2_{a+\alpha m}$ in period $t+1$, we must therefore have

$$\frac{\sigma^2_\epsilon \sigma^2_{a+\alpha m}}{\sigma^2_\epsilon + \sigma^2_{a+\alpha m}} + \sigma^2_\eta + \alpha^2 \sigma^2_\nu = \sigma^2_{a+\alpha m}, \tag{11}$$

yielding

$$\sigma^2_{a+\alpha m} \frac{\sigma^2_{a+\alpha m}}{\sigma^2_\epsilon + \sigma^2_{a+\alpha m}} = \sigma^2_\eta + \alpha^2 \sigma^2_\nu. \tag{12}$$

Second, we derive the agent's incentive to increase output in period t. Clearly, for any $t' > t$ we have $w_{t'} = E[a_{t'} + \alpha m_{t'} | q_1, \ldots, q_{t'-1}]$. In steady state,

$$\frac{\partial w_{t'}}{\partial q_t} = \frac{\sigma^2_{a+\alpha m}}{\sigma^2_{a+\alpha m} + \sigma^2_\epsilon} \left[\frac{\sigma^2_\epsilon}{\sigma^2_{a+\alpha m} + \sigma^2_\epsilon} \right]^{t'-t-1}. \tag{13}$$

Since the agent does not expect her marginal utility to change on average, she uses m_t in evaluating the future return to her effort. Therefore

$$k e_t = m_t \frac{\sigma^2_{a+\alpha m}}{\sigma^2_{a+\alpha m} + \sigma^2_\epsilon} \sum_{i=1}^{\infty} \delta^i \left[\frac{\sigma^2_\epsilon}{\sigma^2_{a+\alpha m} + \sigma^2_\epsilon} \right]^{i-1} = m_t \frac{\delta \sigma^2_{a+\alpha m}}{(1-\delta)\sigma^2_\epsilon + \sigma^2_{a+\alpha m}}. \tag{14}$$

Since $e_t = \alpha m_t$, this reduces to

$$k\alpha = \frac{\delta \sigma^2_{a+\alpha m}}{(1-\delta)\sigma^2_\epsilon + \sigma^2_{a+\alpha m}}. \tag{15}$$

Expressions 12 and 15 are necessary and sufficient for the pair $(\alpha, \sigma^2_{a+\alpha m})$ to constitute a steady state. The following theorem establishes key properties of the steady states of this model.

Theorem 3 *A steady state satisfying expressions 12 and 15 always exists. Furthermore,*

1. *Suppose $k^2(1-\delta)^2\sigma^2_\epsilon < \delta^2 \sigma^2_\nu$. For $\sigma^2_\eta = 0$, there are two steady states, one with $\alpha = 0$ and one with $\alpha > 0$. For any $\sigma^2_\eta > 0$, there is a unique steady state. As $\sigma^2_\eta \to 0$, the steady state approaches the positive steady state corresponding to $\sigma^2_\eta = 0$.*

2. *Suppose $k^2(1-\delta)^2\sigma_\epsilon^2 \geq \delta^2\sigma_\nu^2$. For $\sigma_\eta^2 = 0$, the unique steady state has $\alpha = 0$. For a sufficiently small σ_η^2, the steady state is unique, and as $\sigma_\eta^2 \to 0$, the corresponding α approaches zero.*

Proof: See the Appendix. □

The first part of Theorem 3 demonstrates just how powerful ambition-proving incentives can be. Even with a very small ultimate reason for ambitious people to behave differently from unambitious ones (a small σ_η^2), with an infinite horizon agents exert a significant amount of effort. The intuition is that ambition feeds on itself. Once ambitious people behave differently from unambitious ones because of some reward (in this case career concerns), agents will be willing to work harder not only for the original reward, but also to prove that they are ambitious. Moreover, more ambitious people have stronger incentives to prove their ambition, increasing the difference between the ambitious and unambitious and thus further strengthening the incentive to work, and so on. In the end, people work essentially to prove that they are ambitious, and that matters because they will then want to do so again. Theorem 3 shows that for some parameters this is not only a possibility, but indeed the *unique* steady state equilibrium in a reasonable model of ambition-proving incentives.

As given in Theorem 3, the key condition for bootstrapping to occur is $k^2(1-\delta)^2\sigma_\epsilon^2 < \delta^2\sigma_\nu^2$. In order for bootstrapping to create significant levels of effort, the above self-reinforcing mechanism has to be strong enough. Several factors contribute to the force of bootstrapping. If agents are more responsive to incentives (k is small), the ambition-proving incentive both builds more quickly on itself and more quickly increases the difference between different types of agents (on which ambition-proving incentives depend). If output is accurately observed (σ_ϵ^2 is small), it is easier to prove one's ambition, making it more likely that bootstrapping occurs. Naturally, if ambition changes more from period to period (σ_ν^2 is large), or the agent is more patient (δ is close to 1), bootstrapping is more likely to occur.

Note that for $k^2(1-\delta)^2\sigma_\epsilon^2 < \delta^2\sigma_\nu^2$, bootstrapping creates a discontinuity. If $\sigma_\eta^2 = 0$, $\alpha = 0$ is a steady state, whereas for $\sigma_\eta^2 > 0$, nothing close to it is. Thus, ambition-proving incentives need some other inducement to eliminate the zero-effort steady state, but an arbitrarily small other incentive necessarily creates significant effort.

The bootstrapping result can also be used to demonstrate an important general property of ambition-proving incentives: they act as a multiplier effect that increases the efficacy of other incentives. If, say, σ_η^2 increases, the agent's incentives become stronger. This leads to an increase in α, and thus a further increase in incentives. Once again, the intuition derives from the fact that more ambitious agents react more strongly to the increased incentives, increasing the ambition-proving incentive. This multiplier effect operates in many environments.

Although Theorem 3 derives the bootstrapping result for career concerns, ambition-proving incentives can "attach themselves" to other kinds of incentives as well. As soon as any other incentive creates a small difference between agents of different ambition, the above intuition kicks in, and ambition-proving incentives get a life of their own. We have confirmed that for a vanishingly small exogenous reward for increases in output, bootstrapping relies on the same condition as in Theorem 3.[22]

The bootstrapping nature of ambition-proving incentives indicates that career-concerns type incentives may not decline even after the market has learned a lot about a worker's ability (or cannot provide much in other incentives). As long as people's ambition changes, minimal differences on other dimensions can lead to significant incentives for everybody. And while ability is unlikely to change much in unpredictable ways once a person has finished his education, it is reasonable to expect that one's marginal utility of income shifts regularly due to changing life circumstances.

Although our steady state analysis relies on m changing over time, the basic mechanism behind Theorem 3 survives even when m is not changing. The intuition that more ambitious agents will respond more strongly to ambition-proving incentives just like they respond more strongly to other incentives, and the existence of this force does not depend on m changing over time.[23] Nevertheless, the *strength* of this effect does: if incentives disappear, ambitious people will not exert significantly more effort than unambitious ones, weakening the ambition-proving incentive. And in a model with a constant m, effort is likely to decrease rapidly over the career as the principal learns m. While ambition-proving incentives still bootstrap themselves, the heterogeneity on which they are based

[22] Suppose that there is no heterogeneity in talent, and an outside party gives the agent a payment βq_t in each period. We show that as β approaches zero, there can still be significant effort in steady state.

[23] This can be partially seen from expressions 10 in section 2.4. In period 1, the more ambitious agents work harder not only to prove that they are able, but also to prove that they are ambitious. This is the beginning of ambition-proving building on itself.

gradually disappears.

In addition, when incentives decrease over time, the backward attribution becomes more important. In that case, output provides a stronger signal about past effort than about future effort, so the negative inference about ability is more important relative to the positive inference about future effort. This exacerbates the decrease in effort over time.[24]

These observations reconcile our result that ambition-proving incentives can support a high level of effort even if there is little else to motivate ambitious agents with our earlier claim that heterogeneity in ambition leads to a drastic decrease in effort over time (section 2). In this section, we have assumed that the level of ambition is changing randomly, whereas the analysis in section 2 was based on a constant m. This indicates that the agent's level of effort will tend to be high as long as her needs are changing from time to time, but will decrease rapidly once m is approximately constant.

4 Organizational Design: Hours as Informal Incentives

A central theme running through this paper is that the agent's incentives to work hard derive from his need to prove not only that he is talented, but also that he will work hard. In general, the principal wants to manipulate this incentive as much as she can. We assume that the principal's goal is to increase incentives to work hard,[25] and consider two questions. In this section, we study how the principal might want to tailor the information to observe about the agent's performance. Specifically, we ask whether and when the principal wants to commit to observing a noisy signal of the *effort* the agent expands. A natural example of such a signal is the number of hours an agent puts in everyday: this might provide information about how hard-working he is, but not (directly) about his talent. In section 5, we show that in many situations, the incentive structure of the

[24] Note that steady state analysis, by its very nature, eliminates any time-variance in the balance of forward and backward attribution, and makes forward attribution uniformly stronger. Since the principal only cares about a time-invariant linear combination of ability and ambition, she does not care whether an increase in output is due to ability or hard work.

[25] In a career concerns model, it is theoretically possible that agents work inefficiently hard, so it is not necessarily true that the firm wants to increase their incentives. However, explicit incentives are widely used, and firms almost never provide negative incentives in an attempt to counterbalance career concerns. This indicates that even with career concerns, agents' level of effort is lower than efficient.

organization should be tilted toward ("catered to") better-performing agents.

Suppose, then, that in period 1, the principal can observe $q_1 = a + e_1 + \epsilon_1$ as before, but that now she can also observe a noisy signal of the agent's effort $h_1 = e_1 + \epsilon'_1$, where $\epsilon'_1 \sim N(0, \sigma'^2_\epsilon)$. We keep the assumption that the principal pays the agent a competitive wage. Since the wage now depends on what the principal chooses to observe, this is not such an innocuous assumption anymore. But the crux of career concerns models is in the principal's impression of the agent and the agent's attempts to manipulate it, not in the structure of competition in the market. Thus, we will not be explicit about how a change in impressions is translated into a change in wages.

However, we make a new assumption about the structure of the labor market. We assume that after period 2, the agent changes employers with probability one, and her new employer cannot observe h_1–the second employer is unlikely to have as detailed information on the agent's actions as the first one. This assumption allows us to isolate the phenomenon we are looking for in this section. In addition, as we will see below, even if the agent stays with the first employer, that firm may have an incentive to commit to not using h_1 for setting wages in the third period (if it can).

Once again, we look for the linear equilibrium. Since wages in period 3 depend on the same observables as in section 2, we have $\alpha_1 = \alpha_2 = \alpha^{**} > 0$, and (copying expression 8)

$$k\alpha^{**} = \frac{\sigma_a^2}{\sigma_\epsilon^2 + 2(\sigma_a^2 + \alpha^{**2}\sigma_m^2)}. \tag{16}$$

Also, the level of effort in period 2, not only the difference between agents of different levels of ambition, is the same as before. Turning to period 1, the principal's observations are q_1 and h_1. The variance-covariance matrix of interest is

$$V\left[\begin{pmatrix} a \\ m \\ q_1 \\ h_1 \end{pmatrix}\right] = \begin{pmatrix} \sigma_a^2 & 0 & \sigma_a^2 & 0 \\ 0 & \sigma_m^2 & \alpha^{**}\sigma_m^2 & \alpha^{**}\sigma_m^2 \\ \sigma_a^2 & \alpha^{**}\sigma_m^2 & \sigma_a^2 + \alpha^{**2}\sigma_m^2 + \sigma_\epsilon^2 & \alpha^{**2}\sigma_m^2 \\ 0 & \alpha^{**}\sigma_m^2 & \alpha^{**2}\sigma_m^2 & \alpha^{**2}\sigma_m^2 + \sigma'^2_\epsilon \end{pmatrix}. \tag{17}$$

This leads to the following expectations for ability and ambition:

$$\begin{aligned} E[a|q_1,h_1] &= \frac{\sigma_a^2(\alpha^{**2}\sigma_m^2 + \sigma'^2_\epsilon)(q_1 - \alpha^{**}\mu_m - \underline{e}_1) - \alpha^{**2}\sigma_a^2\sigma_m^2(h_1 - \alpha^{**}\mu_m - \underline{e}_1)}{\alpha^{**2}\sigma_m^2\sigma_a^2 + \alpha^{**2}\sigma_m^2\sigma_\epsilon^2 + \sigma_a^2\sigma'^2_\epsilon + \alpha^{**2}\sigma_m^2\sigma'^2_\epsilon + \sigma_\epsilon^2\sigma'^2_\epsilon} \\ E[m|q_1,h_1] &= \frac{\alpha^{**}\sigma_m^2\sigma'^2_\epsilon(q_1 - \alpha^{**}\mu_m - \underline{e}_1) + \alpha^{**}\sigma_m^2(\sigma_a^2 + \sigma_\epsilon^2)(h_1 - \alpha^{**}\mu_m - \underline{e}_1)}{\alpha^{**2}\sigma_m^2\sigma_a^2 + \alpha^{**2}\sigma_m^2\sigma_\epsilon^2 + \sigma_a^2\sigma'^2_\epsilon + \alpha^{**2}\sigma_m^2\sigma'^2_\epsilon + \sigma_\epsilon^2\sigma'^2_\epsilon}. \end{aligned} \tag{18}$$

As the above indicates, the principal's inference about the agent's ability depends negatively on h_1. The reason is simple: if the principal sees the agent working hard at night every day, given the level of output she attaches a lower ability to the agent. However, this does not mean that the agent will be discouraged from work, because her effort also increases output. The agent's wage in period 2 is $E[a + \alpha^{**}m|q_1, h_1]$. Given this and the above expressions,

$$\frac{\partial}{\partial e_1}E[w_2|e_1] = \frac{\alpha^{**2}\sigma_a^2\sigma_m^2 + \sigma_a^2\sigma_\epsilon'^2 + \alpha^{**2}\sigma_m^2\sigma_\epsilon'^2 + \alpha^{**2}\sigma_m^2\sigma_\epsilon^2}{\alpha^{**2}\sigma_m^2\sigma_a^2 + \alpha^{**2}\sigma_m^2\sigma_\epsilon^2 + \sigma_a^2\sigma_\epsilon'^2 + \alpha^{**2}\sigma_m^2\sigma_\epsilon'^2 + \sigma_\epsilon^2\sigma_\epsilon'^2}. \tag{19}$$

Rearranging gives

$$\frac{\partial}{\partial e_1}E[w_2|e_1] = \frac{1}{1 + \sigma_\epsilon^2 \frac{1}{\sigma_a^2 + \alpha^{**2}\sigma_m^2 + \frac{\alpha^{**2}\sigma_m^2\sigma_a^2 + \alpha^{**2}\sigma_m^2\sigma_\epsilon^2}{\sigma_\epsilon'^2}}}. \tag{20}$$

The corresponding derivative when the principal does not observe h_1 is

$$\frac{\sigma_a^2 + \alpha^{**2}\sigma_m^2}{\sigma_a^2 + \alpha^{**2}\sigma_m^2 + \sigma_\epsilon^2} = \frac{1}{1 + \sigma_\epsilon^2 \frac{1}{\sigma_a^2 + \alpha^{**2}\sigma_m^2}}. \tag{21}$$

From the above two expressions, it is clear that all agents with $m > 0$ exert higher effort in period 1 when the principal observes a measure of their hours in addition to their output. Strikingly, despite the fact that the observation of effort undermines the agent's capacity to "prove" his talents, this always motivates him to work harder in the end! Counterbalancing the negative effect on the principal's inferences about ability is that (due to the extra signal) observation of hours makes it easier to prove one's ambition. The significance of both of these effects–that the agent finds it harder to prove his ability, but easier to prove his ambition–depends on the heterogeneity in ambition. It turns out that the latter effect not only cancels the former, but in fact outweighs it. The intuition is related to one of our earlier points: forward attribution is fully rewarded, while backward attribution is not fully punished. That is, when the principal concludes that the agent is more ambitious from the number of hours worked, she expects the agent to work harder in the next period. She also concludes that the agent is less able, but this attribution is dampened by the inference that the agent must have been less lucky (she exerted high effort but still produced low output).[26]

[26] A different way to see this result is to note that, similarly to section 2.3, $E[a + \alpha_2 m|q_1, h_1] = \frac{\sigma_a^2}{\sigma_a^2+\sigma_\epsilon^2}q_1 + \left(\alpha_2 - \frac{\sigma_a^2}{\sigma_a^2+\sigma_\epsilon^2}\alpha_1\right) E[m|q_1, h_1]$. Since the forward attribution outweighs the backward attribution, and by equations 18 it is easier for the agent to prove her ambition, effort increases.

Since observing h_1 creates a stronger incentive structure than just observing q_1, there is an endogenous reason for signaling about ability to be outweighed by signaling about ambition in period 1. Although increasing incentives overall, observing h_1 does make it harder to signal one's ability, so attention shifts to signaling ambition. This derives from the firm's choice of incentive structure, whereas a similar outcome in section 2 was the result of the agent's overall responsiveness to incentives.

In contrast to the positive role of hours on incentives in our model, in the pure-strategy equilibrium of a standard career concerns model it would not matter whether the principal observes the agent's hours as long as the noise in hours ϵ_1' is of full support. The reason is that in the (unique) equilibrium of the standard career concerns model, the principal knows the equilibrium effort level e_1^*, and any difference $h_1 - e_1^*$ is attributed to the error term. In other words, knowing the agent's strategy, h_1 would not provide the principal with any information she does not already know. However, if the agent's effort is noisy for some exogenous reason, the observation of hours in fact dampens career concerns incentives. For example, the agent might "tremble" and not provide exactly the level of effort she intended. As in the above discussion, observing a higher level of effort would then indicate to the principal that the agent is of lower ability, so the agent would not be willing to put in the work.[27] This makes it all the more striking that in our framework observation of hours unambiguously increases effort.

Since w_3 only depends on the principal's perception of ability, it would be detrimental for incentives if in our setting the employer observed a measure of the agent's period 2 level of effort. Observing $h_2 = \alpha_2 m + \epsilon_2'$ takes away the agent's ability to fool the principal by increasing effort, on which her incentives to prove ability depend. Showing that this undermines effort is now much harder than in a standard career concerns model, however. Formally, the principal can observe three signals about ability before setting the wage in period 3: outputs in periods 1 and 2, and hours in period 2. Therefore, the relevant variance-covariance matrix is

[27] These claims follow easily from Propositions 4.1 and 5.1 of Dewatripont, Jewitt, and Tirole (1999).

$$V\left[\begin{pmatrix} a \\ a+\alpha_1 m+\epsilon_1 \\ \alpha_2 m+\epsilon_2' \\ a+\alpha_2 m+\epsilon_2 \end{pmatrix}\right] = \begin{pmatrix} \sigma_a^2 & \sigma_a^2 & 0 & \sigma_a^2 \\ \sigma_a^2 & \sigma_a^2+\alpha_1^2\sigma_m^2+\sigma_\epsilon^2 & \alpha_1\alpha_2\sigma_m^2 & \sigma_a^2+\alpha_1\alpha_2\sigma_m^2 \\ 0 & \alpha_1\alpha_2\sigma_m^2 & \alpha_2^2\sigma_m^2+\sigma_\epsilon'^2 & \alpha_2^2\sigma_m^2 \\ \sigma_a^2 & \sigma_a^2+\alpha_1\alpha_2\sigma_m^2 & \alpha_2^2\sigma_m^2 & \sigma_a^2+\alpha_2^2\sigma_m^2+\sigma_\epsilon^2 \end{pmatrix}. \tag{22}$$

After a horrendous amount of manipulation, this leads to the following expressions:

$$\begin{aligned} \frac{\partial w_3}{\partial e_1} &= \frac{\alpha_2(\alpha_2-\alpha_1)\sigma_a^2\sigma_m^2\sigma_\epsilon'^2+\alpha_2^2\sigma_a^2\sigma_m^2\sigma_\epsilon^2+\sigma_a^2\sigma_\epsilon^2\sigma_\epsilon'^2}{\sigma_a^2\sigma_m^2\sigma_\epsilon'^2(\alpha_2-\alpha_1)^2+(\sigma_\epsilon'^2+\alpha_2^2\sigma_m^2)(2\sigma_a^2\sigma_\epsilon^2+(\sigma_\epsilon^2)^2)+(\alpha_1^2+\alpha_2^2)\sigma_m^2\sigma_\epsilon^2\sigma_\epsilon'^2} \\ \frac{\partial w_3}{\partial e_2} &= \frac{\alpha_1(\alpha_1-\alpha_2)\sigma_a^2\sigma_m^2\sigma_\epsilon'^2-\alpha_1\alpha_2\sigma_a^2\sigma_m^2\sigma_\epsilon^2+\sigma_a^2\sigma_\epsilon^2\sigma_\epsilon'^2}{\sigma_a^2\sigma_m^2\sigma_\epsilon'^2(\alpha_2-\alpha_1)^2+(\sigma_\epsilon'^2+\alpha_2^2\sigma_m^2)(2\sigma_a^2\sigma_\epsilon^2+(\sigma_\epsilon^2)^2)+(\alpha_1^2+\alpha_2^2)\sigma_m^2\sigma_\epsilon^2\sigma_\epsilon'^2}. \end{aligned} \tag{23}$$

Using the above, we have the following theorem.

Theorem 4 *Suppose that the principal can observe h_2 before setting the wage in period 3. Then, equilibrium exists, and in any equilibrium, $\alpha_1 > \alpha_2$. Furthermore, $\alpha_1 + \alpha_2$ is smaller than it would be if the principal could not observe h_2 in period 3 (section 2), and for agents with $m > 0$, so is effort in period 2 and total effort in periods 1 and 2.*

Proof: See the Appendix. □

While the force that decreases effort in period 2 when h_2 is observed is clear, there are other effects that complicate the analysis. In particular, the change in observability and behavior in period 2 affects behavior in period 1 as well, and effort in period 1 can actually increase. There are two effects on behavior in period 1. If hours in period 2 are observed, backward attribution is weakened. Intuitively, since the principal has a direct measure of effort in period 2, she makes less inference about e_2 from period 1 output. Therefore, for any increase in q_1, the principal does not downgrade her beliefs about the agent's ability given q_2 so much. This increases period 1 effort. Offsetting this is a "precision effect." When h_2 is observed, the principal can back out a more accurate signal about the agent's ability in period 2. And with the principal's beliefs being more precise, it is harder to fool her about ability. This decreases period 1 effort. The net effect could push e_1 either way. But as Theorem 4 demonstrates, the direct effect on e_2 outweighs the possible positive indirect effect on e_1. When the backward attribution that would depress e_1 is muted, so are incentives in period 2–both of these depend on the accuracy with which the principal can observe effort in period 2.[28]

Therefore, we get the following predictions on when a firm would want to obtain information about an employee's effort level that can be used to set later pay. As long as a significant proportion of the agent's career is in front of her, making career concerns important in the future, a firm wants to commit to observing the agent's number of hours, however noisily. The same does not make sense later in the agent's career. Moreover, by that point the firm might even want to "lose" its earlier measure of the agent's effort! Although we have not formally shown this, this could be true for the same reason that observing h_2 before setting w_3 is bad for incentives. If the firm uses h_1 in setting period 3 wages, the agent cannot fool the principal about her ability as much, decreasing incentives.[29]

Although we have no economic evidence about this aspect of organizational structure, it seems that in many occupations hours are emphasized as an informal incentive early in the career, and ignored later. These include law, medicine, consulting, and investment banking.

A possible concern with the principal's observing the agent's hours is that the agent may try to "game" the system. In any setting where the principal can observe multiple signals about the agent's performance, the agent may substitute his effort toward the task that is more strongly rewarded (Holmström and Milgrom 1991). Thus, once the principal decides to observe a measure of the agent's effort, he might go out of his way to show that he works long hours. Examining expressions 18, it is clear that if the agent's hours are very accurately observed ($\sigma_\epsilon'^2$ is small), but his output is not (σ_ϵ^2 is not too small), an increase in h_1 is more rewarded in period 2 than an increase in q_1. Thus, the agent may want to concentrate effort on this signal, for example by unproductively sticking around at night even though he is exhausted.[30]

The above discussion indicates that the principal might not want to observe e_1 too precisely.

[28] The proof of theorem 4, as well as the above discussion, assumes that h_1 is not observed. The statement of the theorem would still be true, and essentially the same proof would work, if h_1 was observed in period 2.

[29] Interestingly, there is a slight difference between whether a firm would want to observe h_1 versus h_2 in period 3. Theorem 4 shows that observing h_2 is unambiguously bad, while observing h_1 could be good in some circumstances. The reason is that observing h_1 can increase the responsiveness of w_2 to q_1: by making $\alpha_2 > \alpha_1$, it makes it attractive to prove one's ambition.

[30] We have confirmed this intuition in a proper model in which the possibility to substitute effort is explicity formulated. Since its presentation would add little, we have omitted it.

This, as well as our earlier observation that it might be good to forget h_1 by the end of the agent's career, provide a rationale for using informal measures to set pay (like a supervisor's "impression" of how hard the agent works). These measures are not very accurate and are available in the short run, but can easily get lost in a dynamically changing firm. In standard models, informal measures are an imperfect substitute for explicit incentives, to be used only when the latter is not available because of some contracting constraint. Here, we have a situation in which the informal measure is better than the formal one even if both are available.[31]

5 Extensions and Discussions

In this section we discuss several variants of our model. First, we study how our basic model is modified when the agent does not know his level of ambition at the beginning of his career. Then we show another consequence of our model for organizational design, and identify a force toward multiple equilibria.

5.1 Learning Ambition

In the model of section 2, we have assumed that the agent knows his ambition from the beginning of his career. In reality, people's ambitions may become known later, perhaps due to changes in needs or circumstances. We now extend the model to account for this possibility. Relative to the setup above, we make two changes. First, the agent is assumed not to know the value of m until the second period, before choosing his effort level in period 2. The fact that he learns m in the second period and knows only the prior distribution before then is common knowledge. Second, to leave the horizon *after* the agent has learned his ambition the same (thus preserving the effects from section 2), we take a four-period model. We continue to limit our attention to linear equilibria.

Since the agent does not know m in period 1, e_1 does not depend on his ambition. However, just as in the proof of Theorem 1, the symmetry of the problem starting in period 2 can be shown

[31] Note that substitution of effort to increase observed hours is only going to be a problem at the beginning of the agent's career. Even if h_2 is observed before setting the wage in period 3, the agent will not exert excessive effort to increase it. In fact, the agent will go out of his way to *decrease* h_2. This allows him to take advantage of a negative backward attribution, so that his earlier outputs seem more of a reflection of ability. If the feigned "slacking" is costly to the principal, this might be another reason not to observe h_2.

to imply that $\alpha_2 = \alpha_3 \equiv \alpha$. Then, the following expressions describe the extent to which future wages respond to past outputs:

$$\begin{aligned}
\frac{\partial w_2}{\partial q_1} &= \frac{\sigma_a^2}{\sigma_a^2+\sigma_\epsilon^2};\ \left(\frac{\sigma_a^2+\alpha^{***}\sigma_m^2}{\sigma_\epsilon^2+\sigma_a^2+\alpha^{***}\sigma_m^2}\right) \\
\frac{\partial w_3}{\partial q_1} &= \frac{\sigma_a^2\sigma_\epsilon^2+\alpha^2\sigma_a^2\sigma_m^2}{(\sigma_\epsilon^2)^2+2\sigma_a^2\sigma_\epsilon^2+\alpha^2\sigma_m^2(\sigma_a^2+\sigma_\epsilon^2)};\ \left(\frac{\sigma_a^2+\alpha^{***}\sigma_m^2}{\sigma_\epsilon^2+2(\sigma_a^2+\alpha^{***}\sigma_m^2)}\right) \\
\frac{\partial w_3}{\partial q_2} &= \frac{\sigma_a^2\sigma_\epsilon^2+\alpha^2\sigma_m^2(\sigma_a^2+\sigma_\epsilon^2)}{(\sigma_\epsilon^2)^2+2\sigma_a^2\sigma_\epsilon^2+\alpha^2\sigma_m^2(\sigma_a^2+\sigma_\epsilon^2)};\ \left(\frac{\sigma_a^2+\alpha^{***}\sigma_m^2}{\sigma_\epsilon^2+2(\sigma_a^2+\alpha^{***}\sigma_m^2)}\right) \\
\frac{\partial w_4}{\partial q_1} &= \frac{\sigma_a^2\sigma_\epsilon^2+2\alpha^2\sigma_a^2\sigma_m^2}{(\sigma_\epsilon^2)^2+3\sigma_a^2\sigma_\epsilon^2+2\alpha^2\sigma_m^2\sigma_\epsilon^2+2\alpha^2\sigma_m^2\sigma_a^2};\ \left(\frac{\sigma_a^2}{\sigma_\epsilon^2+3(\sigma_a^2+\alpha^{***}\sigma_m^2)}\right) \\
\frac{\partial w_4}{\partial q_2} &= \frac{\partial w_4}{\partial q_3} = \frac{\sigma_a^2\sigma_\epsilon^2}{(\sigma_\epsilon^2)^2+3\sigma_a^2\sigma_\epsilon^2+2\alpha^2\sigma_m^2\sigma_\epsilon^2+2\alpha^2\sigma_m^2\sigma_a^2};\ \left(\frac{\sigma_a^2}{\sigma_\epsilon^2+3(\sigma_a^2+\alpha^{***}\sigma_m^2)}\right)
\end{aligned} \tag{24}$$

To facilitate comparison with the model in which m is known from the beginning, we have put the corresponding expressions for that case in parentheses. Similarly to the three-period model, $k\alpha^{***} = \frac{\sigma_a^2}{\sigma_\epsilon^2+3(\sigma_a^2+\alpha^{***}\sigma_m^2)}$. The derivation of these expressions, as well as the proofs of the non-trivial claims in this section, can be found in Theorem 5 in the appendix.

The first noteworthy property of this variation of the model is that w_3 is more responsive to q_2 than it is to q_1. Thus, if period 3 is important enough relative to periods 2 and 4, the agent's level of effort is higher in period 2 than in period 1. In other words, effort is not monotonically decreasing over the career. Given our result that ambition-proving incentives are likely to increase effort early in the career, this wrinkle should not be too surprising: ambition-proving incentives can only affect the agent's behavior once they are known, so effort can increase while the agent learns m.[32] Intuitively, as it becomes known to the principal that the agent has figured out his ambition, she starts making inferences about it from his output. Therefore, the agent is forced to work hard to prove that his newly learned marginal utility of income is high. Thus, our model endogenously generates a feature of career paths in modern society that many find unfortunate. Specifically, the agent has to work hardest in his career in exactly the same period in which he also figures out his personal life. Broadly interpreting our model for the marriage example, as the agent takes his vow and decides on the relative importance of career and family, incentives induce him to prove

[32] Note that the same is not true in a standard career concerns model. There, the agent does not need to know his type to signal it.

his attachment to his job. Evidence on the well-known marriage premium is consistent with our model. Korenman and Neumark (1991) provide evidence (in the form of supervisor evaluations) that the premium is largely due to harder work on the part of married men. Loh (1996) shows that the marriage premium is the same for men with working and non-working wives, and is non-existent for the self-employed. These facts are not consistent with Becker's (1991) division of labor story or models in which marriage simply changes the agent's preferences (and does not lead to the signaling thereof). They indicate that a signaling motive such as ours may be an important part of the explanation.

In addition to the *timing* of incentives, it is interesting to compare the overall *strength* of incentives in this model to one where the agent learns m at the very beginning.[33] Clearly, due to a lack of opportunity to signal ambition, $\frac{\partial w_2}{\partial q_1}$ decreases. At the same time, it is easy to verify that $\frac{\partial w_3}{\partial q_2}$ and $\frac{\partial w_4}{\partial q_1}$ unambiguously increase. Two effects increase $\frac{\partial w_3}{\partial q_2}$. If the agent only learns m in period 2, the principal learns nothing about it from q_1, so more can be proven through output in period 2. In addition, since the agent's period 1 effort cannot depend on m, the backward attribution is also weakened. Similarly, $\frac{\partial w_4}{\partial q_1}$ increases because no backward attribution operates based on output in period 1. In fact, $\frac{\partial w_4}{\partial q_1}$ is higher than it would be in a standard career concerns framework. The reason is that while an increase in q_1 can only be due to ability (besides noise), an increase in q_2 and q_3 can be due to either ability or effort. Therefore, q_1 is more important in the principal's updating of ability.

Expressions $\frac{\partial w_4}{\partial q_2}$ and $\frac{\partial w_3}{\partial q_1}$ are more complicated, and their relationship to the similar derivative in a model with m known from the beginning depends on $\frac{\sigma_a^2}{\sigma_\epsilon^2}$. Take $\frac{\partial w_4}{\partial q_2}$ first. If observation of output is relatively noisy ($\sigma_\epsilon^2 > 2\sigma_a^2$), then w_4 is now more responsive to q_2, but if $\sigma_\epsilon^2 < 2\sigma_a^2$, it is less responsive. This is due to two opposing forces. Since the agent does not know his level of needs m in period 1, he is not punished by backward attribution. Intuitively, even if the principal observes a high output in period 2, she cannot infer that the agent must have exerted a higher level of effort in period 1. Therefore, the principal does not attach less meaning to a high output in period 1. Because his period 2 effort does not feed back into a pessimistic interpretation of q_1, the

[33] The following discussion ignores the fact that in general $\alpha \neq \alpha^{***}$. The appendix shows that this would not change any of the conclusions.

agent works harder in that period. Opposing this is another "precision effect." When the agent does not know his ambition, the principal can back out more about his ability from period 1 output. With less variation remaining, it is harder to prove one's ability in period 2, decreasing incentives. Naturally, the precision effect is relatively stronger when the noise is smaller, so it outweighs the former effect when σ_ϵ^2 is small. $\frac{\partial w_3}{\partial q_1}$ also depends on two opposing forces. The precision effect increases $\frac{\partial w_3}{\partial q_1}$, since it helps the agent prove his ability. On the other hand, since effort in period 1 cannot depend on the agent's ambition, there is no forward attribution coming from q_1, decreasing $\frac{\partial w_3}{\partial q_1}$.

One can prove that the *overall* responsiveness of q_3 and q_4 to previous outputs increases when m is not known in period 1; that is, the sums $\frac{\partial w_3}{\partial q_1} + \frac{\partial w_3}{\partial q_2}$ and $\frac{\partial w_4}{\partial q_1} + \frac{\partial w_4}{\partial q_2} + \frac{\partial w_4}{\partial q_3}$ increase. By eliminating the backward attribution, learning m later in the career effectively allows the agent to signal his ability and his ambition separately–ability first, then ambition. One way to see this is to assume that there is almost no noise in the observation of output ($\sigma_\epsilon^2 \approx 0$). Then, expressions 25 indicate that $\frac{\partial w_2}{\partial q_1}$, $\frac{\partial w_3}{\partial q_1}$, $\frac{\partial w_4}{\partial q_1}$, and $\frac{\partial w_3}{\partial q_2}$ are all close to 1. Variations in q_1 are attributed solely (and completely) to ability, which affects pay in all periods. Then, variations in q_2 are (near-perfect) signals of ambition, which affects pay in period 3. No standard career concerns model can generate so much sensitivity to effort. In short, if periods 3 and 4–the periods after the agent has learned m–are important enough in determining pay, a career structure in which decisions about personal life are delayed generates stronger incentives than one in which they are not.

5.2 Catering to the Best

In the fully competitive models we have considered so far, the firm has almost no leeway in setting its pay policy. But we have also noted that the ambition-proving incentives we have identified would affect behavior in many other situations as well, including in ones where the firm *can* resort to explicit incentives. In these situations, a profit-maximizing firm will design incentives taking into account the ambition-proving motive. Without explicitly setting up the firm's maximization problem, our results in section 2 indicate the general direction in which this concern will change incentive design in organizations.[34] Recall the model analyzed in section 2, and, as before, assume that the firm wants to increase its employees' level of effort. If explicit or implicit incentives are important through most of an employee's career, or if output is observed with a lot of noise, the firm wants to increase α^{**}. That is, in devising incentives, the firm wants to cater to its most ambitious and best-performing workers, setting up an organizational structure that motivates the best employees most strongly. It is easy to see that $\frac{\partial w_2}{\partial q_1}$ is increasing in α^{**}, so if this period is sufficiently important, a higher α^{**} is desirable. Intuitively, by committing itself to a system in which the most ambitious people work hardest, a firm induces *everybody* to try to prove that they are ambitious. This increases effort earlier in the career, even before the explicit incentives kick in. The same argument implies that if the firm can manipulate α_1 and α_2 separately, it will try to increase α_2 most. Therefore, firms should cater to their more ambitious employees more strongly in the middle stages of their career.

We should note that this is a stronger statement than just saying that the firm should reward higher output, creating *income* inequality. Rather, the firm wants to create *effort* inequality with an incentive system that induces the best to also work hardest.

Fast-tracking and up-or-out promotion schemes have this property. Under fast-tracking, employees who are successful early are more carefully mentored and monitored, and are more likely to be promoted again. Under up-or-out, the firm either promotes or fires the employee after a certain time. Both of these systems discourage the less successful by creating a category of "dead-end" jobs with no perspective (Kanter 1977).

If output is very accurately observed, or explicit or implicit incentives do not continue to be important in an employee's career, the firm wants to decrease α^{**}. By decreasing the difference between less and more ambitious people, the firm will be able to tell more easily which employees are more talented. Since this is what ultimately determines their wages, workers will be induced

[34] In order to have a fully fledged model of the optimal contract, we would not only have to specify the firm's problem, but also the competitive situation it is facing. Competition is crucial in determining career concerns as well as ambition-proving incentives, as it determines the agent's payoff from improving the firm's impression of him. Carrying out this exercise is beyond the scope of this paper. Instead, we take a shortcut and assume that the firm can manipulate α^{**}. In our model, one way for it to do so is to change its period 3 marginal payoff to earlier performance. For ways to analyze the interaction of explicit incentives and (standard) career concerns, see Meyer and Vickers (1997) and Gibbons and Murphy (1992).

to work hard in an effort to impress the principal. Since many career paths involve proving oneself over a long period of time, and the output of an individual worker is often not easily observed, this latter case seems to be less plausible.

5.3 A Note on Multiple Equilibria

Finally, we provide an example that observing the number of hours the agent works can generate enough feedback to create multiple equilibria with α_t's being nonnegative. Assume that h_1 is observed in periods 2 and 3, that agents' heterogeneity in ambition extends to all periods as in section 2.4, and that $\sigma_\epsilon'^2 = 0$. Since the agent's effort is perfectly observed, he cannot influence the signal about ability extracted in period 1. This implies that the principal's judgment about the agent's ability does not depend on e_1. Therefore, the agent's incentives in period 1 derive exclusively from ambition-proving. If $\alpha_1 = 0$, no ambition proving is possible, so $\alpha_1 = 0$ is an equilibrium. But there is also an equilibrium featuring $\alpha_1 > 0$. Assuming that this is the case, the principal identifies m from h_1, so incentives in period 2 are the same as with standard career concerns. Then, $k\alpha_1 = \frac{\partial w_2}{\partial q_1} = \frac{\alpha_2}{\alpha_1}$.

Intuitively, multiple equilibria can arise because the ambition-proving incentive depends on what the principal expects different types of agents to do. If she expects much more from ambitious agents, all types will try to prove that they are ambitious, since in that situation a high output actually convinces the principal of ambition. And if more ambitious agents respond more strongly to this incentive, the principal's expectations actually materialize. Although this force exists in all of our models, it is only strong enough to generate multiple equilibria in this last example (other than the less interesting possible multiplicity in section 2.4). In other words, if ambitious people are expected to distinguish themselves from others, observers will take good performance to be a sign of ambition, and this motivates even the less ambitious to work hard. However, if the best-performing agents in the economy are not expected to "lead" in this way, everybody will work less hard.

6 Conclusion

It is now a classic insight in economic theory that an employee's concern for his employer's impression of him can create incentives to work hard even when no explicit contract to reward the agent is at the principal's disposal. Models of this "career concerns" tradition generally assume that agents differ in a productivity-relevant dimension of talent, and the heterogeneity in ability gives rise to the desire to distinguish oneself in the eyes of the principal. While acknowledging that signaling about ability is very important, we argue in this paper that economic theory has ignored an equally crucial dimension of employees' attributes: their ambition. We show that introducing heterogeneity in (a specific aspect of) ambition into an otherwise standard career-concerns model qualifies some of its basic insights, and provides a host of novel predictions for the behavior of agents and the structure of organizations.

Our general insight that agents may want to signal about the extent of their career concerns in addition to just their ability can in principle be applied to a variety of career concerns models. We study the delivery of effort in this paper, but a sizable literature focuses on the career concerns of experts providing information (Scharfstein and Stein 1990, Prendergast 1993, Prendergast and Stole 1996, Ottaviani and Sorensen 2001, Li 2001, among others). Signaling about career concerns do not arise in these models for various reasons. In Prendergast and Stole (1996), the possibility of signaling about career concerns is ruled about by their assumption about the manager's preferences, which is a combination about current profit and *current*, end-of-period reputation. Other reputational cheap talk models typically restrict attention to at most three periods. Signaling about career concerns with reputational cheap talk requires at least four periods, and many of the results would be affected. For example, a prediction of multi-period models is that experts tend to stick to their earlier opinions, because admitting that they were wrong implies that they are not as smart (Prendergast and Stole 1996, Li 2001). However, sticking to his opinion makes it more likely that the expert will stick to his opinion again, so his information can be expected to be less useful. Thus, an expert may make a point of *contradicting* his earlier opinion, or *hiding* his early information in order to signal that he will not have to be worried about career concerns in the next period, and can report his opinion truthfully.

Thus, in contrast to a model with effort delivery, experts in a cheap talk model want to signal that they are *not* ambitious. Naturally, the direction of "concerns about career concerns" depend on the original career concerns' inefficiency. If the original career concerns make the principal better off because the agent would work harder, as in the current model, some agents will want to show that they are particularly sensitive to them, using high effort as a credible signal. On the other hand, if the original career concerns make the principal worse off as in many reputational cheap talk models, some agents may have a strong incentive to show that they are oblivious to the pressure of career concerns.

A Proofs

Theorem 1 *In the unique linear equilibrium, $\alpha_1 = \alpha_2 = \alpha^{**} > 0$, and α satisfies*

$$k\alpha^{**} = \frac{\sigma_a^2}{\sigma_\epsilon^2 + 2(\sigma_a^2 + \alpha^{**2}\sigma_m^2)}. \tag{25}$$

Furthermore, the right-hand side of this equation is the derivative of the agent's period 3 wage with respect to q_1 and q_2.

Proof: Together with a and m, the distribution of the observables $q_1 = a + \underline{e}_1 + \alpha_1 m + \epsilon_1$ and $q_2 = a + \underline{e}_2 + \alpha_2 m + \epsilon_2$ is multivariate normal. In particular,

$$E\left[\begin{pmatrix} a \\ m \\ a + \alpha_1 m + \epsilon_1 \\ a + \alpha_2 m + \epsilon_2 \end{pmatrix}\right] = \begin{pmatrix} 0 \\ \mu_m \\ \underline{e}_1 + \alpha_1 \mu_m \\ \underline{e}_2 + \alpha_2 \mu_m \end{pmatrix} \tag{26}$$

and

$$V\left[\begin{pmatrix} a \\ m \\ a + \alpha_1 m + \epsilon_1 \\ a + \alpha_2 m + \epsilon_2 \end{pmatrix}\right] = \begin{pmatrix} \sigma_a^2 & 0 & \sigma_a^2 & \sigma_a^2 \\ 0 & \sigma_m^2 & \alpha_1 \sigma_m^2 & \alpha_2 \sigma_m^2 \\ \sigma_a^2 & \alpha_1 \sigma_m^2 & \sigma_a^2 + \alpha_1^2 \sigma_m^2 + \sigma_\epsilon^2 & \sigma_a^2 + \alpha_1 \alpha_2 \sigma_m^2 \\ \sigma_a^2 & \alpha_2 \sigma_m^2 & \sigma_a^2 + \alpha_1 \alpha_2 \sigma_m^2 & \sigma_a^2 + \alpha_2^2 \sigma_m^2 + \sigma_\epsilon^2 \end{pmatrix}. \tag{27}$$

Now we use the updating rule for multivariate normals to get:[35]

$$E[a|q_1, q_2] = \begin{pmatrix} \sigma_a^2 & \sigma_a^2 \end{pmatrix} \begin{pmatrix} \sigma_a^2 + \alpha_1^2 \sigma_m^2 + \sigma_\epsilon^2 & \sigma_a^2 + \alpha_1 \alpha_2 \sigma_m^2 \\ \sigma_a^2 + \alpha_1 \alpha_2 \sigma_m^2 & \sigma_a^2 + \alpha_2^2 \sigma_m^2 + \sigma_\epsilon^2 \end{pmatrix}^{-1} \begin{pmatrix} q_1 - \underline{e}_1 - \alpha_1 \mu_m \\ q_2 - \underline{e}_1 - \alpha_2 \mu_m \end{pmatrix}. \tag{30}$$

[35] If

$$\begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{21}' \\ \sigma_{21} & \sigma_{22} \end{pmatrix}\right), \tag{28}$$

then

$$\zeta_1 | \zeta_2 \sim N(\mu_1 + \sigma_{21}' \sigma_{22}^{-1}(\zeta_2 - \mu_2), \sigma_{11} - \sigma_{21}' \sigma_{22}^{-1} \sigma_{21}). \tag{29}$$

Therefore, the agent's period 3 wage is:

$$w_3 = \frac{\sigma_a^2(\alpha_2(\alpha_2-\alpha_1)\sigma_m^2+\sigma_\epsilon^2)q_1+\sigma_a^2(\alpha_1(\alpha_1-\alpha_2)\sigma_m^2+\sigma_\epsilon^2)q_2}{(\sigma_a^2+\alpha_1^2\sigma_m^2+\sigma_\epsilon^2)(\sigma_a^2+\alpha_2^2\sigma_m^2+\sigma_\epsilon^2)-(\sigma_a^2+\alpha_1\alpha_2\sigma_m^2)^2}+\kappa, \tag{31}$$

where κ is a constant the agent cannot control. Note that $(\sigma_a^2+\alpha_1^2\sigma_m^2+\sigma_\epsilon^2)(\sigma_a^2+\alpha_2^2\sigma_m^2+\sigma_\epsilon^2)-(\sigma_a^2+\alpha_1\alpha_2\sigma_m^2)^2 > \sigma_m^2\sigma_a^2(\alpha_1-\alpha_2)^2 \geq 0$. Thus

$$\begin{aligned} sign\left(\frac{\partial w_3}{\partial q_1}\right) &= sign(\alpha_2(\alpha_2-\alpha_1)\sigma_m^2+\sigma_\epsilon^2), \text{ and} \\ sign\left(\frac{\partial w_3}{\partial q_2}\right) &= sign(\alpha_1(\alpha_1-\alpha_2)\sigma_m^2+\sigma_\epsilon^2). \end{aligned} \tag{32}$$

The sign and relative size of these coefficients determines the sign and relative size of α_1 and α_2. We first prove that α_1 and α_2 are positive. If one was positive and one was non-positive, both coefficients would be positive, that is, w_3 would increase in both q_1 and q_2. From the first-order condition of the agent's maximization,

$$k(\underline{e}_1+m\alpha_1) = \frac{\partial w_2}{\partial q_1}+m\frac{\partial w_3}{\partial q_1}, \text{ and } km\alpha_2 = m\frac{\partial w_3}{\partial q_2}. \tag{33}$$

Thus, $k\alpha_1 = \frac{\partial w_3}{\partial q_1}$ and $k\alpha_2 = \frac{\partial w_3}{\partial q_2}$. Since both right-hand sides are positive, it is not possible to have negative α_1 or α_2, a contradiction.

If both α_1 and α_2 were non-positive, then at least one of the coefficients would have to be positive. If $\alpha_1 < \alpha_2$, w_3 increases with q_2 for sure, which contradicts $\alpha_2 < 0$. Similarly, if $\alpha_1 > \alpha_2$, then w_3 increases with q_1, which contradicts $\alpha_1 < 0$.

Second, we prove that $\alpha_1 = \alpha_2$. Given that both are positive, if we have $\alpha_1 > \alpha_2$, then w_3 increases faster with q_2. A high α_1 thus entails higher marginal cost of effort but lower marginal benefit in wage payment, a contradiction. A similar argument rules out $\alpha_1 < \alpha_2$.

Once we have established $\alpha_1 = \alpha_2$, it is easy to derive that α must satisfy equation 8. Finally, for positive α, the left-hand side of equation 8 is increasing in α, while the right-hand side is decreasing. Since the right-hand side is greater at zero but smaller for large α, a unique α satisfies the equation. □

Theorem 2 *Suppose heterogeneity in m extends to all periods. Then*

1. *In any equilibrium, $\alpha_2 > 0$.*
2. *An equilibrium with $\alpha_1, \alpha_2 > 0$ exists.*
3. *In any equilibrium in which α_1 and α_2 are positive, $\alpha_1 > \alpha_2$.*

Proof: We prove each part in turn.

1. We prove by contradiction. First, suppose that $\alpha_1, \alpha_2 < 0$. Adding the two equations 10 we get

$$k(\alpha_1+\alpha_2) = \frac{\sigma_a^2+\alpha_1\alpha_2\sigma_m^2}{\sigma_a^2+\alpha_1^2\sigma_m^2+\sigma_\epsilon^2}+\frac{\sigma_a^2\left((\alpha_1-\alpha_2)^2\sigma_m^2+\sigma_\epsilon^2\right)}{\sigma_\epsilon^2(2\sigma_a^2+\sigma_\epsilon^2)+\sigma_m^2\left((\alpha_1^2+\alpha_2^2)\sigma_\epsilon^2+(\alpha_1-\alpha_2)^2\sigma_a^2\right)}. \tag{34}$$

Now the left-hand side of this equation is negative, while the right-hand side is positive. Next, suppose $\alpha_1 \geq 0$ and $\alpha_2 < 0$. This would make α_2 positive, another contradiction. Finally, $\alpha_2 = 0$ contradicts the condition for α_2.

2. For a positive constant K (chosen to be sufficiently large in a way to be specified below), consider the following set in $\Re^2$ space: $\{(\alpha_1, \alpha_2) | 0 \leq \alpha_1, \alpha_2 \leq K, \alpha_1 \geq \alpha_2\}$. On this set, the equilibrium conditions 10 define a map. Call this map f, and let $f_i(\alpha_1, \alpha_2)$ be the ith component of $f(\alpha_1, \alpha_2)$. It is easy to verify the following properties of f:
   - Whenever $\alpha_1 = \alpha_2$, $f_1(\alpha_1, \alpha_2) > f_2(\alpha_1, \alpha_2) > 0$.
   - Whenever $\alpha_2 = 0$, $f_1(\alpha_1, \alpha_2), f_2(\alpha_1, \alpha_2) > 0$.
   - We can choose K so that $f_1(K, \alpha_2) < K$ and $f_2(K, \alpha_2) > 0$ for any $\alpha_2 < K$.
   - f is continuous.

   These imply that for a sufficiently large K, f defines a continuous inward-pointing map. Thus, by the Halpern-Bergman Theorem (Aliprantis and Border 1994, page 549), it has a fixed point. The fixed point is a linear rational expectations equilibrium.
3. We prove by contradiction: assuming $\alpha_1 \leq \alpha_2$ immediately implies $\alpha_1 > \alpha_2$ from expressions 10. □

Lemma 1 *In the infinite-horizon model, if α_t is a constant α, then the variance of $a_t + \alpha m_t$ is also constant.*

Proof: Denote the variance of $a_t + \alpha m_t$ by $\sigma^2_{a+\alpha m,t}$. From simple updating of normals, for any $t' > t$ we have

$$\frac{\partial w_{t'}}{\partial q_t} = \frac{\partial}{\partial q_t} E[a_{t'} + \alpha m_{t'} | q_1, \ldots, q_{t'-1}] = \frac{\sigma^2_{a+\alpha m,t}}{\sigma^2_{a+\alpha m,t} + \sigma^2_\epsilon} \cdot \prod_{s=t+1}^{t'-1} \frac{\sigma^2_\epsilon}{\sigma^2_{a+\alpha m,s} + \sigma^2_\epsilon}. \tag{35}$$

Thus, the total return to increasing effort in period t is

$$x = \frac{\sigma^2_{a+\alpha m,t}}{\sigma^2_{a+\alpha m,t} + \sigma^2_\epsilon} \sum_{t'=t+1}^{\infty} \delta^{t'-t} \prod_{s=t+1}^{t'-1} \frac{\sigma^2_\epsilon}{\sigma^2_{a+\alpha m,s} + \sigma^2_\epsilon}. \tag{36}$$

For α_t to be constant, the total return to increasing effort in period $t+1$ has to be the same. This leads to the following recursion:

$$x = \frac{\sigma^2_{a+\alpha m,t}}{\sigma^2_{a+\alpha m,t} + \sigma^2_\epsilon} \left[\delta + \delta \frac{\sigma^2_\epsilon}{\sigma^2_{a+\alpha m,t+1}} x \right]. \tag{37}$$

After some manipulation, we can put the above in the following form:

$$\delta \left(\frac{1}{\sigma^2_{a+\alpha m,t}} - \frac{1}{\sigma^2_{a+\alpha m,t+1}} \right) = \text{constant} - \frac{1-\delta}{\sigma^2_{a+\alpha m,t}} \tag{38}$$

Now we prove by contradiction. Suppose the left-hand side of equation 38 is not zero. Suppose first that it is positive. Then, $\sigma^2_{a+\alpha m,t+1} > \sigma^2_{a+\alpha m,t}$. Now, take the corresponding expression for $t+1$:

$$\delta \left(\frac{1}{\sigma^2_{a+\alpha m,t+1}} - \frac{1}{\sigma^2_{a+\alpha m,t+2}} \right) = \text{constant} - \frac{1-\delta}{\sigma^2_{a+\alpha m,t+1}} \tag{39}$$

Since the right-hand side of equation 39 is greater than that of equation 38, we have $\sigma^2_{a+\alpha m,t+2} > \sigma^2_{a+\alpha m,t+1}$. Furthermore, the reciprocal of the variance is decreasing at an increasing rate. But that's impossible, since the reciprocal is bounded from below by zero.

Now suppose that the left-hand side of equation 38 is negative. Then, $\sigma^2_{a+\alpha m,t+1} < \sigma^2_{a+\alpha m,t}$, and the right-hand side of equation 39 is smaller than that of equation 38. Therefore, the reciprocal of $\sigma^2_{a+\alpha m,t}$ is increasing at an increasing rate. This implies that $\sigma^2_{a+\alpha m,t} \to 0$ as $t \to \infty$. But since a_t and m_t change every period (by random variables of given variance), the principal's beliefs cannot become very precise. This completes the proof. $\square$

Theorem 3 *A steady state satisfying expressions 12 and 15 always exists. Furthermore,*

1. *Suppose $k^2(1-\delta)^2\sigma^2_\epsilon < \delta^2\sigma^2_\nu$. For $\sigma^2_\eta = 0$, there are two steady states, one with $\alpha = 0$ and one with $\alpha > 0$. For any $\sigma^2_\eta > 0$, there is a unique steady state. As $\sigma^2_\eta \to 0$, the steady state approaches the positive steady state corresponding to $\sigma^2_\eta = 0$.*
2. *Suppose $k^2(1-\delta)^2\sigma^2_\epsilon \geq \delta^2\sigma^2_\nu$. For $\sigma^2_\eta = 0$, the unique steady state has $\alpha = 0$. For a sufficiently small σ^2_η, the steady state is unique, and as $\sigma^2_\eta \to 0$, the corresponding α approaches zero.*

Proof: For any $\sigma^2_\eta \geq 0$, equations 12 and 15 each define a curve of $\sigma^2_{a+\alpha m}$ as a function of α. Call these curves $f(\alpha, \sigma^2_\eta)$ and $g(\alpha)$, respectively. (f has an extra argument σ^2_η since it depends on σ^2_η, while g does not.)

For $\sigma^2_\eta = 0$, $\alpha = 0$ and $\sigma^2_{a+\alpha m} = 0$ is clearly a steady state. Now consider $\sigma^2_\eta > 0$. At $\alpha = 0$, $f(\alpha) > g(\alpha)$. As α approaches $\frac{\delta}{k}$, $g(\alpha) \to \infty$, while $\lim_{\alpha \to \frac{\delta}{k}} f(\alpha) < \infty$. By continuity, they intersect. This intersection defines a steady state.

1. First, assume that $\sigma^2_\eta = 0$. $\alpha = 0$ and $\sigma^2_{a+\alpha m} = 0$ obviously constitute a steady steady state. To look for a positive steady state, substitute equation 15 into equation 12 and divide by $(\sigma^2_{a+\alpha m})^2$. Letting $x = \sigma^2_{a+\alpha m}$, the resulting equation reduces to

   $$k^2x^2 + \left(2k^2(1-\delta)\sigma^2_\epsilon - \delta^2\sigma^2_\nu\right)x + k^2(1-\delta)^2(\sigma^2_\epsilon)^2 - \delta^2(\sigma^2_\epsilon)^2\sigma^2_\nu = 0. \tag{40}$$

   Since $k^2(1-\delta)^2\sigma^2_\epsilon < \delta^2\sigma^2_\nu$, the constant in the above quadratic is negative. Also, the coefficient on x^2 is positive, so the equation has exactly one positive root. Thus, there is a unique positive steady state. Call it α_0.

   Next, we prove that the steady state is unique for a sufficiently small σ^2_η. f and g are continuously differentiable, so for a sufficiently small positive σ^2_η, any steady state is close to either zero or α_0. Since f is strictly increasing in σ^2_η, there is no steady state near zero.

   Notice that $g(\alpha)$ is convex. Also, $f(\alpha)$ strictly increases and becomes strictly flatter as σ^2_η increases. These two facts, together with the uniqueness of the positive steady state α_0, implies that the steady state is unique for a sufficiently small σ^2_η.

   Finally, we prove uniqueness for any positive σ^2_η. Substituting equation 15 into equation 12 defines the following equation for $x = \sigma^2_{a+\alpha m}$:

   $$k^2\left(\frac{(1-\delta)\sigma^2_\epsilon + x}{x}\right)^2\left(x - \sigma^2_\epsilon - \sigma^2_\eta + 2\frac{(\sigma^2_\epsilon)^2}{x+\sigma^2_\epsilon}\right) = \delta^2\sigma^2_\nu. \tag{41}$$

   We are looking for positive roots of this equation. The derivative of the left-hand side of this is $1 + (1-\delta)\frac{\sigma^2_\epsilon}{x}$ times

   $$\left(1 + (1-\delta)\frac{\sigma^2_\epsilon}{x}\right)\left(1 - \frac{(\sigma^2_\epsilon)^2}{(x+\sigma^2_\epsilon)^2}\right) - 2(1-\delta)\frac{\sigma^2_\epsilon}{x^2}\left(x - \sigma^2_\epsilon - \sigma^2_\eta + \frac{(\sigma^2_\epsilon)^2}{x+\sigma^2_\epsilon}\right) \tag{42}$$

   Clearly, whenever the left-hand side of equation 41 is positive, it is decreasing in σ^2_η, and its derivative is increasing in σ^2_η.

We prove by contradiction that equation 41 has a unique root. For a small x, the left-hand side is negative. Therefore, in order for it to have at least two roots, it has to be decreasing over some range. If that is the case, by the above properties it is also decreasing over some for smaller σ_η^2's. In particular, it is decreasing over some range for any σ_η^2 near zero. Then, we can choose σ_ν^2 so that equation 41 also has at least two roots for σ_η^2 near zero. (To do so, we might have to increase σ_ν^2, but that will not lead to a violation of condition $k^2(1-\delta)^2\sigma_\epsilon^2 < \delta^2\sigma_\nu^2$.) But that is a contradiction, because we have already proved that there is only one steady state for σ_η^2 near zero.

2. Whenever $k^2(1-\delta)^2\sigma_\epsilon^2 - \delta^2\sigma_\nu^2 \geq 0$, we also have $2k^2(1-\delta)\sigma_\epsilon^2 - \delta^2\sigma_\nu^2 \geq 0$. This implies that both roots of equation 40 are non-positive. Therefore, for $\sigma_\eta^2 = 0$, the unique steady state has $\alpha = 0$.

   Since there is no positive steady state for $\sigma_\eta^2 = 0$, for any $\xi > 0$, there is a $\psi > 0$ such that if $\sigma_\eta^2 < \psi$, any steady state has $\alpha < \xi$.[36]

   It is easy to check that when $k^2(1-\delta)^2\sigma_\epsilon^2 \geq \delta^2\sigma_\nu^2$, $\frac{\partial f}{\partial \alpha}(\alpha, 0) \geq \frac{\partial g}{\partial \alpha}(\alpha)$ for a sufficiently small α, with the strict inequality holding for $\alpha > 0$. Since $\frac{\partial f}{\partial \alpha}(\alpha, \eta)$ is strictly decreasing in η, the two curves have exactly one intersection near zero.

Theorem 4 *Suppose that the principal can observe h_2 before setting the wage in period 3. Then, equilibrium exists, and in any equilibrium, $\alpha_1 > \alpha_2$. Furthermore, $\alpha_1 + \alpha_2$ is smaller than it would be if the principal could not observe h_2 in period 3 (section 2), and for agents with $m > 0$ so is effort in period 2 and total effort in periods 1 and 2.*

Proof: To prove existence, we use the same method as in part 2 of Theorem 2. For a sufficiently large positive constant K, consider the set $\{(\alpha_1, \alpha_2) | 0 \leq \alpha_1, \alpha_2 \leq K, \alpha_1 \geq \alpha_2\}$ in $\Re^2$-space. On this set, the equilibrium conditions 23 define a map. Call this map f, and let $f_i(\alpha_1, \alpha_2)$ be the ith component of $f(\alpha_1, \alpha_2)$. It is easy to verify the following properties of f:

- Whenever $\alpha_1 = \alpha_2$, $f_1(\alpha_1, \alpha_2) > f_2(\alpha_1, \alpha_2) > 0$.
- Whenever $\alpha_2 = 0$, $f_1(\alpha_1, \alpha_2), f_2(\alpha_1, \alpha_2) > 0$.
- We can choose K so that $f_1(K, \alpha_2) < K$ and $f_2(K, \alpha_2) > 0$ for any $\alpha_2 < K$.
- f is continuous.

These imply that for a sufficiently large K, f defines a continuous inward-pointing map. Thus, by the Halpern-Bergman Theorem (Aliprantis and Border 1994, page 549), it has a fixed point. The fixed point is a linear rational expectations equilibrium.

Subtracting the second of expressions 23 from the first, we get

$$k(\alpha_1 - \alpha_2) = \frac{-(\alpha_1 + \alpha_2)\left(\alpha_1 - \alpha_2 - \alpha_2 \frac{\sigma_\epsilon^2}{\sigma_\epsilon'^2}\right)\sigma_a^2\sigma_m^2\sigma_\epsilon'^2}{\sigma_a^2\sigma_m^2\sigma_\epsilon'^2(\alpha_2 - \alpha_1)^2 + (\sigma_\epsilon'^2 + \alpha_2^2\sigma_m^2)(2\sigma_a^2\sigma_\epsilon^2 + (\sigma_\epsilon^2)^2) + (\alpha_1^2 + \alpha_2^2)\sigma_m^2\sigma_\epsilon^2\sigma_\epsilon'^2}. \tag{43}$$

Instead, adding the expressions 23 gives

$$k(\alpha_1 + \alpha_2) = \frac{(\alpha_1 - \alpha_2)\left(\alpha_1 - \alpha_2 - \alpha_2 \frac{\sigma_\epsilon^2}{\sigma_\epsilon'^2}\right)\sigma_a^2\sigma_m^2\sigma_\epsilon'^2 + 2\sigma_a^2\sigma_\epsilon^2\sigma_\epsilon'^2}{\sigma_a^2\sigma_m^2\sigma_\epsilon'^2(\alpha_2 - \alpha_1)^2 + (\sigma_\epsilon'^2 + \alpha_2^2\sigma_m^2)(2\sigma_a^2\sigma_\epsilon^2 + (\sigma_\epsilon^2)^2) + (\alpha_1^2 + \alpha_2^2)\sigma_m^2\sigma_\epsilon^2\sigma_\epsilon'^2}. \tag{44}$$

[36] To see this, suppose by contradiction that the set of steady states is bounded away from zero as $\sigma_\eta^2 \to 0$. Since f is decreasing in σ_η^2, we could get a bounded sequence of steady states as $\sigma_\eta^2 \to 0$, which are also bounded away from zero. This sequence has a convergent subsequence. The limit of this subsequence would define a positive steady state for $\sigma_\eta^2 = 0$, a contradiction.

First, we have to prove that α_1 and α_2 are positive. If both were negative, equation 44 would indicate that $\alpha_1 - \alpha_2$ and $\alpha_1 - \alpha_2 - \alpha_2 \frac{\sigma_\epsilon^2}{\sigma_\epsilon'^2}$ have opposite signs, giving that the first one is negative and the second is positive. But that contradicts equation 43. If $\alpha_1 > 0$ and $\alpha_2 < 0$, $\frac{\partial w_3}{\partial e_2}$ would be positive, a contradiction. If $\alpha_1 < 0$ and $\alpha_2 > 0$, $\frac{\partial w_3}{\partial e_1}$ would be positive, another contradiction. This proves that α_1 and α_2 are positive.

From equation 43, $\alpha_1 - \alpha_2$ and $\alpha_1 - \alpha_2 - \alpha_2 \frac{\sigma_\epsilon^2}{\sigma_\epsilon'^2}$ have opposite signs. Since α_1 and α_2 are positive, this can only happen if $\alpha_1 - \alpha_2 > 0$ and $\alpha_1 - \alpha_2 - \alpha_2 \frac{\sigma_\epsilon^2}{\sigma_\epsilon'^2} < 0$. Then, the denominator on the right-hand side of 44 is less than $2\sigma_a^2\sigma_\epsilon^2\sigma_\epsilon'^2$.

Suppose by contradiction that $\alpha_1 + \alpha_2$ is at least as large as when h_2 is not observed in period 3. Then, by convexity of the square function, $\alpha_1^2 + \alpha_2^2$ are also larger than before, since before the two were equal. Then, the denominator on the right-hand side above is strictly greater than $\sigma_\epsilon'^2\sigma_\epsilon^2(\sigma_\epsilon^2 + 2(\sigma_a^2 + \alpha^2\sigma_m^2))$, what it would be when h_1 is not observed. Given that the numerator is smaller than $\sigma_\epsilon'^2\sigma_\epsilon^2$ times that in expression 8, $\alpha_1 + \alpha_2$ must be smaller than before. This also means that the total responsiveness of w_3 to previous output is smaller than in the basic model of section 2.

Finally, we show that $\frac{\partial w_2}{\partial q_1}$ also decreases. We have

$$\frac{\partial w_2}{\partial q_1} = \frac{\sigma_a^2 + \alpha_1\alpha_2\sigma_m^2}{\sigma_a^2 + \alpha_1^2\sigma_m^2 + \sigma_\epsilon^2} < \frac{\sigma_a^2 + \alpha_1^2\sigma_m^2}{\sigma_a^2 + \alpha_1^2\sigma_m^2 + \sigma_\epsilon^2}. \tag{45}$$

If α_1 is less than before, the result is clear from the inequality. Even if it is greater, $\alpha_1\alpha_2$ is less than before, since $\alpha_1 + \alpha_2$ is less. Once agent, the fraction on the right-hand side is smaller in that case. □

Theorem 5

1. *In the four-period model in which m is known from the beginning, $\alpha_1 = \alpha_2 = \alpha_3 \equiv \alpha^{***}$, where α^{***} satisfies*
$$k\alpha^{***} = \frac{\sigma_a^2}{\sigma_\epsilon^2 + 3(\sigma_a^2 + \alpha^{***}\sigma_m^2)}. \tag{46}$$
2. *In the four-period model in which the agent learns m in period 2, the responsiveness of wages to output is given by the expression in 25.*
3. *In the four-period model, $\frac{\partial w_3}{\partial q_1} + \frac{\partial w_3}{\partial q_2}$, $\frac{\partial w_4}{\partial q_1} + \frac{\partial w_4}{\partial q_2} + \frac{\partial w_4}{\partial q_3}$, $\frac{\partial w_3}{\partial q_1}$, and $\frac{\partial w_4}{\partial q_1}$ are higher when m is learned in period 2 than when it is learned in period 1.*

Proof:

1. First, note that after period 1, the principal's beliefs about a and m are multivariate normal, with a and m being negatively correlated. Though the proof of theorem 1 was for uncorrelated a and m, it is easily adapted to a situation with correlation. Thus, $\alpha_2 = \alpha_3$.

   Now,
$$V\left[\begin{pmatrix} a \\ a + \alpha_1 m + \epsilon_1 \\ a + \alpha_2 m + \epsilon_2 \\ a + \alpha_3 m + \epsilon_3 \end{pmatrix}\right] = \begin{pmatrix} \sigma_a^2 & \sigma_a^2 & \sigma_a^2 & \sigma_a^2 \\ \sigma_a^2 & \sigma_a^2 + \alpha_1^2\sigma_m^2 + \sigma_\epsilon^2 & \sigma_a^2 + \alpha_1\alpha_2\sigma_m^2 & \sigma_a^2 + \alpha_1\alpha_3\sigma_m^2 \\ \sigma_a^2 & \sigma_a^2 + \alpha_1\alpha_2\sigma_m^2 & \sigma_a^2 + \alpha_2^2\sigma_m^2 + \sigma_\epsilon^2 & \sigma_a^2 + \alpha_1\alpha_3\sigma_m^2 \\ \sigma_a^2 & \sigma_a^2 + \alpha_1\alpha_3\sigma_m^2 & \sigma_a^2 + \alpha_2\alpha_3\sigma_m^2 & \sigma_a^2 + \alpha_3^2\sigma_m^2 + \sigma_\epsilon^2 \end{pmatrix}. \tag{47}$$

   Let V denote the lower right-hand three-by-three submatrix of the above matrix. By the symmetry of periods 2 and 3 in the determination of period 4 wage, we can write $w_4 = b_0 + b_1 q_1 + b_{23} q_2 + b_{23} q_3$

for some constants b_0, b_1, and b_{23}. Furthermore, $\begin{pmatrix} b_1 & b_{23} & b_{23} \end{pmatrix} = \begin{pmatrix} \sigma_a^2 & \sigma_a^2 & \sigma_a^2 \end{pmatrix} V^{-1}$. Therefore $\begin{pmatrix} b_1 & b_{23} & b_{23} \end{pmatrix} V = \begin{pmatrix} \sigma_a^2 & \sigma_a^2 & \sigma_a^2 \end{pmatrix} I = \begin{pmatrix} \sigma_a^2 & \sigma_a^2 & \sigma_a^2 \end{pmatrix}$. Multiplying out this identity and using that $\alpha_2 = \alpha_3$ gives $b_1\alpha_1(\alpha_1 - \alpha_2)\sigma_m^2 + (b_1 - b_{23})\sigma_\epsilon^2 = 0$.

From the agent's maximization problem, α_1 and b_1 must have the same sign, so $b_1\alpha_1 \geq 0$. This implies that $\alpha_1 - \alpha_2$ and $b_1 - b_{23}$ are either both equal to zero or must have opposite signs. But, from the agent's maximization problem once again, $\alpha_1 - \alpha_2$ and $b_1 - b_{23}$ must have the same sign. This completes the proof of this part.

2. After observing q_1, the conditional distributions of a and m are still independent normals. Therefore, $\alpha_2 = \alpha_3 \equiv \alpha$ follows from the proof of Theorem 1. Then, the variance-covariance matrix for the determination of w_4 is

$$V\left[\begin{pmatrix} a \\ a+\epsilon_1 \\ a+\alpha m+\epsilon_2 \\ a+\alpha m+\epsilon_3 \end{pmatrix}\right] = \begin{bmatrix} \sigma_a^2 & \sigma_a^2 & \sigma_a^2 & \sigma_a^2 \\ \sigma_a^2 & \sigma_a^2+\sigma_\epsilon^2 & \sigma_a^2 & \sigma_a^2 \\ \sigma_a^2 & \sigma_a^2 & \sigma_a^2+\alpha^2\sigma_m^2+\sigma_\epsilon^2 & \sigma_a^2+\alpha^2\sigma_m^2 \\ \sigma_a^2 & \sigma_a^2 & \sigma_a^2+\alpha^2\sigma_m^2 & \sigma_a^2+\alpha^2\sigma_m^2+\sigma_\epsilon^2 \end{bmatrix}. \quad (48)$$

We are looking for $E[a|q_1, q_2, q_3]$. For that, we first need the determinant of the lower right-hand 3 by 3 matrix above. It is

$$(\sigma_a^2 + \sigma_\epsilon^2)((\sigma_\epsilon^2)^2 + 2\sigma_\epsilon^2(\sigma_a^2 + \alpha^2\sigma_m^2)) - 2(\sigma_a^2)^2\sigma_\epsilon^2. \quad (49)$$

Then, it is simple to derive the expressions we need.

3. First, consider the responsiveness of w_4. There are two cases. If $\alpha < \alpha^{***}$, then the result follows from the following consideration. Start from the expression for $\frac{\partial w_4}{\partial q_1} + \frac{\partial w_4}{\partial q_2} + \frac{\partial w_4}{\partial q_3}$ when m is learned early (model I). Subtract $\alpha^{***}\sigma_m^2$ from the denominator, multiply numerator and denominator by σ_ϵ^2, change the α^{***}'s in the denominator to α's, and finally add $2\alpha\sigma_a^2\sigma_m^2$ to both numerator and denominator. All these manipulation increase the value of the fraction, and the end result is $\frac{\partial w_4}{\partial q_1} + \frac{\partial w_4}{\partial q_2} + \frac{\partial w_4}{\partial q_3}$ when m is learned late (model II). Thus, we are done for this case. If $\alpha \geq \alpha^{***}$, then it must be the case that the denominator for $\frac{\partial w_4}{\partial q_1} + \frac{\partial w_4}{\partial q_2} + \frac{\partial w_4}{\partial q_3}$ in model 2 is smaller than in model I. Since the numerator is larger in model II, the result follows.

Now, consider the responsiveness of w_3. If $\alpha \geq \alpha^{***}$, the result is immediate. Now consider $\alpha < \alpha^{***}$. Since α is determined by

$$k\alpha = \frac{\sigma_a^2\sigma_\epsilon^2}{(\sigma_\epsilon^2)^2 + 3\sigma_a^2\sigma_\epsilon^2 + 2\alpha^2\sigma_m^2\sigma_\epsilon^2 + 2\alpha^2\sigma_m^2\sigma_a^2}, \quad (50)$$

$\alpha < \alpha^{***}$ implies that $\alpha^2(2\sigma_a^2 + \sigma_\epsilon^2) > 2\alpha^{***}\sigma_\epsilon^2$. Therefore

$$\frac{2\sigma_a^2\sigma_\epsilon^2 + \alpha^2\sigma_m^2(2\sigma_a^2 + \sigma_\epsilon^2)}{(\sigma_\epsilon^2)^2 + 2\sigma_a^2\sigma_\epsilon^2 + \alpha^2\sigma_m^2(2\sigma_a^2 + \sigma_\epsilon^2)} > \frac{2\sigma_a^2\sigma_\epsilon^2 + 2\alpha^{***}\sigma_\epsilon^2}{(\sigma_\epsilon^2)^2 + 2\sigma_a^2\sigma_\epsilon^2 + 2\alpha^{***2}\sigma_m^2\sigma_\epsilon^2}. \quad (51)$$

The right-hand side of this inequality is exactly $\frac{\partial w_3}{\partial q_1} + \frac{\partial w_3}{\partial q_2}$ for model I, whereas the left-hand side is less then the expression for model II. Since $\frac{\partial w_3}{\partial q_1} + \frac{\partial w_3}{\partial q_2}$ is higher, whereas $\frac{\partial w_3}{\partial q_1}$ is lower in model II, $\frac{\partial w_3}{\partial q_2}$ must be higher in model II. Finally, $\frac{\partial w_4}{\partial q_1}$ is obviously higher in model II. □

References

Aliprantis, C. D., and K. C. Border (1994): *Infinite Dimensional Analysis: A Hitchhiker's Guide.* Berlin: Springer-Verlag.

Aron, D. J. (1987): "Worker Reputation and Productivity Incentives," *Journal of Labor Economics*, 5(4, pt. 2), S87–S106.

Becker, G. S. (1991): *A Treatise on the Family*. Cambridge, MA: Harvard University Press.

Bulow, J. I., and L. H. Summers (1986): "A Theory of Dual Labor Markets with Application to Industrial Policy,Discrimination, and Keynesian Unemployment," *Journal of Labor Economics*, 4(3), 376–414.

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

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

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

Goldin, C. (1986): "Monitoring Costs and Occupational Segregation by Sex: A Historical Analysis," *Journal of Labor Economics*, 4(1), 1–27.

Holmström, B. (1999): "Managerial Incentive Problems: A Dynamic Perspective," *Review of Economic Studies*, 66(1), 169–182, first published in 1982.

Holmström, B., and P. Milgrom (1991): "Multi-Task Principal-Agent Analyzes: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics and Organization*, 7, 24–52, special issue.

Kanter, R. M. (1977): *Men and Women of the Corporation*. New York, NY: BasicBooks.

Korenman, S., and D. Neumark (1991): "Does Marriage Really Make Men More Productive?," *The Journal of Human Resources*, 26(2), 282–307.

Kuhn, P. (1993): "Demographic Groups and Personnel Policy," *Labour Economics*, 1(1), 49–70.

Landers, R. M., J. B. Rebitzer, and L. J. Taylor (1996): "Rat Race Redux: Adverse Selection in the Determination of Work Hours in Law Firms," *American Economic Review*, 86(3), 329–348.

Lazear, E. P. (1989): "Pay Equality and Industrial Politics," *Journal of Political Economy*, 97(3), 561–580.

Levin, J. (2001): "Career Concerns and Collective Reputation," Mimeo, Stanford University.

Li, W. (2001): "Dynamic Problems in Information Transmission," Mimeo, Massachusetts Institute of Technology.

Loh, E. S. (1996): "Productivity Differences and The Marriage Wage Premium for White Males," *Journal of Human Resources*, 31(3), 566–589.

Meyer, M. A., and J. Vickers (1997): "Performance Comparisons and Dynamic Incentives," *Journal of Political Economy*, 105(3), 547–581.

Ottaviani, M., and P. Sorensen (2001): "Professional Advice: The Theory of Reputational Cheap Talk," Mimeo, University College London.

PRENDERGAST, C. (1993): "A Theory of "Yes Men"," *American Economic Review*, pp. 757–770.

PRENDERGAST, C., AND L. STOLE (1996): "Impetuous Youngsters and Jaded Old-Timers: Acquiring a Reputation for Learning," *Journal of Political Economy*, 104(6), 1105–1134.

SCHARFSTEIN, D. S., AND J. C. STEIN (1990): "Herd Behavior and Investment," *American Economic Review*, pp. 465–479.