DRIVE AND TALENT

Botond Kőszegi
University of California, Berkeley

Wei Li
University of California, Riverside

Abstract
We analyze ways in which heterogeneity in responsiveness to incentives ("drive") affects employees' incentives and firms' incentive systems in a career concerns model. On the one hand, because more driven agents work harder in response to existing incentives than less driven ones—and therefore pay is increasing in perceived drive—there is a motive to increase effort to signal high drive. These "drive-signaling incentives" are strongest with intermediate levels of existing incentives. On the other hand, because past output of a more driven agent will seem to the principal to reflect lower ability, there is an incentive to decrease effort to signal low drive. The former effect dominates early in the career, and the latter effect dominates towards the end. To maximize incentives, the principal wants to observe a noisy measure of the agent's effort—such as the number of hours he works—early but not late in his career. (JEL: C70, D82, D23)

1. Introduction

Performance evaluations at most organizations feature judgments about a number of attributes that affect an employee's productivity. Of great importance is a person's inherent "ability" or "talent," typically captured in the existing information-economics literature as a variable shifting output by a constant regardless of incentives or the economic environment. Yet in reality, discussions of this kind of talent usually go hand in hand with assessments about other important attributes, such as an employee's loyalty or his willingness to work hard (Kanter 1977; Landers, Rebitzer, and Taylor 1996). In a survey of British employers, for instance, personnel managers identified reasons that applicants were not qualified for key positions as "lack of technical skills" in only 43% of cases, but as "poor

Acknowledgments: 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, the editor, and two anonymous referees for great comments. We are also grateful for the suggestions given by seminar participants at University of California–Berkeley, University of Chicago Graduate School of Business, Massachusetts Institute of Technology, University of California–Riverside, University of Southern California, Stockholm School of Economics, Universitat Pompeu Fabra, and the 9th World Congress of the Econometric Society.

E-mail addresses: Kőszegi: botand@econ.berkeley.edu; Li: wei.li@ucr.edu

Journal of the European Economic Association March 2008 6(1):210–236

attitude, motivation, or personality" in 62% of the cases (Green, Machin, and Wilkenson 1998). A major component of these "intangibles" is an employee's responsiveness to existing explicit or implicit incentives, with a person being more valuable if a given incentive system is more effective at motivating him. Even upon a casual inspection, such drive clearly seems expected of employees at top investment-banking, consulting, and law firms.[1]

In this paper we analyze a career concerns model in which agents have private information on their responsiveness to incentives, and show that the new dimension of heterogeneity generates implications for the behavior of agents and the design of incentive systems that are different from models focusing only on talent. Although we argue that several reasonable formalizations of drive yield similar models, throughout our formal analysis we identify drive with the agent's marginal utility of income in the final period, m. We examine incentives in an otherwise largely standard career concerns framework: An agent's output depends on his unobservable level of effort; explicit performance-contingent pay, if any, does not fully reward increases in output; and the wage is set period by period in a competitive market.

To study how heterogeneity in drive affects behavior in the simplest possible setting, we begin in Section 2 by considering a two-period model with no heterogeneity in talent. The agent receives a base salary at the beginning of each period, and some performance-contingent pay ("bonus") at the end. Because more-driven agents respond more strongly to the prospect of a period-2 bonus, but the bonus does not return the full product of their higher effort, competition makes the base salary in period 2 increasing in perceived drive. This sensitivity in turn creates *drive-signaling incentives*: It motivates all agents to show that they are driven, raising effort in period 1 above that warranted by the explicit incentives alone. Moreover, drive-signaling incentives are hump-shaped in existing incentives. At low levels of performance-based pay, effort is largely not rewarded by bonuses, so the period-2 base salary is quite sensitive to expected effort. Because an increase in the performance-dependence of pay motivates driven agents more strongly than less driven ones, it increases the extent to which the base salary rewards perceived drive, increasing drive-signaling incentives. At high levels of bonus, however, the base salary is not very sensitive to expected effort. Moreover, because making pay even more dependent on performance "crowds out" the base salary, it decreases drive-signaling incentives.

1. The law firm of Allen and Overy, for example, writes the following to potential recruits: "If you are looking for a career in a premier international law firm, and think you have the skills and drive to succeed, then Allen & Overy offers the ideal environment for realizing your ambitions." Expressing a similar spirit, 51% of law partners in the surveys of Landers, Rebitzer, and Taylor (1996) ranked "ambition for success in the legal profession" as "very important" or "of the utmost importance" in promotion decisions.

To accommodate the seemingly important possibility that a person starts off unsure about how driven he will be, we consider a three-period extension of the above model in which the agent learns m in the second period. Because the principal cannot make inferences about m before the agent knows it, drive-signaling incentives arise only in the second period, inducing the agent to work harder than in the initial period. Because most people presumably learn much about their drive when they are also learning about crucial aspects of their personal lives, they may have to work hardest in their careers in exactly these sensitive times.

In Section 3, we incorporate heterogeneity in talent into our framework. In standard career concerns models with no heterogeneity in drive, because more talented agents produce more on average, high output is taken by the principal as a sign of high ability and hence of high future production. Even without explicit incentives in place, therefore, the agent exerts costly effort in an attempt to show the principal that he is talented (Fama 1980; Holmström 1999). To focus on the interaction between these career concerns incentives and drive-signaling incentives, we assume in this version of the model that there is no performance-contingent pay. And because career concerns do not generate incentives in the last period, to make our model most comparable to the previous one we assume that there are three periods.

Because more-driven agents respond more strongly to career concerns incentives in period 2, the period-2 wage is increasing in perceived drive, so that there is an incentive to increase output in period 1 to show one's drive. We show that under reasonable assumptions, this "forward attribution" increases period-1 effort relative to a standard career concerns model. But if the principal becomes convinced in period 2 that the agent is driven, she must conclude that his past output reflects more toil, and thus less talent, than she had previously thought. Because all types of agents exert zero effort in period 3 and hence the expected output depends only on talent, this "backward attribution" decreases effort in period 2. Heterogeneity in drive therefore increases effort at the beginning of the agent's career, but decreases effort later.

As do most career concerns models, our models assume that the principal bases her inferences about the agent only on his output. In real organizations, however, other information is often available, and the principal may be able to decide what information to observe. In Section 4, we consider the incentive effects of observing a noisy measure of the agent's effort, such as the number of hours he spends at the office. This kind of information allows the principal to make better inferences about the agent's drive, so an increase in effort leads to a greater increase in perceived drive. As we argued previously, higher perceived drive increases the agent's wage in period 2 but decreases his wage in period 3. Hence, incentives are strongest if the principal observes hours early in the agent's career, but not later.

Our theory is in the career concerns tradition in that incentives derive from trying to influence the principal's beliefs. But whereas existing career concerns models typically simplify analysis by assuming symmetric information about all relevant variables (Gibbons and Murphy 1992; Baker, Gibbons, and Murphy 1994; Holmström 1999), we provide a tractable framework in which there is asymmetric information, and such asymmetry has interesting implications due to the role of output and hours as signals of past, present, and future effort levels.[2] In assuming heterogeneity in employee preferences, our work is related to models in Aron (1987) and Landers, Rebitzer, and Taylor (1996). These models do not incorporate heterogeneity in ability and focus primarily on using some measures of effort (such as billable hours) as a screening device. Nevertheless, Landers et al. also argue that in order to select those who will work hard as partners, law firms make long hours a prerequisite for promotion. Although screening by billable hours seems important, even their survey indicates a role for signaling: Hours requirements are usually not explicit in law firms, and when they are, the expectation is for associates to work much more. Kuhn and Lozano (2005) provide further evidence that employees' work hours are driven in part by signaling considerations.[3]

The paper proceeds as follows. In Section 2, we identify how heterogeneity in responsiveness to incentives builds on existing incentives and boosts the agent's effort. Section 3 adds heterogeneity in ability and Section 4 considers whether observing work hours increases the agent's effort. Section 5 discusses the role of some of our modeling assumptions, and Section 6 concludes. All proofs are collected in the Appendix.

2. A Basic Model of Drive-Signaling Incentives

This section presents a simple model of an employment relationship that demonstrates some key implications of heterogeneity in responsiveness to incentives. To isolate the effects of drive-signaling incentives, throughout this section we assume heterogeneity in drive but not in talent. In Section 3, we consider a setting with heterogeneity in both.

---

2. Because a more-driven agent in our model derives more utility from an increase in effort, one may reinterpret our model as a type of standard career concerns model in which ability and effort are complements in producing output. Dewatripont, Jewitt, and Tirole (1999a, 1999b) formalize this possibility by assuming that the agent's effort and talent enter the production function multiplicatively. They show that such complementarity gives rise to an equilibrium multiplicity absent from previous career concerns models. Their model, however, also assumes symmetric information about ability, so that no signaling occurs.

3. The authors document that from 1979 to 2002, the frequency of long work hours (over 50 per week) increased by 14.4% among highly educated, highly paid salaried men. They show that the increase was not due to a change in the demographic composition of the labor force or the level of real hourly earnings. Instead, it was due to an increase in future earnings associated with working beyond 40 hours, suggesting a signaling role of hours.

2.1. Setup

A risk-neutral principal employs an agent for two periods $t = 1, 2$. Output in period t is $q_t = e_t + \varepsilon_t$, where e_t is the agent's effort level and $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ is a noise term, with ε_1 and ε_2 independently distributed. Denoting the agent's wage in period t by w_t, the principal's profit is $\sum_t (q_t - w_t)$, so that each unit of output is valued at 1.

We assume that at the beginning of each period, the principal offers a wage schedule restricted to be of the form $w_t = b_t + \beta_t q_t$. The term b_t can be interpreted as the agent's "base salary" and $\beta_t q_t$ as his performance-contingent "bonus," which serves as a reduced-form representation for some incentives already in place. For simplicity and to focus on the agent's behavior for *given* explicit incentives, we take β_1 and β_2 to be fixed, so that the principal has no control over explicit incentives. Although we later discuss some implications of our findings for the optimal choice of explicit incentives, we do not explore this issue fully in this paper. Following the career concerns tradition, we assume that wages are determined in a perfectly competitive market. Consequently, the principal makes zero expected profit in each period. In the same tradition, we assume that it is impossible (or undesirable) for the principal to use a fully explicit incentive contract: $\beta_1, \beta_2 < 1$.[4]

The key assumption of our model is that risk-neutral agents differ in their marginal utility of income $m \sim N(\mu_m, \sigma_m^2)$, where m is independent of the error terms ε_t and $\mu_m > 0$.[5] The agent knows his marginal utility, but the principal does not. Although heterogeneity in marginal utility may more realistically be modeled as extending to all periods of life, we assume throughout our analysis that it is limited to the last period. This assumption substantially simplifies our derivations, isolates the incentives generated by signaling about a single period's drive, and captures a situation where agents differ mostly in their valuation of some late-career reward, such as becoming a CEO or president. In Section 5, we discuss the limited ways in which our results would be modified if marginal utility of income differed in all periods. The agent also incurs a separable cost of effort $c(e) = e^2/2$. Combining this with his utility from income, out of which he

4. The existing literature identifies a number of compelling reasons not to use fully explicit incentive contracts. It may simply be that some parts of output are unverifiable, so no contract can be written on them. Holmström and Milgrom (1991) show that if a "multitasking" agent is rewarded for one task, she may neglect another task that is important for the principal. Lazear (1989) argues that if workers have strong incentives and are competing for the same reward, they will sabotage others' output.

5. The agent's drive is assumed to be drawn from a normal distribution mostly for technical reasons. Although this first model may be tractable with other distributions, extensions in which the principal updates about multiple attributes (drive and talent) of the agent are not. A normal distribution has the unattractive property that it assigns a positive probability to negative marginal utility. None of the forces and intuitions in our paper seem to rely on the existence of agents with $m < 0$.

cannot save or borrow, his optimization problem is

$$\max_{e_1, e_2(q_1)} E_{q_1,q_2}[w_1 - c(e_1) + mw_2(q_1) - c(e_2(q_1))].$$

We look for the rational-expectations equilibria of this game. Such an 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 her expectations about his behavior. To keep the analysis tractable, we focus on (pure-strategy) linear equilibria, in which the effort level is a non-negative linear function of m in each period: $e_t = \underline{e}_t + \alpha_t m$ with $\alpha_t \geq 0$.[6] Linear equilibria are natural candidates to consider, because the agent's marginal utility of income increases linearly with m and the cost function is quadratic.

Although our model formalizes differences in responsiveness to incentives as deriving from differences in marginal utility of income, several other sources of heterogeneity in this responsiveness would generate qualitatively similar results. One alternative possibility is that agents differ in their marginal cost of effort. Because an agent's behavior only depends on his marginal utility of income *relative to* his marginal cost of effort, heterogeneity in m and heterogeneity in the marginal cost of effort lead to very similar models. Another alternative is that more-driven people plan to stay in their current careers for a long time, whereas less-driven ones intend to switch jobs or quit the labor market altogether. In a way, such heterogeneity is also about taste for income: Those who are not going to work in the future care less about *the income they could earn if they did*. There is, however, a small difference: With heterogeneity in the length of career, the mere fact that the agent shows up for work provides information about his type (it gives a new lower bound for the length of his career). In Section 5, we discuss how this consideration could affect our results.

In contrast to our use of the word, the everyday meaning of "drive" could also encompass an "intrinsic," incentive-independent willingness to work hard. Formally, for instance, agents could have cost of effort $c(e) = (e - k)^2/2$, with heterogeneity in k. Because this kind of willingness to work shifts output by a constant independently of existing incentives, it is formally equivalent to the notion of ability used in much of the economics literature, including this paper.

6. In the model of this section, there may be an equilibrium with $\alpha_1 < 0$. Intuitively, if the principal expects the agent to produce less when he is more driven, he might have an incentive to destroy output to prove his drive. And because agents with higher drive care more about the wage in period 2, they destroy more output, confirming the principal's expectations. Because the existence of this equilibrium relies on the possibility of exerting costly effort to destroy output, we rule it out by assumption.

2.2. Drive-Signaling Incentives

We now analyze the simple model introduced above, showing that explicit incentives in period 2 can generate drive-signaling incentives in period 1 that are hump-shaped in existing incentives.

Because the second period is the last of the agent's career, at that point he is only motivated by the performance-based bonus β_2. Specifically, he chooses effort e_2 to maximize $m\beta_2 E[q_2 \mid e_2] - c(e_2)$. Because $E[q_2 \mid e_2] = e_2$, the optimal choice of effort satisfies the first-order condition $e_2 = \beta_2 m$. Since in a linear equilibrium $e_2 = \underline{e}_2 + \alpha_2 m$, this yields $\underline{e}_2 = 0$ and $\alpha_2 = \beta_2$. Competition implies that from the principal's perspective, the agent's expected output in period 2 must be equal to his expected total compensation, so that $\alpha_2 E[m \mid q_1] = E[w_2 \mid q_1] = b_2 + \beta_2\alpha_2 E[m \mid q_1]$.[7] Using $\alpha_2 = \beta_2$, the base salary in period 2 is

$$b_2 = \alpha_2(1-\beta_2)E[m \mid q_1] = \beta_2(1-\beta_2)E[m \mid q_1].$$

Because b_2 can depend on q_1, in the first period the agent is in general motivated by both the first-period bonus and the second-period base salary. Formally, he chooses effort to solve

$$\max_{e_1}\; E_{q_1}[\beta_1 q_1 + m(\beta_2(1-\beta_2)E[m \mid q_1]) - c(e_1)]. \tag{1}$$

By the updating rule for normals,

$$E[m \mid q_1] = \frac{\sigma_\varepsilon^2}{\alpha_1^2\sigma_m^2 + \sigma_\varepsilon^2}\mu_m + \frac{\alpha_1\sigma_m^2}{\alpha_1^2\sigma_m^2 + \sigma_\varepsilon^2}q_1. \tag{2}$$

Substituting equation (2) into equation (1) and using $E[q_1 \mid e_1] = e_1$ and $e_1 = \underline{e}_1 + \alpha_1 m$, the agent's maximization problem leads to the first-order conditions

$$\underline{e}_1 = \beta_1 \quad \text{and} \quad \alpha_1 = \beta_2(1-\beta_2)\frac{\alpha_1\sigma_m^2}{\alpha_1^2\sigma_m^2 + \sigma_\varepsilon^2}. \tag{3}$$

These two equations reflect two sources of incentives. The first derives from the explicit incentives for period-1 output, which all agents care about equally, and the second derives from the reward implicit in the period-2 base salary, which more-driven agents care about more. Analyzing equation (3) yields the following theorem:

7. In a slight abuse of notation, m represents drive from the perspective of both the agent (who knows its realization) and the principal (who is uncertain about it). The appropriate meaning should be clear in each case.

THEOREM 1. *Consider an agent with positive drive* ($m > 0$). *His effort in period 2 is determined by explicit incentives alone* ($\alpha_2 = \beta_2$). *In the first period,* $\underline{e}_1 = \beta_1$. *To identify further properties, we distinguish two cases:*

1. *If* $\sigma_\varepsilon^2/\sigma_m^2 < \beta_2(1-\beta_2)$, *then there exists a linear equilibrium in which* $\alpha_1 > 0$. *In this equilibrium,* e_1 *is higher than that warranted by the bonus alone* ($e_1 > \beta_1$). *Furthermore,* e_1 *is increasing in* β_2 *on the interval* $(0, 1/2]$, *but decreasing in* β_2 *on the interval* $[1/2, 1)$. *There also exists an equilibrium in which* $\alpha_1 = 0$.
2. *If* $\sigma_\varepsilon^2/\sigma_m^2 \geq \beta_2(1-\beta_2)$, *then there exists a unique linear equilibrium, and in this equilibrium the agent's effort is determined by the bonus alone* ($\alpha_1 = 0, e_1 = \beta_1$).

Equation (3) implies that there is always an equilibrium with $\alpha_1 = 0$, and in fact the second part of Theorem 1 says that if output is observed with too much noise ($\sigma_\varepsilon^2/\sigma_m^2$ is high), this is the only equilibrium. Intuitively, if the principal expects driven and less-driven agents to exert the same effort in period 1, she attributes all variation in production to noise, and none to drive. As a result, performance in period 1 cannot affect future compensation. Because agents have the same marginal utility for period-1 compensation, they work equally hard, confirming the principal's expectations. But as the intuition makes clear, an arbitrarily small correlation of marginal utilities across periods would lead the principal to make inferences about period-2 effort from period-1 output, eliminating this equilibrium. Hence, we do not consider this equilibrium as the compelling one.[8]

The first part of Theorem 1 says that there may also be an equilibrium in which the agent exerts higher effort in the first period than that generated by the bonus alone. This high effort is due to an incentive to manipulate the principal's beliefs about drive, an incentive we shall henceforth call the *drive-signaling incentive*. Given $\alpha_1 > 0$, equation (2) implies that an increase in q_1 increases the mean of the principal's beliefs about m, and because she knows that more driven agents will respond more strongly to explicit incentives in period 2, this increases the effort she expects in period 2. This *forward attribution*, reflected in the base salary in period 2, boosts first-period effort.

As Theorem 1 shows and this intuition explains, the positive equilibrium does not exist if pre-existing incentives in the second period are zero ($\beta_2 = 0$): In this case, all agents put in zero effort in period 2, so perceived drive is not rewarded. This identifies a crucial property of drive-signaling incentives: they *piggy-back* on existing incentives. Moreover, the relationship between drive-signaling incentives and existing incentives is non-monotonic. At low levels of β_2, effort is largely not

8. We have confirmed that the results of the positive equilibrium of our model, including all comparative statics, hold for the unique equilibrium of a corresponding model in which $\beta_1, \beta_2 > 0$ and m differs in both periods.

rewarded by bonuses, so the base salary is quite sensitive to expected effort in period 2. Because an increase in β_2 motivates driven agents more strongly than less-driven ones, it increases the extent to which the base salary rewards perceived drive, increasing drive-signaling incentives.[9] At high levels of β_2, however, the brunt of compensation comes from bonuses, so the base salary is not very responsive to expected effort. And because a further increase in β_2 decreases the share of compensation coming from the base salary, it decreases drive-signaling incentives.

Even though Theorem 1 focuses on the agent's effort supply for exogenously given bonuses β_1, β_2, its results suggest some surprising ways that the principal may want to endogenously choose those explicit incentives to maximize overall incentives. Suppose a firm hires some employees for both periods, and cannot set a fully efficient pay-for-performance incentive in period 1 ($\beta_1 < 1$). Then, it is optimal to commit to an *intermediate* level of explicit incentives in period 2 ($1/2 < \beta_2 < 1$), trading off the power of incentives in period 2 with creating implicit incentives for period 1.

Our model has thus far assumed that the agent knows his drive from the beginning of his career. In reality, many young people do not know how important money and career will be for them until, for instance, they settle down and start a family. To accommodate this scenario in the simplest possible way, we add to our model an early period $t = 0$ when the agent does not know his later drive, and after which he does learn it. In this way, the agent's behavior in periods 1 and 2 is exactly as in the two-period model, immediately yielding the following corollary:

COROLLARY 1 (Non-monotonic effort supply over time). *Suppose that* $\sigma_\varepsilon^2/\sigma_m^2 < \beta_2(1-\beta_2)$. *In a three-period model in which the agent does not learn his drive until period 1, there is a unique linear equilibrium with* $\alpha_1 > 0$. *If explicit incentives are constant and positive* ($\beta_0 = \beta_1 = \beta_2 > 0$)*, in this equilibrium the agent's average effort in period 1 is greater than his average effort in period 0.*

Contrary to predictions of many existing career concerns models, the agent's effort is not monotonically decreasing over his career. Intuitively, the initial output cannot depend on drive, so it cannot form the basis for drive-signaling incentives. Once the agent learns his drive, however, the principal starts making inferences about it from his output, so the agent works hard to signal that it is high. Hence, our model generates a unique feature of career paths in modern society: People have to work hardest in their careers in exactly the same period in which they are also figuring out their personal life—such as around the time of marriage—because

9. In contrast, standard career concerns models predict that if the future power of performance-based pay increases, effort today falls. Intuitively, if a larger part of future compensation comes from bonuses, the base salary is less sensitive to perceived ability, decreasing career concerns incentives.

presumably it is at this time that many learn a lot about their drive. Some evidence on the well-known male marriage premium is consistent with our model.[10]

3. Adding Heterogeneity in Talent

Because most circumstances of interest feature heterogeneity in talent in addition to heterogeneity in drive, we now introduce this possibility into our framework. To focus on the interaction of "talent-signaling" and drive-signaling incentives, we make two additional changes to the previous model. First, we assume that there is no performance-contingent pay. Second, in order to keep the number of periods in which the agent has positive incentives at two, we work with a three-period model.

3.1. Setup

The agent works for three periods $t = 1, 2, 3$. Output in period t is $q_t = a + e_t + \varepsilon_t$, where $a \sim N(0, \sigma_a^2)$ is his time-invariant ability, e_t is his effort level, and $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ is a noise term. Talent a is unknown to the principal. We assume that no explicit incentives exist ($\beta_t = 0$), so that the principal is limited to paying a fixed wage each period. Explicit incentives would complicate our expressions, introduce effects similar to those analyzed in the previous section, and weaken but not qualitatively affect the forces we identify below. The remaining assumptions are identical to those in the previous section. The disutility of effort is $c(e) = e^2/2$, which is additively separable from the utility from consumption. Agents differ in their privately known marginal utility of income in the last period, $m \sim N(\mu_m, \sigma_m^2)$.[11] In Appendix B, we analyze a model in which marginal utility of income differs in all three periods, and in Section 5 we summarize how this qualifies our results. The labor market is perfectly competitive, so that the wage is equal to the agent's expected output. And we still look for linear rational-expectations equilibria, where strategies take the form $e_t = \underline{e}_t + \alpha_t m$.[12]

10. Using data on supervisor evaluations, Korenman and Neumark (1991) provide evidence that the marriage 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 stories based on division of labor within families (such as Becker 1991) and models in which marriage merely 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.

11. If $\sigma_m^2 = 0$, our model becomes equivalent to a standard career concerns model, where the principal makes inferences only about talent.

12. One implicit restriction this imposes on the agent's strategy is that it cannot depend on his beliefs about talent. Even if information about a is symmetric at the beginning, asymmetry of information about effort develops into asymmetry of information about a as well.

The following lemma states some basic properties of the components of the equilibrium strategy; the next subsection analyzes the equilibrium in detail.

LEMMA 1. *In the linear rational-expectations equilibrium of the three-period model,* $e_3 = 0$, $\underline{e}_2 = 0$, $\underline{e}_1 = \partial w_2/\partial q_1$. *Moreover,* $\alpha_1 = \partial w_3/\partial q_1$ *and* $\alpha_2 = \partial w_3/\partial q_2$.

Effort in period 3 is zero because period 3 is the end of the agent's career, and there are no explicit incentives to motivate him. Effort in period 1 is the sum of a constant term $\underline{e}_1$ capturing the agent's response to period-2 incentives, about which all agents care equally, and a term $\alpha_1 m$ coming from period-3 incentives, about which more-driven agents care more. In period 2, all incentives derive from period-3 wage setting, so effort is proportional to m ($\underline{e}_2 = 0$).

3.2. Signaling Drive over the Career

We now show that heterogeneity in drive has opposing implications for career concerns incentives at the beginning of the agent's career (period 1) and towards the end (period 2). We first characterize equilibrium behavior, and then discuss implications of this equilibrium.

THEOREM 2. *Consider an agent with positive drive* ($m > 0$). *In the unique linear equilibrium,* $e_1 > e_2 > e_3 = 0$. *Moreover, the marginal incentive coming from period 3 is positive and equal in the first two periods:* $\alpha_1 = \alpha_2 = \alpha^* > 0$, *where* α^* *satisfies*

$$\alpha^* = \frac{\sigma_a^2}{\sigma_\varepsilon^2 + 2\left(\sigma_a^2 + \alpha^{*2}\sigma_m^2\right)}. \tag{4}$$

Finally, $\underline{e}_1 = (\sigma_a^2 + \alpha^{*2}\sigma_m^2)/(\sigma_\varepsilon^2 + \sigma_a^2 + \alpha^{*2}\sigma_m^2)$.

The agent has no incentive to work in period 3, so $e_3 = 0$ and w_3 is equal to his perceived talent $E[a \mid q_1, q_2]$. The forces determining effort in periods 1 and 2 are readily illustrated by period-2 wage setting by decomposing the zero-profit wage level into two parts:

$$\begin{aligned} w_2 = E[q_2 \mid q_1] = E[a + \alpha_2 m \mid q_1] &= E[E[a \mid m, q_1] \mid q_1] + \alpha_2 E[m \mid q_1] \\ &= \underbrace{\frac{\sigma_a^2}{\sigma_a^2 + \sigma_\varepsilon^2} q_1}_{\text{standard}} + \underbrace{\left(\alpha_2 - \frac{\sigma_a^2}{\sigma_a^2 + \sigma_\varepsilon^2}\alpha_1\right) E[m \mid q_1]}_{\text{drive-signaling incentives}} + C, \end{aligned} \tag{5}$$

where C is a constant. The first term depends only on q_1 and is in fact equal to the term we would have in a version of the model with only standard career concerns ($\sigma_m^2 = 0$). The second term is the drive-signaling incentive, the part of the agent's compensation that depends on the principal's inferences about m. This term reflects the agent's incentives to change the principal's beliefs about his drive, *holding output constant*. Heterogeneity in drive increases the agent's effort relative to the standard career concerns model exactly when this term is positive.

Theorem 2 says that $\alpha_1 > 0$, so an increase in q_1 increases $E[m \mid q_1]$. Moreover, because $\alpha_2 > 0$, an increase in $E[m \mid q_1]$ raises the effort the principal expects in period 2. This *forward attribution* is the analogue of the positive drive-signaling incentives we have identified in the previous section: The principal attributes part of an increase in output in period 1 to drive, leading her to expect higher effort in period 2. In this model, however, there is also an opposing effect: If the principal believes that the agent is driven, given any level of output she downgrades her beliefs about his ability. This *backward attribution* decreases the agent's wage in period 2.

Because Theorem 2 shows that $\alpha_1 = \alpha_2 = \alpha^*$, equation (5) implies that in period-2 wage setting the forward attribution outweighs the backward attribution. Intuitively, holding past output fixed, higher perceived effort due to higher perceived drive does not lead to a one-to-one decrease in perceived talent, because it also leads the principal to believe the agent was less lucky. But higher expected future effort due to higher perceived drive is fully incorporated into the wage.

Whereas the forward attribution outweighs the backward attribution in how the agent's first-period output influences his second-period wage, Theorem 2 shows that the backward attribution dominates later in life: By equation (4), the marginal period-3 benefit of increasing output in the first two periods is smaller than in the standard career concerns model ($\sigma_m^2 = 0$). For part of the intuition, note that when agents differ in their drive, the principal cannot be sure whether higher output is due to harder work or greater inherent ability. Because this makes it more difficult to increase the principal's beliefs about ability, the only determinant of w_3, it diminishes incentives. Even worse, because of the backward attribution inferences about drive made from output in a given period (say, period 2) negatively affect the interpretation of other output (q_1) as well. Intuitively, if an employee well into his career convinces his employer that he is driven, she will conclude that he must have worked hard all these years. Given his past performances, she thus downgrades her opinion of his talent.[13]

13. To see formally that the principal is not merely discounting the additional noise $\alpha^{*2}\sigma_m^2$ introduced by heterogeneity in drive in the second period, note that the coefficient on $\alpha^{*2}\sigma_m^2$ in expression (4) is 2 instead of 1. The effect going the other way, that inferences about drive from earlier output affect the interpretation of later output, is similar to the ratchet effect Laffont and Tirole (1988). If the agent increases output in period 1, the principal concludes that he must be more driven, thus expecting him to work harder in period 2. If he 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.

More generally, Theorem 2 implies that an increase in heterogeneity in drive increases the incentives coming from period-2 wage setting, $\partial w_2/\partial q_1$, and decreases the incentives coming from period-3 wage setting, $\partial w_3/\partial q_1$ and $\partial w_3/\partial q_2$. Intuitively, the forward attribution operating in period-2 wage setting becomes stronger as σ_m^2 increases because the principal makes more of an inference about drive from output. But as σ_m^2 increases, the principal also makes less of an inference about talent, reducing incentives coming from the fully talent-determined period-3 wage. Under the reasonable assumption that period 2—the middle of the career—is quite important in determining agents' incentives, an increase in σ_m^2 increases average effort early in the career, and decreases average effort later.[14]

This analysis and logic also allow us to identify how groups of agents with different perceived average levels of drive, such as men and women, will be paid over time.

Corollary 2. *Consider two identifiable groups of agents with identical prior distributions of talent but different average levels of drive μ_m. Given identical past performances, an agent belonging to the group with lower average drive is paid less in periods 1 and 2, and more in period 3, than an agent belonging to the group with higher average drive.*

Suppose for instance that women are perceived to be less driven on average than men, perhaps because they are thought to balance their careers more evenly with family. Our model then implies that because women will be perceived to respond less strongly to career concerns incentives, they will be paid less in period 1 than men. Because $E[m \mid q_1]$ is increasing in μ_m, equation (5) implies that a woman receives a lower wage in period 2 than a man with the same past performance. Unlike in many previous models, this wage discrimination derives from perceptions about an employee's overall concern for his or her career, not from any firm-specific costs associated with such concern.[15]

More interestingly, our model predicts that the performance-contingent wage gap decreases over time. By equation (5), the gap is smaller in period 2 than in period 1. In fact, a woman will be paid more toward the end of her career than a man whose past performance is the same (so long as the principal's prior beliefs about

14. Although our formal model does not incorporate a variable capturing the "importance" of period 2, this could easily be introduced by adding a weight on period 2 in the agent's utility function, without changing the message of Theorem 2. If this weight is sufficiently large, heterogeneity in drive increases effort early in the career, and decreases effort later.

15. Many existing explanations for discriminatory wage practices rely on some sort of turnover costs, including direct hiring costs or indirect costs such as a loss of investment into the employee's human capital (Kuhn 1993) or a need to resort to costly monitoring (Goldin 1986) or a higher efficiency wage (Bulow and Summers 1986) to ensure that workers are not shirking. In our model, women receive a lower wage than men even if they are not more likely to leave the firm and there is no turnover cost of any kind.

ability were the same). Intuitively, if a woman performs as well as a man despite her lower drive, she is likely to be more talented than he is. Thus, she receives more at the stage when the wage depends largely on ability. This prediction is broadly consistent with the findings of Wellington (1994) that the number of years worked full time, and not the years of training completed or the number of years with the current employer, is the most important variable in accounting for the narrowing gap between men's and women's wages over time.

4. The Incentive Effects of Observing a Measure of Effort

Sections 2 and 3 show how drive-signaling incentives can build on and interact with explicit or implicit incentives to change the agent's effort supply over his career. The analysis in these sections, however, assumes that explicit or implicit compensation is based solely on output, and in reality firms have many other measures of drive or talent at their disposal. Furthermore, whether to use these alternative measures is likely to be a strategic decision firms make to maximize incentives. For instance, the Employee Compensation Survey of the BLS (2004) indicates that in determining pay at large and medium-sized enterprises, the use of measures such as work hours, attendance, and volunteering for difficult assignments has expanded from 17% of employees in 1983 to 42% in 1997.

In this section, we analyze a version of our model with one particular additional measure available to the principal, answering both the positive question of how (an employee's awareness of) a firm's use of this measure affects incentives and the organizational-design question of whether the firm would indeed want to use the measure. The measure we consider is a noisy signal of the agent's effort, such as the number of hours he spends at the office or a supervisor's impression of how hard-working he is.

Similar to the model in Section 3, the agent works for three periods, $t = 1, 2, 3$, and produces output $q_t = a + e_t + \varepsilon_t$ in period t. Agents differ in their marginal utility of income in period 3, $m \sim N(\mu_m, \sigma_m^2)$, and wages are determined competitively. This latter assumption allows us to focus on the essence of a career concerns framework—the agent's incentive to manipulate the principal's beliefs about him—and not on how the set of information the principal chooses to observe may affect market competition. We look for rational-expectations equilibria that are linear in the agent's drive.

In addition to always observing q_t, the principal may now choose to commit to observing $h_t = e_t + \varepsilon'_t$, where $\varepsilon'_t \sim N(0, \sigma_\varepsilon'^2)$ is an independently distributed noise term. For simplicity and to reflect the informal role of hours, we assume that h_t is lost after one period, so that the principal can use h_1 in setting pay in period 2, but not in setting pay in period 3. Observing the agent's hours has the following effects on his effort in the first two periods:

THEOREM 3. *Consider an agent with positive drive* ($m > 0$). *In any equilibrium,* $e_3 = 0$.

1. *Suppose that the principal commits to observing* h_1 *before setting the wage in period 2. Then a linear rational-expectations equilibrium exists. Moreover, in any linear equilibrium, the agent works harder in period 1 than he would if the principal did not observe* h_1. *His effort in period 2 remains unchanged.*
2. *Suppose that the principal commits to observing both* h_1 *and* h_2 *before setting wages in the ensuing periods. Then, a linear rational-expectations equilibrium exists. Moreover, in any linear equilibrium, the agent's effort in period 2 and his total effort in periods 1 and 2 are smaller than they would be if the principal only observed* h_1.

Part 1 of Theorem 3 shows that observing only h_1 increases incentives. As we have argued in the previous section, the forward attribution outweighs the backward attribution in the effect of period-1 output on the period-2 wage, so that the agent benefits from increasing the principal's beliefs about his drive. Since h_1 provides an additional opportunity for the agent to increase those beliefs through increasing effort, his knowledge that the principal observes h_1 motivates him to work harder.

But in contrast to our result for h_1, part 2 of Theorem 3 says that observing h_2 in addition to h_1 *decreases* effort. An increase in h_2 increases the agent's perceived drive and hence decreases his perceived talent. Because only talent matters in period 3, this implies that observing h_2 reduces the agent's incentive to work hard in period 2. Moreover, the fact that the agent is less motivated in period 2 decreases the value of drive, and hence decreases drive-signaling incentives in period 1. These effects guarantee that total effort also falls.

Combining these two results, our framework suggests that using informal measures of effort not (directly) related to output can be effective in motivating employees to work hard, and firms will use such measures mostly early in an employee's career. This result seems consistent with the heavy emphasis on long work hours for young employees in investment-banking, consulting, law, and other professional firms.

5. Discussion

In this section, we comment on how reasonable modifications of our model would affect our results.

Drive as length of career. As mentioned in Section 2.1, our formulation of drive as marginal utility of income delivers results similar but not identical to an alternative, "length-of-career" model in which a driven agent stays in his career longer than a less driven one. To illustrate the differences, consider a modification of our three-period model in which there is no heterogeneity in marginal utility of income, but some agents work for the full three periods, and some only for the first two. The considerations in period-2 wage setting are then similar to those in Section 3. But because the principal learns the agent's drive when he shows up for work in period 3, there is no backward attribution in period-3 wage setting. This implies that so long as $m > 0$, heterogeneity in length of career increases average effort at the beginning of the career without decreasing average effort later.

Drive differs throughout the agent's career. An important simplifying assumption in our model is that the marginal utility of income differs across agents only in the last period. Suppose instead that it differs across agents throughout their careers, so that the agent maximizes $\sum_t [mw_t - c(e_t)]$. As mentioned in our analysis of the two-period model, because an agent with higher drive then benefits more from a bonus in the first period, effort is increasing in drive in both periods, so there is no equilibrium in which drive-signaling incentives are zero.

The assumption that marginal utility of income differs in all periods also qualifies the results of Section 3, where we found that the difference in effort levels between driven and less driven agents is the same in periods 1 and 2 ($\alpha_1 = \alpha_2$). As we show in Theorem B.1 of Appendix B, if m differs in all periods this is no longer true: the difference in effort levels is greater in period 1 than in period 2 ($\alpha_1 > \alpha_2$), because at the earlier time there are two future periods about which more driven agents care more. The fact that $\alpha_1 > \alpha_2$ weakens the forward attribution in period 1 and introduces other subtle effects that can shift effort toward period 2. These effects can sometimes reverse the conclusion that in comparison to a standard career concerns model, heterogeneity in responsiveness to incentives increases effort at the beginning of the career and decreases effort later. As we argue in Appendix B, however, these effects are significant only if α_1 is much higher than α_2, so that they are an artifact of the agent's very short horizon. We also discuss reasons why we conjecture that in a long-horizon model, the effects we have identified in Section 3 are likely to survive.

Drive-signaling over the long term. The present paper analyzes exclusively short-horizon models of two or three periods. In an earlier version (Kőszegi and Li 2004), we have analyzed an infinite-horizon variant of our model in which marginal utility of income evolves over time, and found that this alternative differs from the current models in a major way: Whereas in the short-horizon models drive-signaling incentives vanish as other incentives vanish, with a long horizon drive-signaling incentives can be significant despite arbitrarily low existing incentives. This result derives from a self-reinforcing feature of drive-signaling incentives. Agents with higher drive respond more strongly than others to any existing incentive. Because perceived drive will therefore be rewarded, drive-signaling incentives arise, to which more driven agents also respond more strongly. This creates further drive-signaling incentives, and as a result of this feedback the drive-signaling incentive "bootstraps" itself. Kőszegi and Li provide a necessary and sufficient condition for when an arbitrarily small amount of existing incentive leads to a non-negligible level of effort in steady state.

Catering to the best in organizations. One of the most basic insights of this paper is that if existing incentives motivate driven agents more than others, drive becomes valuable and drive-signaling incentives arise as a result. Although we have not worked out a model in which firms design incentive systems that span agents' careers, one implication of this insight seems to be that—to generate drive-signaling incentives—firms should prefer systems that motivate driven agents disproportionately strongly. Real-life incentive systems that cater to the best-performing employees, such as fast-tracking and up-or-out promotion schemes, might have exactly such a property. Under fast-tracking, employees who are successful early are more carefully mentored and monitored, and are more likely to be promoted again. Under the up-or-out system, the firm either promotes or fires an employee after a certain time. In both cases, the convex (implicit) compensation schemes motivate agents with higher drive more strongly.

6. Conclusion

It is now a classic insight in economic theory that an employee's concern for perceptions of his talent can create incentives to work hard even without explicit incentives. Although talent is very important, we argue in this paper that economic theory has ignored a crucial determinant of employees' productivity: their responsiveness to incentives. Heterogeneity in agents' responsiveness to incentives creates a set of powerful incentives qualitatively different from those derived in the classic framework, and opens up many related questions.

One such question is the effect of drive-signaling incentives in other career concerns settings. We study effort delivery in this paper, but a sizable literature studies the career concerns of various types of experts providing information (e.g., Scharfstein and Stein 1990; Prendergast and Stole 1996; Ottaviani and Sorensen 2006; Li 2007). Because experts often distort their advice to increase their wage, a principal may prefer less driven experts to more driven ones. Hence, in contrast to a model with effort delivery, experts may have an incentive to signal from the beginning that they are not very driven. In general, how heterogeneity in drive affects a career concerns model is likely to depend on the original career concerns' (in)efficiency. If the original career concerns lead to increased output—as in this paper—agents want to show that they are driven. But if the original career concerns lead to decreased output—as with experts—agents may want to show that they are not driven. This distinction is potentially important in understanding why drive is such an important character trait in some industries (such as investment banking) but is considered less important or even detrimental in others (such as expert advisors, judges, and politicians).

Although our paper touched on the implications of drive-signaling incentives for organizational design, it is far from providing a full analysis of this issue. One important question is how explicit incentives interact with and influence implicit incentives. Consider, for instance, an incentive system (such as academics) where "senior" agents have little or no explicit incentives to motivate them. Would the principal want to take advantage of explicit incentives earlier on? Our framework suggests maybe not: Explicit incentives disproportionately motivate agents who are not necessarily talented and do not necessarily like to work so much, but are responsive to incentives. This makes it more difficult to identify the talented and the "inherently" hard-working, and consequently decreases implicit incentives.

Relatedly, if potential employees are heterogeneous in drive, firms may want to design their compensation packages not only to motivate their existing workforce most effectively, but also to attract the workers with the most valuable qualities in the first place. For instance, in professions where a pure pursuit of money may lead to perverse outcomes, firms may want to put greater weight on non-pecuniary compensation, team incentives, or other lower-powered incentives.

Appendix A: Proofs

A.1. Proof of Theorem 1

Recall from the text that the agent's effort in the first period is determined by the following:

$$\underline{e}_1 = \beta_1,$$

$$\alpha_1 = \beta_2(1-\beta_2)\frac{\alpha_1\sigma_m^2}{\alpha_1^2\sigma_m^2 + \sigma_\varepsilon^2}.$$

For any parameter values, $\alpha_1 = 0$ is a solution to this equation. If $\alpha_1 \neq 0$, the equation implies

$$\alpha_1^2 = \beta_2(1-\beta_2) - \sigma_\varepsilon^2/\sigma_m^2,$$

which has a positive solution if $\beta_2(1-\beta_2) > \sigma_\varepsilon^2/\sigma_m^2$. Clearly, the highest α_1 is attained for $\beta_2 = 1/2$.

A.2. Proof of Lemma 1

See equation (A.7) and the arguments afterwards in the following proof.

A.3. Proof of Theorem 2

Together with a and m, the distribution of the observables $q_1 = a + \underline{e}_1 + \alpha_1 m + \varepsilon_1$ and $q_2 = a + \underline{e}_2 + \alpha_2 m + \varepsilon_2$ is multivariate normal. In particular,

$$E\left[\begin{pmatrix} a \\ m \\ a + \alpha_1 m + \varepsilon_1 \\ a + \alpha_2 m + \varepsilon_2 \end{pmatrix}\right] = \begin{pmatrix} 0 \\ \mu_m \\ \underline{e}_1 + \alpha_1 \mu_m \\ \underline{e}_2 + \alpha_2 \mu_m \end{pmatrix} \tag{A.1}$$

and

$$\begin{aligned} V&\left[\begin{pmatrix} a \\ m \\ a + \alpha_1 m + \varepsilon_1 \\ a + \alpha_2 m + \varepsilon_2 \end{pmatrix}\right] \\ &= \begin{pmatrix} \sigma_a^2 & 0 & \sigma_a^2 & \sigma_a^2 \\ 0 & \sigma_m^2 & \alpha_1 \sigma_m^2 & \alpha_2 \sigma_m^2 \\ \sigma_a^2 & \alpha_1 \sigma_m^2 & \sigma_a^2 + \alpha_1^2 \sigma_m^2 + \sigma_\varepsilon^2 & \sigma_a^2 + \alpha_1 \alpha_2 \sigma_m^2 \\ \sigma_a^2 & \alpha_2 \sigma_m^2 & \sigma_a^2 + \alpha_1 \alpha_2 \sigma_m^2 & \sigma_a^2 + \alpha_2^2 \sigma_m^2 + \sigma_\varepsilon^2 \end{pmatrix}. \end{aligned} \tag{A.2}$$

Now we use the updating rule for multivariate normals to obtain[16]

$$\begin{aligned} &E[a \mid q_1, q_2] = \\ &\begin{pmatrix} \sigma_a^2 & \sigma_a^2 \end{pmatrix} \begin{pmatrix} \sigma_a^2 + \alpha_1^2 \sigma_m^2 + \sigma_\varepsilon^2 & \sigma_a^2 + \alpha_1 \alpha_2 \sigma_m^2 \\ \sigma_a^2 + \alpha_1 \alpha_2 \sigma_m^2 & \sigma_a^2 + \alpha_2^2 \sigma_m^2 + \sigma_\varepsilon^2 \end{pmatrix}^{-1} \begin{pmatrix} q_1 - \underline{e}_1 - \alpha_1 \mu_m \\ q_2 - \underline{e}_1 - \alpha_2 \mu_m \end{pmatrix}. \end{aligned} \tag{A.3}$$

Therefore, the agent's period-3 and period-2 wages are, respectively,

$$\begin{aligned} w_3 = {} & \frac{\sigma_a^2 \big(\alpha_2(\alpha_2 - \alpha_1)\sigma_m^2 + \sigma_\varepsilon^2\big)(q_1 - \underline{e}_1 - \alpha_1 \mu_m)}{\big(\sigma_a^2 + \alpha_1^2 \sigma_m^2 + \sigma_\varepsilon^2\big)\big(\sigma_a^2 + \alpha_2^2 \sigma_m^2 + \sigma_\varepsilon^2\big) - \big(\sigma_a^2 + \alpha_1 \alpha_2 \sigma_m^2\big)^2} \\ & + \frac{\sigma_a^2 \big(\alpha_1(\alpha_1 - \alpha_2)\sigma_m^2 + \sigma_\varepsilon^2\big)(q_2 - \underline{e}_2 - \alpha_2 \mu_m)}{\big(\sigma_a^2 + \alpha_1^2 \sigma_m^2 + \sigma_\varepsilon^2\big)\big(\sigma_a^2 + \alpha_2^2 \sigma_m^2 + \sigma_\varepsilon^2\big) - \big(\sigma_a^2 + \alpha_1 \alpha_2 \sigma_m^2\big)^2}, \end{aligned} \tag{A.4}$$

$$w_2 = \frac{\sigma_a^2 + \alpha_1 \alpha_2 \sigma_m^2}{\sigma_a^2 + \alpha_1^2 \sigma_m^2 + \sigma_\varepsilon^2}(q_1 - \underline{e}_1 - \alpha_1 \mu_m) + \underline{e}_2 + \alpha_2 \mu_m. \tag{A.5}$$

16. If

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

then

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

Observe from equation (A.4) that

$$\left(\sigma_a^2+\alpha_1^2\sigma_m^2+\sigma_\varepsilon^2\right)\left(\sigma_a^2+\alpha_2^2\sigma_m^2+\sigma_\varepsilon^2\right)-\left(\sigma_a^2+\alpha_1\alpha_2\sigma_m^2\right)^2 > \sigma_m^2\sigma_a^2(\alpha_1-\alpha_2)^2 \geq 0.$$

Thus

$$\begin{aligned}\operatorname{sign}\left(\frac{\partial w_3}{\partial q_1}\right) &= \operatorname{sign}\left(\alpha_2(\alpha_2-\alpha_1)\sigma_m^2+\sigma_\varepsilon^2\right),\\ \operatorname{sign}\left(\frac{\partial w_3}{\partial q_2}\right) &= \operatorname{sign}\left(\alpha_1(\alpha_1-\alpha_2)\sigma_m^2+\sigma_\varepsilon^2\right).\end{aligned} \tag{A.6}$$

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, the agent's equilibrium effort supply is

$$e_1 = \underline{e}_1 + m\alpha_1 = \frac{\partial w_2}{\partial q_1} + m\frac{\partial w_3}{\partial q_1} \quad \text{and} \quad e_2 = m\alpha_2 = m\frac{\partial w_3}{\partial q_2}. \tag{A.7}$$

Thus, $\alpha_1 = \partial w_3/\partial q_1$ and $\alpha_2 = \partial w_3/\partial q_2$. Because both right-hand sides are positive, it is not possible to have negative α_1 or α_2, a contradiction. Suppose 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, which contradicts $\alpha_2 < 0$. Similarly, if $\alpha_1 > \alpha_2$, then w_3 increases with q_1, which contradicts $\alpha_1 < 0$.

Next, we prove that $\alpha_1 = \alpha_2$. Given that both are positive, if we have $\alpha_1 > \alpha_2$, then $\alpha_2 = w_3/q_2 > w_3/q_1 = \alpha_1$, 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 (4). Finally, for α positive, the left-hand side of equation (4) is increasing in α, while the right-hand side is decreasing. Because the right-hand side is greater at zero but smaller for large α, a unique α satisfies the equation.

A.4. Proof of Theorem 3

Part 1. When only h_1 is observed, w_3 depends on exactly the same observables as in Section 3. Therefore, a linear rational expectations equilibrium exists. In equilibrium, $\alpha_1 = \alpha_2 = \alpha^* > 0$, and α^* satisfies equation (4). In particular, the level of effort in period 2 for an agent with the same drive remains the same whether hours are observed or not. However, at the end of period 1, the principal observes q_1 and h_1 before deciding on the wage w_2, which changes his incentive at $t = 1$. Thus the period-2 return to agent's period 1 effort, $\partial E[w_2 \mid e_1]/\partial e_1$, will determine whether the agent provides more or less effort than when h_1 is not observed.

In period 1, the principal's observes q_1 and h_1, and makes inferences about a and m. The variance-covariance matrix of interest is thus

$$V\left[\begin{pmatrix} a \\ m \\ q_1 \\ h_1 \end{pmatrix}\right] = \begin{pmatrix} \sigma_a^2 & 0 & \sigma_a^2 & 0 \\ 0 & \sigma_m^2 & \alpha^*\sigma_m^2 & \alpha^*\sigma_m^2 \\ \sigma_a^2 & \alpha^*\sigma_m^2 & \sigma_a^2 + \alpha^{*2}\sigma_m^2 + \sigma_\varepsilon^2 & \alpha^{*2}\sigma_m^2 \\ 0 & \alpha^*\sigma_m^2 & \alpha^{*2}\sigma_m^2 & \alpha^{*2}\sigma_m^2 + \sigma_\varepsilon'^2 \end{pmatrix}. \tag{A.8}$$

This leads to the following expectations for talent and drive, up to a constant:

$$E[a \mid q_1, h_1] = \frac{\sigma_a^2\left(\alpha^{*2}\sigma_m^2 + \sigma_\varepsilon'^2\right)(q_1 - \alpha^*\mu_m - \underline{e}_1) - \alpha^{*2}\sigma_a^2\sigma_m^2(h_1 - \alpha^*\mu_m - \underline{e}_1)}{\alpha^{*2}\sigma_m^2\sigma_a^2 + \alpha^{*2}\sigma_m^2\sigma_\varepsilon^2 + \sigma_a^2\sigma_\varepsilon'^2 + \alpha^{*2}\sigma_m^2\sigma_\varepsilon'^2 + \sigma_\varepsilon^2\sigma_\varepsilon'^2},$$

$$E[m \mid q_1, h_1] = \frac{\alpha^*\sigma_m^2\sigma_\varepsilon'^2(q_1 - \alpha^*\mu_m - \underline{e}_1) + \alpha^*\sigma_m^2\left(\sigma_a^2 + \sigma_\varepsilon^2\right)(h_1 - \alpha^*\mu_m - \underline{e}_1)}{\alpha^{*2}\sigma_m^2\sigma_a^2 + \alpha^{*2}\sigma_m^2\sigma_\varepsilon^2 + \sigma_a^2\sigma_\varepsilon'^2 + \alpha^{*2}\sigma_m^2\sigma_\varepsilon'^2 + \sigma_\varepsilon^2\sigma_\varepsilon'^2}.$$

Clearly, the principal's posterior estimate of the agent's talent $E[a \mid q_1, h_1]$ decreases in h_1 and her estimate of his drive $E[m|q_1, h_1]$ increases in h_1. Because the agent's wage in period 2 is $E[a + \alpha^* m \mid q_1, h_1]$, and use the fact that $\partial E[q_1 \mid e_1]/\partial e_1 = \partial E[h_1 \mid e_1]/\partial e_1 = 1$, we obtain

$$\underline{e}_1' \equiv \frac{\partial}{\partial e_1} E[w_2 \mid e_1] = \frac{\sigma_a^2\sigma_\varepsilon'^2 + \alpha^{*2}\sigma_m^2\left(\sigma_a^2 + \sigma_\varepsilon'^2 + \sigma_\varepsilon^2\right)}{\sigma_a^2\sigma_\varepsilon'^2 + \alpha^{*2}\sigma_m^2\left(\sigma_a^2 + \sigma_\varepsilon^2 + \sigma_\varepsilon'^2\right) + \sigma_\varepsilon^2\sigma_\varepsilon'^2}. \tag{A.9}$$

Compare this with the model when hours are not observed; from equation (A.5) in the proof of Theorem 2, we have

$$\underline{e}_1 = \frac{\sigma_a^2 + \alpha^{*2}\sigma_m^2}{\sigma_a^2 + \alpha^{*2}\sigma_m^2 + \sigma_\varepsilon^2} = \frac{1}{1 + \sigma_\varepsilon^2\left(\sigma_a^2 + \alpha^{*2}\sigma_m^2\right)^{-1}}. \tag{A.10}$$

Simple calculations can show that $\underline{e}_1' > \underline{e}_1$, thus the agent works harder in the first period when only h_1 is observed.

Part 2. Note that even though the principal observes both h_1 and h_2, w_3 does not depend on h_1, which is lost by assumption. Thus at the beginning of period 3, the relevant variance-covariance matrix is

$$V\left[\begin{pmatrix} a \\ a+\alpha_1 m+\varepsilon_1 \\ \alpha_2 m+\varepsilon_2' \\ a+\alpha_2 m+\varepsilon_2 \end{pmatrix}\right] = \begin{pmatrix} \sigma_a^2 & \sigma_a^2 & 0 & \sigma_a^2 \\ \sigma_a^2 & \sigma_a^2+\alpha_1^2\sigma_m^2+\sigma_\varepsilon^2 & \alpha_1\alpha_2\sigma_m^2 & \sigma_a^2+\alpha_1\alpha_2\sigma_m^2 \\ 0 & \alpha_1\alpha_2\sigma_m^2 & \alpha_2^2\sigma_m^2+\sigma_\varepsilon'^2 & \alpha_2^2\sigma_m^2 \\ \sigma_a^2 & \sigma_a^2+\alpha_1\alpha_2\sigma_m^2 & \alpha_2^2\sigma_m^2 & \sigma_a^2+\alpha_2^2\sigma_m^2+\sigma_\varepsilon^2 \end{pmatrix}.$$

Simple algebra using this matrix leads to $E[a \mid q_1, h_2, q_2]$, which gives the following expressions:

$$\alpha_1 = \frac{\partial}{\partial e_1}E[w_3 \mid e_1, e_2] = \frac{\alpha_2(\alpha_2-\alpha_1)\sigma_a^2\sigma_m^2\sigma_\varepsilon'^2 + \alpha_2^2\sigma_a^2\sigma_m^2\sigma_\varepsilon^2 + \sigma_a^2\sigma_\varepsilon^2\sigma_\varepsilon'^2}{\sigma_a^2\sigma_m^2\sigma_\varepsilon'^2(\alpha_2-\alpha_1)^2 + \left(\sigma_\varepsilon'^2+\alpha_2^2\sigma_m^2\right)\left(2\sigma_a^2\sigma_\varepsilon^2+\left(\sigma_\varepsilon^2\right)^2\right) + \left(\alpha_1^2+\alpha_2^2\right)\sigma_m^2\sigma_\varepsilon^2\sigma_\varepsilon'^2}, \tag{A.11}$$

$$\alpha_2 = \frac{\partial}{\partial e_2}E[w_3 \mid e_1, e_2] = \frac{\alpha_1(\alpha_1-\alpha_2)\sigma_a^2\sigma_m^2\sigma_\varepsilon'^2 - \alpha_1\alpha_2\sigma_a^2\sigma_m^2\sigma_\varepsilon^2 + \sigma_a^2\sigma_\varepsilon^2\sigma_\varepsilon'^2}{\sigma_a^2\sigma_m^2\sigma_\varepsilon'^2(\alpha_2-\alpha_1)^2 + \left(\sigma_\varepsilon'^2+\alpha_2^2\sigma_m^2\right)\left(2\sigma_a^2\sigma_\varepsilon^2+\left(\sigma_\varepsilon^2\right)^2\right) + \left(\alpha_1^2+\alpha_2^2\right)\sigma_m^2\sigma_\varepsilon^2\sigma_\varepsilon'^2}. \tag{A.12}$$

First, a linear equilibrium exists. For a sufficiently large positive constant K (chosen according to the criteria described subsequently), consider the set $S = \{(\alpha_1, \alpha_2) \mid 0 \le \alpha_1, \alpha_2 \le K, \alpha_1 \ge \alpha_2\}$ in $\Re^2$-space. We define the map $f: S \to \mathbb{R}^2$ in the following way: $f_1(\alpha_1, \alpha_2)$, the first component of $f(\alpha_1, \alpha_2)$, is equal to equation (A.11), and similarly, $f_2(\alpha_1, \alpha_2)$ is equal to equation (A.12). It is easy to verify the following properties of f:

1. whenever $\alpha_1 = \alpha_2$, $f_1(\alpha_1, \alpha_2) > f_2(\alpha_1, \alpha_2) > 0$;
2. whenever $\alpha_2 = 0$, $f_1(\alpha_1, \alpha_2)$, $f_2(\alpha_1, \alpha_2) > 0$;
3. we can choose K so that $f_1(K, \alpha_2) < K$ and $f_2(K, \alpha_2) > 0$ for any $\alpha_2 < K$;
4. f is continuous.

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

Second, the magnitude of α_1, α_2 determines the agent's effort supply when h_1, h_2 are both observed. First we calculate the difference and sum of α_1 and α_2 from equations (A.11) and (A.12):

$$\alpha_1 - \alpha_2 = \frac{-(\alpha_1+\alpha_2)\left(\alpha_1-\alpha_2-\alpha_2\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon'^2}\right)\sigma_a^2\sigma_m^2\sigma_\varepsilon'^2}{\sigma_a^2\sigma_m^2\sigma_\varepsilon'^2(\alpha_2-\alpha_1)^2+\left(\sigma_\varepsilon'^2+\alpha_2^2\sigma_m^2\right)\left(2\sigma_a^2\sigma_\varepsilon^2+\left(\sigma_\varepsilon^2\right)^2\right)+\left(\alpha_1^2+\alpha_2^2\right)\sigma_m^2\sigma_\varepsilon^2\sigma_\varepsilon'^2}, \tag{A.13}$$

$$\alpha_1 + \alpha_2 = \frac{(\alpha_1-\alpha_2)\left(\alpha_1-\alpha_2-\alpha_2\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon'^2}\right)\sigma_a^2\sigma_m^2\sigma_\varepsilon'^2+2\sigma_a^2\sigma_\varepsilon^2\sigma_\varepsilon'^2}{\sigma_a^2\sigma_m^2\sigma_\varepsilon'^2(\alpha_2-\alpha_1)^2+\left(\sigma_\varepsilon'^2+\alpha_2^2\sigma_m^2\right)\left(2\sigma_a^2\sigma_\varepsilon^2+\left(\sigma_\varepsilon^2\right)^2\right)+\left(\alpha_1^2+\alpha_2^2\right)\sigma_m^2\sigma_\varepsilon^2\sigma_\varepsilon'^2}. \tag{A.14}$$

Next, we use the following steps to characterize the basic properties of α_1, α_2 before comparing the agent's effort level.

Step 1. $\alpha_1 > 0, \alpha_2 > 0$. Suppose not. If both α_1 and α_2 were negative, then equation (A.14) shows that $\alpha_1-\alpha_2$ and $\alpha_1-\alpha_2-\alpha_2(\sigma_\varepsilon^2/\sigma_\varepsilon'^2)$ have opposite signs, which contradicts equation (A.13). If $\alpha_1 > 0$ and $\alpha_2 < 0$, then from equations (A.11) and (A.12), $\partial w_3/\partial e_2$ would be positive, a contradiction. Similarly, $\alpha_1 < 0$ and $\alpha_2 > 0$ is not possible.

Moreover, from equation (A.13), $\alpha_1 - \alpha_2$ and $\alpha_1 - \alpha_2 - \alpha_2(\sigma_\varepsilon^2/\sigma_\varepsilon'^2)$ have opposite signs. Because α_1 and α_2 are positive, this can only happen if $\alpha_1-\alpha_2 > 0$ and $\alpha_1 - \alpha_2 - \alpha_2(\sigma_\varepsilon^2/\sigma_\varepsilon'^2) < 0$. Thus the numerator on the right-hand side of (A.14) is less than $2\sigma_a^2\sigma_\varepsilon^2\sigma_\varepsilon'^2$.

Step 2. $\alpha_1 + \alpha_2 < 2\alpha^*$. Suppose by contradiction that $\alpha_1 + \alpha_2 \geq 2\alpha^*$, where $2\alpha^*$ is the total marginal effort when only h_1 is observed. Then $4\alpha^{*2} \leq (\alpha_1+\alpha_2)^2 \leq 2\alpha_1^2 + 2\alpha_2^2$, or $2\alpha^{*2} \leq \alpha_1^2 + \alpha_2^2$. From equation (A.4), $2\alpha^* = 2\sigma_a^2/(2\sigma_a^2 + 2\alpha^{*2}\sigma_m^2 + \sigma_\varepsilon^2)$. Then, the denominator on the right-hand side of equation (A.14) is strictly greater than $\sigma_\varepsilon'^2\sigma_\varepsilon^2(\sigma_\varepsilon^2 + 2(\sigma_a^2 + \alpha^{*2}\sigma_m^2))$ and the numerator is smaller than $2\sigma_a^2\sigma_\varepsilon'^2\sigma_\varepsilon^2$, a contradiction. Thus $\alpha_1 + \alpha_2 < 2\alpha^*$. This also means that the total responsiveness of w_3 to previous output is smaller than in the basic model of Section 2.

Step 3. Finally, note that the return to period 1 effort is

$$\underline{e}_1'' \equiv \frac{\partial w_2}{\partial q_1} = \frac{\sigma_a^2\sigma_\varepsilon'^2 + \alpha_1\alpha_2\sigma_m^2\left(\sigma_a^2+\sigma_\varepsilon'^2+\sigma_\varepsilon^2\right)}{\sigma_a^2\sigma_\varepsilon'^2+\alpha_1^2\sigma_m^2\left(\sigma_a^2+\sigma_\varepsilon^2+\sigma_\varepsilon'^2\right)+\sigma_\varepsilon^2\sigma_\varepsilon'^2}.$$

Compare with equation (A.9), which measures the return to period-1 effort when h_2 is not observed. We have

$$\underline{e}_1'' - \underline{e}_1' = \frac{(\alpha_1\alpha_2 - \alpha^{*2})\sigma_m^2\big(\sigma_a^2 + \sigma_\varepsilon^2 + \sigma_\varepsilon'^2\big)\big(\alpha^{*2}\sigma_m^2\big(\sigma_a^2 + \sigma_\varepsilon^2 + \sigma_\varepsilon'^2\big) + \sigma_\varepsilon^2\sigma_\varepsilon'^2\big)}{\big(\sigma_a^2\sigma_\varepsilon'^2 + \alpha_1^2\sigma_m^2\big(\sigma_a^2 + \sigma_\varepsilon^2 + \sigma_\varepsilon'^2\big) + \sigma_\varepsilon^2\sigma_\varepsilon'^2\big)\big(\sigma_a^2\sigma_\varepsilon'^2 + \alpha^{*2}\sigma_m^2\big(\sigma_a^2 + \sigma_\varepsilon^2 + \sigma_\varepsilon'^2\big) + \sigma_\varepsilon^2\sigma_\varepsilon'^2\big)},$$

which is negative because $\alpha_1\alpha_2 \leq (\alpha_1 + \alpha_2)^2/4 < \alpha^{*2}$. Therefore the total average effort of an agent in both periods is smaller when h_1, h_2 are observed than when only h_1 is observed.

Appendix B: When *m* Differs throughout the Agent's Career

In this appendix, we consider a variation of the model in Section 3 in which m is the agent's privately known marginal utility of income in *all* periods, not just in period 3. All other assumptions of the model remain the same, and we still look for linear rational-expectations equilibria.

When m differs across the agent's career, in equilibrium $e_t = \alpha_t m$. Moreover, calculations very similar to those in the proof of Theorem 2 lead to the following equilibrium conditions for α_1 and α_2:

$$\alpha_1 = \underbrace{\frac{\sigma_a^2 + \alpha_1\alpha_2\sigma_m^2}{\sigma_a^2 + \alpha_1^2\sigma_m^2 + \sigma_\varepsilon^2}}_{=\partial w_2/\partial q_1} + \underbrace{\frac{\sigma_a^2\big(\alpha_2(\alpha_2 - \alpha_1)\sigma_m^2 + \sigma_\varepsilon^2\big)}{\sigma_\varepsilon^2\big(2\sigma_a^2 + \sigma_\varepsilon^2\big) + \sigma_m^2\big(\big(\alpha_1^2 + \alpha_2^2\big)\sigma_\varepsilon^2 + (\alpha_1 - \alpha_2)^2\sigma_a^2\big)}}_{=\partial w_3/\partial q_1},$$

$$\alpha_2 = \underbrace{\frac{\sigma_a^2\big(\alpha_1(\alpha_1 - \alpha_2)\sigma_m^2 + \sigma_\varepsilon^2\big)}{\sigma_\varepsilon^2\big(2\sigma_a^2 + \sigma_\varepsilon^2\big) + \sigma_m^2\big(\big(\alpha_1^2 + \alpha_2^2\big)\sigma_\varepsilon^2 + (\alpha_1 - \alpha_2)^2\sigma_a^2\big)}}_{=\partial w_3/\partial q_2}. \tag{B.1}$$

The major difference relative to our basic model is the first term in the expression for α_1, which is the derivative $\partial w_2/\partial q_1$. Unlike in the basic model, more driven agents respond more strongly to this incentive. The properties of equilibrium are summarized in the following theorem.

THEOREM B.1. *If marginal utility of income differs across agents in 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. Part 1. We prove this by contradiction. First, suppose that $\alpha_1, \alpha_2 < 0$. Adding the two equations in expressions (B.1), we obtain

$$\begin{aligned}\alpha_1+\alpha_2 = &\frac{\sigma_a^2+\alpha_1\alpha_2\sigma_m^2}{\sigma_a^2+\alpha_1^2\sigma_m^2+\sigma_\varepsilon^2}\\ &+\frac{\sigma_a^2\big((\alpha_1-\alpha_2)^2\sigma_m^2+\sigma_\varepsilon^2\big)}{\sigma_\varepsilon^2\big(2\sigma_a^2+\sigma_\varepsilon^2\big)+\sigma_m^2\big((\alpha_1^2+\alpha_2^2)\sigma_\varepsilon^2+(\alpha_1-\alpha_2)^2\sigma_a^2\big)}.\end{aligned} \tag{B.2}$$

Now the left-hand side of this equation is negative, and the right-hand side is positive, which is impossible. For claim 3, suppose that $\alpha_1 \geq 0$ and $\alpha_2 < 0$. From expressions (B.1), α_2 must be positive, another contradiction. Finally, $\alpha_1 = \alpha_2 = 0$ contradicts both conditions in expressions (B.1).

Part 2. Steps very similar to the proof of Theorem (3.2) can show that the equilibrium exists.

Part 3. Assume that $\alpha_1 \leq \alpha_2$. Then from expressions (B.1), we can see that if α_1, α_2 are positive, it must be that $\alpha_1 > \alpha_2$. □

The new result that $\alpha_1 > \alpha_2$ in the (positive) equilibrium of this model introduces caveats to our discussion in Section 3. Namely, it constitutes a force that acts against our claim that an increase in heterogeneity in drive increases effort at the beginning of the agent's career, but lowers effort later. Comparing $\partial w_2/\partial q_1$ in the two models (equation (B.1) versus $\underline{e}_1$ in Theorem 2), the difference is that this derivative now features $\alpha_1\alpha_2\sigma_m^2$ in the numerator instead of $\alpha_1^2\sigma_m^2$, tending to decrease incentives in period 1. Intuitively, a small α_2 means the forward attribution is weak: If the agent is expected to slack off in the next period, there is less of a point in signaling drive, because this will be rewarded less generously.

In addition, the asymmetry between α_1 and α_2 introduces some subtle effects through new terms in the numerators for $\partial w_3/\partial q_1$ and $\partial w_3/\partial q_2$. The new term in $\partial w_3/\partial q_1$ is negative, $\alpha_2(\alpha_2-\alpha_1)\sigma_m^2 < 0$, whereas the new term in $\partial w_3/\partial q_2$ is positive, $\alpha_1(\alpha_1-\alpha_2)\sigma_m^2 > 0$. To see why the new term in $\partial w_3/\partial q_2$ is positive, suppose for a moment that $\alpha_1 > 0$ and $\alpha_2 = 0$. Then, any increase in q_2 is attributed to ability, not drive. Given q_1, *this decreases the principal's belief about the agent's drive* (because 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 drive decreases, this leads to a further increase in perceived ability, as effort in period 2 is perceived to be smaller.[17] A mirror image of this effect decreases incentives in period 1. With a smaller

17. In reality the principal does all the updating together. The above discussion is merely a heuristic argument that helps to understand why the responsiveness of the wage to period-2 output can remain high.

α_2, more inference is made about drive 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.

If α_2 is considerably smaller than α_1, these effects can be so strong that the agent works less in the first period, and more in the second period, than he would in a standard career concerns model (with no heterogeneity in drive). We conjecture, however, that such a large decrease in the effort gap between driven and less driven agents from one period to the next can only occur when the agent's horizon is short, and in fact the effects we have identified in Section 3 would survive in a more realistic, long-horizon setting. To understand the intuitions that lead us to these conjectures, note that α_t decreases over time because as the number of periods on the horizon—and hence the number of periods driven agents care about more—decreases, the difference in behavior between agents of different drive decreases. In a three-period model, a decrease in the horizon from two periods to one is drastic, so α_2 can be much smaller than α_1. When the agent is relatively far from the end of his career, however, a one-period decrease in the horizon should have a relatively small effect on overall incentives, so α_t decreases slowly. As a result, drive-signaling incentives are likely to increase effort early in the career. But as α_t starts decreasing drastically toward the end of the career, the forward attribution weakens, decreasing effort.

References

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

Aron, Debra J. (1987) "Worker Reputation and Productivity Incentives." *Journal of Labor Economics*, 5, S87–S106.

Baker, George, Robert Gibbons, and Kevin Murphy (1994). "Subjective Performance Measures in Optimal Incentive Contracts." *Quarterly Journal of Economics*, 109, 1125–1156.

Becker, Gary S. (1991). *A Treatise on the Family*. Harvard University Press.

BLS (2004). "Employee Benefits Survey, U.S. Department of Labor." Data available at http://stats.bls.gov/ncs/ebs/home.htm.

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

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

Dewatripont, Mathias, Ian Jewitt, and Jean Tirole (1999b). "The Economics of Career Concerns, Part II: Application to Missions and Accountability of Government Agencies." *Review of Economic Studies*, 66, 199–217.

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

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

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

Green, Francis, Stephen Machin, and David Wilkenson (1998). “The Meaning and Determinants of Skills Shortages.” *Oxford Bulletin of Economics and Statistics*, 60, 165.
Holmström, Bengt (1999). “Managerial Incentive Problems: A Dynamic Perspective.” *Review of Economic Studies*, 66, 169–182.
Holmström, Bengt, and Paul Milgrom (1991). “Multi-Task Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design.” *Journal of Law, Economics and Organization*, 7, 24–52.
Kanter, Rosabeth Moss (1977). *Men and Women of the Corporation*. BasicBooks.
Korenman, Sanders, and David Neumark (1991). “Does Marriage Really Make Men More Productive?” *The Journal of Human Resources*, 26, 282–307.
Kőszegi, Botond, and Wei Li (2004). “Drive and Talent.” Working paper, University of California, Berkeley.
Kuhn, Peter (1993). “Demographic Groups and Personnel Policy.” *Labour Economics*, 1, 49–70.
Kuhn, Peter, and Fernando Lozano (2005). “The Expanding Workweek: Understanding Trends in Long Work Hours Among U.S. Men, 1979–2002.” Working paper, University of California, Santa Barbara.
Laffont, Jean-Jacques, and Jean Tirole (1988). “The Dynamics of Incentive Contracts.” *Econometrica*, 56, 1153–1175.
Landers, Renee M., James B. Rebitzer, and Lowell J. Taylor (1996). “Rat Race Redux: Adverse Selection in the Determination of Work Hours in Law Firms.” *American Economic Review*, 86, 329–348.
Lazear, Edward P. (1989). “Pay Equality and Industrial Politics.” *Journal of Political Economy*, 97, 561–580.
Li, Wei (2007). “Changing One’s Mind when the Facts Change: Incentives of Experts and the Design of Reporting Protocols.” *Review of Economic Studies*, 74, 1175–1194.
Loh, Eng Seng (1996). “Productivity Differences and the Marriage Wage Premium for White Males.” *Journal of Human Resources*, 31, 566–589.
Ottaviani, Marco, and Peter Sorensen (2006). “Professional Advice.” *Journal of Economic Theory*, 126, 120–142.
Prendergast, Canice, and Lars Stole (1996). “Impetuous Youngsters and Jaded Old-Timers: Acquiring a Reputation for Learning.” *Journal of Political Economy*, 104, 1105–1134.
Scharfstein, David S., and Jeremy C. Stein (1990). “Herd Behavior and Investment.” *American Economic Review*, 80, 465–479.
Wellington, Allison (1994). “Accounting for the Male/female Wage Gap among Whites: 1976 and 1985.” *American Sociological Review*, 59, 839–848.