

Ferrall, Christopher

**Working Paper**

## Promotion & Incentives in Partnerships: Theory & Evidence

Queen's Economics Department Working Paper, No. 808

**Provided in Cooperation with:**

Queen's University, Department of Economics (QED)

*Suggested Citation:* Ferrall, Christopher (1991) : Promotion & Incentives in Partnerships: Theory & Evidence, Queen's Economics Department Working Paper, No. 808, Queen's University, Department of Economics, Kingston (Ontario)

This Version is available at:

<https://hdl.handle.net/10419/189132>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



Queen's Economics Department Working Paper No. 808

# Promotion & Incentives in Partnerships: Theory & Evidence

Christopher Ferrall

Department of Economics  
Queen's University  
94 University Avenue  
Kingston, Ontario, Canada  
K7L 3N6

3-1991

**DISCUSSION PAPER #808**

**PROMOTION & INCENTIVES IN PARTNERSHIPS:  
THEORY & EVIDENCE**

**by**

**Christopher Ferrall**

**Queen's University**



## PROMOTION AND INCENTIVES IN PARTNERSHIPS:

### THEORY AND EVIDENCE

Christopher Ferrall\*  
Department of Economics  
Queen's University  
Kingston, Ont. K7L 3N6  
Canada

March 1991

I analyze the use of promotions as an incentive and screening device in professional partnerships. Partners make production decisions and share in profits. Incentives are modeled as a multi-agent tournament. Associates provide effort to the firm by competing for promotions. Promotions also screen associates by selecting the most skilled to be partners. Either or both of these aspects of promotions lead to endogenous long run growth of firms. Competition to hire associates leads firms to offer equal expected utility to incoming workers. This constraint is estimated using cross-sectional data on major U.S. law firms. Tournament effects explain the data significantly better than a pure screening model. A uniform distribution of underlying skill is chosen by the data. In this case incentives alone can not generate interior solutions for a firm's growth rate. Screening and incentives jointly explain firm growth and inter-firm variation in compensation.

---

\* This paper is a much revised adaptation of portions of my Ph.d. thesis. I would like to thank Ken Wolpin for his support and encouragement. Avner Ben-Ner, Dan Bernhardt, Michael Gibbs, Hide Ichimura, Taka Ito, Ed Prescott, Mark Rosenzweig and Scott Thompson have also made valuable comments. Comments are appreciated on earlier versions from seminar participants at British Columbia, Buffalo, Minnesota, Penn State, Purdue, Queen's, San Diego, Toronto and the *Econometrica* Workshop "Empirical Applications of Structural Models" at Madison. Remaining errors are my own.



## I. INTRODUCTION

The process of promotion from associate to partner is an integral component of law firm behavior. Legal trade journals discuss the pros and cons of different promotion methods. The historical promotion rates within firms are reported to law school graduates in a survey of employers. In both economics and law academics have denoted much effort to explain up-or-out promotions. The simple structure of law firms and their reliance on intense skilled work make them an important laboratory for studying promotion rules. This is borne out by the application of screening, matching, and incentive theories of promotion to law firm data (Spurr (1987), Gifford and Kenney (1986) and O'Flaherty and Siow (1990)).

Three observations on law firm dynamics stand out.<sup>1</sup> First, the use of up-or-out promotion rules has been ubiquitous, although adoption of other policies has become more frequent of late.<sup>2</sup> Second, law firms experience persistent and rapid growth.<sup>3</sup> While demand considerations may explain part of this trend, factors internal to firms appear to be important as well. Finally, firms vary greatly in size, promotion rates and profits. Data presented here show that these variables are correlated across firms and that variation is large even among top firms.

This paper explores the equilibrium consequences of a joint model of incentives and screening within partnerships. Incentives are modeled as a multi-agent tournament among associates that compete for partner.<sup>4</sup> The tournament model has been used to link promotions and the structure of wages within firms (e.g. Malcomson (1984), Rosen (1986), and Gibbs (1989)). This paper extends the tournament model by embedding it in a model of firm dynamics. Partners make

---

<sup>1</sup>Most of the stylized facts and trends concerning law firms reported here are discussed in Gilson and Mnookin (1985,1989) and Galanter and Palay (1990). Issues of sharing rules and incentive problems for partners themselves are not addressed here. For discussions of partnership stability and efficiency see Ben-Ner (1984), Miyazaki (1984), Farrell and Scotchmer (1988), Legros (1989), Gilson and Mnookin (1985) and Pauly and Gaynor (1990).

<sup>2</sup>Several papers have shown that up or out contracts can be optimal in incomplete and asymmetric information environments, e.g., Carmichael (1983), Kahn and Huberman (1988), O'Flaherty and Siow (1988), Waldman (1990) and Bernhardt (1990).

<sup>3</sup>Galanter and Palay (1991) reports growth dating back to the 1920s for a sample of firms. See also Leibowitz and Tollison (1980).

<sup>4</sup>The tournament literature begins with Lazear and Rosen (1981).

sequential decisions. They take as given the effort generated by tournament incentives and the market reservation utility level of associates. The firm's path is subgame perfect: associates and partners anticipate that optimal decisions will be made in the future. The optimal hierarchy and growth rate are determined endogenously.

In addition to motivating associates, promotions screen ability by promoting those who reveal better ability to perform as partners.<sup>5</sup> Screening takes place during a fixed period spent as an associate. The screening models presented in Prescott and Vischer (1980) and O'Flaherty and Slow (1989) provide similar explanations for firm growth. Firm specific skill accumulates and interacts with associate effort. Prescott and Boyd (1987) also discusses firm growth through a similar mechanism without turnover and asymmetric information.

Two main conclusions are offered. First, endogenous long run growth can be generated in either a screening or a tournament environment. Higher promotion rates are preferred by associates, *ceteris paribus*, and thus lower the wage required to attract them to the firm. Higher promotion rates, however, lower incentives for effort and lower average skill of promoted partners. Combined with constant returns to scale in production, optimal promotion policies generate growth proportional to current firm size. Average effort of associates and average skill of partners reach stationary levels and the growth rate becomes constant.

The second conclusion is based on cross-sectional data from major U.S. law firms. Although either screening or incentives is consistent with steady growth, the data suggest that both are necessary to explain cross-sectional variation as an equilibrium outcome. A pure screening equilibrium equates expected income of associates across firms. Tournament competition generates unobserved cost of effort as a function of observed promotion rates and per-partner profits. These differences are statistically significant and consistent with other restrictions implied by the tournament model.

Alone this empirical result would suggest that tournament effects are sufficient to explain both growth and cross-firm variation. The data, however, are best explained by uniformly distributed monitoring errors, which are skill dif-

---

<sup>5</sup>Although not discussed formally, two features of this environment make up or out contracts optimal. First, younger workers will defer more compensation to the future than older workers because they live one more period for certain. Second, skill is unimportant in the job of associate.



ferences when screening occurs. The uniform distribution precludes interior solutions to the firm's decision problem in a pure tournament environment because promotion rates do not affect the Nash equilibrium level of effort. Only if promotions also determine the skill level of partners in the future can a corner solution to growth be avoided. In this sense, then, screening and incentives are necessary to explain simultaneously the stylized facts of steady growth and the "hard" facts of cross sectional variation.<sup>6</sup>

The next section develops the effort function of associates in a tournament environment. Section III tests whether implied differences across firms can be explained without tournament effort. Section IV presents the models of optimal firm decisions with screening and incentives. Some properties of both environments are established, although the screening model is easier to analyze. Simulations demonstrate existence in the tournament case directly. They also show the effect of asymmetric information on firm behavior. The qualitative effect, in terms of levels of choices and signs of elasticities, appears small. This suggests that only the use of the exact equilibrium conditions may distinguish theories of promotion and compensation. Patterns within firms and simple correlations across firms may be consistent with either symmetric or asymmetric information. Section V concludes.

## II. PROMOTIONS AND INCENTIVES

A partnership firm is modeled as a fixed technology used by a sequence of overlapping generations. Workers have an infinite horizon and after their second period they remain in the firm for one more period with constant probability  $\gamma \in [0,1)$ . In their first period, workers are hired by a firm as associates. In the second period some associates are promoted to partner. Partners make decisions that maximize the expected per-partner value of the firm at each point in time. They act sequentially because the entry and exit of partners inhibits commitment to policies that are not subgame perfect. Partners choose a fixed wage to pay associates in a period. They also choose the number of associates to hire and the promotion rules in each period.

Associates take as given the choices made by current partners. They choose

---

<sup>6</sup>Even with screening, corner solutions to growth may occur. Adjustment costs and other constraints on growth which have been emphasized may be important as well. The point is that these additional factors are necessary to avoid unbounded returns or corner solutions with a uniform skill distribution and incentive effects alone. They are not sufficient to rule out unbounded or decreasing changes in firm size.

a symmetric Nash equilibrium effort level which determines the probability of making partner. Within this structure current effort depends on the expected future performance of the firm. An associate at time  $t$  has an expected utility function of the form:

$$E_t U = w_t - c(x_t) + \beta E_t I_{t+1},$$

where  $w_t$  is the wage received as an associate,  $x_t$  is the chosen effort level,  $\beta < 1$  is a discount factor, and  $I_{t+1}$  is the value of discounted expected income from time  $t+1$  on. Effort  $x_t$  is non-negative. Effort is supplied inelastically and costlessly after the first period of life. Associates have a reservation expected utility level  $\bar{U}$ , which constrains the firm's promotion and wage choices.<sup>7</sup>

The cost function satisfies

- A1 [1]  $c$  is convex and  $c''$  is continuous  
 [2]  $c(0) = 0$   
 [3]  $c'(0) = 0$ .

$I_{t+1}$  depends on whether the lawyer makes partner or not. If offered partnership,  $I_{t+1} = V$  where  $V$  represents per-partner value of the firm as of time  $t+1$ .  $V$  is a function of the state of the firm at time  $t+1$ . The state is determined by current partners so associates take  $V$  parametrically when making their effort decisions. If not promoted,  $I_{t+1} = R$ , where  $R > 0$ .  $R$  represents discounted expected returns in a secondary job. To motivate effort a firm must have  $V > R$ .

Let a worker's promotion probability at time  $t$  be  $P_t$ , whose arguments will be specified later. The objective of associates is

$$\max_{x_t \geq 0} E_t U = w_t - c(x_t) + \beta \left( P_t V + (1-P_t)R \right) \quad (1)$$

An interior solution,  $x^*$ , satisfies the first order necessary condition (subscripts are dropped for convenience):

$$c'(x^*) = \beta \frac{\partial P}{\partial x} (V-R), \quad (2)$$

and the second order necessary condition:

---

<sup>7</sup>Law firms hire every year so these values are yearly flows. However, the period spent working for a wage is roughly 8 years for law firms (from the National Association for Law Placement). Implicitly individuals and firms smooth effort and wages over the course of the associate period. With  $\beta$  appropriate for the time before the promotion decision all values can then be expressed in terms of yearly flows. Information on wage increases during the associate period are not available, but the assumption that it does not depend on performance appears to be accurate (see previous sources).

$$c''(x^*) \geq \beta \frac{\partial^2 P}{(\partial x)^2} (V-R). \quad (2')$$

The second order condition is required because this problem is generally quasi-concave rather than concave.

Individual effort is monitored imperfectly. Let effort of each associate  $i$ ,  $i$  going from 1 to  $a_t$ , be measured by the firm as  $q_t^i$ , in the form

$$q_t^i = x_t^i + \varepsilon_t^i. \quad (3)$$

The error  $\varepsilon_t^i$  can be interpreted as the associate's skill which is unknown to all agents when effort is chosen.<sup>8</sup> It satisfies

- A2
- [1]  $\varepsilon_t^i \sim F(\varepsilon)$  i.i.d across  $i$  and  $t$  with density  $f(\varepsilon)$ .
  - [2]  $f'(\varepsilon)$  exists and  $f'(\varepsilon) = -f'(-\varepsilon)$  and  $f'(\varepsilon) \geq 0 \forall \varepsilon < 0$ .
  - [3]  $\sum_{i=1}^{a_t} \varepsilon_t^i \approx \bar{\varepsilon} = 0$ .

The second assumption implies  $\varepsilon$  is distributed symmetrically around a mean and unique mode of 0. Risk in production is assumed unimportant by A2.[3], although technically A2.[1] and A2.[3] are incompatible. Skill may affect productivity of an associate, but for convenience the values for an unselected cohort of associates cancel out each period.

Asymmetry across individuals exists if each knows his own skill level or if  $\varepsilon_t^i$  is not identically distributed across  $i$  for a given  $t$ . In either case solving for Nash equilibrium tournament effort decisions is much more difficult than in the symmetric case.<sup>9</sup> The other assumptions in A2.[1], independent distributions across  $i$  and identical distributions across  $t$ , are made for convenience. Tournament contracts efficiently shelter risk averse agents from correlation in  $\varepsilon_t^i$ . In a sense, A2.[1] assumes away the commonly held rationale for tournaments. This paper, however, focuses upon the interaction between incentives and dynamics, rather than between incentives and risk management.

Whether the skill of associates retained as partners enters the firm's revenue determines whether promotions act as screens. Production will be discussed in section IV. In the mean time note that the firm is at least indifferent to the set off associates promoted. If partner skill is an input the only credible rule is one in which highest ranked associates are promoted. Assumption A2.[2]

---

<sup>8</sup>Asymmetry introduced when individuals know their own skill level makes tournament effort decisions much less tractable. See Battacharya and Guasch (1988), Lazear (1989) and Gibbs (1990).

<sup>9</sup>See Battacharya and Guasch (1988), Lazear (1989) and Gibbs (1990).

guarantees that average skill of promoted associates has a positive mean.<sup>10</sup> An assumption similar to A2. [3] will be made to eliminate uncertainty in  $\varepsilon_t^1$  conditional on making partner.

Given F and (4) we can describe the probability of making partner under different promotion rules. Partners create a tournament environment when they choose the number of promotions, denoted  $\eta_t$ , before measuring effort. Another simple rule is an individual promotion standard. If  $q$  is private to the firm such rules may not be enforceable directly, although, as Malcomson (1984) points out, a tournament among a large number of workers converges to a promotion standard. The effort function in a tournament is easier to derive and underscores that small firms can extract more information from a larger group of workers (see Green and Stokey (1983)).

Begin with an associate who faces a common effort level  $\tilde{x}$  chosen by his competitors. Let  $\varepsilon^{(j)}$  stand for the  $j$ th-order statistic of the other  $a_t - 1$  associates. ( $\varepsilon^{(1)}$  is the lowest ranked  $\varepsilon_t^1$ ). If a person chooses effort  $x$  and receives  $\varepsilon$ , the probability of beating exactly  $k$  other associates in the ranking of  $q$ 's is

$$\begin{aligned} P(x + \varepsilon > \tilde{x} + \varepsilon^{(k)}, x + \varepsilon < \tilde{x} + \varepsilon^{(k+1)}) &= \\ P(\varepsilon^{(k)} > \varepsilon + x - \tilde{x}, \varepsilon^{(k+1)} < \varepsilon + x - \tilde{x}) &= \\ P(\varepsilon^{(k)} < \varepsilon + x - \tilde{x}) &= \\ \binom{a_t - 1}{k} F(\varepsilon + x - \tilde{x}) [1 - F(\varepsilon + x - \tilde{x})]^{a_t - 1 - k} \end{aligned}$$

where  $\binom{a_t - 1}{k}$  is the binomial coefficient. For a given  $\varepsilon$ , the probability of making partner is the sum of the probabilities of being exactly one of the top  $\eta_t$  ranked  $q$ 's. That is, beating  $a_t - 1$  others or  $a_t - 2$  or . . .  $a_t - \eta_t$  will in each case imply partnership. Integration over  $\varepsilon$  gives the probability of promotion:<sup>11</sup>

$$P(x, \tilde{x}) = \int \left\{ \sum_{k=a_t - \eta_t}^{a_t - 1} \binom{a_t - 1}{k} F(\varepsilon + x - \tilde{x}) [1 - F(\varepsilon + x - \tilde{x})]^{a_t - 1 - k} \right\} f(\varepsilon) d\varepsilon. \quad (4)$$

Differentiating  $P$  in (4) with respect to  $x$  results in:

<sup>10</sup>The symmetry assumption A3. [2] implies that  $E(\varepsilon_t^1 | \varepsilon_t^1 \geq \varepsilon^{(j)}) \geq 0$  for any order statistic  $\varepsilon^{(j)}$ . In the extreme case everyone is promoted, so  $E(\varepsilon_t^1) = 0$ .

<sup>11</sup>This probability of promotion function is also derived in Gibbs (1989).

$$\frac{\partial P(x, \tilde{x})}{\partial x} = (a_t - 1) \binom{a_t - 2}{a_t - \eta_t - 1} * \int F^{a_t - \eta_t - 1} (\varepsilon + x - \tilde{x}) [1 - F(\varepsilon + x - \tilde{x})]^{\eta_t - 1} f(\varepsilon + x - \tilde{x}) f(\varepsilon) d\varepsilon. \quad (5)$$

This is well-defined only for  $a_t \geq 2$  and  $1 \leq \eta_t \leq a_t - 1$ . When  $\eta_t$  is 0 or  $a_t$  then  $P$  is 0 or 1, respectively, so the derivative is zero. If  $a_t = 1$  then  $\eta_t$  must be 0 or  $a_t$ . All terms inside the integral are non-negative, so  $\frac{\partial P}{\partial x}$  is non-negative, as accords with intuition: more effort cannot lower the probability of high rank. (Also note that  $\frac{\partial P}{\partial x}$  is differentiable even if  $f'$  does not exist.)

$A_1$  and  $\frac{\partial P}{\partial x}$  non-negative ensure an interior solution because at  $x=0$  a small amount of extra effort yields higher expected income at no marginal cost. Thus the conditions in (2) are necessary for any solution.<sup>12</sup> Sufficiency of (2) for a global maximum is not guaranteed because  $\frac{\partial^2 P}{\partial x^2}$  can be positive for low values of  $x$ . Thus, there may be multiple solutions to the first order condition, although one solution is the global maximum.

A symmetric Nash solution among associates at a given  $t$ , is an effort level  $x^*$  such that Eq (2) - (2') hold and  $\partial P$  is evaluated at  $x^* = \tilde{x}$ . In this case, (6) collapses to:

$$\frac{\partial P(x^*)}{\partial x^*} = \int (a-1) \binom{a-2}{a-\eta-1} F^{a-\eta-1}(\varepsilon) [1-F(\varepsilon)]^{\eta-1} f^2(\varepsilon) d\varepsilon. \quad (6)$$

The second derivative of the probability an individual faces at a Nash equilibrium reduces to

$$\frac{\partial^2 P(x^*)}{\partial x^2} = (a-1) \binom{a-2}{a-\eta-1} * \int F^{a-\eta-2} [1-F]^{\eta-2} \left[ F[1-F]f' + f^2[a-\eta-1+(2-a)F] \right] f d\varepsilon. \quad (6')$$

In the above expressions the implicit choice of others and time subscripts have been dropped. The portion of these derivatives that depends on integer promotions and associates is the binomial coefficient. The integer constraint will generally be ignored, but some results concerning the shape of the  $P$  function are easier to show in the exact integer case.

Let  $B^{-1}(a-\eta, \eta) = (a-1) \binom{a-2}{a-\eta-1}$ , where  $B$  is the beta-function. With the

<sup>12</sup>This is true for all  $\tilde{x}$  when  $\varepsilon$  has infinite support. If the support is finite, then the marginal benefit of zero effort is positive only if  $\tilde{x}$  is less than 2 times the range of  $\varepsilon$ .

transformation  $y = F(\varepsilon)$ , (7) can be written as

$$\frac{\partial P(x^*)}{\partial x^*} = \int_0^1 B^{-1}(a-\eta, \eta) y^{a-\eta-1} [1-y]^{\eta-1} f(F^{-1}(y)) dy. \quad (7)$$

*Proposition I.* With assumptions A1 and A2 and  $a_t$  and  $\eta_{t+1}$  non-negative integers such that  $1 \leq \eta_t \leq a_t$ , then

[1].  $x^*$  is unique for a given  $(a_t, \eta_t)$ .

[2].  $P(x^*) = \frac{\eta_t}{a_t}$ .

[3].  $\left. \frac{\partial P(x^*)}{\partial x^*} \right|_{\eta_t=\eta} = \left. \frac{\partial P(x^*)}{\partial x^*} \right|_{\eta_t=a_t-\eta}$ .

[4].  $\left. \frac{\partial^2 P(x^*)}{(\partial x^*)^2} \right|_{\eta_t=\eta} \begin{matrix} \geq \\ < \end{matrix} 0 \text{ as } \eta \begin{matrix} \leq \\ > \end{matrix} a_t/2$

*Proof:* Appendix.

The fact that at equilibrium the probability of making partner reduces to the number of openings over the number of associates is clear (result I. [2]). Result I. [3] says that the derivative depends on  $\eta_t$ 's distance from  $a_t/2$ . I. [3] and I. [4] imply that  $\frac{\partial P}{\partial x}$  is maximized at  $\eta_t = a_t/2$ . A firm that wants to maximize effort for a given  $a_t$  would promote 1/2 of its associates. Promotions affect the future of the firm and any cost of effort must be offset by other forms of compensation (wages or future income). Therefore, a promotion rate of 1/2 is generally not profit-maximizing.

Result I. [4] also implies that a solution to the first order conditions is guaranteed to be a local maximum when a firm promotes more than half the associates. If it hires less than half, a local maximum occurs only if  $c$  is more convex at the Nash equilibrium effort level than the right hand side of (2').

The following distributions are useful because they yield closed form solutions for the Nash equilibrium value of  $\frac{\partial P}{\partial x}$ .

#### Example 1

Let  $\varepsilon$  be distributed uniformly in  $[-b/2, b/2]$  for some parameter  $b$ . Then  $f(F^{-1}(x)) = \frac{1}{b}$  for all  $x$  in  $[0, 1]$ . The integral in (8) reduces to  $B(a-\eta, \eta)/b$ . After canceling  $B(a-\eta, \eta)$  and  $B^{-1}(a-\eta, \eta)$ ,

$$\frac{\partial P(x^*)}{\partial x^*} = \frac{1}{b} \text{ for all } a > n > 0. \quad (8)$$

In the case of a uniform monitoring error, the promotion rate  $\eta / a$  does not

affect effort: only profits and the monitoring parameter  $b$  affect the Nash equilibrium effort level. Promotions still affect effort indirectly through future profits. It also can be shown from (8') that

$$\frac{\partial^2 P(x^*)}{\partial x^2} = 0. \quad (8')$$

That is, the first order conditions are sufficient to describe an interior solution.

### Example 2

A mean-zero logistic distribution implies  $f(F^{-1}(x)) = sx(1-x)$  for the precision parameter  $s$ . That is, with

$$\begin{aligned} F(\epsilon) &= \frac{e^{s\epsilon}}{1 + e^{s\epsilon}} \\ 1 - F(\epsilon) &= \frac{1}{1 + e^{s\epsilon}} \\ f(\epsilon) &= sF(\epsilon)(1-F(\epsilon)), \end{aligned} \quad (9)$$

the integral in (8) is  $sB(a-\eta+1, \eta+1)$ . Canceling this with  $B^{-1}(a-\eta, \eta)$  reduces to

$$\frac{\partial P(x^*)}{\partial x^*} = \frac{s\eta(a-\eta)}{a(a+1)} \quad (10)$$

This expression is not a function only of  $\lambda = \eta/a$ , the promotion rate.

However, notice that for large  $a$ , (10) converges to  $s\lambda(1-\lambda)$ . For large firms incentives are not affected by firm size. In this example the symmetry of  $\frac{\partial P}{\partial x}$  around  $a-\eta$  is non-trivial. The value of  $\partial P$  is bounded by  $f(0) = s/4$  (see Gibbs (1989)). The second derivative of  $P$  at the Nash equilibrium is

$$\frac{\partial^2 P(x^*)}{\partial x^2} = \frac{2s^3 \eta(a-2\eta)(a-\eta)(a-\eta+1)(\eta+1)}{a(a+1)(a+2)(a+3)(a+4)} \quad (10')$$

## III. A SHRED OF EMPIRICAL EVIDENCE

This section estimates  $R$ ,  $f(\epsilon)$  and  $c(x)$  from cross-sectional data on U.S. law firms by imposing the restriction that firms offer equal expected utility to incoming workers. The procedure is quite simple, but demonstrates the empirical content of equations (1) and (2). The tournament model nests a screening model so a test of the null hypothesis that promotions have no incentive effect is feasible.

Three assumptions about associate expectations are made. First, young lawyers form expectations about the promotion rate in a firm based on its partner to associate ratio. They take into account growth and the fact that more time

is spent as partner than associate. Second, associates expect the per partner value of firms to be constant over time. This implies that the value of partnership is equal to the present value of current per-partner profits. The next section demonstrates that constant per-partner value in a growing firm is optimal. The null hypothesis is that promotions are not incentive devices, whether they act as screens or not.<sup>13</sup>

Two important aspects of the reality of partnership firms may appear to throw doubt on the assumption that the reward to making partner is the current value of a partnership. Partners typically buy-in to the firm and are bought out when they leave. Most firms appear not to share profits evenly across cohorts of partners. Such clauses apparently address incentive and liability issues among partners which are not the focus of this paper. If, however, it is safe to assume that lawyers are not liquidity constrained then such details do not affect the lifetime utility of a partnership offer. That is, only the present value of the share in the firm matters, not the timing of compensation. It seems unlikely that law partners face serious credit market constraints.

The third assumption made is that measurement error or real variation in per-partner profit around its long term mean exists. This error is i.i.d normal across firms and time. This reasonable assumption is necessary: otherwise the only random feature of the model is which associates are promoted. The model would have to fit the data perfectly.

### Data

Cross-sectional data on U. S. law firms are available on various details of firms. One survey (conducted by *American Lawyer*) reports profits for 100 major firms. One would expect self-reported profits for one year to be open to considerable measurement error and real variation. This justifies inclusion of a mean zero random component in profits. However, it should be noted *American Lawyer* uses outside sources close to each firm to corroborate the responses. Data from the fourth such survey is used because earlier surveys had smaller sample sizes. See *American Lawyer* for details of the methodology.

Other surveys gather various, less sensitive, pieces of information about U.S. law firms. *Of Counsel* conducts a survey of the 500 largest firms in the country. Its values for numbers of lawyers and associate salaries are used in

---

<sup>13</sup>Given the lack of a long time series on profits (law firms are not publicly traded) the assumption of steady per-partner value remains untested.



the estimation, because it also reports the number of new hires and promotions in the last year. A regression of promotion and hiring flows in one year on the partner/associate ratio shows a strong linear relationship. Total promotions expected for all associates at work in a year was almost exactly 25% of the partner/associate ratio.<sup>14</sup> This is used as the correction factor for the promotion rate.

Tables 1 and 2 present summary statistics from the data. Of note is the large variation in all variables except starting salaries. As a percentage of means, the standard deviation of profits is 60%, associates 51%, promotion rates 33%, and partners 43%. For salaries it is 14%. The distribution of all the variables are skewed positively, but it is largest for profits and associates: median profits are \$287,000 while the mean is \$364,000; for associates the median is 163 and the mean is 191.

Table 2 shows that the variables in Table 1 exhibit strong and fairly complicated correlation patterns. Per-partner profits and promotion rates are negatively correlated, which is consistent with any equilibrium in which the probability of a promotion is the promotion rate. However, salary differentials and implied levels of unobservables, such as effort or ability, enter into any equilibrium condition. Also note that firms with more partners are less profitable (per lawyer) and offer higher promotion rates.

---

<sup>14</sup>That is, let  $\hat{\lambda}$  be the flow of promotions in a year divided by the number of associates hired in a year. Suppose that  $\hat{\eta}_i = \hat{\lambda} a_i$  measures the "true" number of promotions expected from all associates this period in firm  $i$  with mean zero error  $v_i$ . This error is due to year-to-year fluctuations which do not affect firm's growth rate. Posit that total promotions from a group of associates in a given year is linearly related to the number of partners, so that

$$\hat{\eta}_i = \gamma_0 + \gamma_1 n_i + v_i.$$

The table below presents OLS estimates of this equation. The strength of the relationship and the fact that  $\gamma_0 = 0$  can not be rejected support the hypothesis of equal growth rates across firms. The value  $\hat{\eta}_i$  is gross promotions and  $(\gamma_1 - \gamma)$  is the promotion rate net of the exit rate  $\gamma$ . Correcting for heteroscedasticity did not alter the estimates significantly.

	Estimate	t-ratio
$\gamma_0$	0.838	.22
$\gamma_1$	0.249	8.06
$R^2 = .45$		

### Estimation of the Equal Utility Restriction

Observed per-partner profits take the form

$$\bar{\pi} = E \bar{\pi} + \nu.$$

where  $E \bar{\pi}$  stands for expected profits each period. The term  $\nu$  is normally distributed with mean 0 and standard deviation  $\sigma_\nu$ . This term can have the interpretation of either measurement error or an additive real shock to per-partner profits. The value of  $E \bar{\pi}$  and other choices depend on the technology of the firm. Firms vary in their technology endowment, but hire identical workers.<sup>15</sup>

If per-partner profits are stationary across time, then the value of joining the firm is

$$V = \frac{1}{1-\beta\gamma} E \bar{\pi} = \frac{1}{1-\beta\gamma} (\bar{\pi} - \nu).$$

The utility restriction across firms can be now be expressed as

$$\bar{U} = w - c(x) + \beta\lambda(V - R) + \beta R \quad (11)$$

In this equation  $\lambda$  stands for the promotion rate in the firm.  $\bar{U}$ ,  $\beta$ , and  $R$  are constant across firms. The cost of effort  $c(x)$ , is determined in the tournament environment as

$$c = (c')^{-1} \left( \beta \frac{\partial P}{\partial x} (V - R) \right).$$

The following forms for the error in monitoring effort and the cost function are posited:

$$\begin{aligned} (1) \quad c(x) &= e^{\xi x} - 1 & \xi > 0 \\ (2) \quad f(\epsilon) &= 1/b \text{ for } \epsilon \in [-b/2, b/2] & b > 0 \\ (3) \quad \bar{U} &= \beta R + R(1-\beta\gamma) \end{aligned} \quad (12)$$

The first two equations in (12) yield the indirect cost function

$$c = \frac{\beta}{\xi b} (V - R) - 1. \quad (13)$$

Assumption A1.[3] is not satisfied, because  $c'(0) = 1$ . The marginal net prize to each firm's contest,  $\frac{\beta}{\xi b}(V-R)$ , must be greater than 1, a testable restriction. Exponential cost yields a linear indirect cost function which identifies a unique random component  $\nu$ . Other convex cost functions result in

---

<sup>15</sup>Ferrall (1990) uses the restrictions from optimal decisions in a two-period model to estimate underlying firm specific parameters. This procedure is not repeated here because the results were not robust to the longer decision horizon.

a non-linear equation that may have multiple solutions for  $v$ .

The psychic cost of effort parameter  $\xi$  and the parameter of the skill distribution  $b$  are assumed constant across firms. More precisely, firms are assumed to be hire from the same distribution of associate skill. This assumption is clearly violated across all law firms. The critical assumption is that the firms in the *American Lawyer* survey can all attract the best associates conditional on law school records. Clearly  $\xi$  and  $b$  are not separately identified in (13). Define  $z \equiv 1/(b\xi)$ .

The third equation in (23) says that an associate's alternative at the time of entry is to enter a non-partnership firm where a spot wage is paid and no extraordinary effort is required. There is no wage growth in this sector, so the present value of the alternative if not promoted after the first period is still  $R$ . The term  $R(1-\beta\gamma)$  is the implied one period value of the alternative given up to become an associate in a firm. This is a simple equilibrium restriction on the parameters when law firms must compete with ordinary firms for law school graduates.

Using this assumption (11) can now be written as

$$R(1-\beta\gamma) = w - \beta z(V - R) + 1 + \beta\lambda(V - R). \quad (14)$$

Before estimating (14) the following flexible form for  $\frac{\partial P}{\partial x}$  was estimated:

$$\frac{\partial P}{\partial x} = \alpha_0 \left( \lambda(1-\lambda) \right)^{\alpha_1}$$

With  $\alpha_1 = 1$ ,  $f(\epsilon)$  is logistic. With  $\alpha_1 = 0$ ,  $f(\epsilon)$  is uniform. For intermediate values the properties in Proposition I still hold, but the form of  $f$  that generates  $\frac{\partial P}{\partial x}$  is not known. With this used form in (11) the likelihood of the data was maximized with  $\hat{\alpha}_1$  pushed to the boundary at 0. A uniform assumption was adopted.<sup>16</sup>

A simpler model is also contained in the form (14). If  $z = 0$  and  $\bar{U}$  is redefined to include the 1 unit constant offset, then a simple expected income restriction results:

$$R(1-\beta\gamma) = w + \beta\lambda(V - R).$$

---

<sup>16</sup>  $\hat{\alpha}_1 < 0$  does not satisfy the maintained hypothesis that unobserved differentials across firms are caused by tournament effort. There may be other systematic differences, but no known theory would predict that promotions affect utility as with  $\alpha_1 < 0$ .

This is the observable restriction of a screening model with no effort differences across firms. Firms and individuals begin without information on ability. Firms offer equal expected utility and then measure ability during the associate period during which time there are no unobservable components to effort or compensation. Firms then promote the best individuals.

Table 3 presents the results of maximum likelihood estimates of (14) for  $\beta = .95^8$  and  $\gamma = .9$ . (Recall that associates expect 8 years before partnership, hence the choice of  $\beta$ . The value of  $\gamma$  implies that 10% of the current partners will leave in the next 8 years). If these parameters are estimated then likelihood is maximized with  $\beta = 1$ , but the main effect of increasing  $\beta$  is increasing  $R$ . The estimate of  $z$  is insensitive to these parameters. The choice of  $\gamma$  does not affect any of the estimates or likelihood significantly.

The estimated value of  $R$  is reasonable. It says that the alternative to a lawyer outside these firms has a lifetime value of \$262,000, discounted by  $\beta$  and  $\gamma$ . This translates to a wage of \$105,000 each year, which is above the maximum observed wage (\$74,000) and below the lowest observed profit (\$155,000). The corresponding lifetime utility is \$280,000. The restriction that firms offer feasible tournaments reduce to three inequalities: (i)  $0 < z < \lambda$  (ii)  $R > \lambda/z + w$  and (iii)  $R < E(V)$  for each firm. Each of these is satisfied by the estimated parameters.

The estimate of  $z$  is significantly greater than 0, so the tournament model explains the data as an equilibrium across firms better than a pure screening model. A likelihood ratio test of  $z=0$  rejects the hypothesis at any conventional level of significance. As stated above, the value of  $\hat{z}$  is insensitive to  $\beta$  and  $\gamma$ . The estimated average cost of effort due to tournament effects is \$15,500 per year per associate. Of course, there may be unobserved effort differences that are not tied to promotions. For instance, firms may be able to choose an enforceable effort level directly. A firm's revenue parameters determine optimal effort, but the distribution of these parameters across firms would have to be specified to identify the equal utility restriction.

One measure of the fit of the model is the degree of profit variation that is explained by the error  $v$ . Estimated  $\sigma_v^2$  is 73% of the variance in profits (85% in terms of standard deviations). The other 27% is "explained" by variation in wages and promotion rates according to the tournament model. Figure 1 shows the relationship between promotion rates and both observed profits and estimated expected profits. One firm is a clear outlier - it reports the

highest profits and nearly the highest promotion rate.<sup>17</sup> This firm accounts for a large part of the variance in error. It is difficult to provide a more concise assessment of the model because revenue parameters that make observed choices optimal are not identified by (14).

The tournament model helps explain variation in direct and indirect compensation across top U.S. law firms. This data has been used previously in similar studies (Gifford and Kenney (1986) and Gilson and Mnookin (1989)), but the approach taken here uses exactly the equilibrium restrictions implied by the model. A viable nested alternative, that promotion is purely a screening device, can be rejected. The data also suggest that the promotion rate within a firm does not directly affect Nash equilibrium effort levels. Promotion and effort are indirectly related because firms which choose relatively high per-partner value do so partly by acquiring more able partners through lower promotion rates.

#### IV. DYNAMICS OF PARTNERSHIP FIRMS

This section presents the firm's problem with symmetric and asymmetric information on worker effort. Its main goal is to show that a constant rate of growth is an optimal policy in a partnership firm when promotions select and motivate associates. First, the technology of the firm is introduced. Profit of the firm at time  $t$  takes the form

$$\pi_t = n_t \psi \left( \frac{\epsilon_t^p}{n_t}, \frac{a_t}{n_t} \bar{x}_t \right) - w_t a_t \quad (15)$$

The function  $\psi$  represents per-partner value of output. Inputs to production are the number of partners  $n_t$ , the total skill of partners  $\epsilon_t^p$  and total effort of associates  $a_t \bar{x}_t$ . (15) assumes constant returns to scale in partners in  $n_t$ ,  $\epsilon_t^p$ , and  $a_t \bar{x}_t$ . Decreasing returns imply finite long run firm size. While this may certainly be the case, the objective here is to rationalize observed growth of many law firms in this century as a long run phenomenon. Per-partner revenue satisfies

- A3      [1]  $\psi$  is strictly concave and increasing in its two arguments  
           [2]  $\bar{U} > \beta R$   
           [3]  $\psi$  is twice continuously differentiable.

---

<sup>17</sup>The firm is Wachtell, Wachtell and Lipton. Gilson and Mnookin (1989) argues that this firm's practice is substantially different than others.

$$[4] \quad \psi(0, \cdot) = \psi(\cdot, 0) = 0.$$

Notice that numbers of partners and their skill enter total revenue separately. For example, partner skill may affect quality of output but not quality of associate supervision. Or, more partners may allow the firms to use support staff and information more efficiently. A3.[1] requires that this independent effect of the number of partners be positive. This is necessary but not sufficient to guarantee a concave objective function.

Numbers of associates, on the other hand, do not have an effect separate from effort. If associates were not just effort producers, a long run growth path would require average effort to grow with the number of partners. With convex cost of effort, this is not sustainable indefinitely. Total effort, however, can grow indefinitely simply by hiring more associates. Technologies that do not satisfy this restriction can not support long run growth even if they exhibit constant returns to scale. A3.[2] guarantees that the marginal cost of hiring each associate is positive.

Per-partner profits are defined from (3) as

$$\bar{\pi}_t \equiv \pi_t / n_t = \psi - \frac{a_t w_t}{n_t}$$

When effort is not directly observed it is determined by the incentives inherent in the promotion rate and the future behavior of the firm. In this case, partners at time  $t$  takes as given the tournament effort function  $\bar{x}_t$ . Partners in the firm at any point in time face a constant probability of exit ("death") equal to  $\gamma \in [0,1]$ . Recall that  $\eta_t$  denotes the number of promotions made at time  $t$ . Then the number of partners next period is determined by

$$n_{t+1} = \gamma n_t + \eta_t.$$

A constant probability of exit after the second period of life is a valuable simplification. The special case of two period lives is captured by  $\gamma = 0$ . Finite lives of more than two periods complicate the problem in two ways. First, voting rules among partners with different horizons must be specified. Second, under any voting rules except when the oldest partners have full control, partners deciding today must take into account how their decisions affect optimal decisions made later in their lifetimes. Even if partners of different ages have equal power a young partner today must calculate the return expected as an older partner from decisions made today. With a random risk of death all partners have the same horizon conditional on surviving until today. With  $\gamma < 1$  current partners do have shorter horizons than associates anticipating partnership because associates live for certain until the second period.

Perhaps the hardest component of the model to specify is how skill of a new partner,  $\varepsilon_t^i$ , is aggregated first with the rest of the cohort and then with the skill of current partners. Few models attempt to make explicit the process by which firm specific capital aggregates across generations.<sup>18</sup> Here aggregation of the vector of  $\varepsilon_t^i$ 's for a given  $t$  is summarized by a function  $Q$ . From the firm's point of view,  $Q$  depends on the promotion rate, because the promotion rate determines the  $\varepsilon_t^i$  required to make partner. Assume that skills within a cohort strongly interact, so that  $Q$  depends only on the promotion rate  $\frac{\eta_t}{a_t}$  and not on the number of promotions as well. That is, more partners can not substitute for lower average skill of partners. Revenue already depends on numbers of partners and  $\varepsilon_t^p$  separately, so this assumption simply maintains the conceptual separation between quality and quantity.

For simplicity, aggregation of  $Q(\frac{\eta_t}{a_t})$  across  $t$  is completely smooth. Each partner in the firm at a point in time embodies the average quality of all cohorts. In essence, the exit of any partner from the firm diminishes total skill equally. Although motivated by its convenience, this assumption is not unreasonably given both the team aspect of production and the fiduciary responsibilities of partners. In addition, this is consistent with equal sharing of profits, because the firm is technologically egalitarian. Of course, equal sharing across cohorts is not observed. As discussed in section III, the timing of compensation solves incentive problems that are immaterial to the problem existing between partners and associates.

Together, the two assumptions on skill aggregation imply

$$\varepsilon_{t+1}^p = n_{t+1} \left( \gamma \varepsilon_t^p + Q\left(\frac{\eta_t}{a_t}\right) \right) \quad (15)$$

where  $Q$  satisfies

- A4
- [1]  $Q(0) = b$ ,  $0 < b < \infty$
  - [2]  $Q(1) = 0$
  - [3]  $Q' < 0$
  - [4]  $Q'' \leq 0$ .

Assumptions contained in A4 are sufficient for  $Q$  to be concave in  $a_t$  and  $\eta_t$  jointly. A4.[3] says that higher promotions rates keep less able associates and thus lower average skill. A4.[1] simply says that a firm can not increase skill

---

<sup>18</sup> An exception is Prescott and Boyd (1987)

indefinitely by lowering its promotion rate. The depreciation of average skill at the same rate that partners exit the firm reflects smoothed skill across cohorts.

The choice of  $Q$  used in simulations is the expected value of  $\varepsilon$  conditional on being promoted. That is  $Q(\lambda) = E(\varepsilon | \varepsilon > F^{-1}(1-\lambda))$ . With a promotion rate of  $\lambda$ , the firm selects skill from the  $1-\lambda$  tail of the unscreened skill distribution. Another plausible  $Q$  is the minimum promoted skill.

### Firm Dynamics With Symmetric Information

To simplify the analysis without incentives, associate effort is assumed constant and normalized to unity with 0 cost. To make the environments comparable, simulations presented below allow the firm to choose an enforceable effort level when only screening is important.

With symmetric information, partners at time  $t$  choose the number of associates to hire, how many to promote to partner next period and the wage. Because associate effort is observed the firm chooses associates with high  $\varepsilon_t^1$  to promote. Thus  $\eta_{t+1}/a_t$  is the probability of promotion faced by associates. Recall  $\bar{U}$  is the reservation utility level of associates. At time  $t$  the firm's objective function is the (stationary) present value of per-partner profits

$$\bar{\pi}_t + \beta\gamma V[\bar{\varepsilon}_{t+1}^p].$$

This anticipates the result that per-partner value  $V$  depends on per-partner skill,  $\bar{\varepsilon}_{t+1}^p$ , but not the number of partners. Use the associates's objective (1) to rewrite the wage as a function of  $\bar{U}$ ,  $a_t$ , and  $\eta_t$ :

$$w_t = \bar{U} - \beta \left( \frac{\eta_t}{a_t} V[\bar{\varepsilon}_{t+1}^p] + \left(1 - \frac{\eta_t}{a_t}\right) R \right) \quad (16)$$

Use  $n_{t+1} - (1-\gamma)n_t = \eta_t$  to eliminate  $\eta_t$  and substitute (12) into the profit function (15) to write  $\bar{\pi}_t + \beta\gamma V[\bar{\varepsilon}_{t+1}^p]$  as

$$\frac{1}{n_t} \left[ n_t \psi \left( \frac{\varepsilon_t^p}{n_t}, \frac{a_t}{n_t} \right) - a_t (\bar{U} - \beta R) + n_{t+1} \beta \left( V[\bar{\varepsilon}_{t+1}^p] - R \right) + \gamma n_t \beta R \right]. \quad (17)$$

An important feature of (17) is that the wage decomposes into two parts. One part is the marginal cost of hiring each associate:  $\bar{U} - \beta R$ . The other part is the net return from the size of the firm next period:  $\beta \left( V[\bar{\varepsilon}_{t+1}^p] - R \right)$ . This term includes wages deferred as promised promotions as well as the direct interest current partners have in the future.

The choices of  $a_t$  and  $n_{t+1}$  enter (17) in proportion to  $n_t$  or as a ratio. Define per-partner values of  $a_t$  and  $n_{t+1}$  as  $\sigma_a$  and  $\sigma_n$ , respectively. Define as  $B$  the interval  $[0, \frac{b}{1-\gamma}]$ . With assumption A4. [1],  $\bar{\varepsilon}_t^p \in B$  implies  $\bar{\varepsilon}_{t+1}^p \in B$ , so



attention can be restricted to  $V: B \rightarrow \mathbb{R}$ . The result is the functional equation that  $V$  solves:

$$\begin{aligned} \forall \bar{\varepsilon}_t^p \in B, \\ V[\bar{\varepsilon}_t^p] &\equiv \max_{\{\sigma_a, \sigma_n\}} \psi(\bar{\varepsilon}_t^p, \sigma_a) - \sigma_a(\bar{U} - \beta R) + \sigma_n \beta \left( V[\bar{\varepsilon}_{t+1}^p] - R \right) + \gamma \beta R \quad (18) \\ &\text{subject to} \\ (i) \quad &\gamma \leq \sigma_n \leq \min\{\sigma_a + \gamma, \bar{\sigma}_n\}, \quad \bar{\sigma}_n < 1/\beta \\ (ii) \quad &\bar{\varepsilon}_{t+1}^p = \gamma \bar{\varepsilon}_t^p + Q \left( \frac{\sigma_n - \gamma}{\sigma_a} \right) \end{aligned}$$

Constraint (i) guarantees a promotion rate between 0 and 1 and that  $V$  remains bounded. Constraint (ii) determines  $\bar{\varepsilon}_{t+1}^p$  as a function of choices and the current state. The variable  $\sigma_a$  is simply the associate to partner ratio. The variable  $\sigma_n$  is the firm's growth rate (in terms of partners) and  $\frac{\sigma_n - \gamma}{\sigma_a}$  is the implied promotion rate. Each is a function of the skill level which exists in the firm,  $\bar{\varepsilon}_t^p$ . The growth rate  $\sigma_n$  must balance skill accumulation and the current wage bill. Let  $T_V(\varepsilon)$  stand for the maximization operator in (18) and let  $C$  be the set of continuous non-decreasing functions on  $B$ . Then  $T: C \rightarrow C$ .

*Proposition II:* A unique, continuous non-decreasing  $V$  satisfying (18) exists. *Proof:* For continuous  $V$  the choice of  $\sigma_a$  is effectively bounded by A3. [1], A3. [2], and A4. [1]. The bound is continuous in  $\bar{\varepsilon}_t^p$ .  $B$  is compact so a uniform finite bound on  $\sigma_a$  exists. Therefore  $T$  is the maximization of a continuous function over a compact subset of  $\mathbb{R}^2$ . Blackwell's sufficient conditions for a contraction mapping are met.  $C$  is a complete metric space (with the sup norm), so  $T$  has a unique fixed point in  $C$ . ■

Assumptions made do not quite guarantee that the objective function is concave. Although  $Q$  is concave in  $\sigma_a$  and  $\sigma_n$ ,  $\sigma_n Q$  is not concave. If  $V$  is close to linear in  $\bar{\varepsilon}_t^p$  the objective may be convex in  $\sigma_n$  in some regions. Conditions for which  $V$  is sufficiently concave so that  $T_V$  is the maximization of a concave function are not known.

*Proposition III:* If  $V$  is sufficiently concave, a steady-state level of skill,  $\varepsilon^*$ , exists. If  $\beta V(0) > \bar{U}$  and  $V$  is differentiable, then  $\varepsilon^* > 0$ .

*Proof.* With  $V$  sufficiently concave, optimal choices are continuous single-valued functions, so the optimal amount of skill next period,  $\bar{\varepsilon}_{t+1}^p = g(\bar{\varepsilon}_t^p)$ , is

continuous in  $\bar{\varepsilon}_t^p$ . The function  $g$  maps  $B \rightarrow B$ , so a fixed point of  $g$  exists. Assumption A3. [4] implies that at  $\bar{\varepsilon}_t^p = 0$  the first order conditions on  $\sigma_a$  and  $\sigma_n$  reduce to

$$\sigma_a = \frac{\beta[V(Q)-R]}{\bar{U} - \beta R}(\sigma_n^{-\gamma})$$

where  $Q = Q\left(\frac{\sigma_n^{-\gamma}}{\sigma_a}\right)$ . If  $\beta V(0) > \bar{U}$ , then  $\beta V(x) > \bar{U}$  for all  $x$  because  $V$  is non-decreasing. But this implies  $\sigma_a > (\sigma_n^{-\gamma})$ , or  $\frac{\sigma_n^{-\gamma}}{\sigma_a} < 1$ . That is,  $g(0) > 0$ , so  $\varepsilon^*$  is not 0. ■

This result does not depend on the form of  $Q$ . It demonstrates that at least one locally stable steady-state skill level exists. At such an  $\varepsilon^*$  the growth rate  $\sigma_n$  and associate/partner ratio  $\sigma_a$  are constant. If firms differ in the revenue function  $\psi$  then their growth rates are potentially different but growth in each firm converges to a constant rate. Of course, the theory does not predict this rate must be greater than 1: firms may shrink at a constant rate.

Without the bound  $\bar{\sigma}_n$  there is no guarantee that a solution exists. Notice however, that if skill does not enter  $\psi$  then  $V$  is constant and optimal  $\sigma_n$  will either be  $\gamma$  or  $\bar{\sigma}_n$  depending on the sign of  $V - r$ . Screening in itself does not guarantee an interior solution to firm growth, but without screening (or incentives) a corner solution is assured.

### Firm Dynamics in a Tournament Environment

This section analyzes the choices of the current partners choosing tournament promotion rules. Attention is restricted to  $\varepsilon$  uniformly distributed (Example 1). This is for two reasons. First, a uniform distribution seems to explain cross sectional variation well compared to other possible distributions. Second, with  $f'(\varepsilon) \neq 0$  promotion rules affect effort both through the change in future value of the firm and the marginal probability of promotion function (7). The set of feasible promotion rules must be specified because second order conditions on the effort decision may fail to hold. These complications increase notation significantly.

The firm knows that equation (7) is used by associates to choose their symmetric Nash effort. Furthermore, the current partners take as given the set of possible choices of  $(a_t, \eta_{t+1})$  that yield Nash equilibrium effort values. If A1 is satisfied and  $\varepsilon$  is uniform, any promotion rules in which the future value of the firm is greater than  $R$  is feasible. The effort function is then simply

$$x(a_t, \eta_t) \equiv \begin{cases} (c')^{-1} \left( \frac{\beta}{b} (V-R) \right) & 0 \leq \eta_t \leq a_t \text{ and } V(\bar{\varepsilon}_{t+1}^p) \geq R \\ 0 & \text{Otherwise} \end{cases}$$

At times below,  $x(a_t, \eta_t)$  and the corresponding indirect costs  $c(x(a_t, \eta_t))$  will be written simply as  $x$  and  $c$ , respectively. Technically, (8) does not hold for  $\eta_{t+1} = 0$  or  $\eta_{t+1} = a_t$ . The definition of  $x$  ignores this.

From Proposition I.1, the probability of promotion in the Nash equilibrium is the promotion rate. Therefore, the wage the firm must pay can be written as

$$w_t = \bar{U} + c - \beta \left( \frac{\eta_t}{a_t} V - \left( 1 - \frac{\eta_t}{a_t} \right) R \right) \quad (19)$$

Substitution of this into (15) as done in (18) gives the functional equation that defines the firm's behavior:

$$V^*[\bar{\varepsilon}_t^p] \equiv \max_{\{\sigma_a, \sigma_n\}} \psi(\bar{\varepsilon}_t^p, \sigma_a x) - \sigma_a (\bar{U} + c(x) - \beta R) + \sigma_n \beta \left( V^*[\bar{\varepsilon}_{t+1}^p] - R \right) + \gamma \beta R$$

subject to

(20)

$$(i) \gamma \leq \sigma_n \leq \min\{\sigma_a + \gamma, \bar{\sigma}_n\}$$

$$(ii) \varepsilon_{t+1}^p = \gamma \varepsilon_t^p + Q \left( \frac{\sigma_n - \gamma}{\sigma_a} \right)$$

Define  $T^*$  as the operator in (20). A critical difference between  $T$  and  $T^*$  is that  $T^*$  is not guaranteed to be a contraction mapping even with a bound on  $\sigma_n$ . Future value of the firm feeds into  $x$  and  $c(x)$ .  $T^*$  is not necessarily monotone in  $V$  because increasing  $V$  increases  $c(x)$  as well as  $x$ . Two results do provide some characteristics of the tournament firm.

*Proposition IV.* If  $\frac{\partial \psi}{\partial (\bar{\varepsilon}_t^p)} = 0$ , then the choice of  $\sigma_n$  is on one of the boundaries in 20(i) depending on the sign of  $V^* - R$ .

*Proof.* In this case the promotion rate affects neither  $V$  nor effort, so  $\sigma_n$  has a constant return  $\beta(V - R)$ . If  $V > R$  then  $\sigma_n = \bar{\sigma}_n$ , if  $V < R$  then  $\sigma_n = \gamma$ .

*Proposition V.* Define  $\bar{v}$  as the set of constant functions on  $B$  with value  $< R$ . Then  $T^*$  maps  $\bar{v}$  to  $\bar{v}$  and has a unique fixed point in  $\bar{v}$  at  $V = 0$ .

*Proof.* If  $V < R$  for all skill levels, then associate effort is always 0. The optimal choices are  $\sigma_n = \sigma_a = \gamma$ , so that  $T^*V = \gamma\beta V$ , which converges to  $V = 0$ .

Although quite simple, these results show that screening is required to avoid non-trivial growth paths with a uniform distribution and that a tournament firm will unravel if the value of the firm is ever expected to be less than  $R$ . A trivial solution to (20) exists. This does not rule out other solutions, because  $T^*$  need not have a unique fixed point. Indeed, simulations reported below found solutions for  $V^*$  through iteration on (20) despite the fact that this process is not guaranteed to work.

Proportional growth of the firm in the long run does not depend on the assumption that  $\varepsilon$  is uniformly distributed. The incentive effect of promotions at a Nash equilibrium level,  $\frac{\partial P}{\partial x}$ , depends upon only the promotion rate for firms. The set of feasible promotion rates is smaller and harder to define than with a uniform distribution because low promotion rates can not support symmetric Nash effort levels.

### The Effect of Incentives on Firm Behavior

Both the tournament and screening environments are complicated enough to make strong characterization of optimal decisions difficult. This section discusses simulations based on a Cobb-Douglas specification for revenue. That is

$$\psi = A\varepsilon^\alpha(ax)^\theta.$$

The skill aggregator  $Q$  is specified as the mean of  $\varepsilon_t^1$  conditional on making partner. For the uniform case  $Q(\lambda) = (\frac{b}{2})(1-\lambda)$ . To maintain some connection with the data, the estimated values of  $b$  and  $R$  were used in the simulations.<sup>19</sup> (Note that (18) was modified to allow the firm to choose and enforceable effort level.) The steady-state equations were solved for randomly chosen values of  $A$ ,  $\alpha$  and  $\theta$ . These equations are the first order conditions for choices and the requirements that partner skill and the value of the firm be constant. Approximately 10% of the sets of parameters yielded positive and bounded growth rates for both the screening and tournament environments. Convergence to solutions requires very good starting values, which partly accounts for the low success rate.<sup>20</sup>

---

<sup>19</sup>The value of  $\beta$  was lowered to .40 ( $=.89^8$ ) to speed convergence, and  $\gamma$  was set at .1 to increase the range of  $\sigma_n$ . Iteration on the value function used piecewise linear approximation based on 50 values of  $\bar{\varepsilon}_t^p$ .

<sup>20</sup>As a check on these results, iteration on (18) and (20) was also carried out.

Table 4 reports how long run behavior of this firm responds to the market parameters  $b$  and  $R$ . The screening model provides unambiguous sign predictions on elasticities of steady-state values. The tournament model tends to have ambiguous predictions, but in only two cases do the average elasticities conflict. For example, the first two columns show that an increase in  $b$  (a mean-preserving spread in the skill distribution) increases firm growth ( $\sigma_n$ ) and the steepness of the firm hierarchy ( $\sigma_a$ ). A steeper hierarchy would suggest lower promotion rates, but the increase in growth allows promotion rates to increase as well. The value of the firm is larger, but the higher promotion rates lower the long run skill level of partners. A larger  $b$  lowers incentives because effort is measured with more noise. This effect outweighs the increase in  $V$ , so effort falls in the tournament case. In essence, firms substitute toward numbers of associates and partners and away from skill and effort, which are relatively more costly to motivate and select with larger  $b$ .

The last two columns of Table 4 present elasticities of firm choices to a change in  $R$ . This increases  $\bar{U}$  as well. In the screening environment firm response is opposite to a change in  $b$  except for wages. Alternatives of associates outside the firm are better so they are more expensive to hire. Firm value is decreased, but firms substitute toward tougher partner selection through lower promotion rates. The responses in the tournament environment are ambiguous. On average they agree with the screening responses except with firm value and the wage.

Table 5 compares the levels of the steady-state values in the two environments. Although only a few draws (77) survive the criterion that both sets of equations be solved for the set of parameter values. Firms with asymmetric information on worker effort grow slower, have lower promotion rates and are less valuable. Again, a flatter firm structure accompanies the lower promotion rate because of slower growth. The percentage change in firm value is small, however.

These results suggest that individual firm structure may not reveal whether promotions have incentive effects or not. Time series data on the response to market conditions and levels of inputs may be similar with or without asymmetric

---

Iteration can demonstrate existence of a tournament  $V^*$  directly but existence is not ruled out if the process converges to  $V^* = 0$ . Solutions to  $TV^* = V^*$  were found over the range of revenue parameters, but not for every case. This suggests, but does not prove, that solutions to the steady-state equations typically describe actual optimal values, but starting values are important for both approaches.

information. Restrictions from the equilibrium condition across firms at a given point in time may be more pertinent. A limitation of this approach is that the cost of effort in a screening model with variable effort is not determined by observable features of the firm alone. The firm's problem must be solved to obtain effort. For this reason the null hypothesis in section III was a pure screening model.

## V. CONCLUSION

This paper has explored the relationship between promotion rules and information in a dynamic model of partnerships. Promotions allocate workers and skill to levels of the firm and generate incentives for workers below the partner level. The tournament model of incentives generates an equilibrium restriction across firms that nests a pure screening model. Tournament effects explain data on major U.S. law firms significantly better than the pure screening hypothesis. Both the screening and incentive aspects of promotions create costs to firm growth. A bounded long run growth rate is then optimal with constant returns to scale in production. This provides an explanation for steady growth in the size of partnerships based on the nature of information within the firm.

Attention has been restricted to partnerships and, in particular, law firms, but the analysis extends to ordinary firms. Use of data from firms to test models of promotions and firm organization is increasing (e.g. Baker and Jensen (1990)). Screening and motivating are logically distinct tasks that promotion rules may or may not perform. Simulations suggest that they affect optimal firm behavior in similar ways. That is, analysis of patterns within one firm or many firms separably probably can not gauge the relative importance of screening and incentives. The models also predict similar patterns in simple correlations across firms. This paper has suggested that the equilibrium restrictions across firms at a point in time identify the importance of incentive effects of promotion rules. Similar work with data from other types of firms is clearly suggested.

## APPENDIX

### Proof of Proposition I

- [1]. From (6), the right hand side of (2) does not depend upon  $x^*$ . Given the strict convexity of  $c$ , only one  $x$  can solve (2).
- [2]. This is obvious because  $a_t$  people who choose identical effort competing for  $\eta_t$  spots will have a promotion probability of  $\eta_t/a_t$ . This can also be shown analytically by noticing that the term in the sum in (5) reduces to  $1/a_t$  for each  $k$  once we make the transformation to uniform variates. Because (5) sums over  $\eta_t$  terms, the result obtains.
- [3]. Replace  $n$  by  $a-n$  in (6) and use symmetry of  $f$  and the binomial coefficient and the fact that  $F(\epsilon) = 1 - F(-\epsilon)$ , to show:

$$\begin{aligned} \left. \frac{\partial P(x^*)}{\partial x^*} \right|_{\eta_t = a-n} &= \\ &= \int (a-1) \binom{a-2}{n-1} F^{n-1}(\epsilon) [1-F(\epsilon)]^{a-n-1} f^2(\epsilon) d\epsilon = \\ &= \int (a-1) \binom{a-2}{a-n-1} [1-F(-\epsilon)]^{n-1} F(-\epsilon)^{a-n-1} f^2(-\epsilon) d(-\epsilon) = \\ &= \int (a-1) \binom{a-2}{a-n-1} [1-F(\epsilon)]^{n-1} F(\epsilon)^{a-n-1} f^2(\epsilon) d(\epsilon) = \\ &= \left. \frac{\partial P(x^*)}{\partial x^*} \right|_{\eta_t = n} \end{aligned}$$

- [4]. Define the part of the integrand in (6') as  $T(p, \epsilon)$ , with  $\epsilon$  suppressed for compactness:

$$T(n, \epsilon) = F[1-F]f' + f^2[a-n-1+(2-a)F].$$

Then, using the properties in [3] and that  $f'(\epsilon) = -f'(-\epsilon)$ :

$$\begin{aligned} T(a-n, -\epsilon) &= [1-F]Ff'(-\epsilon) + f^2[n-1+(2-a)(1-F)] \\ &= (-1)[1-F]Ff'(\epsilon) + f^2[-(a-n-1)-(2-a)F] \\ &= (-1) \left[ [1-F]Ff'(\epsilon) + f^2[a-n-1+(2-a)F] \right] \\ &= -T(n, \epsilon). \end{aligned}$$

Combined with the equivalence, as in the proof of [3], of other components in (6'), this implies negative symmetry of  $\partial^2 P$  with respect to  $\eta_t$  around  $a_t/2$ .

## REFERENCES

- Battacharya, Sudipto and Guasch, J. Luis. "Heterogeneity, Tournaments, and Hierarchies." *Journal of Political Economy* 96 (August 1988): 867-881.
- Bernhardt, Dan. "Skill Profiles, Observability and Firm Hierarchies: A Theory of Promotion and Compensation." manuscript, Dept. of Economics, Queen's University, 1990.
- Carmichael, Lorne. "Firm-Specific Human Capital and Promotion Ladders." *Bell Journal of Economics* 14 (Spring 1983): 251-258.
- Farrell, Joseph and Scotchmer, Susanne. "Partnerships." *Quarterly Journal of Economics* 103, (May 1988): 279-297.
- Ferrall, Christopher. "Agents Becoming Principals: Theory and Estimation of a Tournament Model of Promotions in Partnerships." Ph.d. thesis, University of Minnesota, December 1990.
- Galanter, Marc and Palay, Thomas. "Why the Big Get Bigger: The Promotion-to-Partner Tournament and the Growth of Large Law Firms." *Virginia Law Review* 76 (May 1990): 747-811.
- . *Tournament of Lawyer: The Transformation of the Big Law Firm*. Chicago: University of Chicago Press, 1991.
- Gaynor, Martin and Pauly, Mark. "Alternative Compensation Arrangements and Productive Efficiency in Partnerships: Evidence from Medical Group Practice." *Journal of Political Economy* 98, (June 1990): 544-573.
- Gibbs, Michael. "Promotion, Compensation and Firm Organization." Phd. thesis, University of Chicago, June 1989.
- . "Promotions and Incentives: Tournaments vs. Standards." manuscript, Graduate School of Business, Harvard University, April 1990.
- Gifford, Jr., Adam and Kenney, Roy. "The Production of Information through Labor Market Contests." *Journal of Law, Economics, and Organization* (Fall 1986): 305-313.
- Gilson, Ronald and Mnookin, Robert. "Sharing Among the Human Capitalists: An Economic Inquiry into the Corporate Law Firm and How Partners Split Profits." *Stanford Law Review* 37 (January 1985): 313-92.
- . "Coming of Age in a Corporate Law Firm: The Economics of Associate Career Patterns." *Stanford Law Review* 41, (February 1989): 567-95.
- Green, Jerry and Stokey, Nancy. "A Comparison of Tournaments and Contracts." *Journal of Political Economy* 91, (June 1983): 349-64.
- Jensen, Michael and Murphy, Kevin J. "Performance, Pay and Top Management Incentives." *Journal of Political Economy* 98 (April 1990): 225-264.
- Kahn, Charles and Huberman, Gur. "Two-sided Uncertainty and 'Up-or-Out' Contracts." *Journal of Labor Economics* 6 (October 1988): 423-444.



- Lazear, Edward. "Pay Equality and Industrial Politics." *Journal of Political Economy* 97 (June 1989): 561-80.
- , and Rosen, Sherwin. "Rank-Order Tournaments as Optimal Labor Contracts." *Journal of Political Economy* 89 (October 1981): 841-64.
- Legros, Patrick. "Sustainability in Partnerships." manuscript, Dept. of Economics, Cornell University, October 1989.
- Leibowitz, Arleen and Tollison, Robert. "Free Riding, Shirking and Team Production in Legal Partnerships." *Economic Inquiry* XVIII (July 1980): 380-94.
- Malcomson, James. "Work Incentives, Hierarchy and Internal Labor Markets." *Journal of Political Economy* 92 (June 1984): 486-507.
- Miyazaki, Hajime. "On Success and Dissolution of the Labor-Managed Firm in a Capitalist Economy." *Journal of Political Economy* 92 (December 1984): 909-931.
- MacLeod, W. Bentley and Malcomson, James. "Reputation and Hierarchy in Dynamic Models of Employment." *Journal of Political Economy* 96 (August 1988): 833-854.
- O'Flaherty, Brendan and Siow, Alysius. "On the Job Screening, Up or Out Rules and Firm Growth." manuscript, Department of Economics, Columbia University, February 1989.
- , "Up or Out Rules and Firm Growth in the Market for Lawyers." manuscript, Department of Economics, University of Toronto, October 1990.
- Prescott, Edward and Visscher, Michael. "Organization Capital." *Journal of Political Economy* 88 (June 1980): 446-61.
- and Boyd, John. "Dynamic Coalitions, Firm Size, Growth, and the Firm." in *Contractual Arrangements for Intertemporal Trade*, ed Edward Prescott and Neil Wallace, Minneapolis: Univ. of Minnesota Press, 1987.
- Rosen, Sherwin. "Prizes and Incentives in Elimination Tournaments." *American Economic Review* 77, (September 1986): 701-15.
- Spurr, Stephen. "How the Market Solves an Assignment Problem: The Matching of Lawyers with Legal Claims." *Journal of Labor Economics* 5 (December 1987): 502-32.
- Waldman, Michael. "Up-Or-Out Contracts: A Signaling Perspective." *Journal of Labor Economics* 8 (March 1990): 230-50.

TABLE 1. PROMOTIONS, PROFITS, STARTING SALARY,  
AND NUMBERS OF LAWYERS FOR MAJOR U.S. LAW FIRMS 1987

(99 firms)

Variable	Mean	Std. Dev.	Min	Max
promotion rate*	.1660128	.0561528	.063	.2949329
per-partner profits	363.9495	218.9245	155	1405
starting salary	59.93939	8.417448	39	74
number of associates	191.2626	98.338	41	670
number of partners	113.6869	49.09854	43	370

Sources: American Lawyer and Of Counsel. \*0.25 X partners/associates.  
See Appendix. Profits and salary are in thousands. Salaries rounded to  
\$1000, profits to \$5000

TABLE 2. CORRELATIONS ACROSS VARIABLES IN TABLE 1

	promo. rate	profits	salary	asso.	part.
promotion rate	1.00				
per-partner profits	-0.43	1.00			
starting salary	-0.61	0.54	1.00		
number of associates	-0.49	0.24	0.33	1.00	
number of partners	0.24	-0.21	-0.16	0.67	1.00

TABLE 3. ESTIMATES FROM UTILITY EQUATION

Parameter	Estimate	asymptotic t-ratio
R	262.885	24.24
z	0.03789	6.42
$\sigma$ $\nu$	187.080	73.62
ln like.		-567.42
ln like. for z=0		-571.88
lnlk. ratio [ $\chi^2_{.99}(1)$ ]		8.95 [6.63]

Parameters held constant:  $\beta = .95^8$ ;  $\gamma = .9$

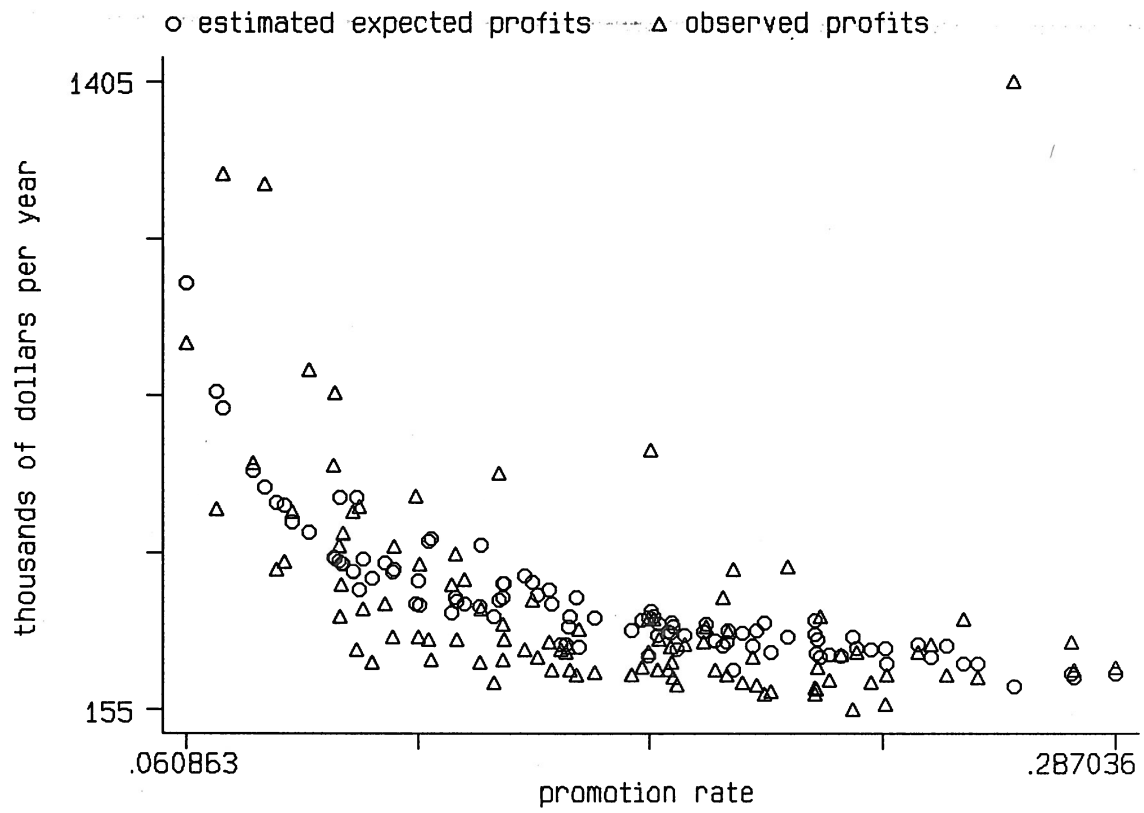


Fig 1. promotion rates and profits

STATA™

TABLE 4. ELASTICITY OF FIRM BEHAVIOR TO MARKET PARAMETERS

	b		R	
	tournament	screening	tournament	screening
$\sigma_a$	1.79 (85)	3.30 (100)	-1.50 (31)	-7.22 (100)
$\sigma_n$	2.40 (79)	3.44 (100)	-1.97 (31)	-7.83 (100)
V	0.48 (87)	1.24 (100)	0.91 (69)	-2.01 (100)
x	-0.17 (76)	-0.45 (100)	0.55 (69)	1.11 (100)
$\lambda$	0.79 (77)	0.42 (100)	-3.58 (31)	-1.23 (100)
$\lambda(V-R)$	1.47 (77)	2.35 (100)	-4.55 (31)	-5.77 (100)
w	0.71 (95)	0.65 (100)	-0.32 (31)	1.43 (100)
$\epsilon$	-1.32 (82)	-1.89 (100)	2.41 (27)	4.88 (100)

1. Average elasticities of steady-state values over revenue parameters. Based on 2455 draws of Cobb-Douglas revenue parameters which yielded positive bounded growth rates ( $1/\beta > \sigma_n > 1$ ). Only 203 and 179 draws for the tournament and screening models, respectively, satisfied the criterion. Many of the unselected draws are due to bad starting values.
2. Percentage of elasticities with same sign as mean in ().
3. x = effort,  $\lambda$  = promotion rate, w = wage,  $\epsilon$  = partner skill level.

TABLE 5. RELATIVE DIFFERENCES BETWEEN TOURNAMENT AND SCREENING VALUES

Variable	Mean	Min	Max
$\sigma_n$	-0.25	-1.47	0.10
$\sigma_a$	-1.26	-8.29	0.07
V	-0.10	-0.50	0.01
$\lambda$	-1.12	-9.80	0.001

1. Change in steady-state levels from screening to tournament environments relative to tournament levels.
2. Based on 77 draws of revenue parameters for which growth rates are positive and bounded and solutions to both sets of equations found.