ALLOCATING EFFORT AND TALENT IN PROFESSIONAL LABOR MARKETS

GADI BARLEVY
FEDERAL RESERVE BANK OF CHICAGO
AND
DEREK NEAL
UNIVERSITY OF CHICAGO & NBER

ABSTRACT. In many professional service firms, new associates work long hours while competing in up-or-out promotion contests. Our model explores why these firms require young professionals to take on heavy work loads while facing significant risks of dismissal. We argue that the productivity of skilled partners in professional service firms, e.g. law, consulting, investment banking, public accounting, etc, is quite large relative to the productivity of their peers who are competent and experienced but not well-suited to the partner role. Therefore, these firms adopt personnel policies that facilitate the identification of new partners. In our model, both heavy work loads and up-or-out rules serve this purpose. Market participants learn more about new workers who perform more tasks, and when firms replace experienced associates with new less productive workers, they gain the opportunity to identify talented professionals who will have long careers as partners. Both of these personnel practices are costly. However, when the gains from increasing the number of talented partners exceed these costs, firms employ both practices in tandem. We present evidence on life-cycle patterns of hours and earnings among lawyers that support our claim that both heavy work loads and up-or-out rules are screening mechanisms.

Preliminary: Do Not Cite or Circulate

We thank John Kennan, Kevin Lang, Kevin Murphy, Canice Prendergast, Chris Taber, and Fabrice Tourre for useful discussions. We thank seminar participants at the University of Chicago, UCL, Edinburgh University, UW-Madison, the 2014 North American Summer Meetings of the Econometric Society, and the NBER Organizational Economics working group for valuable feedback. We thank Nicole Bazer, Ted Froum, Hank Kelly and Stephanie Scharf for useful conversations about personnel practices in law firms. We thank Andrew Jordan, Caitlin McCarthy, Jorge L. García G.-Menéndez, and Max Samels for excellent research assistance.

Date: February, 2016.

Introduction

Many professional service firms employ two personnel practices that are uncommon in other labor markets. First, these firms assign heavy work loads to young professionals. Second, these firms often employ up-or-out promotion policies. These policies dictate that newly hired professionals expect to either progress to a position like equity partner within a relatively fixed number of years or leave the firm. They know that if the existing partners decide not to promote them to partner, they will likely not give them the option to remain in a non-partner role. We develop a model that explains why many professional service firms require their new associates to work long hours while competing in up-or-out promotion contests.

Our work fills a hole in the existing literature on professional labor markets. The literature on why young professionals work long hours does not overlap with the somewhat larger literature that seeks to understand why many of the same young professionals face up-or-out promotion policies. Rat race models and the literature on career concerns provide reasons that young professionals may work long hours, but these literatures do not directly address why many professional service firms adopt up-or-out promotion rules. The literature on up-or-out explains how this policy can be used to solve commitment problems or to facilitate the identification of talented professionals who will succeed as partners, but these models of commitment and screening often ignore worker effort and make no clear predictions about the effort levels of young professionals relative to those of other young workers.

While heavy work loads and up-or out promotion rules may serve many purposes in professional labor markets, our results suggest that they are not separate phenomena. Both heavy work loads and up-or-out rules serve a common purpose. These practices facilitate the identification of the talented professionals who will lead their organizations in the future. Market participants learn more about new workers who perform more tasks, and when firms replace experienced associates with new workers, they gain the opportunity to identify talented professionals who will have long careers as partners.

Both practices are costly. Work loads beyond statically optimal levels reduce the current surplus generated by new associates, and replacing competent, experienced associates with new associates lowers current output. However, in some professional labor markets, gains from increasing the number of talented professionals who occupy partner positions exceed both of these costs, and here, we expect to see both practices used in tandem. In recent decades, large law and public accounting firms have begun to create a small number of non-partner track positions for experienced professionals with special expertise. We argue that consolidations and mergers in these fields may have created new, productive roles for specialists and thus induced less strict adherence to up-or-out rules.

In the next section, we review the literature on personnel practices in professional labor markets more carefully. Then, we present our model of work loads and job assignment. We solve a planner's problem that illustrates how both heavy work loads for young professionals and up-or-out promotion rules facilitate the identification of talented partners. Many market mechanisms decentralize the planner's solution that we describe. We describe one mechanism and discuss its implications for how to interpret observed relationships between long hours and up-or-out rules.

In the penultimate section, we document several patterns in data on earnings and hours worked among lawyers that support our contention that professional

service firms require new associates to take on heavy work loads while participating in up-or-out promotion contests because both policies speed the discovery on new partners. Our conclusion reviews our contribution and discusses future research that may shed more light on the evolution of personnel policies in professional markets.

1. Literature Review

A significant literature documents the fact that new associates in elite professional service firms, e.g. leading firms in law, consulting, investment banking[1] and public accounting, often work much longer hours than most white collar workers who have similar levels of education. Landers et al. [1996] offer a possible explanation for this pattern. They analyze data on law firms but suggest that their insights may apply to other professional labor markets as well.

Law firms and other professional services firms are often organized as partnerships. Building on the work of Akerlof [1976], Landers et al. [1996] treat law firms as teams and assume that, while team output is observed, individual output is not. In their model, teams share output according to the following rules. Partners pay associates a fixed salary and share remaining profits equally. Associates work schedules that partners dictate, and at the end of their terms as associates, they bid for shares in the partnership. Retiring partners sell their shares to the next generation of partners. Partners and associates have heterogeneous effort costs and also possess private information about these costs.

Partners benefit from hiring associates with low effort costs because this allows them to sell their equity shares to more productive lawyers in the future. Thus, partners desire some screening mechanism that allows associates to reveal their type. In the separating equilibrium that Landers et al. [1996] describe, a menu of employment contracts specifies hours requirements and compensation for new associates in each law firm. These contracts also describe auction mechanisms that dictate how existing partners in various firms will sell their ownership stakes to their associates in the future. Lawyers who share the same effort costs select the same contracts and work together in the same firms.

Ex post, associates in all firms work more than the efficient number of hours.[2] As in the canonical rat race model, Akerlof [1976], hours distortions are the equilibrium mechanism that sorts heterogeneous workers to heterogeneous teams. This model of hours requirements for new associates in professional labor markets does not directly address retention or promotion decisions. In the separating equilibrium that Landers et al. [1996] describe, no one leaves law as a profession, no one changes law firms, and all associates become partners.

Holmstrom [1999] provides a different reason that young professionals may work long hours. In his model, output is not contractible, so firms pay workers ex ante based on their reputation. Workers benefit from having strong reputations but possess no private information about their ability levels. All market participants have the same prior beliefs about all workers and learn about all workers at the same

[1]Here, we refer to the mergers and acquisitions services that investments banks provide, not all functions of investment banks.

[2]Further, partners take on work loads that are below efficient levels since they share the returns of their efforts with other partners.

rate by observing public output signals. However, workers do have private information about their effort levels, and Holmstrom [1999] shows that young workers may work more than efficient levels to increase the output signals that firms use to form beliefs about their abilities. In equilibrium, firms infer the workers' equilibrium effort choices and adjust their inferences about worker ability accordingly. However, as in rat race models, no individual worker has an incentive to deviate from the inefficient equilibrium.[3] Further, the model in Holmstrom [1999] does not address promotion or retention decisions in any way.

Gicheva [2013] presents a model of variation in hours worked among young, well-educated workers. She notes that, if workers differ in their preferences for leisure, young workers who work the most are also most likely to enjoy promotions and high rates of wage growth in the future. Gicheva [2013] does not address why some firms make commitments to dismiss all workers they do not promote or why different promotion and retention rules arise in different professional labor markets.

While the literature on long hours among young professionals does not address up-or-out rules directly, the literature on up-or-out promotion rules has little to say about the long hours that young professionals work while participating in up-or-out promotion contests. The up-or-out literature contains several variations on two themes, but neither literature addresses why young professionals in up-or-out firms often work much more than other workers with similar levels of education.

One literature characterizes up-or-out rules as commitment devices that solve a double moral hazard problem between workers and firms. In these papers, firms have private information about either the output of a worker or a worker's ability. Workers have private information about their actions. Firms want to provide workers with incentives to take efficient actions, but workers know that, ex post, firms may have an incentive to renege on payments linked to performance measures that only the firm observes. Firms solve this double moral hazard problem by making verifiable commitments to up-or-out promotion rules. These rules force firms to dismiss all workers they do not promote. So, if a firm were to make an unfavorable report about a worker who produced a positive output signal, the firm would expect to incur a loss. Because up-or-out rules create an incentive for firms to make truthful reports, they allow firms to credibly promise to reward hidden actions, and this credibility allows firms to elicit more efficient actions from workers.[4]

This literature begins with Kahn and Huberman [1988] who argue that up-or-out allows firms to induce workers to make investments in firm-specific skills.[5] Prendergast [1993] argues that up-or-out rules are not always needed to solve the double moral hazard problem that Kahn and Huberman [1988] identify. If firms have

[3]In contrast to rat race models, equilibrium effort levels in career concerns models need not be excessive relative to efficient levels under full information. These models highlight one reason that work effort may decline over a worker's life cycle, but not all parameterizations yield the result that young workers begin their careers working too hard relative to efficient levels of effort.

[4]Levin and Tadelis [2005] suggest partnerships solve a commitment problem between professional service firms and their customers. Professional service firms promise to supply talented professional who perform quality work, but clients of professional service firms may find it difficult to judge the talent of different professionals ex ante. Revenue sharing within partnerships makes promises concerning the quality of professional services more credible because clients know that the other partners in the firm have agreed to share revenue with the partners who are directing their cases.

[5]See Gilson and Mnookin [1989] for an application of the Kahn and Huberman [1988] model to law firms.

positions for skilled workers that sufficiently leverage their skills, they incur costs when they fail to make deserved promotions, and these costs may make contingent promises concerning raises and promotions credible. Thus, firms can induce workers to invest in firm-specific skills without employing up-or-out rules as long as the firm benefits from promoting all workers who do invest.

Waldman [1990] extends the logic of Kahn and Huberman [1988] to an environment where firms have private information about general worker productivity as opposed to firm-specific capacities. He shows that private information about worker productivity creates the same moral hazard problems that Kahn and Huberman [1988] describe even when all human capital is completely general. When firms have private information about how productive their workers would be in other firms, they are tempted to deny promotions to deserving workers in order to maintain their information rents. In this scenario, firms may use up-or-out rules to commit to efficient promotion decisions, and this commitment induces young workers to invest more efficiently in general skills.

Ghosh and Waldman [2010] extend Waldman [1990] to an environment where new professionals take hidden actions that influence output signals and a portion of worker productivity is firm-specific. They conclude that up-or-out is more likely in professional labor markets where the promotion of workers to senior positions has relatively small effects on their productivity and most human capital is not firm-specific. In these settings, the ex post surplus generated by efficient promotions is relatively low. Therefore, firms demand a mechanism that allows them to credibly commit to contingent promises concerning raises and promotions.[6]

A second literature on up-or-out promotion rules links up-or-out rules to optimal screening procedures. OFlaherty and Siow [1992] develop a model of professional partnerships where partners work with one associate and receive signals about the suitability of the associate for promotion. They argue that partnerships grow by identifying people who are talented enough to be partners and show that the optimal screening rule in their environment involves two cutoffs. When the posterior belief about an associate crosses the upper cutoff, the candidate becomes a partner and takes on a new associate. When the posterior falls below the lower cutoff, the existing associate is dismissed and replaced by a new associate. Since beliefs about all associates eventually cross one of these thresholds, each associate either goes up or out.

Demougin and Siow [1994] link up-or-out rules to screening in a model of hierarchies. In this model, firms decide what portion of their new workers they will train to be potential managers. This training may be interpreted as on-the-job learning or as a screening process that determines the suitability of workers for the management position. When the outside wage for new workers is high enough, all firms in a given industry choose to train or screen all new workers and dismiss all who are not deemed worthy of promotion. Those without the talent required to work as managers leave the industry, and if a given firm identifies more managers

[6]Ghosh and Waldman [2010] also model worker effort. However, in contrast to our model, worker actions are hidden in their model. As in Holmstrom [1999], workers may expend effort to influence signals that determine their reputation, but this is true in firms that employ standard promotion practices as well as those that adopt up-or-out rules. Further, effort levels among new professionals may be high or low in both up-or-out firms and firms that follow standard promotion practices.

than it needs, these excess managers are hired away by firms that failed to identify enough managers.

OFlaherty and Siow [1992] and Demougin and Siow [1994] describe an up-or-out equilibrium where new professionals are no more productive than the experienced professionals they replace, but these new professionals have more option value than their more experienced counterparts. Below, we derive results that link up-or-out rules with this same option value logic.

Yet, our results differ in several ways. First, we model worker effort and introduce a signaling technology such that the market learns more about new professionals when these professionals perform more tasks. This allows us to explain why young professionals work long hours in the same sectors where up-or-out promotion rules are most common. Second, our comparative static results concerning when up-or-out regimes exist in professional labor markets do not deal with changes in outside options but rather changes in the relative productivities of experienced professionals of different abilities who occupy different roles within the professional sector. Changes in technology or organizational structure that raise the relative productivity of experienced professionals who are skilled but not partner material make up-or-out less attractive while changes that raise the relative productivity of partners make up-or-out rules more productive. Third, our model provides new insights concerning the interpretation of outcomes in up-or-out firms. Some scholars argue that up-or-out rules are puzzling because they require the dismissal of workers who may be doing a competent job in their current positions. Our model demonstrates that, in up-or-out settings, associates never want to continue in their associate positions once they realize that they are not going to become partners. The long hours that associates work are a cost that young professionals pay to learn whether or not they are well-suited to become partners. If these young professionals learn they are not going to become partners, they are no longer willing to pay this cost.

Finally, because we model professional work loads, our model provides new insights concerning life-cycle patterns of changes in hours worked among professionals who follow different career trajectories. We document that, among lawyers with roughly ten years of experience in private law firms, those who leave private law or leave the partnership track within private law reduce their hours significantly even though their wage rates often rise. This pattern is difficult to understand in most models of life-cycle labor supply. However, in our model, the heavy work loads that young associates tackle in private law firms are one component of an ordeal that young associates endure in order to audition for partnership positions. When an associate learns that she is not going to make partner, her audition is over, and she may well reduce her hours even though her skill set continues to command a high wage rate.

Our work seeks to connect the literatures on hours and up-or-out rules in professional service firms, but we are not offering a direct challenge to the existing theoretical literatures that address why up-or-out exists. We assume that all market participants in a given professional labor market learn symmetrically about all other market participants, but we do not address the costs of verifying information for courts. Thus, nothing in our work eliminates the possibility that up-or-out rules do help firms solve important commitment problems. In addition, our results support a key idea in both OFlaherty and Siow [1992] and Demougin and Siow

[1994], since we too conclude that up-or-out rules should be interpreted as optimal screening procedures.

2. Model Setup

Our model describes production, learning, and job assignment in professional labor markets. Individuals may work in the professional sector or in an outside sector. In the outside sector, there is one job, and output does not vary with worker ability. There are two jobs in the professional sector, associate and partner, and output does vary with ability. We use the term partner because professional service firms use this title for their leaders. However, we do not model the organizational structure of these firms. In our model, the term partner simply refers to a position where the most skilled professionals are most productive.[7]

We begin by describing preferences and production. We then describe our learning technology.

2.1. Preferences and Production.

Time is measured in discrete periods, and the time horizon is infinite. Each period, a unit mass of workers is born and lives two periods. Thus, in any period, a mass two of workers exists.

Workers are ex ante identical in this model. Thus, we suppress individual subscripts as we describe the preferences and production possibilities that characterize all workers.

Our model contains no information asymmetries or hidden actions. Therefore, we find it expedient to present our model as a planning problem. Below, we describe a planner who assigns workers to jobs and work loads. These work loads are the number of tasks that the planner assigns to each worker.

Workers are risk neutral with the following utility function

$$U = x - c(n)$$

where x is expected income. n is the number of tasks performed, and $c(n)$ is the disutility of performing n tasks. We assume $c(0) = 0$, $c'(0) = 0$. Further, $\lim_{n\to\bar{n}} c'(n) = \infty$, $c''(n) > 0 \quad \forall n \in [0, \bar{n}]$. All workers pay the same utility cost to complete any task.

Let θ denote worker ability, which is either high or low, i.e. $\theta \in \{0, x\}$, with $x > 0$. If a worker has high ability, the expected output generated by each task she completes is greater than the expected output generated by a low ability worker who performs the same task. At birth, the ability of workers is not known, but in each cohort, a constant fraction, π, is high ability, and the rest are low ability. All market participants know the distribution of ability, but no one has private information about their own ability or the ability of others.

There are two sectors in the economy. In the outside sector, output, y^o, is a deterministic, linear function of worker effort, and the mapping between effort and output does not vary with worker experience or ability. The production function in the outside sector is

[7] In public accounting, investment banking, consulting, and a growing number of law firms, the path to partner involves more than two job titles. Here, we use only two titles for convenience. A model with many potential up-or-out decisions points on the path to partner would be much more cumbersome, but the basic insights from our model would remain.

$$y^o = w^o n \tag{2.1}$$

Output in the professional sector is determined by worker ability, worker experience, and job assignment. Define y_s^j as the output of a worker assigned to professional job j given s periods of professional experience, where $j \in \{a, p\}$ for associate and partner, and $s \in \{0, 1\}$ for inexperienced and experienced. In contrast to the outside sector, output is stochastic in the professional sector. Nature draws i.i.d. production shocks, ϵ, that are mean zero for all professional workers in each period. The production function for new associates is

$$y_0^a = (1 + \theta)n + \epsilon \tag{2.2}$$

The production function for experienced associates is

$$y_1^a = z^a(1 + \theta)n + \epsilon \tag{2.3}$$

Here, the parameter $z^a > 1$ captures the idea that associates who have experience are able to perform more productive tasks.

Finally, the production function for partners is

$$y_1^p = \begin{cases} z^p(1 + \theta)n + \epsilon & \text{if } \theta = x \\ -\infty & \text{if } \theta = 0 \end{cases} \tag{2.4}$$

The parameter z^p, where $z^p > z^a > 1$, captures the idea that partners perform tasks that more fully leverage professional skill. We assume that skill levels are functions of both experience and talent. Further, we assume that, if low ability workers of any experience level were to act as partners, the mismatch between their skills and their task assignments would create losses. To facilitate our exposition, we set the value of these losses to $-\infty$. Likewise, we assume that $y_0^p = -\infty$ for all workers. This assumption captures the idea that, regardless of their ability, workers with no experience would also make costly mistakes if they were to act as partners.

Our planner must allocate workers between the professional labor market and all other employments. For now, we cap employment in the professional sector at $q < 1$ to capture the idea that only a fraction of highly-educated agents work in the professional sector. Later, we treat q as an endogenous variable that is determined by the costs of maintaining professional jobs and the productivities of positions in the professional sector.

We are interested in assignment decisions. These decisions involve interesting trade-offs if the following productivity relationships hold

$$z^a < w^o < z^a(1 + \pi x) \tag{2.5}$$

The first inequality in equation 2.5 implies that an experienced associate who has low ability is more productive in the outside sector than the professional sector. The second inequality implies that the expected productivity of an experienced associate with unknown ability is greater in the professional sector than the outside sector. If this were not true, the planner would never retain any experienced professionals who were not known to possess high ability. Given $z^a > 1$, both inequalities, taken

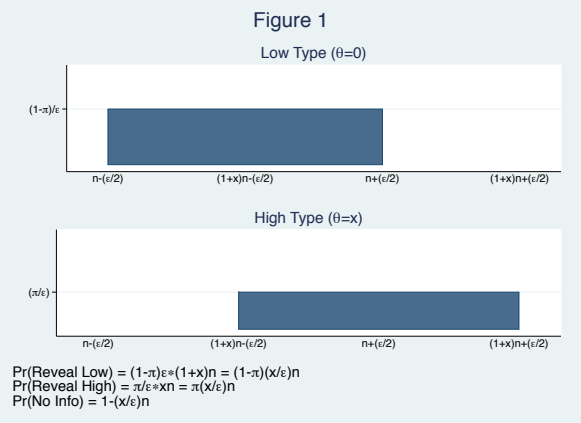
together, imply that the expected productivity of an experienced associate with uncertain ability exceeds the expected productivity of a new associate.

2.2. **Learning.** In our framework, the ability of each worker is drawn from a common distribution. During the first period, the market receives a signal about the ability of each new associate that either fully reveals her talent or provides no information about her talent. We assume that the market is more likely to receive this fully revealing signal if the associate performs more tasks, n.

Define $\phi(n)$ as the probability that the market observes a revealing signal. We assume that $\phi(n)$ is an increasing and concave function of n, the number of tasks a worker performs, i.e. $\phi'(n) > 0$ and $\phi''(n) \leq 0$. Here, we provide one specific micro-foundation for this learning technology, but it is easy to construct others, and our key results would remain.[8]

Suppose the market observes the work load, n, and the resulting output for each professional worker, y_s^j. Then, given n, output provides a signal about worker ability. We follow Pries [2004] and assume that shocks to output are uniformly distributed, $\epsilon \sim U[\frac{-\varepsilon}{2}, \frac{\varepsilon}{2}]$. Figure 1 illustrates why this assumption generates all or nothing learning given our previous assumptions about y_s^j.

The two panels in the figure present two joint densities. Both densities describe output realizations for an associate who takes on a given work load, n. A given area under the density in the top panel equals the joint probability that a new associate both has low ability and produces output in a given interval. An area under the density in the bottom panel gives the corresponding joint probability of being high ability and producing output in a given range.



The regions of non-overlap between these joint densities contain signals that fully reveal the ability of new associates because only one ability type can produce the signals found in each of these regions. If a new associate produces less than $(1+x)n - \frac{\varepsilon}{2}$, the associate must be of low ability because a high ability associate would always produce at least this much. Further, if a new associate produces more than $n + \frac{\varepsilon}{2}$, the associate must be of high ability because a low ability associate would always produce this much or less.

[8]See Bonatii and Horner [2016] for a framework where revealing signals arrive at rates determined by worker effort levels.

Any signal in the region where these densities overlap provides no information about the ability of a new associate. For output values in this region, the joint density function in the bottom panel is $\frac{\pi}{\varepsilon}$ while the joint density in the top panel is $\frac{1-\pi}{\varepsilon}$. Thus, Bayes' rule implies that

$$Pr(\theta = x | y_o^a \in [(1+x)n - \frac{\varepsilon}{2},\ n + \frac{\varepsilon}{2}]) = \frac{\pi/\varepsilon}{\pi/\varepsilon + (1-\pi)/\varepsilon} = \pi$$

Given an output signal in the overlap region, the probability that a new associate is high ability is π, which is the prior probability that each new associate has high ability.

The length of this region of overlap, $[(1+x)n - \frac{\varepsilon}{2},\ n + \frac{\varepsilon}{2}]$, is $\varepsilon - xn$. Multiply this length by the density of the production shock, $\frac{1}{\varepsilon}$, to get, $1 - \frac{x}{\varepsilon}n$, which is the probability that the output signal reveals no information about associate ability.

Thus, $\phi(n) = \frac{x}{\varepsilon}n$ is the probability that an output signal reveals the type of a worker who takes on work load n. We assume that $\phi(\bar{n}) < 1$ to create an environment where it is not possible to achieve complete information about the ability of associates simply by working them "hard enough." Nonetheless, heavier work loads do create more information in this model, i.e. $\phi'(n) > 0\ \forall n \in [0, \bar{n}]$. New associate effort, n, not only produces output but also reveals information new associates.

Given assumptions 2.1 and 2.5, associates who reveal that they are low ability should always be re-assigned to the outside sector, and associates who reveal they are high ability should always be promoted to partner. Optimal second period assignments for associates of uncertain ability are more subtle. Our analysis below highlights how z^p and z^a interact to determine both optimal work loads for new associates and whether or not experienced associates of uncertain ability face dismissal.

3. The Planner's Problem

Here, we describe the planner's problem for our economy. The planner seeks to maximize the present discounted value of the sum of present and future differences between per period output and effort costs by assigning workers to jobs and work loads.

In each period, the planner's problem involves ten choices. Table 1 demonstrates that there are five different types of workers in this economy. The planner must choose a job assignment and effort level for each type. We proceed by showing that the optimal job assignment for three of these types is immediate. We then argue that the optimal effort levels for four of these five types are solutions to standard static optimization problems. Thus, 7 of the planner's 10 choices are immediate given our assumptions. We devote our analysis below to the two assignment decisions and one effort choice that remain.

Table 1
Optimal Assignment for Worker Types

History	Ability: $\theta = 0$	Ability: $\theta = x$	Ability: $Pr(\theta = x) = \pi$
New	n.a.	n.a.	Associate / outside
Experienced Outside	n.a.	n.a.	Outside
Experienced Professional	Outside	*Partner*	?

Notes: The rows delineate three types of workers: new, experienced in the outside sector, and experienced in the professional sector. The columns spell out the three possible information states about worker ability.

The rows of Table 1 describe three different types of workers with respect to their previous work experience. Recall that workers live two periods, and there are two sectors. Thus, workers may have no experience in either sector, one period of experience in the outside sector and no experience in the professional sector, or no experience in the outside sector and one period of experience in the professional sector. The columns of Table 1 describe three different information states that may apply to workers. The market may know that a worker is low ability, $\theta = 0$. The market may know the worker is high ability, $\theta = x$, or the market may be uncertain about the worker's ability and believe that there is a probability π that the worker is high ability.

The intersections of these three experience types and three information sets yield nine cells in Table 1. We begin by explaining why the first two columns of the first two rows are marked, n.a., for not applicable. These rows describe workers who have no professional experience yet are known to be either high or low ability. Since all workers are born with uncertain ability, and all learning takes place in the professional sector, no one can know the true ability of any worker who has no professional experience. Thus, these four cells describe types that never exist. The five cells in the last column and last row of Table 1 describe types that may exist. The planner must assign these types to jobs and work loads.

Three of the planner's five job assignment decisions are trivial. First, new workers in the top right corner of Table 1 have uncertain ability. So, they never work as partners. Given our linear production technologies, the planner either assigns all new workers to the outside sector, or he assigns new workers to associate positions in the professional sector until the constraint on professional employment, q, binds and then assigns the remaining new workers to the outside sector. We assume that the later case holds. If not, the planner never assigns new workers to the professional sector, and all workers spend their entire career in the outside sector.

Next, consider workers in the bottom left hand corner. These workers have one period of professional experience as an associate, and their output signals have revealed that they are low ability. The planner clearly assigns these workers to the outside sector. Because they have only one period of life remaining, option value

considerations cannot affect their assignment, and by assumption, they are more productive in the outside sector, since $z^a < w^0$.

Finally, turn to the second row and last column. These workers have one period of experience in the outside sector. It is straightforward to see that the planner keeps these workers in the outside sector. To begin, the planner never assigns these workers to be partners in the professional sector because their abilities are uncertain. In addition, the planner never assigns these workers to associate positions. Each new worker has the same expected associate productivity as an experienced outside worker plus the prospect of being promoted to partner in the next period, and since employment in the professional sector is capped at $q < 1$, the planner could fill the whole sector with new associates.

This leaves two job assignment decisions for the planner. The planner must allocate experienced professionals who do not have known low ability, i.e. the professionals in the second and third columns of the bottom row. Our assumptions dictate that once the planner decides how many of these the experienced professionals to retain in the professional sector, the planner fills the remaining slots in the professional sector by assigning new workers to the associate position. He then assigns the remaining new workers to the outside sector. Thus, the allocations of experienced professionals who have either unknown ability or high ability pin down the allocations of new workers and completes the assignments of workers to positions.

We quickly establish below that the planner promotes high ability types, $\theta = x$, to partner. We then turn to the optimal assignment rule for experienced professionals of uncertain ability. We refer to the case where the planner assigns these workers to the outside sector as an up-or-out environment because the planner only retains those high-ability professionals that he promotes to partner. All other experienced professionals are assigned to the outside sector.

Now, consider optimal work loads for the five types described in the bottom row and last column of Table 1. Four of these five types are experienced workers. These workers have only one period of life remaining, so the planner assigns work loads to them that equate expected marginal products of effort with the marginal costs of effort.

Finally, the planner must assign work loads to the new workers in the top right hand corner of Table 1. Among those who begin in the outside sector, output levels produce no information about worker ability, and we have shown above that the planner will not assign these workers to the professional sector in next period. Thus, optimal effort for these workers also corresponds to the solution of a simple static maximization problem.

In contrast, the optimal effort choice for new workers who start their careers as associates in the professional sector is more interesting. Among these workers, heavier work loads generate more output and more information about worker ability. Because this information guides assignment decisions in the next period, the optimal work load for new associates must reflect the fact that effort today affects assignments and work loads in the future.

To review, our planner's problem involves five job assignment decisions and five effort choices. In seven of ten cases, the planner's optimal policies are immediate. Further, we quickly establish below that the planner always assigns experienced professionals with known high ability to the partner position. Thus, our planner

must confront two key questions: Should experienced associates whose abilities remain uncertain continue working as associates in the professional sector, and what is the optimal work load for new associates? In the following section, we analyze these questions and demonstrate how the answers to these two questions are connected.

3.1. **Recursive Formulation.** In this section, we describe the planner's optimal policies as solutions to a specific Bellman equation. Before describing this equation, we introduce some additional notation.

- v^o is the per period surplus created by an outside sector worker.

Our assumptions imply that all workers in the outside sector produce the same amount and incur the same effort costs regardless of their past work experiences or talent level.

For professional workers, v_s^j describes the per period surplus created by a professional worker with $s \in \{0, 1\}$ periods of professional experience in position $j \in \{a, p\}$.

- $v_0^a(n)$ - the expected surplus created by a new associate who takes on a work load of n.
- v_1^a - the expected surplus created by an experienced professional of uncertain ability who works as an associate
- v_1^p - the expected surplus created by an experienced professional of high ability who works as a partner

We omit explicit notation for effort in v^o,v_1^a, and v_1^p, because optimal effort levels for these workers are solutions to simple, static maximization problems. We include n in $v_0^a(n)$ because the effort of new associates both produces output and creates information that influences future payoffs. Our notation does not specify beliefs about the abilities of experienced associates or partners because, given our production function assumptions, the planner only assigns workers with known high-ability to the partner position, and the planner only retains experienced workers of uncertain ability as associates.

Now, consider the stock variables for our problem. Each period, there is a mass one of new workers. There are experienced workers who spent their first period of life in the outside sector, and there are three types of experienced workers who spent their first period of life in the professional sector: those with known low ability, those with known high ability, and those with uncertain ability.

In our formulation, we need only keep track of two of these five stock variables.

- ρ^u - the mass of experienced professionals who have uncertain ability.
- ρ^x - the mass of experienced professionals who have known high ability.

Recall that the planner retains all experienced outside workers in the outside sector and also assigns all professionals with known low ability to the outside sector. This implies that the mass q of workers in the professional sector is divided among new associates and two types of experienced professionals. The planner's assignment decisions concerning the stocks of experienced professionals with uncertain ability or known high ability pin down the mass of new associates in the professional sector, and this pins down allocations of new workers among the outside and professional sectors.

The key control variables for our planner involve job assignments for two types of experienced associates and the work load assignment for new associates.

- α^u - the fraction of experienced professionals with uncertain ability, ρ^u, that the planner retains in the professional sector, $\alpha^u \in [0,1]$
- α^x - the fraction of experienced professionals with known high ability, ρ^x, that the planner retains in the professional sector, $\alpha^x \in [0,1]$
- n - the work load for new associates

The per period surplus flow in this model is

$$s(n,\alpha^u,\alpha^x) = (2-q)v^o + (q-\alpha^u\rho^u-\alpha^x\rho^x)v_0^a(n) + \alpha^u\rho^u v_1^a + \alpha^x\rho^x v_1^p \tag{3.1}$$

The control variables (α^u,α^x) pin down the entire allocation of workers to positions. Given an allocation of workers to positions, work loads determine expected output. The planner assigns statically optimal work loads to all workers except new associates. The work load, n, for new associates not only influences expected output but, given the current stock of new associates, n also determines the key stock variables, ρ^u and ρ^x, next period. Thus, the planner's policies concerning α^u, α^x, and n drive the evolution of stocks and output flows over time.

The choice variable α^x does not explicitly involve the choice to assign high-ability workers to the partner position. However, we note above that the condition $z^p > z^a > 0$ implies that whenever the planner retains an experienced professional with known high ability in the professional sector, the planner also assigns this worker to the partner position. Likewise, the variable α^u does not explicitly involve the choice to assign the uncertain-ability professional to the associate position, but our assumptions about productivity in the partner position ensure that professionals of uncertain ability are never promoted to partner.

The planner's objective is to choose work loads and job assignments that maximize the discounted present value of the infinite stream of per period surplus generated in this economy. If we assume that the planner discounts the future using $\beta < 1$, we can write the planner's problem using the following recursive formulation

$$\begin{aligned} V(\rho^u,\rho^x) &= \max_{\alpha^u,\alpha^x,n} \; s(n,\alpha^u,\alpha^x) + \\ &\quad \beta V\left(\left[q-\alpha^u\rho^u-\alpha^x\rho^x\right]\left[1-\phi(n)\right],\; \left[q-\alpha^u\rho^u-\alpha^x\rho^x\right]\pi\phi(n)\right) \end{aligned} \tag{3.2}$$

Let $\{\hat{\alpha}^u,\hat{\alpha}^x,\hat{n}\}$ denote the solution to planner's problem. If the planner begins with positive stocks of uncertain and high ability professionals, $\rho^u > 0, \rho^x > 0$, the planner faces a trade off. These workers are more productive than new associates. The uncertain types are more productive because they have more professional experience. The high ability types are not only more skilled but also able to work in partner positions that exploit their skills. However, for each experienced professional that the planner retains in the professional sector today, he will have one less experienced professional next period. Further, the planner's effort choice for new associates, n, interacts with these retention decisions because the probability that the planner observes the actual ability of a new associate is $\phi(n)$.

3.2. **Promotion to Partner.** Now, we establish that $\hat{\alpha}^x = 1$, i.e. the planner retains all experienced professionals with known high ability. This implies, as we note in our discussion of Table 1, that the planner promotes all professionals with known high ability to partner.

There are no productivity spillovers among workers who occupy different positions in this model, and the output generated in one position is not a function of total employment in the position. This suggests that $V(\rho^u, \rho^x)$ is linear. We confirm this result and use it to establish that the planner always chooses $\hat{\alpha}^x = 1$. See Appendix A for proofs of this result and all other results in sections three and four.

Claim 1. $V(\rho^u, \rho^x) = K_1 + K_2\rho^u + K_3\rho^x$ for some constants K_1, K_2, and K_3.

This claim establishes that $V(\rho^u, \rho^x)$ is linear in the two stocks of experienced professionals, which implies that the value created by experienced professionals of high ability is not influenced by the stock of experienced professionals of uncertain ability, and vice versa. Further, the value of an addition to either stock of experienced professionals does not depend on the current level of either stock. This feature of our model implies our second claim.

Claim 2. $\hat{\alpha}^x = 1$

Since the value generated by an experienced professional of high ability is independent of the current stock variables, the planner gains nothing from trying to smooth the stock of experienced, high-ability professionals over time. This implies that he never replaces a high-ability, experienced professional with a new associate. If the planner were to make such a replacement, there would be a probability, $\pi\phi(\hat{n}) < 1$, that the new associate would be revealed to have high-ability at the end of the period, and even in this case, she would be no more productive next period than a high-ability, experienced professional would be this period.

4. The Link Between Work Loads and Up-or-Out

We have now determined optimal rules for eight of the ten choices the planner makes each period. There are two choices that remain: the work load for new associates, n, and the job assignment for experienced associates of uncertain ability, α^u. In this section, we demonstrate how these choices are related.

Substitute the expression for $V(\rho^u, \rho^x)$ in Claim 1 into equation 3.2. Then, take the derivative with respect to n to get the first order condition that defines the optimal work load for new associates, $\hat{n}$:

$$c'(\hat{n}) = (1 + \pi x) + \beta\phi'(\hat{n})(\pi K_3 - K_2)$$

This equation highlights an important property of optimal work loads for new associates. The marginal cost of new associate effort must equal the sum of two marginal returns. The first, $(1 + \pi x)$, is the expected marginal product of new associate effort. The second, $\beta\phi'(n)(\pi K_3 - K_2)$, is the marginal information return from worker effort. Note that K_2 and K_3 represent incremental values to the planner of replacing a new associate with an experienced professional who possesses either uncertain ability or known high ability, respectively. When the planner increases a new associate's work load, n, the probability that the associate's output signal reveals her ability increases by $\phi'(n)$. If the signal is revealing, there is a probability π that the signal will reveal high ability, and the planner's value function will increase by K_3 instead of K_2. Thus, $\phi'(n)(\pi K_3 - K_2)$ is the marginal information rent generated next period by new associate work this period.

We show in Appendix A that πK_3 is always greater than K_2. Thus, the information rents created by new associate effort are always positive. Given this result, we prove the following proposition:

Proposition 1. *The optimal work load for new associates, $\hat{n}$, exceeds the static optimum implied by the expected per period output of new associates.*

New associates take on work loads that are greater than those that would maximize the current surplus generated by their positions in order to produce information that improves professional job assignments in the future.

A similar trade off between current output and future information shapes the planner's decision concerning the retention of experienced associates of uncertain ability. Given equation 3.2, the first order condition that defines $\hat{\alpha}^u$ is

$$\hat{\alpha}^u = \begin{cases} 1 & \text{if } v_1^a - v_0^a(\hat{n}) - \beta\left[(1-\phi(\hat{n}))\,K_2 + \pi\phi(\hat{n})K_3\right] \geq 0 \\ 0 & \text{if } v_1^a - v_0^a(\hat{n}) - \beta\left[(1-\phi(\hat{n}))\,K_2 + \pi\phi(\hat{n})K_3\right] < 0 \end{cases}$$

In Appendix A, we solve for K_2 and K_3 and use these solutions to express the first order condition for α^u in a more informative way.

Proposition 2. $\hat{\alpha}^u = 0$ *if* $v_1^a - v_0^a(\hat{n}) < \beta\pi\phi(\hat{n})(v_1^p - v_1^a)$

The planner imposes an up-or-out rule, i.e. he promotes experienced associates with known high ability to partner and assigns all other experienced associates to the outside sector, when the expected future return from adding a new associate this period is greater than the expected returns to experience among associates of uncertain ability.

The left-hand side of the inequality in Proposition 2 is the current surplus cost of replacing an experienced associate of unknown ability with a new associate. The right-hand side gives the expected future returns that these replacements create. Experienced associates with uncertain ability are never going to be partners, but new associates may become partners next period. The probability of this event is $\beta\pi\phi(\hat{n})$, and the additional surplus generated by known high-ability professionals who work as partners is $v_1^p - v_1^a$.

Thus, the planner's decision concerning $\hat{\alpha}^u$ reflects a trade off between the return to worker experience in the associate position and the value of identifying high-ability workers and promoting them to the partner position. As in our discussion of heavy work loads among new associates, up-or-out rules reduce current surplus but increase expected surplus in the future by increasing the number of high-ability professionals who become partners.

We are interested in how the relative productivities of different types of workers in different roles within organizations shape optimal personnel policies. Thus, we want to understand how z^p and z^a shape the value of identifying candidates for promotion to partner and the surplus cost associated with replacing experienced professionals. Our key comparative static results spell out how both parameters affect $\hat{n}$ and $\hat{\alpha}^u$. Appendix A proves the following result concerning optimal work loads:

Proposition 3. *The optimal work load for new associates, $\hat{n}$, is increasing in z^p and non-increasing in z^a.*

The probability that a new associate's output signal reveals her true ability increases with $\hat{n}$. Thus, if new associates work more this period, the planner will be able to identify and promote more partners next period. For parameter values such that $\hat{\alpha}^u = 1$, the additional surplus generated by these promotions increases with z^p and decreases with z^a. Therefore, optimal effort, $\hat{n}$, increases with z^p and decreases with z^a. If $\hat{\alpha}^u = 0$, z^a does not enter these surplus calculations because no one works as an experienced associate. However, $\hat{n}$ still increases with z^p.

We argue above that many professional service firms employ both heavy work loads for new associates and up-or-out promotion rules as tools that facilitate their search for talented partners. Thus, the effects of z^p and z^a on firm decisions concerning up-or-out should be similar to their effects on work loads for new associates. Our second comparative static result confirms this:

Proposition 4. *$\hat{\alpha}^u$ is non-increasing in z^p and non-decreasing in z^a.*

Up-or-out is optimal when the option value associated with a new associate exceeds the productivity gains that come from associate experience. The returns from finding talented professionals and promoting them to partner are increasing in z^p, and z^a determines the gains from associate experience.

Up-or-out promotions rules and heavy work loads for new associates go together. Given any initial pair (z^a, z^p), $\hat{n}$ increases and an up-or-out rule is either retained or adopted as we increase z^p. Further, $\hat{n}$ decreases and an up-or-out rule may be abandoned as we increase z^a.

In environments where z^p is large enough relative to z^a, the planner chooses heavy work loads and up-or-out retention rules. These policies make no sense as static allocation rules. Yet, both policies are optimal because the planner is not just producing surplus for the current period. He is also conducting a search for talent, and the results of this search impact future surplus.

Several authors point out that strict adherence to up-or-out rules became less common in law firms during the 1980s and 1990s.[9] Gorman [1999] argues that, as the market for legal services grew, private law firms began doing relatively more work that required special expertise and experience. Thus, experienced, skilled lawyers who were not well-suited to the role of partner became more valuable. If we phrase these claims in terms of the language of our model, Gorman [1999] is arguing that, at least for some lawyers with special expertise, z^a increased over time and law firms responded by moving away from strict up-or-out rules.

As the market for professional services has grown in recent decades, the relative demand for specialists has produced similar movements away from strict up-or-out policies in public accounting.[10] Still, in most law and public accounting firms, the majority of young professionals begin their careers expecting that they will either move up to partner or out to another employer.

4.1. **Stay or Go.** A large literature on screening rules points out that, in models with one position, the option value associated with bringing in a new worker determines the stringency of the rule that governs the retention decision for the

[9]See Galanter and Palay [1991], Galanter and Henderson [2008], Gorman [1999], Gilson and Mnookin [1985]

[10]Press accounts concerning changes in the use of up-or-out rules in public accounting firms echo Gorman's claims about the rising value of specialists. See **New York Times**, May 17, 1990. See Almer [2004] for most recent data.

incumbent workers. Thus far, we have assumed $z^p > z^a$, but if $z^p = z^a$, the professional sector in our model effectively contains only one position. Once again, the planner will dismiss all experienced professionals who are known to possess low-ability, $\theta = 0$, and retain all experienced professional with known high-ability, $\theta = x$. Given some parameters, $\hat{\alpha}^u = 1$ and given others, $\hat{\alpha}^u = 0$, but in both settings the planner is only making a retention decision. There is no promotion decision.

In this setting, some may describe the more stringent screening rule, $\hat{\alpha}^u = 0$, as an up-or-out policy, but this is not how we employ the term. When $z^a = z^p$, one can interpret $\hat{\alpha}^u = 0$ as a decision to search for stars, but our goal is to examine the links between personnel policies and the relative productivities of different positions in an organization.

Partners are not simply persons who do the work that new associates do well. Partners in professional service firms do more than provide quality professional services. They also perform the business development and client management activities that allow their firms to survive and grow. Our assumptions concerning z^p, z^a, and the losses that would occur if low ability types were to occupy the partner position are all attempts to create a planner's problem that characterizes a search for persons who can succeed in the demanding role of partner.

5. Learning About Rates of Learning

Above, we analyzed a model where the planner assigns tasks to an associate in order to produce output and to learn whether or not the associate is talented enough to work as a partner. In the following section, we will show that the planner's solution can be implemented as an equilibrium of a decentralized market for professionals, and in this equilibrium, new associates willingly pay an up-front utility cost to learn whether or not they are well-suited to partner jobs. However, young professionals surely gain more than information about their talents through work experience. They also gain human capital. Further, both the academic and trade literatures on professional labor markets assert that young professionals enter elite professional service firms, in part, to acquire valuable skills.[11]

For these reasons, we now consider human capital accumulation, but just as we assume that workers do not begin their careers with a complete understanding of their own skill levels, we also assume that workers do not begin their careers with certainity about their capacity to aquire new skills. Thus, they learn ex post whether or not their work experiences have produced the skills required to earn promotions or raises. In this section, we flesh out the many parallels between learning about worker skill and learning about the ability of workers to acquire skills. We show that both frameworks allow us to think about professional service firms as organizations where new associates audition for partner positions.

In our model above, new associates possess different ability levels, and three information states are possible for an associate who has just completed her first period of professional work. The planner may know that she has high ability. The planner may remain uncertain about her ability, or the planner may know she has low ability.

[11]See Rosen [1972] for an early model of the market for training opportunities. See Wilkins and Gulati [1998] for a discussing of how partners train law associates. See Batchelor [April 20, 2011] for a discussion of training in management consulting.

Now consider a learning by doing model in which new associates begin their careers at a common skill level but learn by doing at different rates that are unknown to them or the planner.[12] Appendix B presents a formal treatment of this model. Here, we summarize key features and results.

To facilitate comparisons with our screening model, again assume that three information states are possible for an associate who has just completed her first period of professional work. The planner may know that she has aquired the skills needed to function well as a partner. The planner may know that she has acquired the skills required to function effectively as an experienced associate but not as a partner, and the planner may know she has obtained few skills and belongs in the outside sector.

In our screening model, with probability$\pi\phi(n)$, associates who perform work load n learn that they have high-ability. With probability $1-\phi(n)$, these associates learn nothing and maintain their priors beliefs about their abilities. With probability $(1-\pi)\phi(n)$, they revise their beliefs downward because they discover that they possess low-ability. In our learning by doing model, no one observes, ex ante, the different rates at which new associates learn by doing. Yet, ex post, all market participants do observe the skill level that each experienced associate has obtained. Since skills grow through learning by doing at uncertain rates, the work load, n, that the planner assigns to new associates determines the distribution of realized skill levels among experienced associates, and $\phi_h(n)$, $\phi_m(n)$, and $\phi_l(n) = [1 - \phi_m(n) - \phi_h(n)]$ are the probabilities that a new associate who performs n tasks achieves a high, medium, or low skill level, respectively.

In our screening model, our key assumption about the information technology is that the probability of a revealing signal, $\phi(n)$, is an increasing and concave function of n. This implies that the probability of identifying a high-ability professional, $\pi\phi(n)$, is increasing and concave, while the probability of remaining uncertain about the ability of an experienced professional is decreasing in n. In our learning by doing model, the analgous conditions are that $\phi_h(n)$ is increasing and concave and $\phi_m(n)$ is decreasing in n.

Given these conditions, our learning by doing model is essentially isomorphic to our screening model, and we are able to establish results that parallel all of our propositions above.[13] Our model is a model of auditions. Young professionals enter associate positions and work long hours as part of a process that reveals whether or not they either have the skills or can acquire the skills they need to perform well as partners.

Although Appendix B demonstrates that all the results from our screening model carry over to our model of learning by doing if $\phi_h(n)$ is increasing and concave and

[12]Demougin and Siow [1994] present another model of a training and screening where new professionals train for management positions, but ex post, only some acquire the skills required to be effective managers.

[13]Our learning by doing model is not identical to our screening model. In the screening model, the probability that the planner learns that a new associate has high ability is linearly related to the probability that the planner receives no information about the associate and thus maintains his prior belief about the associate's ability, but in our learning by doing model, there is no linear relationship between the corresponding probabilities $\phi_h(n)$ and $\phi_m(n)$. Moreover, in our screening model, the probability that the planner learns identifies an associate with low ability is increasing in n while the corresponding probability in our learning by doing model, $\phi_l(n)$, is increasing in n Nonetheless, these differences do not impact the arguments we use to prove Propositions 1 through 5 above.

$\phi_m(n)$ is decreasing, there are plausible versions of our learning by doing model where $\phi_m(n)$ is an increasing function. As young associates put in more effort, they acquire more skills, and in some versions of our learning by doing model, this increase in skills lowers $\phi_l(n)$ while increasing both $\phi_m(n)$ and $\phi_h(n)$.[14] In these cases, results that parallel Propositions 1 through 4 still go through as long as the ratio $\phi_h(n)/\phi_m(n)$ increases with n sufficiently fast. If this ratio does not rise fast enough with n, the comparative static results in Proposition 3 may flip, i.e. n may decrease with z^p or increase with z^a.

In sum, Appendix B establishes parallel versions of Propositions 1,2, and 4 for our learning by doing model given weak restrictions on $\phi_l(n)$, $\phi_m(n)$, and $\phi_h(n)$, but we require an additional restriction on how the ratio $\phi_h(n)/\phi_m(n)$ changes with n to establish a parallel version of Proposition 3. This restriction requires that changes in works loads for new associates must increase the future stocks of high-ability partners relative to future stocks of experienced but merely compotent professionals.

6. Endogenous Sector Size and Decentralization

So far, we have described solutions to a planner's problems given the constraint that professional employment cannot exceed a cap of q, and we have assumed production technologies such that this cap on professional employment always binds. Here, we show that our results hold in a more general setting where the planner also chooses the optimal size of the professional sector. Further, we characterize an equilibrium of a decentralized economy that implements the solution of this more general planner's problem.

6.1. **Endogenous Sector Size.**

We assume that maintaining a professional position is costly. Professional workers require support staff that facilitate their capacity to interact efficiently with clients. Further, we assume that the supply curve for effective support services is upward sloping because each potential support worker has a different outside option in another sector. Thus, the resource cost of supporting a given professional position is determined by the outside option of the staff who support that position.

Suppose the per-period cost of supporting the q-th position is given by $\kappa(q)$, where $\kappa(\cdot)$ is an increasing function such that $\lim_{q\to 0}\kappa(q) = 0$ and $\lim_{q\to 1}\kappa(q) = \infty$. Given our previous assumptions, these restrictions ensure that some but not all of the workers in a cohort begin their careers in the professional sector. The planner faces the same problem as before, but now, he must also determine $\hat{q}$, the optimal number of professional workers. Recall equation 3.1 and note that expected surplus in the current period depends on q, which is now a choice variable for the planner. The new planner's problem is

$$\begin{aligned} V(\rho^u, \rho^x) \quad = \quad & \max_{n,\alpha^u,\alpha^x,q} \quad s(n,\alpha^u,\alpha^x,q) - \int_0^q \kappa(y)dy + \\ & \beta V\left(\left[q - \alpha^u\rho^u - \alpha^x\rho^x\right]\left[1-\phi(n)\right]\right),\ \left[q - \alpha^u\rho^u - \alpha^x\rho^x\right]\pi\phi(n)\right) \end{aligned} \tag{6.1}$$

Appendix C analyses this problem. It is straightforward to show that the three first-order conditions for n, α^u, α^x are the same as the first-order conditions in our

[14] The distribution of new associate skills at the end of period one is stochastically increasing in n. This implies $\phi_l'(n) \leq 0$ and $\phi_h'(n) \geq 0$, but this imposes no restriction on $\phi_m(n)$.

original formulation. This result holds because there are no productivity interactions between positions in our model. The productivity of a worker in a given position is not influenced by how other workers are allocated to positions. The planner's optimal sector size choice does not impact optimal work loads for new associates or optimal retention or promotion rules. Rather, the shape of $\kappa(q)$ and the expected surplus that optimal work loads and retention rules generate pins down the planner's choice for $\hat{q}$.

6.2. **Decentralization.** We next turn to the question of whether the planner's allocation can be achieved in a decentralized market economy. Observe that our model does not involve any information asymmetries. Workers have no private information about their abilities or their actions. Further, all output signals and actions are public. Thus, many different market mechanisms could implement the solution to our planner's problem. Appendix C proves that one particular mechanism does. Here, we discuss how and why this mechanism would work.

The most straightforward way to decentralize our planner's problem is to assume that all workers choose whether to work in the outside sector at a fixed wage, w^o, or work in the professional sector as independent contractors, i.e. they choose whether to work as an associate or as a partner and then receive the output they create. We treat workers as independent contractors to facilitate exposition. The same results would hold in a competitive labor market where identical professional service firms posted a menu of employment contracts that specified optimal work loads and wages equal to expected marginal products for all possible combinations of worker types and positions.

In this decentralization, each professional must hire support services at a cost $\kappa(\tilde{q})$, where $\tilde{q}$ is the size of the professional sector induced by the sector choices of market participants.[15] Given any $\tilde{q}$, our assumptions on productivity imply that experienced professionals with known low-ability choose the outside sector and that experienced outside workers choose to remain in the outside sector. Further, the work load choices for all outside workers and experienced professionals are immediate. Thus, we consider the sector choices of experienced professionals who do not possess known low ability, the intitial sector choice of new workers, and the work load choices of new workers who begin their careers as professional associates.

Given a professional sector of size $\tilde{q}$, define $\upsilon(n)$ as the expected lifetime utility of a new associate in the professional sector who chooses work load n.[16]

(6.2)
$$\upsilon(n) = v_0^a(n) - \kappa(\tilde{q}) + \beta\pi\phi(n)\max[v^o, v_1^p - \kappa(\tilde{q})] + \beta(1-\pi)\phi(n)v^o + \beta(1-\phi(n))\max[v^o, v_1^a - \kappa(\tilde{q})]$$

The first two terms are the utility of working as an associate. The next term is the worker's expected discounted utility from learning that she has high ability. The fourth term is the expected discounted value of learning that she has low ability and then switching to the outside sector in the second period. The last term is the expected discounted value of being uncertain about her ability at the end of the first period.

[15]Competition for staff workers implies that all professionals must pay their staff the outside option of the marginal staff worker.

[16]Here, we are assuming that professional workers expect the size of the professional sector to remain constant in both periods of their life.

Appendix C demonstrates that the market equilibirum in our independent contractor scenario implements the planner's solution to the problem described by equation 6.1. We show that, if $\tilde{q} = \hat{q}$, the following are true: new associates choose the planner's work load,$\hat{n}$, to maximize $v(n)$, experienced associates with known high ability stay in the professional sector and work as partners, and experienced associates of uncertain ability leave the professional sector and choose outside work if and only if the production environment is such that the planner would make the same assignment, i.e. $v^o > v_1^a - \kappa(\hat{q}) \Leftrightarrow \hat{\alpha}^u = 0$.

We know from our discussion of the planner's problem that, in order to implement the planner's choice for sector size, $\hat{q}$, in each period, some new workers must choose the outside sector and some must enter the professional sector.[17] Hence, new workers must be indifferent between beginning their careers in the professional sector or the outside sector, i.e. $v(\hat{n}) = v^0(1+\beta)$. If $\tilde{q} > \hat{q}$, new workers strictly prefer the outside sector. If $\tilde{q} < \hat{q}$, new workers strictly prefer the professional sector. Thus, the condition $v(\hat{n}) = v^0(1+\beta)$ implies $\tilde{q} = \hat{q}$.

Appendix C also proves another result that helps us understand how this decentralization works.

Proposition 5. *All else equal, $\hat{q}$ and $\kappa(\hat{q})$ are increasing functions of z^p.*

When partners become more productive, the professional sector grows, and the costs of maintaining professional positions grows. Because $v^o > v_1^a - \kappa(\hat{q}) \Leftrightarrow \hat{\alpha}^u = 0$, up-or-out is optimal when z^p is high enough, given z^a, to drive $\kappa(\hat{q})$ beyond $v_1^a - v^o$. In these cases, all associates who realize they are not going to become partners willingly choose the outside sector. In up-or-out regimes, the search for talented partners drives the costs of maintaining professional positions so high that each position is occupied by someone who either is a partner or could become one.

New associates willingly accept lower expected utility in their first period of employment because they are willing to pay to learn about their talent levels.[18] Among experienced professionals, there is no next period, and therefore information gained this period has no value.

The fact that working as a new associate involves paying an up-front utility cost in exchange for the possibility of becoming a partner may shed light on survey evidence concerning the job satisfaction of young professionals. Young professionals often report low job satisfaction, and in particular, they report that they would be willing to accept lower earnings in exchange for less demanding work schedules.[19] While advocates of rat race models cite these responses as evidence that young professionals take on work loads that are inefficient, our model offers another interpretation.

Consider the possibility that young professionals who respond to such surveys may be reporting that, holding all else constant, they are willing to accept lower

[17] We know that, in each period, some professional workers must be new workers. Note that $\hat{q}$ does not depend on the current stocks of experienced professionals, and this means that even if all the $\hat{q}$ workers in the previous period were new professionals, the market would we need some new professionals this period to reach $\hat{q}$ professional workers again, since all of those who would have learned during last period that they have low-ability would have chosen the outside sector this period.

[18] This follows directly from $v(\hat{n}) = v^0(1+\beta)$, and in part, reflects the fact that $\hat{n}$ is greater than the work load that maximizes static surplus.

[19] See Landers et al. [1996] for survey responses from young laywers.

earnings in exchange for less demanding work loads. Further, suppose that one of the things they hold constant when answering these questions is their future prospects for promotion. Given this scenario, new associates in our model would express the same willingness to exchange current salary for reduced work loads. The problem is that there is no way to make such an exchange while holding all else constant. If new associates did perform fewer tasks, the market would learn less about them, and they would be less likely to become partners.[20]

Although many young professionals report that, relative to their current terms of employment, they would be willing to exchange money for leisure, these reports are not necessarily evidence of market failure. Information is costly, and these reports may simply mean that workers would rather live in a world that allowed them to discover and reveal their abilities at no cost.

Scholars have argued that up-or-out policies are puzzling because surely some experienced associates who are not well suited to the partner role are nonetheless competent professionals. Why would firms refuse to negotiate a retention package for these associates?[21]

Our model provides an answer. In our framework, young professionals pay a utility cost to acquire information about whether or not they are well-suited to lucrative partner positions. Thus, when a young professional learns that she is not "partner material," she is not willing to pay this cost any longer. Further, although she would be more productive, in expectation, than the new associate who replaces her, her firm cannot profitably make a retention offer that she would accept because it can no longer offer her the opportunity to learn about her fitness for the partner role.

If up-or-out is one of several tools that professional service firms use to search efficiently for partners, then the decision of firms not to negotiate a new deal with skilled professionals who turn out not to be well-suited for the partner role is not a riddle that needs to be solved. Associates enter these firms because they hope to become partners. Once they realize that they will not become partners, they also know that their best options are elsewhere.

7. Empirical Patterns Concerning Hours and Promotions

In this section, we discuss some empirical patterns in the market for lawyers. We focus on differences in hours worked, billing rates, and total compensation among lawyers who occupy different positions during the first twelve years of their careers. We highlight one pattern that appears in two different data sources. On average, lawyers who begin their careers as associates in private law firms and subsequently accept non-partner track positions or leave private law completely appear to reduce their hours worked significantly even though their wage rates appear to be constant or increasing. We can easily pick parameters for our model that produce this pattern because, in our model, associates take on heavy work loads as part of an audition process, and once their auditions are over, they no longer earn the same information returns from their work. This is true for those who make partner as well as those

[20]Likewise, in the training model we discuss in section 5, associates who work less would be less likely to acquire the skills needed to become partners.

[21]See Kahn and Huberman [1988]. Gilson and Mnookin [1989] discuss this puzzle in the context of experienced legal associates. Batchelor [April 20, 2011] discusses how leading consulting firms often place those who fail to make partner in prestigious jobs in large companies that are clients of the firm.

who do not, but those who earn partner positions work more because they take on valuable leadership and business development roles that increase their productivity.

Before describing our results, we must discuss several issues that arise concerning how to map our results into data. In our model, there are only two positions, associate and partner. For much of the 20th century, most private law firms created only these two positions. However, in recent decades, many firms have created additional positions. Many firms now have non-equity partner positions, and a smaller number have "Of Counsel" or counsel positions.

Our reading of the literature and our work with one of the data sets described below leads us to conclude that most counsel positions are occupied by persons who are not being consider for promotion to partner. However, lawyers who occupy non-equity partnership positions relatively early in their careers are still being evaluated for promotion to equity partner, and some already function somewhat like traditional partners since they develop new business, share in the profits of the firm, and vote on matters of firm governance. Appendix D describes the information we have gathered about these positions in more detail.

We focus on lawyers who are in roughly the first twelve years of their careers. We proceed under the assumption that persons who are promoted from associate to non-equity partner early in their careers are are likely persons who are still trying to earn promotions to full partner positions. We assume that persons who transition from associate to counsel positions early in their careers are no longer being considered for promotion to partner.[22] Thus, firms that retain associates by promoting them to non-equity partner are often gathering more information about whether or not these lawyers should be full partners, but firms that retain associates by promoting them to counsel are deviating from strict adherence to up-or-out.

Appendix D demonstrates that new associates almost never move to counsel positions in the original firm and rarely transition directly from being an associate in one firm to a counsel in another. We therefore assume that lawyers who begin their careers as associates but transition into counsel positions are individuals who became associates hoping to become partners, yet over time, learned that they were not well suited to the partner role. In addition, both they and the market learned over time that they are particularly well-suited to a senior, non-partner role that likely requires special expertise.[23]

Given this scenario, consider Table 2. This table describes data from the Survey of Law Firm Economics (SLFE), which is conducted annually by ALM Legal Intelligence. The data come from eight annual surveys taken during the period 2007-2014. Some firms appear in more than one annual survey, but these are not observations from a panel data set. Rather, the data come from eight repeated cross-sectional surveys.[24]

Table 2 describes outcomes for lawyers who are between eight and twelve years into their careers. We chose this experience interval because firms make many crucial retention and promotion decisions about progression to partner in this interval.

[22]Among more experienced lawyers, some counsels and non-equity partners are former partners in their firms or other firms who voluntarily or involuntarily left their partnership positions because they were unwilling or unable to meet the expectations of other partners. See Richmond [2010].

[23]However, we are assuming that these specialists are still significantly less productive than the typical partner. In our framework, firms adopt up-or-out rules because partners are so productive relative to the professionals who work for them.

[24]Some lawyers may appear in two different cross-sections, but we cannot link these records.

The table presents results for associates, equity partners, non-equity partners, and counsel attorneys. Because these counsel attorneys are less than 12 years into their careers, it seems reasonable to assume that most of them are attorneys who recently left an associate position in their current or previous firms and now occupy a senior role off the partnership track. Thus, it is interesting to note that, compared to the other three groups, counsel attorneys bill fewer hours but charge clients higher rates for their time. It is particularly noteworthy that counsels bill their time at rates 17 percent greater than associates yet bill 25 percent fewer hours.[25]

These data contain information about hours billed but do not measure total hours worked. Therefore, we cannot use these data to recover actual wage rates for workers in different positions. Still, given the large gaps in hours billed and billing rates, it seems highly unlikely that the associates in Table 2 are earning higher hourly wages than counsel attorneys who have comparable total experience. So, why would these associates bill many more hours than their peers who work as counsels? Our model suggests that this hours gap reflects the fact that counsel attorneys are no longer auditioning for partnerships. Their past work experience has revealed their type, and they are no longer producing signals about their suitability for the partner position.

Some readers may conjecture that these results for counsel attorneys are a "mommy track" outcome, but these results are not driven by women taking counsel positions in order to spend more time with young children. We have created a similar table that contains only male lawyers, and the results are quite similar.[26]

[25]Given the experience restrictions we impose on our sample, the associates in Table 2 are persons who are reaching the end of their tenure as associates, and in some cases, may already know that they are not going to be promoted and are therefore engaged in searches for new positions. Nonetheless, these associates still bill more hours than counsel attorneys who have similar levels of experience and bill their time at higher rates.

[26]See Harrington and Hsi [2007] for a discussion of links between gender differences in promotion rates and gender differences in time spent caring for young children. Gender is missing for many records in the SLFE data. However, in the sample of respondents who report being male, the same patterns exit, and the hours billed gap between Associates and Counsels is even slightly larger than the one reported in Table 2 for the full sample.

Table 2
Hours Billing, Billing Rates, and Total Compensation By Position

	Equity Partner	Non-Equity Partner	Associate	Counsel
Avg Hours Billed	1,645	1,611	1,494	1,128
(Std Deviation)	(417)	(539)	(592)	(632)
N	3,004	3,139	7,177	559
Avg Hourly Rate	289	298	258	304
(Std Deviation)	(77)	(96)	(84)	(104)
N	3,012	3,135	6,964	544
Avg Compensation	230,773	196,035	140,098	138,453
(Std Deviation)	(99,221)	(79,366)	(56,236)	(71,739)
N	3,075	3,192	7,316	575

Notes: This table reports job characteristics for lawyers included in the Survey of Law Firm Economics taken by ALM Legal Intelligence between the years 2007 and 2014. Within each panel of this table, the first number in each cell is the sample mean, the numbers in parentheses are standard deviations, and N is the sample size. This table considers only attorneys with between 8 and 12 years of experience, defined as years since passing the bar. Compensation is defined as take home salary and retirement contributions plus year end bonus plus benefits. Hours Billed is the annual number of billable hours for each attorney. Hourly rate is the typical hourly rate charged by each attorney. Sample sizes differ by cell due to different frequencies of item non-response.

Partners and non-equity partners also work more than counsel attorneys, but this pattern is easy to understand. The implied hourly wage rate for new partners may be much greater than the implied wage rate for counsels, and many non-equity partners are still auditioning for full partnerships, receiving performance-based profit sharing, or both.

The results in Table 2 are only suggestive because they do not contain panel data on particular lawyers and therefore do not allow us to know with certainty that these counsel attorneys recently left associate positions. Tables 3 present results from a second data set that provides much smaller samples but does allow us to track individual lawyers over time.

Table 3
AJD Full Sample
Hours Worked and Compensation By Wave 3 Position

	Equity Partner	Non-Equity Partner	Associate	Counsel	Other
Avg Hours Wave 1	51.7	51.9	49.5	49.7	49.9
(Std Deviation)	(12.7)	(9.4)	(12.4)	(11.2)	(13.1)
N	161	102	75	31	343
Avg Salary Wave 1	118,669	124,887	94,081	144,533	134,503
(Std Deviation)	(50,008)	(49,496)	(40,660)	(48,442)	(71,778)
N	165	106	74	32	360
Avg Hours Wave 3	51.2	54.3	51.9	46.2	47.0
(Std Deviation)	(12.8)	(12.2)	(10.7)	(12.2)	(12.8)
N	171	108	77	32	365
Avg Compensation Wave 3	240,153	226,289	134,187	215,974	190,565
(Std Deviation)	(220,199)	(111,049)	(75,675)	(136,151)	(197,825)
N	121	89	72	26	316

Notes: This table reports job characteristics for lawyers included in the After the JD Survey (AJD), which was sponsored by the American Bar Association. This panel survey involved three rounds of data collection that began in 2002, 2007, and 2012. The first number in each cell is the sample mean, the numbers in parentheses are standard deviations, and N is the sample size. The sample for this table includes only attorneys who passed the bar in 1998 or later and were associates at private law firms in the first wave of the survey. Average Hours Wave 1 and Average Hours Wave 3 are average hours worked by attorneys in the weeks before completing the first and third surveys, respectively. Average Salary Wave 1 is defined as the salary including bonus. Average Compensation Wave 3 is the average sum of salary, bonus, profit sharing, and other income received by respondents in the third wave. Sample sizes differ by cell due to different frequencies of item non-response.

The *After the JD* (AJD) study conducted three rounds of interviews with a cohort of lawyers who passed the bar around 2000. These interviews took place between May, 2002 - March, 2003, July, 2007 - May, 2008, and May, 2012 - December, 2012. We chose a subsample of these lawyers who reported in wave 1 that they worked as an associate in a private law firm and who also participated in the wave 3 interviews. By wave 3, most of these lawyers should have had eleven or twelve years of experience as lawyers. Thus, they are slightly more experienced that the average lawyer in our Table 2 sample.

Each column in Table 3 describes outcomes for lawyers who were associates in wave 1 and occupy a specific position in wave 3. As in Table 2, we present results for equity partners, non-equity partners, associates, and counsels, but we also present

an "other" category. This category includes lawyers who have left private law practice or are self-employed.[27]

The first row of the Table 3 reports hours worked in wave 1. All of the lawyers in this sample were working as associates in private law firms in wave 1. Young associates who are going to make partner or non-equity partner do work slightly longer hours in wave 1. Although this difference is not statistically significant, Gicheva [2013] notes that, if yong workers differ in their costs of effort and also learn by doing, those with low costs of effort work more and are more likely to acquire skills that earn promotions. Further, since the wave 1 interview takes place in year two or three of these young lawyers' careers, the small differences in wave 1 hours worked recorded in the first row of Table 3 may indicate that partners give more work to second and third year associates who appear most likely to make partner.[28]

In wave 3, we do see more noteworthy differences in hours worked among these lawyers. The small number who took counsel positions and the much larger sample who no longer work in private law firms work about three hours less per week than they worked as new associates. In contrast, those in Associate and Non-Equity Parnter positions who are still trying to become equity partners work more than they worked in wave 1, while those who are partners in wave 3 work thirty minutes less per week than they worked in wave 1.

The contrast between associates and other lawyers in wave 3 is striking. In wave 1, when both groups were working as associates, those in the wave 3 other category enjoyed higher annual earnings and worked a few minutes more per week. Between waves 1 and 3, annual earnings grew for both groups by about 42 percent, but in wave 3, those who remain in associate positions work roughly five hours per week more than those who have left private law. The results for the small sample of attorneys who are counsels in wave 3 parallel those for attorneys in the other category, except they enjoy slightly higher earnings and earnings growth while reporting slightly larger declines in hours between wave 1 and wave 3.

Taken as a whole the comparisons between wave 3 associates and those who have left the partnership track by wave 3 provide additional support for our claim that associates work long hours as part of a screening process. Further, it is difficult to imagine an explanation for these patterns that does not involve new associates being uncertain about the output that their heavy work loads will produce.[29]

We contend that associates are uncertain about whether or not their work as associates will reveal that they are going to have lucrative careers as partners. When they learn they are not going to make partner, the information value produced by their effort is diminshed. The same is true for those who make partner, but when associates become partners, they take on new leadership and business development roles that raise the value of their work effort.

Table 4 summarizes the association between leaving the partnership track and changes in hours worked between waves 1 and 3 for several different samples. Here, we define terms as follows: within our sample of wave 1 associates in private law firms, those who are partners, non-equity partners, or associates in wave 3 are still

[27] add details

[28] Wilkins and Gulati [1998] discuss differences in work assignments among associates.

[29] In any life-cycle model where workers have perfect forsight about their future wage rates, it is hard to generate sharp drops in hours over intervals where wages are constant or growing.

on the partnership track, but those who have left private law or accepted counsel positions by wave 3 are off the partnership track. Table 4 reports the results form a series of bivariate regressions of hours worked in wave 3 minus hours worked in wave 1 on a dummy variable indicating whether or not a lawyer is off the partnership track in wave 3.

The statistical association is noteworthy. Lawyers who leave the partneship track reduce their hours by by roughly 3 to 7 hours per week relative to lawyers who remain on the partnership track. The differences among the estimated effects in the five columns are not statistically significant, but there is some indication that the effect is stronger among lawyers who began their careers in large private law firms. There is also no evidence that this result is driven purely by women seeking to reduce their hours in order to spend more time at home. The estimated effect among women is slighlty larger in the full sample but smaller in the sample of lawyers who began their careers at law firms with more than 150 attorneys.

Table 4

The Association Between Leaving the Partnership Track and Changes in Hours

Regressions of Changes in Hours (Wave 3 - Wave 1) on Off-Track in Wave 3 (Counsel, Other)					
	Full Sample	Male	Female	Male Firm Size Wave1 = 150+	Female Firm Size Wave1 = 150+
Off-Track Wave 3	-3.70	-3.26	-4.25	-7.43	-5.01
(Std Error)	(1.07)	(.80)	(1.78)	(2.75)	(2.92)
N	707	424	279	178	114

Notes: See notes to Table 3 for description of the AJD sample. Once again, we restrict our samples to persons who report in Wave 1 that they are associates in private law firms. We also eliminate respondents who report in Wave 3 that they are self-employed and five respondents who report working in staff or contract positions. Those who are Equity Partners, Non-equity Partners, or Associates in a private law firm are On-Track. Those who are solo practioners, counsels in a private law firm or employess of an organization that is not a private law firm are Off-Track. The entries here are regression coefficients on a dummy variable indicating that, in Wave 3, a lawyer is Off-Track. All regressions are bivariate regressions of changes in hours worked per week, i.e. wave 3 hours minus wave 1 hours, on the Off-Track indicator. Standard errors are in parantheses. N denotes sample size.

8. Conclusion

The literature on up-or-out rules and the literature on long work hours among young professionals often reference the same labor markets but, to date, economists have sought to understand these features of professional labor markets as separate phenomena. We argue that both practices serve the same purpose. Professional service firmsprosper when they identify talented partners who can attract and maintain valuable clients. Thus, the work in professional service firms is performed primarily by partners or by new associates who are participating in strenuous auditions for partner positions.

In recent decades, many law firms and public accounting firms have moved away from strict adherence to up-or-out, but the majority of new professionals in law, public accounting, consulting, and investment banking firms still begin their careers expecting to either progress to partner or leave the firm during the next 10 to 15 years. Our reading of trade publications and our interviews with professionals from these industries indicate that less than 15 percent of those who enter elite professional service firms seeking to become partners actually make partner.[30] This promotion rate is difficult to square with the idea that young professionals work long hours to signal private information about their cost of effort. We argue that young professionals take on heavy work loads so that both they and their employers can learn more quickly whether or not they should be partners, and they do so as part of up-or-out promotions contests because these firms adopt personnel policies that facilitate their searches for talented partners.

In future work, we hope to explore how developments in information technology have shaped the productivities of both partners and experienced specialists in professional service firms. We want to examine how these developments have shaped the work lives of new professionals and the personnel policies they face in different industries.

[30]The notes to Table 3 explain that roughly one fifth of our AJD sample achieved the rank of equity partner before the wave three data collection. These results do not offer a precise point of comparison because many of the AJD firms are not elite firms. Further, the follow-up response rate in the AJD is less than 50%, and we do not know if those who made partner are more or less likely to participate in follow-up surveys, although we conjecture that persons who remained with the same firm are easier to locate than those who switched employers. Finally, some of the persons who report being equity partners in wave three may have left their original firms to accept positions at less prestigious firms.

References

George Akerlof. The economics of caste and of the rat race and other woeful tales. *The Quarterly Journal of Economics*, 90(4):599–617, 1976.

Elizabeth Dreike Almer. *AICPA Work/Life and Women's Initiative 2004 Research: Executive Summary*, 2004.

Charles Batchelor. Up or out is part of industry culture. *Financial Times*, April 20, 2011.

Alessandro Bonatii and Johaness Horner. Career concerns with exponential learning. *Theoretical Economics*, page forthcoming, 2016.

Dominique Demougin and Aloysius Siow. Careers in ongoing hierarchies. *The American Economic Review*, 84(5):1261–1277, 1994.

Marc Galanter and Thomas Palay. *Tournament of Lawyers: The Transformation of the Big Law Firm*. The University of Chicago Press, Chicago, 1991.

Marc S Galanter and William D. Henderson. The elastic tournament: The second transformation of the big law firm. *Stanford Law Review*, 60(1867):1867–1930, 2008.

Suman Ghosh and Michael Waldman. Standard promotion practices versus up-or-out contracts. *The RAND Journal of Economics*, 41(2):201–325, 2010.

Dora Gicheva. Working long hours and early career outcomes in the high-end labor market. *Journal of Labor Economics*, 31(4):785–824, 2013.

Ronald J. Gilson and Robert H. Mnookin. Sharing among the human capitalists: An economic inquiry into the corporate law firm and how partners split profits. *Stanford Law Review*, 37(2):313–392, 1985.

Ronald J. Gilson and Robert H. Mnookin. Coming of age in a corporate law firm: The economics of associate career patterns. *Stanford Law Review*, 41(3):567–595, 1989.

Elizabeth H. Gorman. Moving away from up-or-out: Determinants of permanent employment in law firms. *Law and Society Review*, 33(3):637–666, 1999.

Mona Harrington and Helen Hsi. Women lawyers and obstacles to leadership. MIT Workplace Center, 2007.

Bengt Holmstrom. Managerial incentive problems: A dynamic perspective. *The Review of Economic Studies*, 66(1):169 – 182, 1999.

Charles Kahn and Gur Huberman. Two-sided uncertainty and up-or-out contracts. *Journal of Labor Economics*, 6(4):423–444, 1988.

Renee M. Landers, James B. Rebitzer, and Lowell J. Taylor. Rat race redux: Adverse selection in the determination of work hours in law firms. *The American Economic Review*, 86(3):329–348, 1996.

Jonathan Levin and Steven Tadelis. Profit sharing and the role of professional partnerships. *Quarterly Journal of Economics*, pages 131–171, 2005.

Brendan OFlaherty and Aloysius Siow. On the job screening, up or out rules, and firm growth. *The Canadian Journal of Economics*, 25(2):346–368, 1992.

Canice Prendergast. The role of promotion in inducing specific human capital acquisition. *The Quarterly Journal of Economics*, 108(2):523–534, 1993.

Machael J. Pries. Persistence of employment fluctuations: A model of recurring job loss. *The Review of Economics Studies*, 71(1):193–215, 2004.

Douglas Richmond. The partnership paradigm and law firm non-equity partners. *Kansas Law Review*, 58:507 – 551, 2010.

Sherwin Rosen. Learning and experience in the labor market. *Journal of Human Resources*, 7(3):326 – 342, 1972.

Michael Waldman. Up-or-out contracts: A signaling perspective. *Journal of Labor Economics*, 8(2):230–250, 1990.

David B. Wilkins and G. Mitu Gulati. Reconceiving the tournament of lawyers: Tracking, seeding, and information control in the internal labor markets of elite law firms. *Virginia Law Review*, 84:1581–1681, 1998.