

Managers, Coordination, and the Firm Age - Wage Relationship

Peter Thompson

Florida International University

March 2009

This paper describes a version of Lucas' span of control model, in which managers of younger and smaller firms are less able than managers of older firms to provide precise instructions to employees. Employees differ in their propensity to follow instructions, and those least likely to follow instructions are said to be high- α types. In equilibrium, younger [older] firms employ high- α [low- α] types, and wages exhibit a U-shaped relationship in which the lowest wages are paid by firms of *intermediate* age. A natural extension of the model, in which employee ability also varies, is developed to examine the effect of employer age and size on entry into self-employment.

JEL Classification Codes: J24, J31, L26.

Keywords: Management, firm age and wages, self-employment.

Department of Economics, Florida International University, Miami, FL 33199. Email: peter.thompson2@fiu.edu.

1. Introduction

Analysis of establishment data has shown that wages are higher in older firms, and that this relationship survives controlling for size, industry, location, union status and other observable firm characteristics [Dunne and Roberts (1990a, 1990b), Davis and Haltiwanger (1991), Troske (1998)]. One explanation is that older firms are perceived to have a greater ability to pay higher wages, and doing so is perceived as only fair; as a result, older firms find it costly to pay low wages [Kahneman, Knetsch, and Thaler (1986), Blinder and Choi (1990)]. A second explanation is that older firms expect to survive longer than younger firms, so they invest more in training [Idson (1996)]. If some of the benefits of this training are transferable to outside employment, the result is an increase in wages. These explanations are readily tested using establishment data.

Other explanations have focused on the composition of workers in young and old firms, the contributions of which cannot be identified from establishment data. In particular, employees in older firms have longer tenure and more work experience. Worker characteristics might explain all the age effects but, as Brown and Medoff (2003) note, they do not constitute true firm-age differentials. Using data on 1,410 individuals employed in the private sector Brown and Medoff measure the fraction of observed firm-age differentials that can be explained by readily observed worker characteristics. As expected, tenure and experience play an important role, explaining almost half of the firm-age differential. Variations in education, gender, marital status, and occupation class, together explain much of the other half.

After accounting for observable worker characteristics, there is no monotonic relationship between wages and firm age. Instead, Brown and Medoff identify a U-shaped relationship in which the lowest wages are paid by firms of *intermediate* age. This finding remains something of a puzzle.

What might account for such a pattern? We suspect that no one factor can do so. Working backward from data to theory, we should be looking for a relationship that can account for new firms paying higher wages that becomes less important as firms age, and/or a positive wage-age relationship that is muted among young firms.

Brown and Medoff (2003, p. 694)

This paper suggests a single factor that might explain the U-shaped firm-age differential. Our explanation builds on the following production technology, which is developed in Section 2. A manager hires n workers. Each worker carries out exactly one

unit of work per unit of time, so that n units of work are done in the firm. The value of this work depends on the ability of the manager to direct the work that is done, each unit of which has a well-defined direction. Conditional on the intended direction, the actual direction of individual workers depends on the success with which the manager organizes his employees; the less directed are the employees, the less progress they make in the intended direction. For reasons given later, we assume that managers of younger firms are (stochastically) less able than managers of older firms to provide precise instructions to employees. At the same time, employees differ in their propensity to implement managers' instructions and to act instead on their own assessment of the direction work should take. The propensity to implement managers' instructions is described by a parameter, α ; those least likely to follow instructions are high- α types.

We show in Section 3 that the model yields Brown and Medoff's U-shaped function in the following manner. In equilibrium, firms with the least ability to direct workers are keen to hire high- α types, while firms most able to direct workers strictly prefer low- α types. No firm has a strong preference for intermediate- α types. As a consequence, the least able firms bid up the wages at one end of the distribution of α , while the most able bid up the wage at the other end. Hence, wages are a U-shaped function of α , α is inversely related to managerial ability, and managerial ability is stochastically increasing in firm age.

Firm size is determinate in the model in large part because the success with which a manager of given ability coordinates work declines as the number of employees under his direction rises. This notion has several well-known antecedents. Williamson (1967) was the first to introduce the idea that only a fraction, α , of the intentions of a manager are carried out by his employees, and that this loss of control sets a limit to firm size.¹ The model is also related to Becker and Murphy (1992), who explore how increasing coordination costs limit the degree of specialization in firms and economies.² Becker and Murphy do not consider a distinct coordination role for managers, although there is nothing in their model specifically to preclude it. The model is perhaps most closely related to Lucas (1978). Indeed, the present model can be viewed as both a particular interpretation of, and an extension of, the span of control

¹ The setting is a monitoring problem. See Mirrlees (1976), and Calvo and Wellisz (1978) for further refinements. Keren and Levhari (1979) study hierarchies in the context of coordination.

² Becker and Murphy provide numerous examples of how excessive specialization induces excessive coordination costs and reduces productivity.

model. It is an interpretation of Lucas when α does not vary across employees: in this case, like Lucas, the model predicts that the best managers operate the largest firms, and that labor productivity is invariant to firm size. But, as we show in Section 4, these central characteristics of Lucas' model are not robust to our extension in which α varies across individuals. In this latter case, managers of intermediate ability may under certain conditions employ the most workers, while labor productivity is a U-shaped function of firm size.³

In Section 5 of the paper, we develop our model further to explore its ability to explain some other intriguing relationships between firm age and size, and labor markets. Recent research has revealed the following empirical regularities:

R1. The cross-sectional variance of wages in younger firms is greater than the variance in older firms, but less than the variance of self-employment earnings [Elfenbein, Hamilton, and Zenger (2008)].⁴

R2. Employment in small and young firms is positively associated with entry into self-employment [Boden (1996), Wagner (2004), Elfenbein *et al.* (2008)], and with self-employment success [Elfenbein *et al.* (2008)].⁵

R3. Entrants into self-employment are drawn predominantly from the upper and lower ends of the wage distribution. This finding holds overall, as well as separately in large [older] and small [younger] firms. [Elfenbein *et al.* (2008)].

To explore whether our model is consistent with these regularities, we add a second dimension to employee heterogeneity, and we add some simple dynamics with a three-period model. When employees ignore their manager's instructions they vary in the quality of their own decisions. High- α employees who improve on the instructions they are given are high ability – colloquially, they can be said to exhibit *initiative*. High- α employees who do worse are lower ability and considered as merely *insubor-*

³ Others, such as Jovanovic (1994), have extended Lucas' model in reasonable ways and overturned some key results.

⁴ Elfenbein *et al.* relate the variances to firm size but they do not condition on firm age, so the relationship also holds unconditionally for age. The greater variance of self-employment earnings has been documented in numerous other studies, including Hamilton (2000), Rosen and Willen (2002), Moskowitz and Vissing-Jørgensen (2002), and Åstebro, Chen and Thompson (2008),

⁵ Elfenbein *et al.* (2008) again analyze only size effects. Wagner (2004) analyzes size and age effects and concludes that the latter is more important.

dinate. In the first period, incumbent firms employ workers knowing only the employee's α ; they can identify the worker's propensity to follow instructions, but they cannot distinguish between initiative and insubordination. At the end of the first period, however, each employee's ability is revealed, so in the second period the offered wage is adjusted. In the third period, workers choose whether to enter self-employment.

The predictions of the model relate to the regularities given above as follows:

R1. In the second period, the largest wage increases are captured by employees with initiative, while the largest declines are suffered by the insubordinate. The ability of low- α employees is unimportant to the firm, so their wage does not adjust much. As high- α employees are concentrated in young firms, wage adjustments are larger in young and small firms. In the second period, therefore, the variance of wages is greater in young firms.

R2. Leaving the current employer is preferred only by the insubordinate in small firms, and by workers with initiative employed by large firms. Numerical calculations suggest that small firms produce more self-employed than do large firms, and both produce more than intermediate-sized firms.

R3. Entrants into self-employment are drawn predominantly from the lower end of the distribution of wages in small firms and from the upper end of the distribution in large firms.

2. The Production Technology

A firm employs n workers, each of whom does one unit of work per period. How productive this work is depends upon how well each employee's activity is aligned with a particular strategy. The degree of alignment is a random variable, and the expected distance between the strategy and the activity depends upon the firm's ability to align its employee activities with its preferred strategy. The firm's ability to do so depends upon four factors: the number of employees, the manager's ability to provide precise directions to his employees, the employees' propensity to follow these directions, and the ability of employees to divine the correct activity whenever they fail to follow directions as given.

Figure 1 illustrates the implementation of this idea for $n = 3$. Each of three workers carries out one unit of work, indicated by segments x_1 , x_2 and x_3 . We shall normalize the length of a unit of work to be equal to $2A$. The contribution of this work toward firm output depends upon the relation between the direction of strategy, which is

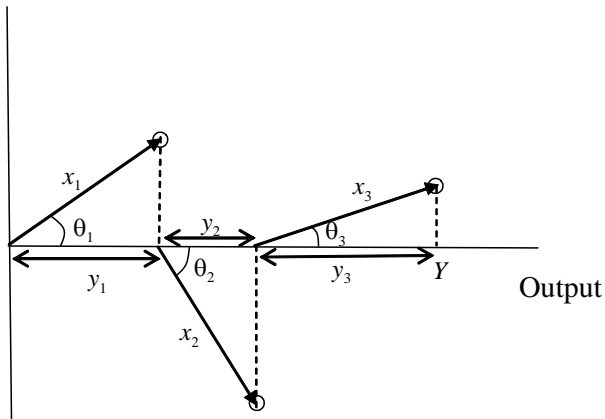


FIGURE 1. The production technology.

normalized to consist of an attempt to move as far along the horizontal axis as possible, and the direction of the worker's effort. The contribution of worker 1, for example, is given by the length of the horizontal segment y_1 , and firm output is given by $Y = \sum_{i=1}^3 y_i$. As each of the line segments x_i has length $2A$, the contribution of worker i to output is given by $y_i = 2A \cos(\theta_i)$. In order that no worker contributes a negative amount, we assume that (in radians) $-\pi/2 \leq \theta_i \leq \pi/2 \forall i$, so $y_i \in [0, 2A] \forall i$.

Assume that each $\cos(\theta_i)$ is a random draw from the uniform distribution with positive support on the interval $[0, a]$. The parameter a is our measure of the firm's ability to align employee activities with strategy. If it did not vary with the number of workers to be directed, expected output for a firm with n workers would be $E[Y; a, n] = nE[y; a]$, indicating constant returns to scale. Consequently, the best manager would employ all the workers and produce all the output. We introduce diminishing returns with the simple parameterization $a(n) = \lambda^\gamma / n^\beta$, where $\lambda > 0$ is an index of the quality of decisions being made in the firm. Both γ and β lie in the unit interval.

Our assumption that $\cos(\theta_i)$ is a draw from the uniform distribution yields a familiar production function. Expected profit is

$$E[\pi | s] = \int_0^{\lambda^\gamma n^{-\beta}} \frac{2Anx}{\lambda^\gamma n^{-\beta}} dx - w(\alpha)n,$$

$$= A\lambda^\gamma n^{1-\beta} - w(\alpha)n, \quad (1)$$

Assume that each firm has one manager/owner, and let $s \in [0,1]$ denote his ability. There is a continuum of firms of unit mass, with abilities drawn from the distribution $F_s(s)$, with density $f_s(s)$. There is also a continuum of workers of mass m . Workers vary in their propensity to follow their own ideas rather than their manager's instructions. We index this propensity by $\alpha \in [0,1]$, and denote its population distribution as $F_\alpha(\alpha)$. We shall refer to α as a measure of the worker's character. Just like managers, workers also vary in their ability to make decisions. We denote their ability by $v \in [0,1]$, and assume that it is a draw from the known distribution $F_v(v)$. A worker's type is therefore defined by the 2-tuple $\{\alpha, v\}$. All distributions are assumed to be continuous, so there are no mass points.

The parameters s , α , and v are related to firm performance in the following way. Assume a firm has a manager with ability s , and hires n workers with characters α_i and abilities v_i . Then the firm type is

$$\lambda = sn^{-1} \sum_{i=1}^n (1 - \alpha_i) + n^{-1} \sum_{i=1}^n \alpha_i v_i. \quad (2)$$

We assume throughout the paper that each firm does not employ workers with different characters, so (2) can be written as

$$\lambda = (1 - \alpha)s + \alpha n^{-1} \sum_{i=1}^n v_i. \quad (3)$$

While there is perhaps an appearance of novelty to the production function, it is a quite standard technology. Expected firm output is governed by a Cobb-Douglas aggregator with firm human capital, λ , and raw labor hours, n , as the only inputs. Firm human capital depends upon managerial ability and the workers' types. One would suppose that part of the firm's human capital, and hence its expected performance, can be predicted from observable characteristics of managers and founders [e.g., Bates (1990), Colombo, Delmastro, and Grilli (2004), Åstebro and Bernhardt (2005)], part from observable characteristics of the firm itself [e.g., Dunne, Roberts, and Samuelson (1988, 1989), among many others], and part from observable characteristics of employees [e.g., Abowd, Kramarz and Margolis (1999), Brown and Medoff (2003)]. Of course, the actual performance of the firm and the earnings of its employees will contain a significant unpredictable component, because output contains a stochastic component and because s and $\{\alpha, v\}$ are not predetermined by observables. Indeed, Abowd *et al.* conclude that person effects, especially effects unrelated

to observables, are a very important source of wage variation.

3. The U-shaped Relationship between Firm Age and Wages

In this section, we shall also assume that $F_v(v)$ is degenerate, so $v = \bar{v}$ for all employees, allowing us to define the firm's type by

$$\lambda = (1 - \alpha)s + \alpha\bar{v}. \quad (4)$$

Let $w(\alpha)$ denote the wage paid to α workers, and normalize the price of output to unity. The firm maximizes (1) by choosing the number of workers, n , and employee character, α .

The first-order condition for firm size is⁶

$$(1 - \beta)\lambda^\gamma n^{-\beta} = w(\alpha), \quad (5)$$

If all workers had the same character, $\bar{\alpha}$, earning the same wage, $w(\bar{\alpha})$, optimal employment would be proportional to $\left((1 - \bar{\alpha})s + \bar{\alpha}\bar{v}\right)^{\gamma/\beta}$. But as $F_\alpha(\alpha)$ is not degenerate, two countervailing forces induce a deviation of firm size from this simple solution. On the one hand, firms can raise α to ameliorate the consequences of a low s or reduce α to enhance the advantages of a high s . This effect tends to convexify $n(s)$. On the other hand, as we shall see, wages are higher in the tails of the distributions (of α and s), and this force operates on employment in the opposite direction. In general, however, $n(s)$ is an increasing function, consistent with Lucas' (1978) result that firm size is increasing in managerial ability. However, the precise form that $n(s)$ takes depends upon assumptions about the form of the distributions.

The first-order condition for employee character is

$$\gamma\lambda^{\gamma-1}n^{1-\beta}(\bar{v} - s) = w'(\alpha)n. \quad (6)$$

A sufficient condition for (6) to be optimal is that $w(\alpha)$ is convex; we will now show that this is the case in an equilibrium assignment of workers to managers.

The worst [best] managers prefer to employ high [low] α employees, and this implies a non-monotonicity in the wage schedule. To see why, note first that continuity of all distributions imply that an interior solution for α must hold for almost all s . Imagine that, to the contrary, a manager with ability s interior to the domain of $F_s(s)$ preferred to hire workers of type $\alpha = 0$. Then every manager with ability $s' \geq s$ also

⁶ Given the restriction on β , the second-order condition is always satisfied.

prefers $\alpha = 0$. This creates a mass point of demand for workers of type $\alpha = 0$, but continuity of $F_\alpha(\alpha)$ precludes the existence of such a mass point. Similarly if manager s prefers $\alpha = 1$, so do all managers with $s' \leq s$ creating an untenable mass point of demand at $\alpha = 1$. Next, for (5) to define an interior solution it must be the case that $\text{sgn}[w'(\alpha)] = \text{sgn}[\bar{v} - s]$. That is, wages are locally increasing in α for the types of workers hired by low-ability managers, and locally decreasing in α for the types of workers hired by high-ability managers. The intuition here is also straightforward: if wages were not locally increasing, all managers of type $s < \bar{v}$ with employees of given character α could increase profits by raising α ; the wage must increase at a sufficient rate to make the current α optimal. Similarly, all managers of type $s > \bar{v}$ can increase profits by reducing α unless wages fall with α at a sufficient rate. Finally, firms with low s employ high- α workers, and this yields the first proposition:

PROPOSITION 1. (i) $w(\alpha)$ is a U-shaped function; (ii) $w(s)$ is a U-shaped function.

To go from here to Brown and Medoff's (2003) result, we need only relate s to firm age. We shall not do so formally: we shall simply assume that the distribution of s can be written as $F_s(s | a)$, where a is firm age and F_s is decreasing in a . We offer two justifications for this assumption; they do not strike us as especially contentious. The first is a simple managerial learning-by-doing story. As is the case for employees, managers in young firms have less tenure and less experience, and their relative inexperience makes them less able to manage effectively. The second is a *firm* learning story. Young firms, especially in high-tech industries, frequently have less well-defined strategies and fewer established routines than older firms. In such environments, precise direction of workers is more difficult, even holding managerial ability constant.⁷

COROLLARY 1. *The wage is a U-shaped function of firm age.*

We can say a little more about the wage function if we are willing to make a narrow assumption about the assignment of worker characters to manager ability. Eliminating n from (5) and (6) yields a first-order differential equation for the wage function,

⁷ The absence of well-defined routines may provide young firms with a comparative advantage in activities for which routines are less valuable. For example, Berger *et al.* (2002) note that small banks are more likely to specialize in lending to businesses in which information is "soft".

$$w'(\alpha) - \frac{\gamma(\bar{v} - s)}{A(1 - \beta)((1 - \alpha)s + \alpha\bar{v})} w(\alpha) = 0, \quad (7)$$

the solution to which depends upon the relationship between s and α . We proceed by assuming an assignment in which there is a one to one mapping, $s \rightarrow \alpha$. Both s and α are defined on the unit interval, so for purposes of illustration we let this assignment be $s = 1 - \alpha$.⁸

Substituting $s = 1 - \alpha$ into (7) and solving the differential equation yields a wage function,

$$w(\alpha) = c \cdot \exp \left\{ \frac{\gamma\sqrt{\bar{v}} \tan^{-1} \left(\frac{\bar{v} + 2(\alpha - 1)}{\sqrt{\bar{v}(4 - \bar{v})}} \right)}{(1 - \beta)\sqrt{4 - \bar{v}}} + \frac{\gamma \ln(\alpha\bar{v} + (1 - \alpha)^2)}{2(1 - \beta)} \right\}, \quad (8)$$

where the constant c will depend, *inter alia*, on the parameters of the model and the measure of available workers, m . The function is, of course, U-shaped. Figure 2 illustrates some plots for various values of $x = \gamma / (1 - \beta)$. The figure plots $w(\alpha) / c(\cdot)$ for an arbitrarily fixed value of c ; as c in fact varies when the parameters change, the

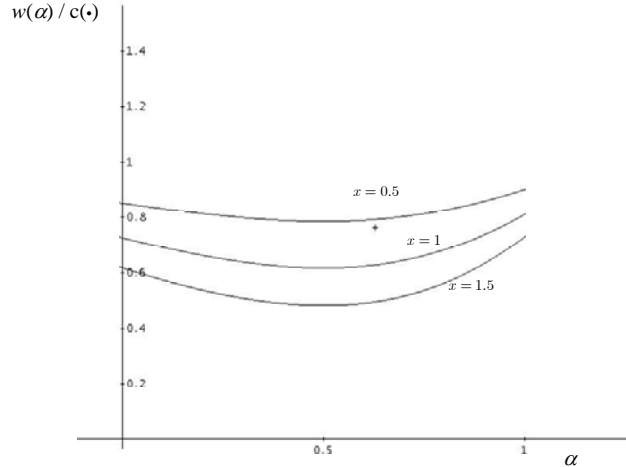


FIGURE 2. Numerical plots of the wage function: Parameter values are $\bar{v} = 0.5$, $c = 1$, and three values of $\gamma / (1 - \beta)$.

⁸ This is not an innocuous assumption: it requires that $f_n(1 - s) = n(s)f_s(s)$ for all s . More gen-

apparent shifts in the function as $\gamma / (1 - \beta)$ changes are uninformative. The functions have their minima at $\alpha = 0.5$, because given our simple assignment this is the point at which $s = \bar{v}$. When s and \bar{v} are equal, the firm is indifferent about the value of α ; it will therefore choose to hire the workers with the lowest wage. The figure also shows that the wage function is steeper when α is high than when it is low, and this asymmetry is stronger when $\gamma / (1 - \beta)$ is larger. The intuition is straightforward: when s is small, the marginal product of λ is greater, so changes in α have a greater effect on output. Finally, note that if s is distributed reasonably uniformly on the unit interval, then the inter-firm variance of wages among the group of small firms (say, those with $\alpha > 2/3$), exceeds the variance among large firms ($\alpha < 2/3$) which in turn exceeds the variance of firms of intermediate size.

Figure 3 plots the relationship between firm size and wages for two values of A and c . Managerial ability s , is zero at the left-hand end of each curve, and equal to one at the right. Employment in every firm is proportional to $(A/c)^{1/\beta}$ while from (8) the wage function is proportional to c but independent of A . Thus, if the measure of available workers, m , is fixed, any technological advance that increases A is matched by an equal increase in c ; the wage function shifts upward keeping employment constant in every firm. In contrast, if the supply of workers is perfectly elastic, then c is

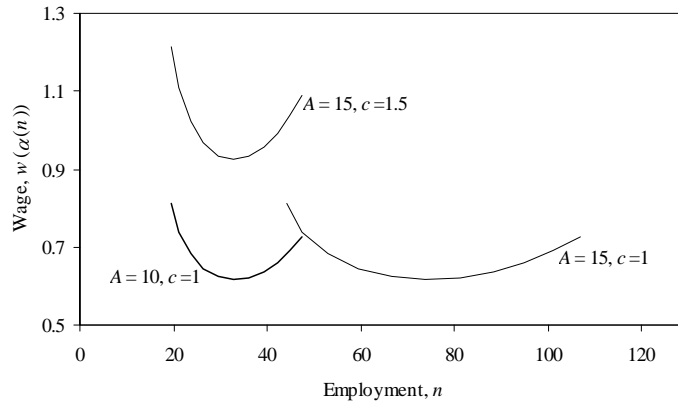


FIGURE 3. Numerical plots of the wage as a function of firm size. Parameter values are $\bar{v}=0.5$, $\gamma=0.5$, $\beta=0.5$, and two values of A and c .

erally, the assignment takes the form $s = g(\alpha)$ for some decreasing function g that depends on the distributions of abilities and worker characteristics.

unchanged and the wage-employment curve shifts to the right.

4. Coordination and the Span of Control Model

When there is no variation in α , the production technology is a particular implementation of Lucas' (1978) span of control model. We know from Lucas that, in this case, the most able managers employ more workers and their firms earn greater profits. We also know that output per worker is invariant to managerial ability, because good managers exploit their skills by increasing the number of workers they manage rather than maintaining higher productivity of workers. These results need not hold when α varies across workers and firms. This section shows how the implications of the present model differ from the well-known results in Lucas, continuing with the assignment $s = 1 - \alpha$.

- *Managerial ability and firm size.* A key result in Lucas is that the best managers operate the largest firms as measured by employment. Somewhat surprisingly, this result does not always hold in the present model. Using (4), (5), and (8), firm size satisfies

$$n(s) = \frac{A^{1/\beta}(1-\beta)^{1/\beta} \left(s(1-s) + (1-s)\bar{v} \right)^{1/\beta}}{c^{1/\beta} \exp \left\{ \frac{\gamma\sqrt{\bar{v}} \tan^{-1} \left(\frac{\bar{v}-2s}{\sqrt{\bar{v}(4-\bar{v})}} \right)}{\beta(1-\beta)\sqrt{4-\bar{v}}} + \frac{\gamma \ln \left((1-s)\bar{v} + s^2 \right)}{2\beta(1-\beta)} \right\}}, \quad (9)$$

from which it is easy to verify that

$$\text{sgn} \{n'(s)\} = \text{sgn} \{ \bar{v}\beta - s(2\beta - 1) \}. \quad (10)$$

If $\beta \leq (2 - \bar{v})^{-1}$, it follows that $n(s)$ is a strictly increasing function for all $s \in [0, 1]$, so in this case the model replicates Lucas' central result. As $1 - \beta$ is the elasticity of output with respect to employment, we would typically expect $\beta > 1/2$, so this restriction on β usually holds. But when it does not, there exists an $s^* < 1$ such that $n(s)$ is increasing for any $s \in [0, s^*)$ but decreasing in s for any $s \in (s^*, 1]$. Figure 4 plots some examples.⁹ When α is invariant and the wage is therefore constant, the standard span of control model predicts that employment is strictly increasing in s ;

⁹ One's intuition should be that n rises more rapidly with managerial ability when β is small, and the plots are consistent with intuition.

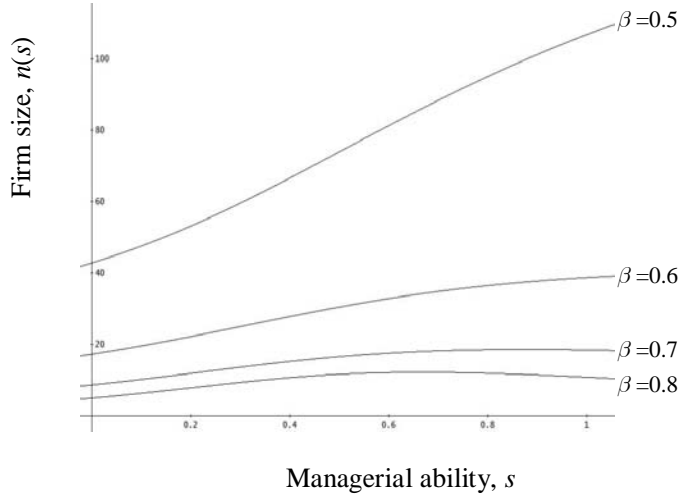


FIGURE 4. Managerial ability and optimal employment. Parameter values are $\bar{v}=0.5$, $\gamma=0.5$, $c=1$, $A=15$, and four values of β .

at the same time, a decline [rise] in the wage induces an increase [a reduction] in firm size. Thus, in the present model, employment rises quite rapidly with managerial ability when $s \leq \bar{v}$. When $s > \bar{v}$, increases in employment in response to increases in managerial ability are offset by the countervailing force of a wage that rises as α declines. For sufficiently large s , the countervailing force may dominate, and this is more likely when the elasticity, $1 - \beta$, is small.

- *Managerial ability and firm profits.* Despite the potential non-monotonic relationship between managerial ability and optimal employment, it is easy to verify that expected profits are an increasing function of s regardless of the value of β and \bar{v} . This is an intuitive result: if it were not, high-ability entrepreneurs could match the firm size of low-ability entrepreneurs and still outperform them. Figure 5 illustrates. The plots show convex functions, but this depends on parameter values. If \bar{v} is small and β sufficiently larger than γ , the function may be s shaped, with a concave portion appearing when $s > \sqrt{\bar{v}\beta / (\beta - \gamma)}$.

- *Labor productivity and firm size.* Expected output per worker is

$$n^{-1}E[Y | s] = A[n(s)]^{-\beta} (s(1-s) + s\bar{v})^{\gamma}, \quad (11)$$

where $n(s)$ is given by (9). It is easy to verify that

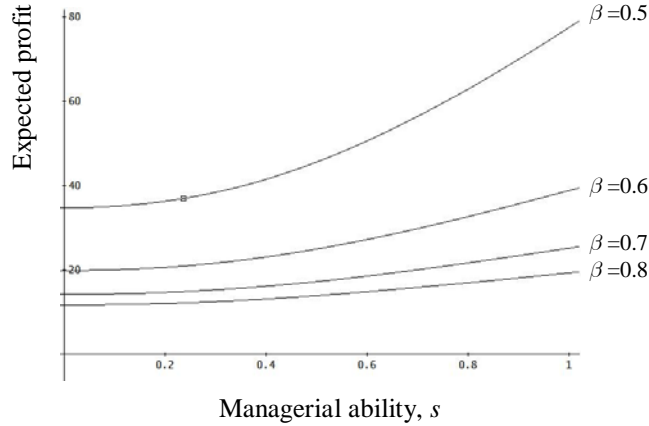


FIGURE 5. Managerial ability and expected profits. Parameter values are $\bar{v}=0.5$, $\gamma=0.5$, $c=1$, $A=15$, and four values of β .

$$\text{sgn} \left\{ \frac{d[n^{-1}E[Y | s]]}{ds} \right\} = \text{sgn} \{s - \bar{v}\}, \quad (12)$$

so output per worker exhibits a U-shape with respect to managerial ability. Except for the (infrequent) cases in which $\beta > (2 - \bar{v})^{-1}$, this implies that labor productivity in the largest and smallest firms, and in the youngest and oldest firms, is greater than in firms of intermediate age and size.

5. Variations in Ability and Self-Employment Choices

In this section, we allow ability, v , to vary across workers in addition to with their character, α . However, at the time workers are hired, managers are unable to observe ability, which is revealed only after the passage of some time. If worker i is revealed to have low [high] ability, he is revealed to be less [more] valuable than is merited by his initial wage. A firm would then prefer to dismiss a low-ability worker or reduce his wage, and it may be willing to increase the wage it pays to high-ability workers if doing so is necessary to retain them. In this subsection, we explore how information about worker ability affects the maximum wage a firm is willing to pay, focusing in particular on how revisions in the wage vary with s .

We shall first proceed with some simplifying assumptions, leaving a more complete analysis to a numerical example. First, we restrict attention for the moment to the

special case in which all workers in the economy have the same character, $\alpha = \bar{\alpha}$, and we let \bar{w} denote the wage paid to a worker whose ability is known only to be a draw from $F_v(v)$, with mean \bar{v} . Second, we assume that firms initially hire workers as though they all had ability $v_i = \bar{v}$, so that we can make use of the solution $n = \mu \bar{\lambda}^{\gamma/\beta} \equiv \mu \left((1 - \bar{\alpha})s + \bar{\alpha}\bar{v} \right)^{\gamma/\beta}$, where $\mu = \left((1 - \beta)\bar{w} \right)^{1/\beta}$.¹⁰ After hiring under the assumption that $\lambda = \bar{\lambda}$, firms observe individual abilities and they note that this changes their realization of λ .

To see how realizations of an individual's ability affects revenues, consider first a small departure of v_i from \bar{v} . That is, while workers were hired under the assumption that $\bar{\lambda} = (1 - \bar{\alpha})s + \bar{\alpha}\bar{v}$, we then define

$$\lambda = (1 - \bar{\alpha})s + \bar{\alpha} \left(\frac{(n-1)\bar{v} + v_i}{n} \right), \quad (13)$$

and allow v_i to change. Substituting $n = \mu \bar{\lambda}^{\gamma/\beta}$ into the expected revenue function,

$$E[R | \lambda] = \int_0^{\lambda^\gamma \mu^{-\beta} \bar{\lambda}^{-\gamma}} \frac{2A \left(\mu \bar{\lambda}^{\gamma/\beta} \right)^{1-\beta} x}{\lambda^\gamma} dx. \quad (14)$$

and differentiating with respect to individual i 's ability yields

$$\left. \frac{dE[R | \lambda]}{dv_i} \right|_{v_i = \bar{v}} = \kappa \bar{\lambda}^{-(\gamma+1+2\gamma/\beta)}, \quad (15)$$

where κ is a positive constant. The absolute change in revenues occasioned by a change in v_i is larger when s is small, reflecting the fact that the marginal productivity of employee ability is higher when managers are less able. As managerial ability is inversely related to firm size and age, it follows that the variance of offered wages after observing ability is larger in smaller and younger firms, consistent with the evidence summarized in regularity R1.

PROPOSITION 2. After individual abilities are revealed, the variance of wages is greater in smaller and younger firms.

¹⁰ The correct approach expresses expected output as the n -fold convolution of individual worker outputs. With n endogenous (and treated as a continuous variable in this paper), the correct approach is intractable.

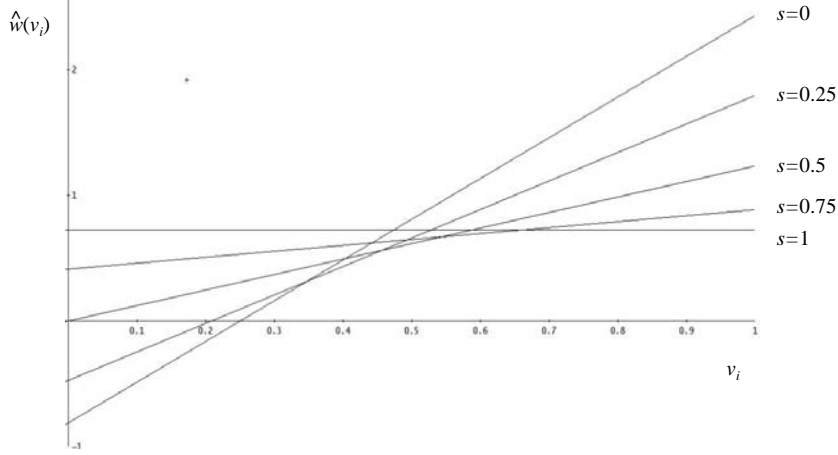


FIGURE 6. The maximum offered wage as a function of v_i for firms with differing managerial ability. Parameter values are $\bar{v}=0.5$, $\gamma=0.5$, $\beta=0.5$, $A=10$, and $c=1$.

Proposition 2 is Wages are more sensitive to individual ability when α is large. As α is larger in smaller firms, one would expect that the inverse relationship just obtained between the variance of wages and firm size would be enhanced when we allow α to vary in a manner consistent with optimal behavior. Complicating this effect, however, is the fact that wages adjust with α . We examine the net effects of these countervailing forces by means of our numerical example. Define

$$\Delta(v_i | s) = E[R | \lambda(v_i, s, \alpha(s))] - E[R | \bar{\lambda}(s, \alpha(s))] \quad (16)$$

as the change in revenue induced by a change in employee i 's ability from \bar{v} to v_i , when α varies across individuals, individuals are matched to firms according to $s = 1 - \alpha$, and both $w(\alpha)$ and n are determined endogenously. Equation (16) compares the revenues that are obtained from having individual i with his revealed abilities, or replacing him with a new employee of the same character but with ability \bar{v} .

The maximum wage the firm is willing to pay after observing v_i is

$$\hat{w}(v_i | s) = \Delta(v_i | s) + w(\alpha(s)). \quad (17)$$

That is, the firm will pay a wage that differs from that paid to a worker with average ability by at most the difference between the revenues that are expected from employing each of the two workers. Figure 6 plots examples of equation (17) for firms with differing managerial abilities. When $s = 1$, individual i 's ability has no bearing

on his wage, because in this case individual i has character $\alpha = 0$ and simply does what he is instructed to do. Hence, in the largest firms, there is no variance in wages. As s declines (and hence as firm size and age decline), the sensitivity of the wage to individual ability increases, just as in the special case analyzed previously. Perhaps surprisingly, the maximum wage the firm is willing to offer to the least able workers becomes negative for the least able managers (and hence in the smallest and youngest firms).¹¹ This result turns on the difference between the marginal products of ability for fixed n and of n for fixed ability. The former generally exceeds the latter, but it is the latter that determines the wage that is paid to a new hire.

The analysis of entry into self-employment is now straightforward. After experiencing a change in wages, the worker reassesses his options. He may remain in his current job, seek new employment, or enter self-employment. However, seeking new employment is unattractive. Firms only hire workers of identical character, so employment prospects exist only with managers having the same ability as the worker's current manager. If his current wage is observable to outside firms, then the worker can do no better elsewhere. If his current wage is not observable, there is an adverse selection problem. For any wage offered by an outside firm, the only applicants for the position are workers who will do better by moving. But these are workers that are unprofitable to hire at the offered wage, so the offered wage is reduced. Then, the only wage offered is equal to the wage paid to a worker with ability known to be zero, and no workers move. Thus, the only alternative to remaining with the current employer is to enter self-employment.

If the agent enters self-employment, he draws a random ability from the distribution $F_s(s | a = 0, v_i)$, where F_s is decreasing in v_i . Suppose that entry into self-employment requires financing such that the agent retains only a fraction, ϕ , of the profits, and assume that $\phi E[\pi | a = 0, \bar{v}] < w(\alpha | v = \bar{v})$ for all α ; this ensures that wage employment was initially more attractive than self-employment. After learning v_i , workers enter self-employment if $\phi E[\pi | a = 0, v_i] > \hat{w}(v_i | s)$. Figure 7 superimposes the function $\phi E[\pi | a = 0, v_i]$ onto the maximum offered wages from Figure 6 for the simple case where each agent's ability as a manager is identical to his ability as a worker. The profit function is convex and bounded above zero for all abilities. Unsurprisingly, the insubordinate -- low-ability workers in small and young firms -- choose self-employment over continued wage work. But in addition, some high-ability workers in

¹¹ That is, workers who prove to be sufficiently insubordinate may be driven out of the labor market even in the absence of alternatives offering positive income.

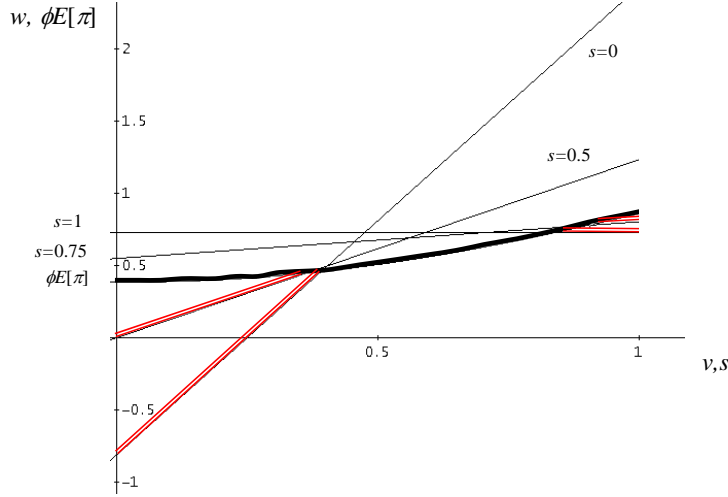


FIGURE 7. The wages and expected self-employment earnings as a function of v_i . Parameter values are $\bar{v}=0.5$, $\gamma=0.5$, $\beta=0.5$, $A=10$, $c=1$, $\phi=0.025$, and $s_i=v_i$. The bold line is expected self-employment earnings. Double-lines represent segments of wages for which self-employment is preferred.

large firms choose self-employment because they are low- α types and therefore are not rewarded for their ability. These workers can make better use of their ability by managing their own firms. Thus, given the correlation between age and size, the self-employed are drawn from the tails of the firm age and size distributions, which is consistent with the evidence in Elfenbein *et al.* (2008)

PROPOSITION 3. *Entry into self-employment is more likely from the tails of the employer size and distributions, and especially from small and young firms.*

Selection into self-employment induces a somewhat subtle relationship between wages and subsequent entry into self-employment than was reported in Elfenbein *et al.* The self-employed that came from small firms were previously low paid, and had previously suffered a decline in their wages. The self-employed that came from large firms were previously more highly paid than their colleagues and had enjoyed an increase, albeit a modest one, in their wages. However, they were not the highest paid workers overall, which position is occupied by workers with initiative who are employed by small firms.

PROPOSITION 4. (i) *Self-employed individuals from small firms previously were low wage earners and had suffered a decline in wages* (ii) *Self-employed individuals from large firms previously were high wage earners among employees of large firms, but intermediate earners among all workers; they had previously experienced increases in their wage.*

6. Conclusions

We have modeled a simple production technology in which managers must coordinate the activities of employees, and their success depends on both the manager's ability and the type of workers he employs. We have shown how this simple framework provides a unified explanation for the U-shaped relationship between firm age and wages, and for some recent and somewhat subtle relationships between firm age, size, and entry into entrepreneurship. The model builds on some well-established ideas, most notably on how limited control of employees can limit the size of the firm [cf. Williamson (1967) and Lucas (1978)]. Nonetheless, the ability of this paper's implementation of these ideas to explain some recent empirical results on earnings and employment choices suggests there remains much more that we might learn from continuing to model production technologies with explicit managerial functions.

References

- Abowd, John M., Francis Kramarz, and David N. Margolis (1999): "High wage workers and high wage firms ." *Econometrica*, 67(2): 251-333.
- Åstebro, Thomas, and Irwin Bernhardt (2005): "The winner's curse of human capital." *Small Business Economics*, 24(1):63-78.
- Åstebro, Thomas, Jing Chen, and Peter Thompson (2008): "Stars and misfits: A theory of occupational choice." Working paper, Florida International University.
- Bates, Timothy (1990): "Entrepreneur human capital inputs and small business longevity." *Review of Economics and Statistics*, 72(4): 551-559.
- Becker, Gary S., and Kevin M. Murphy (1992): "The division of labor, coordination costs, and knowledge." *Quarterly Journal of Economics*, 107(4):1137-1160.
- Berger, Allen N., Nathan H. Miller, Mitchell A. Petersen, Raghuram G. Rajan, and Jeremy C. Stein (2002): "Does function follow organizational form? Evidence from the lending practices of large and small banks." NBER working paper no. 8752.
- Blinder, Alan, and Don Choi (1990): "A shred of evidence on theories of wage stickiness." *Quarterly Journal of Economics*, 105(4):1003-1016.
- Boden, R.J. (1996): "Gender and self-employment selection: An empirical assess-

- ment.” *Journal of Socioeconomics*, 25:671-682.
- Brown, Charles, and James L. Medoff (2003): “Firm age and wages.” *Journal of Labor Economics*, 21(3):677-697.
- Calvo, Guillermo, and Stanislaw Wellisz (1978): “Supervision, loss of control, and the optimum size of the firm.” *Journal of Political Economy*, 85(5):943-952.
- Colombo, Massimo G., Marco Delmastro, and Luca Grilli (2004): “Entrepreneurs' human capital and the start-up size of new technology-based firms.” *International Journal of Industrial Organization*, 22(8-9):1183-1211.
- Davis, Steven, and John Haltiwanger (1991): “Wage dispersion between and within U.S. manufacturing plants, 1963-86.” *Brookings Papers on Economic Activity: Microeconomics*, no. 1: 115-180.
- Dunne, Timothy, and Mark Roberts (1990a): “Plant, firm, and industry wage variations.” Manuscript.
- Dunne, Timothy, and Mark Roberts (1990b): “Wages and risk of plant closing.” Manuscript.
- Dunne, Timothy, Mark J. Roberts, and Larry Samuelson (1988): “Patterns of Firm Entry and Exit in U.S. Manufacturing Industries.” *RAND Journal of Economics*, 19(4):495-515.
- Dunne, Timothy, Mark J. Roberts, and Larry Samuelson (1989): “The Growth and Failure of U.S. Manufacturing Plants.” *Quarterly Journal of Economics*, 104(4):671-698.
- Elfenbein, Daniel W., Barton H. Hamilton, and Todd R. Zenger (2008): “The entrepreneurial spawning of scientists and engineers: Stars, slugs, and the small firm effect.” Working paper, Ohlin Business School, Washington University in St. Louis.
- Hamilton, Barton (2000): “Does entrepreneurship pay? An empirical analysis of the returns to self-employment.” *Journal of Political Economy*, 108(3):604-631.
- Kahnemann, Daniel, Jack Knetsch, and Richard Thaler (1986): “Fairness as a constraint on profit seeking: Entitlements in the market.” *American Economic Review*, 86(4):728-741.
- Keren, Michael, and David Levhari (1979): “The optimum span of control in a pure hierarchy.” *Management Science*, 25(11):1162-1172.
- Idson, Todd (1996): “Employer size and labor turnover.” *Research in Labor Economics*, 15:273-304.
- Jovanovic, Boyan (1994): “Firm formation with heterogeneous management and labor skills.” *Small Business Economics*, 6:185-191.
- Lucas, Robert E. (1978): “On the size distribution of business firms.” *Bell Journal of Economics*, 9:508-523.
- Mirrlees, James A. (1976): “The optimal structure of incentives and authority within an organization.” *Bell Journal of Economics*, 7(1):105-131.

- Moskowitz, Tobias J., and Annette Vissing-Jørgensen (2002): "The returns to entrepreneurial investment: A private equity premium puzzle?" *American Economic Review*, 92(4):745-778.
- Rosen, Harvey, and Paul Willen (2002): "Risk, return, and self-employment." Working paper, Department of Economics, Princeton University.
- Troske, Kenneth (1998): "The worker-establishments characteristics database." In John Haltiwanger and Marilyn Manser (eds.) *Labor Statistics Measurement Issues*. Chicago: University of Chicago Press, pp. 371-404.
- Wagner, Joachim (2004): "Are young and small firms hothouses for nascent entrepreneurs? Evidence from German microdata." Institute of Labor Studies (IZA), Bonn, Discussion Paper no. 989.
- Williamson, Oliver E. (1967): "Hierarchical control and optimum firm size." *Journal of Political Economy*, 75(2):123-138.