Prediction Ability

Katsuya Takii*
Department of Economics
University of Pennsylvania
3718, Locust Walk
Philadelphia, PA
19104 U.S.A.
E-mail ktakii@ssc.upenn.edu

Abstract

This paper models managerial human capital as the ability to predict future events. My assignment model shows that a manager who has high prediction ability goes into a risky industry, because risk increases the marginal productivity of prediction ability. This conclusion contrasts with that of Lucas (1978). In Lucas's model talented managers simply manage bigger firms. The data supports my view: talented B-school graduates choose to work in risky industries, and the correlation between an ability measure and a risk measure is 0.75. The simulated assignment model fits B-school placement data quite well, and a 1 percent increase in the GMAT score of a B-school graduate implies a 158 percent increase in the risk of the firm to which the graduate is assigned.

I also employ a dynamic analysis, which shows that prediction ability increases a firm's expected Tobin's Q and allows a firm to attain a higher expected growth rate. The COMPUSTAT dataset confirms these points as well.

*I am grateful to Andrew Abel, John Core, Jason Cummins, Joao Gomes, Johannes Horner, Boyan Jovanovic, Richard Kihlstrom, Rafael Rob and Masako Ueda for important comments. I also wish to thank seminar participants at the 1999 annual meeting of Econometric Society, the 1998/1999 annual meeting of the Society of Economic Dynamics, the Kansai Macroeconomics Workshop, Kobe University of Commerce, Nagoya City University, Osaka City University, Tokyo University and the University of Pennsylvania.

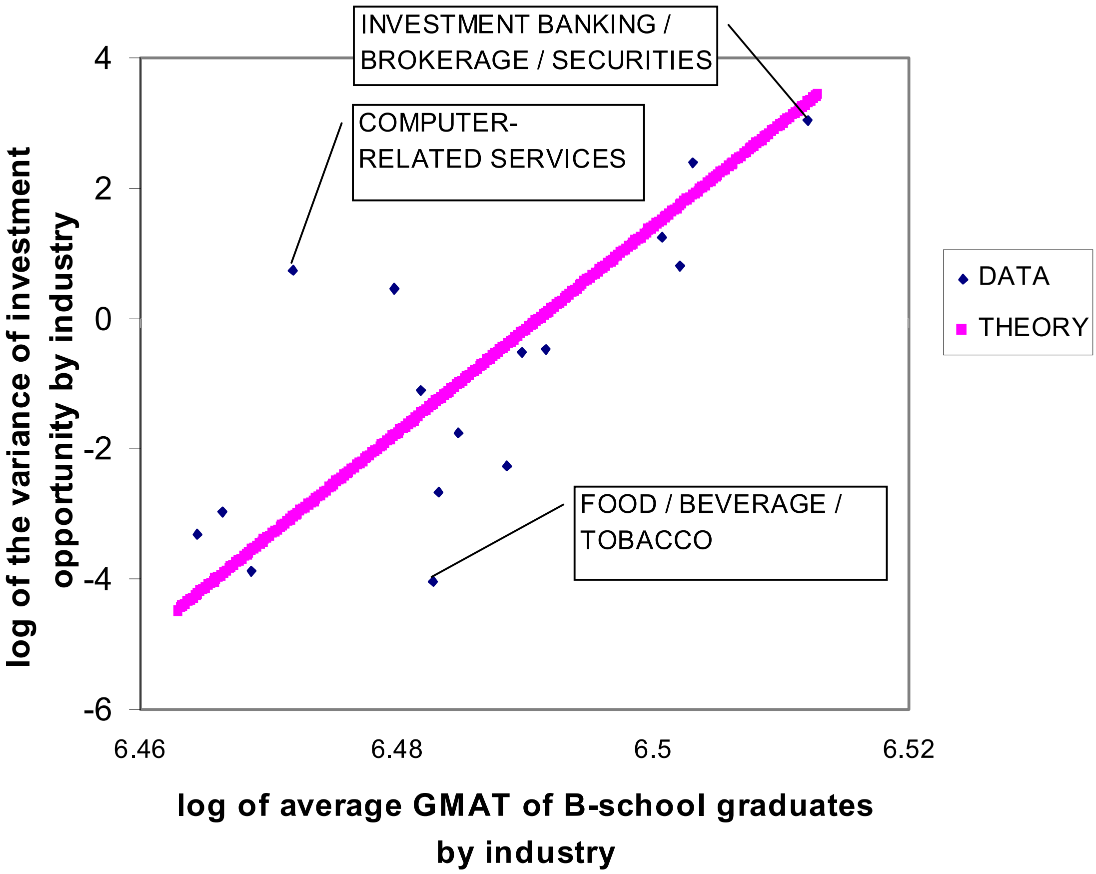


Figure 1: Risk vs Ability

1 Introduction

This paper lays a micro foundation for one important component of a manager's human capital: the ability to predict future events that affect the profits of a firm. The idea is simple. In an uncertain world, managers receive noisy signals about the market the firm faces. It is natural to assume that a good manager discerns better signals than does a bad manager – a good manager is one who predicts accurately. Hence this paper measures a manager's prediction ability by the quality of the information that the manager observes and uses to predict future profitability.

A puzzle: Figure 1 depicts a puzzling phenomenon. *Top B-school graduates get jobs in risky industries.*

The vertical axis measures the risk of an industry. To measure this risk, I calculate the sample variance of a firm's operating income over its net capital stock during the period 1988-1997 and then compute a weighted average of the variance by industry with the firm's net capital stock in 1997 as its weight. I take the logarithm of this weighted average. The data come from COMPUSTAT.[1] The horizontal axis measures the cognitive skill of managers who entered an industry in 1998. On their website, US News & World Report provides placement data of the top 50 business schools in the US. From that site I obtained the average GMAT score of students in a school, the fraction of graduates who entered each industry and the total number of graduates. I, then, estimate the log of a weighted average of GMAT score by industry, using the number of graduates who chose a job in an industry as the weight. The figure shows that a smart graduate prefers to go into a risky industry. The correlation between risk and ability is 0.75.

An answer: My theory predicts that a manager who can predict future investment opportunities goes into a risky industry, *because risk increases the marginal productivity of prediction ability, and therefore a risky firm demands a manager with high prediction ability.* Think about a project which has no risk. Since everyone knows what will happen, additional information has no value. That is, prediction ability is of no use. We need prediction ability because there is uncertainty. If a manager has a clearer vision under uncertainty, his ability is appreciated. In other words, the return to prediction ability is higher when a project is riskier. Hence, as long as a cognitive ability increases prediction ability, it is natural that a talented person chooses a risky industry.

Based on the presented model of prediction ability, I show that there exists a positive assortative assignment equilibrium between firms' risk and managers' prediction ability. The simulation of my assignment model fits B-school placement data quite well. Consider Figure 1 again. The solid line represent the fit of my assignment model. My simulation results predict that a 1 percent increase in the GMAT score of a B-school graduate brings about a 158 percent increase in the risk of the firm to which the graduate is assigned. That is, a slightly higher difference in ability pushes a manager into a much more complex environment.

Related papers: This prediction contrasts with that of Lucas (1978), whose span of control model implies that a talented manager will run a large firm. This may well be true. Murphy (1998) provides robust results that show a strong positive correlation between the size of a firm and the compensation of its CEO. The span of control model, however, is silent about risk. Kihlstrom and Laffont (1979) emphasize the importance of the risk taking behavior of an entrepreneur, but they focus on the attitudes of an entrepreneur but not on an entrepreneur's ability.

[1] For a more detailed description of the data construction, please see Appendix 3.

My view is close to Schultz (1975, 1980). Based on rich empirical studies such as Welch (1971), he emphasizes that education raises the ability "to interpret new information and to decide to reallocate their resources to take advantage of new and better opportunities".[2] He called it "the ability of entrepreneurs to deal with disequilibria".[3] Since my measure evaluates a manager's reaction to future profitability, as I show later, *my measure can be interpreted as a measure of entrepreneurial ability.*[4]

My results also explain certain aspects of CEO compensation. The finance and accounting literatures found not only that firm size has a positive impact on CEO compensation, but also that the existence of investment opportunities has a positive effect on CEO compensation (Smith and Watts [1992] and Gaver and Gaver [1993]). Typically, greater opportunity to invest makes a firm more risky. In fact, researchers sometimes use the variance of total return on a firm as a measure of investment opportunity (Smith and Watts [1992] and Gaver and Gaver [1993]). These literatures usually explain this fact using contract theory: in order to ensure that a manager takes more risk rewards must be higher. My theory provides a different explanation: *Since the results of a talented manager are easier to observe in a risky firm, it is not surprising to see higher rewards for such a manager.*

Other contributions

A dynamic investment model in this paper makes contributions to two other research topics: (1) the empirical estimation of the value of information in a firm and (2) the theory of investment under uncertainty.

Empirical estimation of the value of information in a firm: *I construct an observable measure of a manager's prediction ability.* Although many economists apply the Blackwell theorem (1953) to analyze the economic impact of accurate information, nobody estimates it.[5] I show that, given a particular production function, the quality of information can be estimated by the correlation coefficient between future profitability and current decisions. Since Tobin's Q reflects the future profitability of capital, I construct a measure of prediction ability from the correlation between a firm's future Q and its current growth rate. Using this measure I show that *a manager's prediction ability has a positive impact on a firm's expected Tobin's Q with evidence from the COMPUSTAT dataset.*

Investment under uncertainty: My dynamic investment model shows that *a*

[2] Schultz (1980).

[3] Schultz (1980).

[4] Holmes and Schmitz (1990) also formalize Schultz's view of entrepreneurship. They emphasize the importance of the division of labor based on the comparative advantage of an entreprenur.

[5] Marschak and Miyasawa (1968) and Kihlstrom (1984) have excellent discussions about Blackwell's theorem. Arthey and Levin (1998) intensively investigate the ordering of information structures in a particular class of utility functions and distribution functions.

manager who expects to receive more valuable information attains a higher expected growth rate. If current investment decreases future adjustment costs, a good manager has incentive to invest more on average in order to prepare for future investment opportunities.[6] This result is in contrast to that of Demers (1991), who argued that, when investment is irreversible, the more valuable information a manager expects to receive, the less on average he invests. *Using my measure of prediction ability, the COMPUSTAT data conforms to my theory.*

The organization of this paper

This paper is organized as follows. The next section develops a basic static model that describes why information is valuable. Here I will introduce a measure of prediction ability and show that risk and ability are complementary. The complementarity provides the basis for an assignment model. Section 3 provides my assignment theory based on the model of prediction ability. Here I will show that a talented person will be assigned to a risky industry. I will also provide some simulation results. Section 4 formalizes the dynamic investment model. Section 5 extends the results of the static model to the dynamic model under i.i.d. random shocks. Section 6 extends the measure of the static model so that the analysis applies under a Markov process with stationary transitions. Section 7 provides assumptions that facilitate empirical application of the model. Here I discuss the issue of heterogeneous managers. Section 8 offers some empirical evidence for the effect of accurate information on expected Tobin's Q and the expected growth rate. The last section concludes the main results and discuss possible extensions.

2 Preliminaries: The Complementarity Between Risk and Prediction Ability

In this section I develop a static investment decision model that describes why information is useful and how it is captured.[7] The model will show that risk and prediction ability are complementary, which will provide the groundwork for the assignment analysis in section 3.

[6]The logic is clearer if one assumes that the adjustment cost represents the cost of training new workers on new machines: a superior manager has more incentive to keep more skilled workers to prepare for future investment opportunities.

[7]After completing this section, I found Nelson (1961). The structure of the model in this section is the same as that in Nelson (1961). Although his emphasis is different from mine, he established some of the results in this section.

2.1 Static formulation

Consider a manager's investment decision problem with investment adjustment costs under uncertainty. I focus on an isolated firm's behavior in this section.

Suppose that a representative firm solves the following problem:

$$\pi\left(E\left(z|s\right)\right)=\max_{I}\left[\frac{\int zdG\left(z|s\right)I}{r}-I-\frac{AI^{2}}{2}\right], \tag{1}$$

where r is the interest rate, I is the amount of investment, z is a random shock, s is a signal and $G\left(z|s\right)$ is the conditional distribution of z given s. The term $\left[\int zdG\left(z|s\right)I\right]/r$ is the present value of the expected revenue from investment and $I+AI^{2}/2$ is a cost of investment. Specifically, I is investment expenditure and $AI^{2}/2$ is the associated adjustment cost. I assume that the price of investment is 1 and that $z\geq r$. A manager observes a signal s, infers z and decides how much he will produce. It can be shown that the profit function is given by

$$\pi\left(E\left(z|s\right)\right)=\frac{\left[\int zdG\left(z|s\right)-r\right]^{2}}{2r^{2}A}. \tag{2}$$

Notice that the profit function is a convex function of the conditional expectation of the random shock. This is key to understanding why prediction ability brings about more profit. The following extreme example explains the importance of the convexity.

2.2 Extreme case

Now assume that z takes only two values $\underline{z}$ and $\bar{z}$, where each occurs with probability 1/2. Suppose that the joint distribution is given by $G\left(z,\ s_{g},\ s_{b}\right)$. Suppose that Mr. Gates observes a signal s_{g} that perfectly predicts the realization of the random shock. Hence,

$$\begin{aligned}\int zdG^{G}\left(z|s_{g}\right) &= \underline{z},\ if\ \ \underline{z}\ \ occurs,\\ &= \bar{z},\ if\ \ \bar{z}\ \ occurs,\end{aligned}$$

where $G^{G}\left(z|s_{g}\right)$ is the conditional distribution given s_{g}. On the other hand, suppose that Mr. Bean observes a signal s_{b} that has no predictive power. Then

$$\int zdG^{B}\left(z|s_{b}\right)=\frac{\underline{z}+\bar{z}}{2}\ always,$$

where $G^{B}\left(z|s_{b}\right)$ is the conditional distribution given s_{b}. Figure 2 illustrates the two managers' profit functions.

The key point is that each profit function is a convex function of the conditional expectation of the random shock. Gates can increase his investment when he expects that z is high and decrease it when he expects that z is low. On the other hand, Bean cannot exploit this benefit, since his signal does not reveal anything about z. Figure 2 shows that since the profit function is convex, Gates can make more profit than Bean on average. The difference between Gate's expected profits and Bean's represents the value of information.[8]

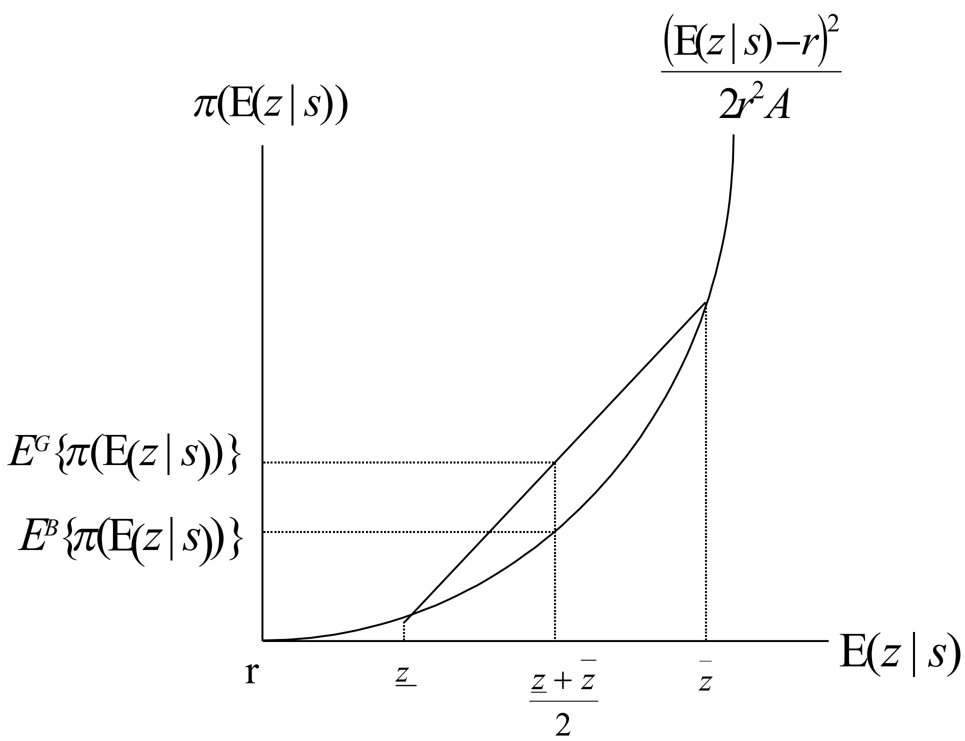


Figure2: The Value of Information

2.3 The measure of ability

In order to analyze the effect of prediction ability under a more general information structure, I need to construct a suitable measure to capture the value of the signal.

[8]The importance of the convexity of the profit function in investment problems is emphasized by Hartman (1972) and Abel (1983).

The above extreme example indicates that the more the conditional expectation of a random shock given the signal varies, then the more accurate the signal will be. I will show that the measure developed in this paper has this property. The measure is as follows.

Definition 1 *The basic measure of prediction ability, h, is defined by*

$$h \equiv 1 - \frac{\int Var\left(z|s\right) dG_s\left(s\right)}{\sigma_z^2}, \tag{3}$$

where $\sigma_z^2 = \int \left(z - \int z dG_z\left(z\right)\right)^2 G_z\left(z\right)$, $Var\left(z|s\right) = \int \left\{z - \int z dG\left(z|s\right)\right\}^2 dG\left(z|s\right)$, *and* $G_z\left(z\right)$ *and* $G_s\left(s\right)$ *are the marginal distribution of* z *and* s, *respectively.*[9]

This measure captures how well a signal is able to predict z on average. The unconditional variance, σ_z^2, is just an adjustment factor that makes it possible to compare prediction ability under different environments. I subtract $\left[\int Var\left(z|s\right) dG_s\left(s\right)\right]/\sigma_z^2$ from 1, which is the maximum value that $\left[\int Var\left(z|s\right) dG_s\left(s\right)\right]/\sigma_z^2$ can attain. In this way, the measure attains its highest value for the best manager. Thus, if a manager can perfectly predict z, then $h = 1$; if the signal is useless for prediction, then $h = 0$. This is my proxy for human capital in this paper.

Example: Suppose $z = z^e + s_g + s_b$, where s_g and s_b are independent of each other. Assume that $E\left(s_i\right) = 0$, $Var\left(s_i\right) = \sigma_i^2$, where $i = g$ or b, and $\sigma_g^2 > \sigma_b^2$. Assume that a manager can observe only one signal, say s_g. For this manager, s_b is just noise. In this case,

$$h = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_b^2}. \tag{4}$$

Since everyone knows the unconditional variance of z, $\sigma_g^2 + \sigma_b^2$, this measure implies that the variance of the good signal explains a large proportion of the unconditional variance of z. In other words, the noisy term becomes insignificant when one observes a good signal.

The following theorem provides an important property of this measure.

Theorem 1 *The above basic measure of prediction ability can be written as follows*

$$h = \frac{Var\left(E(z|s)\right)}{\sigma_z^2}, \tag{5}$$

where $Var\left(E\left(z|s\right)\right) = \int \left[\int zG\left(z|s\right) - \int z dG_z\left(z\right)\right]^2 dG_s\left(s\right)$.

[9]This measure can be also found in Nelson (1961).

Proof. Using the identity equation,

$$\sigma_z^2 = E\left(Var\left(z|s\right)\right) + Var\left(E\left(z|s\right)\right), \tag{6}$$

the result is immediate. ■

This new expression clearly shows that the measure has a high value when the variance of the conditional expectation is high. This is exactly the property that the previous extreme example indicates.

2.4 The complementarity between σ_z^2 and h

Using this measure I can analyze a more general information structure. The next theorem shows that prediction ability has a positive effect on expected profit, and that it can be estimated by the correlation between z and I. It will also show that σ_z^2 and h are complements.

Theorem 2 *Expected profit is an increasing function of prediction ability given by*

$$V\left(h : \sigma_z^2\right) \equiv \int \pi\left(E\left(z|s\right)\right) dG_s\left(s\right) = \frac{\left(z^e - r\right)^2 + \sigma_z^2 h}{2r^2 A}, \tag{7}$$

where $z^e = \int zG_z\left(z\right)$. *Moreover prediction ability* h *can be estimated by*

$$h = \left(\rho_{zI}\right)^2, \ \rho_{zI} \geq 0, \tag{8}$$

where $\rho_{zI} = \frac{\int [z - z^e][I(s) - I^e] dG(z,\, s)}{\sqrt{\int [z - z^e]^2 dG_z(z) \int [I(s) - I^e]^2 dG_s(s)}}$ *and* $I^e = \int I\left(s\right) dG_s\left(s\right)$.

Proof. The proof is similar to the proof of Theorem 10 in Appendix 1. I do not repeat it here. ■

The theorem shows that expected profit is increasing in prediction ability. The theorem has four important implications.

1. All the effects of the noisy signal in theorem 2 are captured by h. This means that the more accurate the information in the sense of Blackwell, the larger the value of h. However, a higher h does not imply more accurate information in the Blackwell sense.

2. Prediction ability has a strictly increasing relationship with the correlation coefficient between investment and the random shock. Since a good manager has an accurate signal with which to predict z, he knows the optimal time at which to invest. That is, a good manager can increase his investment during a boom and reduce it during a recession. This is the basic value of prediction ability.

3. Prediction ability h has a stronger effect on expected profits when σ_z^2 is larger. That is, σ_z^2 and h are complements. This complementarity brings a positive assortative assignment between σ_z^2 and h in the following assignment model.

4. Prediction ability h has a larger effect on expected profit, when A is small. Since the advantage of a good manager is to time investment well, if A is small, a good manager can easily change his decision and take advantage of his ability. This observation provides intuition into why a manger attains a high expected growth rate in the dynamic model. I will explain this in section 4.

3 The Equilibrium Assignment of Ability to Risk

Consider the labor market for managers. To which firm (or industry) does a good manager go, and how much will a good manager earn? I will show that a person who has good prediction ability not only becomes a manager, but also prefers to work in a risky firm (industry). My proof is based on assignment theory. Because σ_z^2 and h are complements, assignment theory predicts that a talented person (high h) will be assigned to a risky firm.[10]

Suppose that each firm (industry) is characterized by risk σ_z^2 and that each manager is characterized by prediction ability h. Suppose that σ_z^2 is distributed on $\left[\underline{\sigma}_z^2,\ \bar{\sigma}_z^2\right]$ with distribution function $\Psi\left(\sigma_z^2\right)$, where $\Psi' > 0$. For simplicity, suppose that each firm has the same z^e.[11] Assume that h is distributed over $\left[\underline{h},\ \bar{h}\right]$ with a distribution function $\Gamma\left(h\right)$, where $\Gamma' > 0$.

Suppose that a person can obtain a reservation wage of w_l as a worker if he does not become a manager. On the other hand, a firm's owner will get 0 profit if he does not employ a manager. Consider the following problem of the firm which is motivated by theorem 2:[12]

$$\max_{h \in \left[\underline{h},\ \bar{h}\right]} \left\{V\left(h : \sigma_z^2\right) - w\left(h\right)\right\},$$

$$\textit{where } V\left(s : \sigma_z^2\right) = \frac{\left(z^e - r\right)^2 + \sigma_z^2 h}{2r^2 A} \textit{ and}$$

$w\left(\cdot\right)$ is a wage function.

[10]Refer to Koopmans and Beckmann (1957), Becker (1973) or Sattinger (1993).

[11]The result follows even with the more general assumption that $z^e = f\left(\sigma_z^2\right)$, where $f' \geq 0$.

[12]This formulation implies that "manager" in this paper includes analysts. This notion of a manager is suitable to B-school data.

Define an assignment function $a(\cdot)$ so that $\sigma_z^2 = a(h)$. An assignment function is a mapping which dictates which person should be assigned to which firm. Using this assignment function, I will define a positive assortative assignment equilibrium. In this equilibrium, a talented person not only becomes a manager, but also prefers to work in a risky firm.

Definition 2 *A positive assortative assignment equilibrium consists of an assignment function $a(\cdot)$, a wage function $w(\cdot)$ and a unique cutoff point h^* which satisfy:*

1. *A positive assortative assignment condition:*

$$1 - \Gamma(h) = 1 - \Psi(a(h)), \text{ for all } h \geq h^*; \tag{9}$$

2. *A firm's maximization problem must be consistent with the assignment:*

$$\arg \max_{h \in \left[\underline{h}, \bar{h}\right]} \left\{ V\left(h : \sigma_z^2\right) - w(h) \right\} = a^{-1}\left(\sigma_z^2\right), \text{ for all } \sigma_z^2; \tag{10}$$

3. *A cut off point condition:*

$$V(h^* : a(h^*)) = w(h^*) = w_l, \tag{11}$$

and if $h \geq (\leq) h^*$, then a person becomes a manager (a worker). Moreover a firm enters the economy only if $\sigma_z^2 \geq a(h^*)$.

The following theorem not only shows that there exists a positive assortative assignment equilibrium, but also characterizes the manager's wage function.

Theorem 3 *Assume that $\frac{(z^e - r)^2 + \underline{\sigma}_z^2 \underline{h}}{2r^2 A} \leq w_l \leq \frac{(z^e - r)^2 + \bar{\sigma}_z^2 \bar{h}}{2r^2 A}$. Then there exists a positive assortative assignment equilibrium. Moreover a manager's wage satisfies:*

$$w(h) = \int_{h^*}^{h} \frac{a(\tau)}{2r^2 A} d\tau + w_l,$$

where $a(h) = \Psi^{-1}(\Gamma(h))$.

Proof. See Appendix 1. ■

There are two important properties of this equilibrium. First, $a'(h) > 0$. That is, if you have more ability to predict z, then you will be assigned to a firm which has more risk. Second, the wage function is increasing and convex in h. This implies that the manager's wage will be skewed right.

3.1 Fitting a log-uniform example to data

I want to know how an assignment model fits B-school placement data. In order to do so, I need some additional assumptions.

Assumption 1: Suppose that $\Gamma(h)$ and $\Psi(\sigma_z^2)$ are log uniform.
Assumption 2: $h = \delta \times (score)$, where $\delta > 0$ and *score* represents cognitive skill, which is measured by GMAT score or GPA score.

Assumption 1 and 2 imply that the log of GMAT (GPA) score is also log uniform. Consider the figures in Appendix 4. They show estimates of the cumulative distribution of $\log(\sigma_z^2)$ and $\log(score)$ by industry, where σ_z^2 is a weighted average of the variance (of operating income over net capital stock) by industry, and *score* is the average GMAT or GPA score of B-school graduates by industry. I estimate $\log(\sigma_z^2)$ and $\log(score)$ as explained in the introduction. The figures show that the distributions of $\log(\sigma_z^2)$ and $\log(score)$ can be approximated by a uniform distribution.

An assignment function: I want to investigate the fit of the assignment model. First, I need to show the quantitative results of the assignment function. Next, I will investigate the wage function. Let $m_{\log\sigma}$ and $\sigma^2_{\log\sigma}$denote the mean and variance of $\log\sigma_z^2$, respectively, and let $m_{\log(score)}$ and $\sigma^2_{\log(score)}$ denote the mean and variance of $\log(score)$, respectively. The next theorem shows that $\log a(h)|_{h=\delta(score)}$ is an affine transformation of $\log(score)$.

Theorem 4 *Suppose that Assumptions 1 and 2 are satisfied.*[13] *Then*

$$\log a(h)|_{h=\delta(score)} = \xi \log(score) + \theta,$$

where

$$\begin{aligned} \xi &= \sqrt{\frac{\sigma^2_{\log\sigma}}{\sigma^2_{\log(score)}}} \text{ and} \\ \theta &= m_{\log\sigma} - \xi m_{\log(score)}. \end{aligned}$$

Proof. See Appendix 1. ■

Theorem 4 shows that there is a clear relationship between the cognitive ability of a manager and the risk of the industry to which he is assigned. Since it is easy

[13] For the more general case, Assumptions 1 and 2 are not necessary for Theorem 4. For simplicity, I construct an assignment model between prediction ability and risk. But for this empirical study we need an assignment between GMAT (GPA) score and risk. If we assume that prediction ability is an increasing twice differentiable function of GMAT (GPA) score, the log uniformity of the risk distribution and the log uniformity of GMAT (GPA) score is sufficient for Theorem 4.

to estimate ξ and θ from data, it is possible to see the quantitative results of the assignment model.

Table 1 provides simple statistics of the data. It shows quite a high correlation between the risk measure and the ability measures; the correlation between the log of σ_z^2 and the log of an average GMAT (GPA) is 0.75 (0.68).

Table 1: Simple Statistics of the Data

	$\log \sigma_z^2$	log GMAT	log GPA
mean	-0.893	6.486	1.210
standard deviation	2.213	0.014	0.008
correlation coefficient			
$\log \sigma_z^2$	1	0.751**	0.679*
log GMAT		1	0.937***
log GPA			1
ξ		158	267
θ		-1029	-324

* significant at 1 % level.

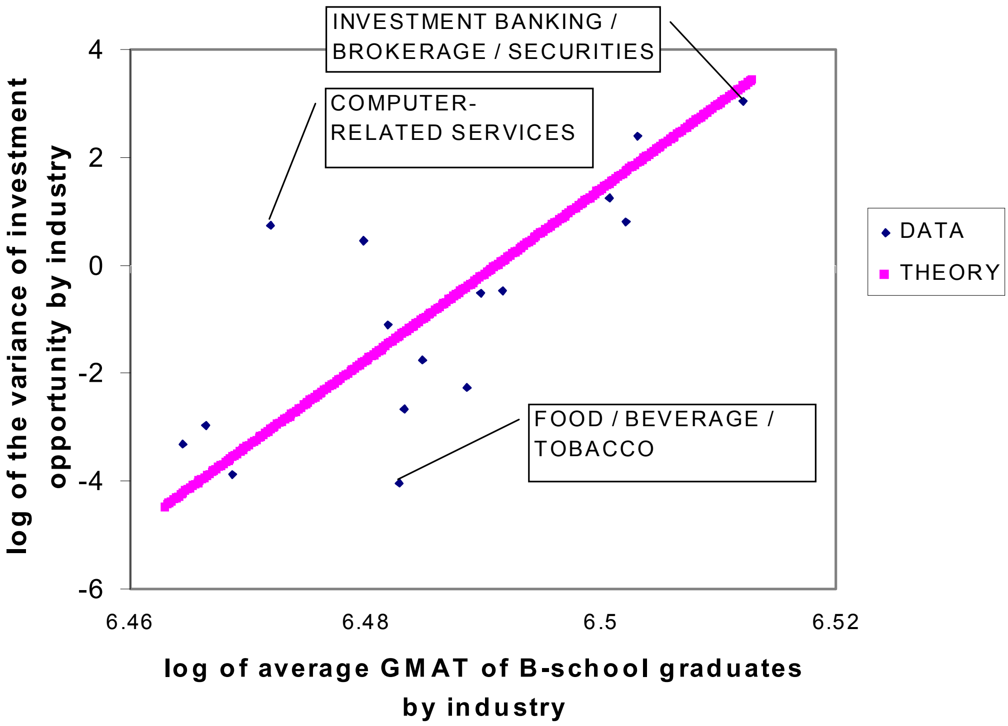
** significant at 0.1 % level.

*** significant at 0.01 % level.

Using estimated ξ and θ, I simulate the model. Figure 3 shows that the assignment model fits the data extremely well. Different measures of cognitive ability, GMAT and GPA, show quite similar results.[14] The simulation results strongly support an assignment between cognitive ability and risk.

Table 1 also shows that the relatively large standard deviation of $\log \sigma_z^2$ as compared to the standard deviation of log GMAT or log GPA induces a high ξ. A high ξ has one important implication: A 1 percent increase in GMAT score of a B-school graduate brings about a 158 percent increase in the risk of the firm to which he is assigned. If I use GPA as a measure of cognitive ability, the result becomes even more extreme: a 1 percent increase in GPA score of a B-school graduate brings about a 267 percent increase in the risk of the firm to which he is assigned. This implies that a small increase in a manager's ability causes a large change in the risk of his work.

[14] I also repeated the same exercise computing the sample variance of the shock over a different time period. I also tried the same exercise using an estimated idiosyncratic shock. Although I do not report them here, the results are quite similar. Hence the conclusion that there is a positive assignment between risk and ability is robust.



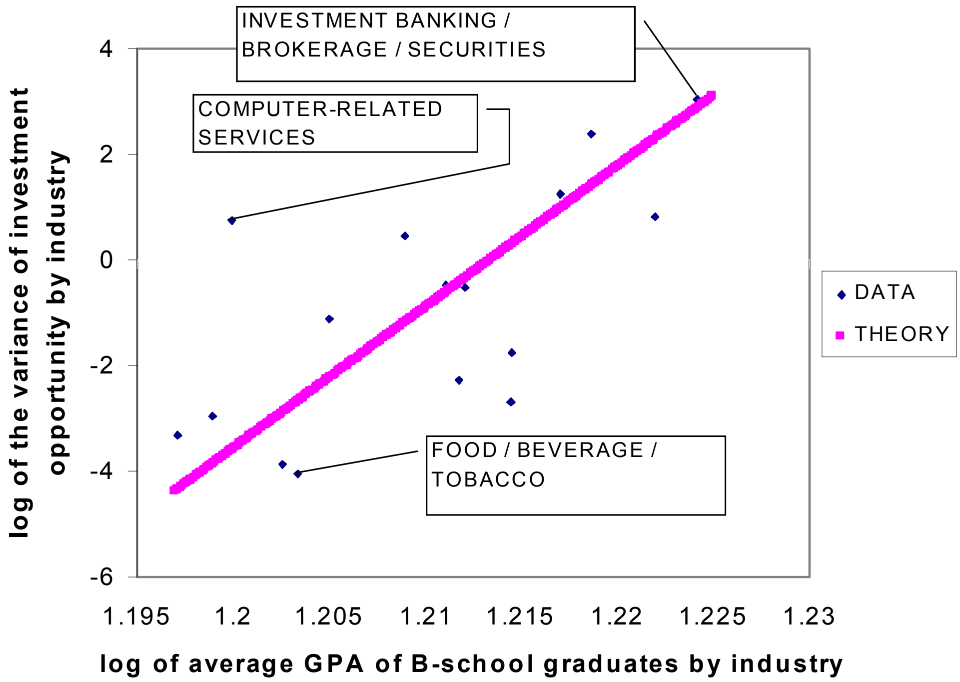


Figure 3: The fit of the log uniform assignment model

The wage function: Now I will examine the quantitative results regarding the wage function in the assignment model. The following theorem derives a closed form solution for the wage function.

Theorem 5 *Suppose that Assumptions 1 and 2 are satisfied. Then*

$$w(h)\,|_{h=\delta(score)} = \frac{\delta e^{\theta}\left[(score)^{\xi+1} - (score^{*})^{\xi+1}\right]}{2r^{2}A(\xi+1)} + w_{l},$$

where

$$\begin{aligned} \xi &= \sqrt{\frac{\sigma^2_{\log\sigma}}{\sigma^2_{\log score}}} \text{ and} \\ \theta &= m_{\log\sigma} - \theta m_{\log score}. \end{aligned}$$

Proof. See Appendix 1. ■

In order to simulate this function I need to estimate r, A, δ, $score^{*}$ and w_l. I simply assume $r = 0.05$. This value has little effect on the results. The adjustment cost parameter, A, according to the first order condition for the investment decision, is given by

$$A = \frac{z^e - r}{rI^e}.$$

I estimate a firm's z^e and I^e by the sample means of shocks and investment over the period 1988-1997. The above equation provides the estimation of the firm's adjustment cost. I calculate the weighted average of the cost where the weight is net capital stock in 1997.

The parameter δ can be estimated by $h/(score)$ where h is average prediction ability and *score* is average GMAT score or GPA score. To estimate the average h, I first estimate a firm's correlation coefficient between investment and operating income/net capital stock over 1988-1997. Then I take a weighted average of the correlation with net capital stock as the weight. Finally, I square the results as implied by Theorem 2.[15] Average *score* at the aggregate level is estimated by a simple average of an industry's average GMAT (GPA).

[15] Estimated h is quite a rough measure of prediction ability at the aggregate level. First, the investment decision will depend not only on the current shock, but also on the future shock. I consider this possibility more seriously in the dynamic context. As I will show in next section, as long as I assume that the shock is close to i.i.d., this measure has no problem. Second, you may wonder whether the measure is biased by a liquidity constraint. This is possible, but less likely. Since h is weighted by the net capital stock, a bigger firm's correlation is weighted more heavily. Since a bigger firm is less likely to suffer from a liquidity constraint, I expect that my measure will roughly reflect the prediction ability of a firm's manager at the aggregate level.

Since the chemical industry has the lowest average GMAT and GPA scores, I choose $score^*$ and w_l from the chemical industry. In other words, I choose $score^*$ and w_l so that the theory and the data coincide for the chemical industry. This does not mean that the average GMAT (GPA) score of the chemical industry must be the cutoff point between a manager and a worker. In fact, Theorem 5 implies that

$$\begin{aligned} w(h)\,|_{h=\delta(score)} &= \frac{\delta e^{\theta}\left[(score)^{\xi+1} - \left(score^{\#}\right)^{\xi+1}\right]}{2A(\xi+1)} + w(h)\,|_{h=\delta\left(score^{\#}\right)}, \\ \forall score^{\#} &\geq score^*. \end{aligned}$$

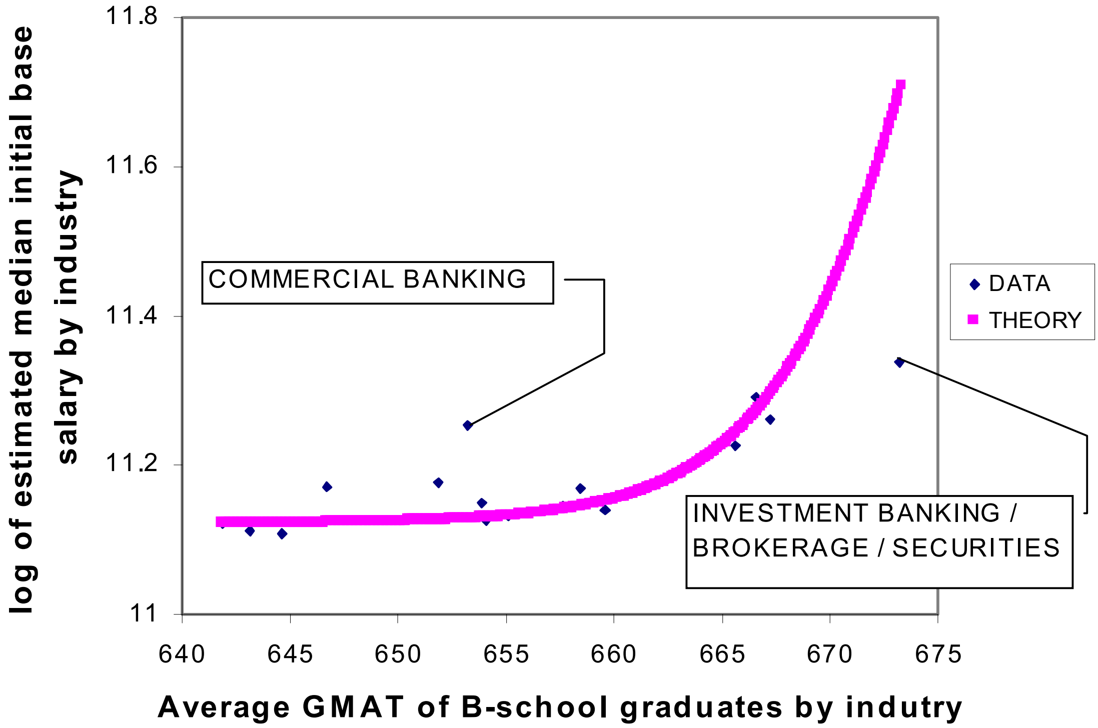
I estimate the median initial year base salary by industry from 1998 placement data at US News & World Report website. It, however, provides the median base salary only for the service and manufacturing industries for each B-school. I assume that each industry's median wage by B-school is the same as the service or manufacturing median wage by B-school. Then I calculate a weighted average of the median wage by industry with the weight taken to be the number of graduates who go into the industry. Table 2 shows my estimate of h, A, δ, $score^{\#}$ and $w(h)\,|_{h=\delta\left(score^{\#}\right)}$.

Table 2: Parameter estimates

	h	A	β	$score^{\#}$	$w(h)\,\vert_{h=\delta\left(score^{\#}\right)}$
$GMAT$	0.0004	2×10^{-7}	6×10^{-7}	642	67642
GPA	0.0004	2×10^{-7}	1×10^{-4}	3.31	67642

A small h implies that the manager's average prediction ability is small. If h is small, δ is as well. This implies that increases in GMAT (GPA) score have little effect on prediction ability. That is, the ability differences among people are small. This, however, does not imply that differences in ability have no economic impact. Remember that ξ is large, and therefore a small ability difference may have a large impact.

In fact, it does. Figure 5 compares the wage function of the assignment theory to the data. The theory fits quite well. In particular, it captures the convexity of the wage/talent relationship. It also shows that a small difference in ability may correspond to a large income difference due to the large value of ξ. Although the difference in ability itself may be small, because a talented person works in a risky industry where his ability can be better exploited, the wage difference is huge.



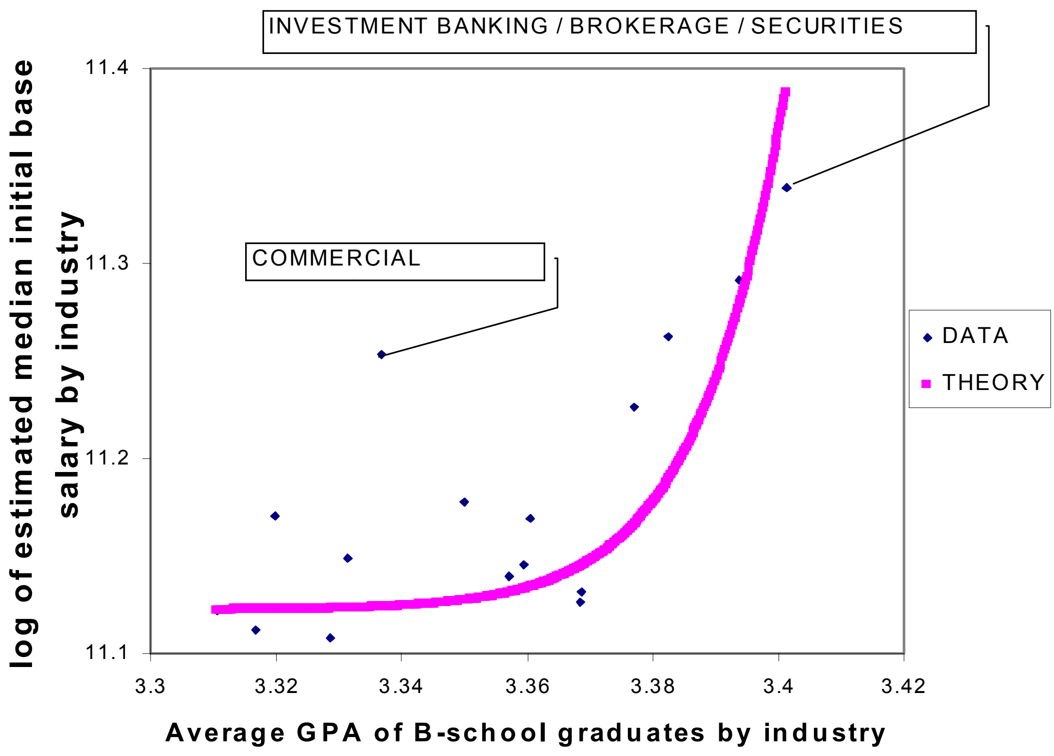


Figure 4: The fit of the log uniform assignment model 2

Although I tried several estimations, the theory often produces a more convex wage function than does the data. Part of the reason may come from my measurement of the compensation data. I assume that each industry's median wage by B-school is the same as the service sector or manufacturing sector by B-school. Typically, the higher the average GMAT (GPA), the larger the fraction of students going into the service sector. There is higher variance of the wage in the service sector than in the manufacturing sector. As long as we expect that graduates will go to the industry that offers the highest wage, we should expect that more accurate data will bring about more convexity in the median wage.

3.2 Size vs Risk

Table 3 compares the results of this section with those of Lucas (1978). The theory presented in this paper predicts that a person who has higher prediction ability can time investment well, and that his ability will be of more use in a more risky firm. This result contrasts with that of Lucas (1978). In his model talented managers can increase the productivity of a firm, and this ability is appreciated more by bigger firms.

Table 3: Comparison between Lucas (1978) and this paper

	Lucas (1978)	This paper
Role of a manager	To increase productivity	To time investment well
Allocation of talent	Able persons run a big firm	Able persons run a risky firm
Compensation	Increasing and convex in ability	Increasing and convex in ability

The CEO compensation literatures support both views. Murphy (1998) summarizes this literature and insists that there is a strong, robust positive correlation between the size of a firm and the compensation of the firm's CEO. On the other hand, Smith and Watts (1992) and Gaver and Gaver (1993) find that investment opportunities have a positive effect on CEO compensation. This paper highlights that these two different effects may come from two different types of ability.

In fact, my data does not show any size effect. Table 3 shows that a firm's size has a negative impact on risk, the talent of B-school graduates and initial base salary. A firm's size is measured by average sales or net capital stock by industry. A negative correlation between a firm's size and a firm's risk is not surprising. For instance, Mansfield (1962) shows a negative relationship between the size of a firm and the variance of the firm's growth rate. However, a negative correlation between the size of the firm and the compensation of B-school graduates contrasts with the positive size effect on CEO compensation. This may indicate that talented B-school graduates first choose risky projects to determine their ability to manage projects or to learn the skills to handle difficult tasks, and then end up with a large firm that compensates them generously. This is an interesting future research project.

Table 4: **Size Effect**

	$\log(risk)$	$\log(GMAT)$	$\log(GPA)$	$\log(wage)$
$\log(Capital)$	-0.51**	-0.33	-0.27	-0.44*
$\log(Sales)$	-0.56**	-0.39	-0.36	-0.54**

* significant at 10 % level.

** significant at 5 % level.

4 Dynamic Model

In this section I formally describe a dynamic investment model. In order to understand the difference between the static model and the dynamic model, I will first show that prediction ability does not affect the level of investment on average in the static model. This result is in contrast to that of the dynamic model.

Theorem 6 *Suppose a firm solves problem (1). The expected amount of investment does not depend on prediction ability:*

$$I^e = \frac{z^e - r}{rA}.$$

Proof. This follows directly from the first order condition. ■

This theorem confirms that prediction ability does not affect the level of investment on average in the static model.[16] However, a good manager invests more on average in the dynamic model. In fact, I will show that a manager who has high prediction ability attains a high expected growth rate. A key assumption is that a larger current capital stock reduces adjustment costs. I now describe the dynamic model.

Production function: Assume that the production function is linear in the capital stock:

$$y_t = z_t k_t,$$

[16] This result is not robust even in the static model. Takii (1999 a) shows that the extent to which prediction ability affects expected investment depends on the third derivative of the adjustment cost function. When the adjustment cost function is quadratic, the third derivative is 0, and prediction ability has no impact on expected investment. Since I want to emphasize the dynamic effect in this paper, a quadratic specification is an appropriate benchmark.

where y_t is output, $z_t \in \left[\underline{z},\ \bar{z}\right]$ is a random shock and k_t is the capital stock in period t. You can think of z_t as the marginal productivity of capital at period t in this dynamic context.

Signals: Assume that a manager observes a signal, s_t, at date t, and infers the stream of future profitability, $\{z_s\}_{s=t+1}^{\infty}$. Define a vector $u_t = (z_t,\ s_t)$. Assume that $\{u_t\}$ follows a Markov process with stationary transition function $F\left(u_{t+1}|\ u_t\right)$. Let $F_u\left(u\right)$ denote the marginal distribution of u.

Adjustment costs: Assume that adjustment costs, x_t, take the following form:

$$x_t = \frac{AI_t^2}{2k_t}, \tag{12}$$

where I_t is investment in period t, and A is the adjustment cost parameter. Adjustment costs are evaluated by the investment price, which is assumed to be 1 over time to simplify the analysis. This adjustment cost function has two common properties: it is convex in I_t and exhibits constant returns to scale in k_t and I_t.

A key assumption here is that adjustment costs are decreasing in the current capital stock. The static model implied that the marginal productivity of prediction ability is larger when A is small. Hence I expect that a good manager has more incentive to accumulate capital in order to reduce adjustment costs.

Firm's Problem: The firm's profit maximization problem is

$$\begin{gathered} V^*\left(u_0,\ k_0\right) = \max\nolimits_{\{k_{t+1}\}_{i=i}^{\infty}} E\left[\sum_{t=0}^{\infty}\left(\tfrac{1}{1+r}\right)^t\left[z_t k_t - I_t - x_t\right]|u_0\right] \\ s.t.\ x_t = \tfrac{AI_t^2}{2k_t}\ and \\ I_t = k_{t+1} - k_t, \end{gathered} \tag{13}$$

where r is a constant interest rate. For simplicity, I temporarily assume that the depreciation rate is 0. This assumption does not affect the main results at all. I discuss the treatment of the depreciation rate for empirical purposes in Appendix 2.

Since the production function and adjustment cost function exhibit constant returns to scale in k_t and I_t, I can divide both sides of (13) by k_0:

$$\frac{V^*\left(u_0,\ k_0\right)}{k_0} = \max_{\{g_t\}_{i=0}^{\infty}}\left\{E\left[\sum_{t=0}^{\infty}\left(\Pi_{s=0}^{t}\beta_s\right)\left[z_t - g_t - \frac{A}{2}g_t^2\right]|u_0\right]\right\},$$

where $\beta_s = \frac{1+g_{s-1}}{1+r}\ for\ s \geq 1,\ \beta_0 = 1$ and $g_t = \frac{k_{t+1}-k_t}{k_t}$.

I can now define the Bellman equation, which is expected to be equivalent to (13):

$$Q\left(u\right) = \max_{g}\left[z - g - \frac{A}{2}g^2 + \beta\int Q\left(u'\right)dF\left(u'|u\right)\right], \tag{14}$$

where $z \in \left[\underline{z},\ \bar{z}\right]$ and $\beta = \frac{1+g}{1+r}$. I will show that $Q(u)$ is equivalent to Tobin's Q.

Maximization conditions: Now I present two basic equations that characterize the investment decision. In order to do so, I need two technical conditions:

$$g_t \in [0,\ \alpha],\ \alpha < r \tag{15}$$

and

$$\begin{aligned} \underline{z} &> r, \\ \bar{z} &< \frac{Ar^2}{2} + r. \end{aligned} \tag{16}$$

The upper bound of equation (15) prevents the optimal solution from exploding. The lower bound of equation (15) does not need to be 0, but for simplicity I assume that it is. Equation (16) guarantees that the solution is interior.

The following well-known theorem simply restates the results of Lucas and Prescott (1971), Hayashi (1982) and Hayashi and Inoue (1991) in this special formulation.

Theorem 7 *Suppose that equation (15) and some other technical conditions (described in Appendix 1) are satisfied. Then equation (14) has a unique solution $Q(\cdot)$ and*

$$Q(u) = \frac{V^*(u,\ k)}{k} \quad for\ any\ \ (u,\ k). \tag{17}$$

Moreover, suppose that assumption (16) is also satisfied. Then $Q(\cdot)$ and the associated unique policy function $g(\cdot)$ satisfy

$$g(u) = \frac{1}{A}\left[\frac{1}{1+r}\int Q(u')\,dF(u'|u) - 1\right] \tag{18}$$

and

$$Q(u) = \left[z + 1 + Ag(u) + \frac{A}{2}g^2(u)\right]. \tag{19}$$

Proof. See Appendix 1. ■

This $Q(u)$ is nothing more than Tobin's average Q.[17] Equation (18) is the first order condition, which says that all future information that affects the firm's investment decision is summarized by the expected value of Tobin's Q. Equation (19) is the Bellman equation, which tells how Q is determined.

[17]Tobin's Q is usually defined as $Q^{\#}(u) = \frac{1}{1+r}\int Q(u')\,dF(u'|u)$. That is, $Q^{\#}(u)$ is evaluated before the random shock is realized; $Q(u)$ is evaluated after the shock is realized. You can also refer to Ueda and Yoshikawa (1986), who show that investment is positively related to the expectation of future Tobin's Q rather than current Tobin's Q when there are time-to-build or delivery lags.

5 The Case when z_t is I.I.D.

Let us assume that a sequence $\{(z_{t+1},\ s_t)\}$ is i.i.d.. Then I can rewrite the transition function such that $F(u_{t+1}|u_t) = G(z_{t+1}|s_t)\, G_s(s_{t+1})$, where $G(z|s)$ is the conditional distribution of z given s and $G_s(s)$ is the marginal distribution of s. In this case, given s_t, the manager can only predict z_{t+1}. Hence, I can directly apply the basic measure in the static model to the dynamic model.

Theorem 8 *Suppose $\{(z_{t+1}, s_t)\}$ is an i.i.d. sequence. The present value of expected profits and the expected growth rate are increasing in h:*

$$\int\int V^*(z,\ s,\ k_0)\, dG_z(z)\, G_s(s) = Q^e k_0, \tag{20}$$

$$Q^e = (1+r)\,[Ar+1] - \sqrt{(1+r)^2 A\left[Ar^2 + 2(r-z^e)\right] - \sigma_z^2 h},$$

and

$$g^e = r - \sqrt{r^2 + 2\frac{r-z^e}{A} - \frac{\sigma_z^2 h}{A^2 (1+r)^2}}, \tag{21}$$

where $Q^e = \int\int Q(z,\ s)\, dG_z(z)\, G_s(s)$ and $g^e = \int g(s_t)\, dG_s(s_t)$. Moreover prediction ability can be estimated by

$$h = (\rho_{zg})^2,\ \rho_{zg} \geq 0, \tag{22}$$

where $\rho_{zg} = \frac{\int [z_{t+1}-z^e][g(s_t)-g^e]dG(z_{t+1},\ s_t)}{\sqrt{\int (z_{t+1}-z^e)^2 dG_z(z_{t+1}) \int [g(s_t)-g^e]^2 dG_s(s_t)}}$.

Proof. The proof is a similar to the proof of Theorem 10 in Appendix 1. I do not repeat it here. ■

This theorem says that prediction ability has a positive effect on expected Tobin's Q. Since Tobin's Q is the shadow price of capital under a constant returns to scale assumption, prediction ability has a positive impact not only on the present value of expected profits but also on the expected growth rate.

Since current investment not only increases future profits but also reduces future adjustment costs, a superior manager has more incentive to invest today in order to establish a flexible position in the future. This is why we have a growth effect in the dynamic context.

- The implications of the static model can be extended to the dynamic model:

  1. All of the effects of the signals in equations (20) and (21) are captured by h. That is, more accurate information in the Blackwell sense must have a larger value of h.

2. Prediction ability h has a stronger effect on expected profits when σ_z^2 is larger and when A is smaller.

- The dynamic model brings about a profitability effect and a scale effect: prediction ability has a large impact on the present value of expected profit when z^e is large and the initial capital stock is large.

- The dynamic model also implies that h has a stronger effect on the expected growth rate when z^e and σ_z^2 are large and when A is small.

6 The Case when z_t is a Markov Process

6.1 A generalized measure of prediction ability

Let us return to the general case. If a random process $\{u_t\}$ follows a Markov process with stationary transitions, then predicting next period profitability is not enough to determine the investment decision. A manager must predict the whole path of future profitability. Since Theorem 7 suggests that the entire path of future profitability is captured by only one variable, Tobin's Q, it is natural to construct a measure of ability to predict Tobin's Q in the next period in a fashion similar to the way I derived h.

One difficulty comes from the fact that Tobin's Q is an endogenous variable. A signal not only helps to predict next period's Tobin's Q, but also affects the degree of fluctuation of Tobin's Q. In order to take care of this problem, I need one more definition.

Definition 3 *The benchmark Q, $Q^*(z)$, is defined by the $Q(u)$ that solves equation (14) along with technical conditions (15) and (16) without observing any signal. That is, $Q^*(.)$ is the unique function that satisfies*

$$g^* = \frac{1}{A}\left[\beta \int Q^*(z)\, dF_m(z) - 1\right] \tag{23}$$

and

$$Q^*(z) = z + 1 + Ag^* + \frac{A}{2}(g^*)^2 . \tag{24}$$

Using this benchmark Q, I can define the generalized measure of prediction ability as follows.

Definition 4 *The generalized measure of a manager's ability to predict Q, h_Q, is defined by*

$$h_Q = \frac{\sigma_Q^2}{\sigma_{Q^*}^2} - \frac{\int Var\left(Q\left(u'\right)|u\right) dF_u\left(u\right)}{\sigma_{Q^*}^2}, \tag{25}$$

$$\begin{aligned} \textit{where } \sigma_Q^2 &= \int \left[Q\left(u\right) - \int Q\left(u\right) dF_u\left(u\right)\right]^2 dF_u\left(u\right), \\ \sigma_{Q^*}^2 &= \int \left[Q^*\left(z\right) - \int Q^*\left(z\right) dF_m\left(z\right)\right]^2 dF_m\left(z\right) \textit{ and} \\ Var\left(Q\left(u'\right)|u\right) &= \int \left[Q\left(u'\right) - \int Q\left(u'\right) dF\left(u'|u\right)\right]^2 dF\left(u'|u\right). \end{aligned}$$

The crucial difference between this and the previous measure is that here I use the unconditional variance of the benchmark Q instead of the observable Q as an adjustment factor. The reason is that Tobin's Q is an endogenous variable and the unconditional variance of the observable Q already reflects the effect of the signal. In order to separate the effect of the signal from the adjustment factor, I use the benchmark Q. As a result, the maximum value of $\left[\int Var\left(Q\left(u'\right)|u\right) dF_u\left(u\right)\right]/\sigma_{Q^*}^2$ is not 1 but $\sigma_Q^2/\sigma_{Q^*}^2$. That is why I subtract $\left[\int Var\left(Q\left(u'\right)|u\right) dF_u\left(u\right)\right]/\sigma_{Q^*}^2$ from $\sigma_Q^2/\sigma_{Q^*}^2$.

Identity of h_Q and h: I want to show that h_Q is a natural extension of h.

Theorem 9 *If the random sequence $\{(z_{t+1},\ s_t)\}$ is i.i.d., then $h_Q = h$.*

Proof. See Appendix 1. ■

The theorem says that if the random sequence is i.i.d., the value of two measures coincides. Hence it is fair to say that the generalized measure is a natural extension of the basic measure.[18]

[18] I must agree that this is just one possible practical treatment. In fact, I cannot claim that the generalized measure is adjustment cost parameter free. That is, since Tobin's Q is an endogenous variable, a change in the adjustment cost parameter varies the value of the generalized measure. Hence, from now on, I must assume that every firm has the same adjustment cost function.

I have three comments on this problem. First, empirical studies in the investment literature usually assume that every firm has the same adjustment cost parameter (Summers [1981], Salinger and Summers [1983], Fazzari, Hubbard and Petersen [1988] and Cummins, Hassett and Hubbard [1994]). Second, if I assume a quadratic adjustment cost function,

$$x_t = \frac{A}{2} I_t^2,$$

6.2 The effects of prediction ability

Now I am ready to analyze prediction ability under a Markov Process with stationary transitions. The following theorem summarizes the main results of the dynamic model.

Theorem 10 *Suppose that a random shock and a signal follow a Markov process with stationary transitions. The present value of expected profits and the expected growth rate are increasing in h:*

$$\int V^* (u,\ k_0)\, dF_u (u) = Q^e k_0,$$

$$Q^e = (1+r)\,[Ar+1] - \sqrt{(1+r)^2 A\,[Ar^2 + 2\,(r - z^e)] - \sigma_z^2 h_Q}$$

and

$$g^e = r - \sqrt{r^2 + 2\frac{r - z^e}{A} - \frac{\sigma_z^2 h_Q}{A^2\,(1+r)^2}}.$$

Moreover h_Q *is estimated by*

$$h_Q = a\,(\rho_{Qg})^2\,,\ a = \frac{\sigma_Q^2}{\sigma_z^2},\ \rho_{Qg} \geq 0, \tag{26}$$

where $\rho_{Qg} = \frac{\int [Q(u_{t+1}) - Q^e][g(u_t) - g^e] dF(u_{t+1},\, u_t)}{\sqrt{\sigma_Q^2 \sigma_g^2}}$.

Proof. See Appendix 1. ■

The results are the same as in the i.i.d. case except that h_Q is estimated by a weighted correlation coefficient between future Tobin's Q and the current growth rate, as opposed to a simple correlation coefficient. This weight reflects the fact that Tobin's Q is an endogenous variable. Since a good manager has the ability to change his investment decision aggressively based on his own signal, he will also change the value of Tobin's Q. This weight reflects this effect. Notice that more accurate information in the Blackwell sense must a have high value of h_Q in this formulation.

then I can prove that marginal Q does not depend on the parameter A. In this case, the generalized measure assesses the value of information for any adjustment cost parameter A. Third, as Theorem 10 will show, it is still true that more accurate information in the Blackwell sense must have a larger value of h.

7 Assumptions for an Empirical Analysis: The Issue of Heterogeneous Managers

If managers observe different signals, I immediately encounter several problems. A manager may try to develop a good reputation for his ability. A manager may ignore his own signal and mimic other managers' behavior. Also it is difficult to see how the capital market will clear. An empirical study needs to avoid these problems. I make the following assumptions:

1. Everyone knows every other manager's value of h.
2. The random shocks include firm specific-shocks. Hence prediction ability includes a firm-specific ability.
3. A manager's wage in period t is a fixed proportion χ of profits in period t.
4. Managers and shareholders have linear utility functions.
5. People make their decisions in the following order:
   (a) A random shock of the ith firm, z_t^i, and a signal for the jth manager, s_t^j, are realized.
   (b) Managers announce their investment decisions simultaneously. After the announcements are made they cannot change their decisions.
   (c) Shareholders make their investment decisions.
6. A manager does not know what signals managers in other firms observe. Each manager incurs a sufficiently large educational fixed cost of learning the relationship between an alternative signal and a random shock.
7. There is an externality among managers in a firm, and they can share the same signal within a firm.

Given these assumptions I claim:

- *No reputation problem*: No manager has an incentive to create a reputation since everyone knows his ability (Assumption 1).[19]
- *No "herd" behavior*: Managers announce their investment decisions simultaneously (Assumption 5). Therefore, no manager can mimic other managers' behavior.

[19] The reputation problem is a concern of Scharfstein and Stein (1990) and Prendergast and Stole (1996). They consider the case where a manager's investment decision reveals his prediction ability.

- *No learning problem*: Since a manager does not know the signals that managers in other firms observe (Assumption 6), other firms' behavior does not reveal which signal they observe. Moreover a large fixed cost of learning an alternative signal (Assumption 6) discourages a manager from finding better signals.

- *Capital market condition*: Since shareholders are risk neutral (Assumption 4), in equilibrium the return to their shares must be equal to r:

$$\begin{aligned} r &= \frac{Div_t + \int V\left(u_{t+1},\ k_{t+1}\right) dF\left(u_{t+1}|u_t\right) - V_t}{V_t} \\ &= \frac{Div_t + (1+r)\left[Ag_t + 1\right](1+g_t)\, k_t - V_t}{V_t}, \end{aligned}$$

  where Div_t is the dividend at period t. The second equality holds because shareholders invest after knowing the growth rate of capital (Assumption 5). Therefore, if a manager maximizes the present value of expected profits, shareholders can calculate conditional expected profits from the growth rate of capital by equations (17) and (18). In fact, a manager maximizes the present value as will be shown later.

- *A firm keeps the same ability*: Assumption 2 prevents a manager from changing jobs. Assumption 7 says that a firm can keep the same level of prediction ability even after the replacement of a manager. Hence, a firm maintains the same level of prediction ability over time.

- *A manager maximizes profits*: Assumption 3 and 4 ensure that a manager maximizes the present value of expected profits. Therefore, I can derive equations (18) and (19).

8 Empirical Evidence

This empirical study has two purposes. One is to examine whether or not prediction ability has a positive effect on profit; the other is to examine whether or not more accurate information increases investment.

Theorem 10 derives the following empirical equations:

$$\begin{aligned} Q^e &= a + bz^e + c\sigma_z^2 h_Q + \varepsilon, \\ g^e &= d + ez^e + f\sigma_z^2 h_Q + \mu, \end{aligned}$$

where a, b, c, d, e and f are parameters, Q^e is expected Tobin's Q, g^e is the expected growth rate of the capital stock, z^e is the expected random shock, which measures the profitability of exogenous investment opportunities, σ_z^2 is the variance of the random shock, which measures the riskiness of investment opportunities, and h_Q is the generalized measure, which is estimated by equation (26). The theory predicts that all parameters should be positive. I am especially interested in the parameters c and f, which show the impact of prediction ability.

The data consists of a 20 year (1975-1994) panel of firms from the COMPUSTAT data base. Appendix 3 shows how I construct variables Q, g and z. Using these three variables I can estimate Q^e, g^e, z^e, σ_z^2 and h_Q by calculating sample means, variances and correlation coefficients over time. I regress over cross sections of the data using these estimates.

8.1 Econometric issues

Selection bias: For the measure of prediction ability, I use the generalized measure h_Q. Constructing the measure requires a positive correlation coefficient between the current growth rate and future Tobin's Q. But 29% of the firms in the sample do not satisfy this condition. These omissions may cause a serious selection bias. In order to address the selection bias problem, I also investigate the regression using the simple correlation between future Q and the current growth rate: $\rho_{Qg} = \left[\int \left[Q\left(u_{t+1}\right) - Q^e\right]\left[g\left(u_t\right) - g^e\right] dF\left(u_{t+1},\ u_t\right)\right] / \sqrt{\sigma_Q^2 \sigma_g^2}$. Although this measure does not have any direct connection with the theory, it has the benefit of allowing us to use every observation.

Measurement error and simultaneous equation bias: Since my measure of prediction ability is constructed using endogenous variables, a simple OLS may have a problem. Moreover, the sample means may not be accurate proxies of expected values. Hence, my prediction measure and the error term may be correlated. In order to consider this issue, I also apply two stage OLS over the last ten years of the data, 1985-1994. I use the sample mean over the period 1975-1984 of each variable as an instrument of the corresponding variable, which is estimated by the sample mean over 1985-1994.

8.2 Results

Summary statistics: Table 5 shows summary statistics of Q^e, g^e, z^e, σ_z^2, h_Q and ρ_{Qg}. My estimation of the prediction measure h_Q has a mean of 14, a standard deviation of 38 and a median of 3.07 in 1975-1994 and a mean of 20, a standard deviation of 45 and a median of 4 in 1985-1994. The large standard deviations indicates a huge difference in prediction ability across firms. Moreover, the large gap between the mean and median implies a huge skewness of ability.

Table 5: Summary statistics

		Q^e	g^e	z^e	σ_z^2	h_Q	ρ_{Qg}
1975-1994	mean	1.57	0.09	0.27	0.09	14.49	0.22
	standard deviation	1.28	0.07	0.25	0.30	37.94	0.39
	median	1.09	0.07	0.21	0.01	3.07	0.24
	# of observations	1059	1059	1059	1059	756	1059
1985-1994	mean	1.30	0.05	0.21	0.04	19.58	0.22
	standard deviation	1.35	0.06	0.22	0.27	45.17	0.43
	median	0.74	0.04	0.13	0.003	4.32	0.28
	# of observations	441	441	441	441	316	737

Q^e is expected Tobin's Q.
g^e is the expected growth rate of capital stock.
z^e is the expected random shock (measure of profitability).
σ_z^2 is the variance of the random shock (measure of risk).
h_Q is the measure of prediction ability.
ρ_{Qg} is the simple correlation between future Q and the current growth rate.

Table 6.1: The profit effect (1975-1994)

Dependent variable is expected Tobin's Q

intercept	0.636***	0.700***	0.664***
	(0.057)	(0.051)	(0.052)
z^e	3.118***	3.020***	2.708***
	(0.148)	(0.127)	(0.141)
h_Q	0.008***		
	(0.001)		
ρ_{Qg}		0.208*	
		(0.084)	
$\sigma_z^2 \times h_Q$			0.474***
			(0.038)
$Adj - R^2$	0.385	0.349	0.450
obs	755	1058	755

* significant at 5 % level.
** significant at 0.5% level.
*** significant at 0.05% level.
standard error in parentheses.

Table 6.2: The profit effect (1985-1994)

Dependent variable is expected Tobin's Q

OLS

intercept	0.299***	0.350***	0.338***
	(0.069)	(0.050)	(0.066)
z^e	4.683***	4.776***	4.47***
	(0.214)	(0.148)	(0.22)
h_Q	0.002*		
	0.001		
ρ_{Qg}		-0.003	
		(0.077)	
$\sigma_z^2 \times h_Q$			0.221***
			(0.063)
$Adj - R^2$	0.604	0.586	0.614
obs	315	736	315

* significant at 5 % level.
** significant at 0.5% level.
*** significant at 0.05% level.
standard error in parentheses.

2SLS

intercept	-2.598	-1.546*	0.328
	(2.090)	(0.732)	(0.402)
z^e	6.049	4.539**	-0.617
	(2.547)	(1.430)	(3.374)
h_Q	0.136		
	(0.096)		
ρ_{Qg}		8.923*	
		(3.960)	
$\sigma_z^2 \times h_Q$			5.106*
			(2.503)
$Adj - R^2$	0.014	0.046	0.042
obs	315	736	315

* significant at 5 % level.
** significant at 0.5% level.
*** significant at 0.05% level.
standard error in parentheses.

Prediction ability raises expected Q: Table 6.1 reports the effect of prediction ability on expected Tobin's Q. The second and third column show the effect of prediction ability. The last column shows the effect of the product of prediction ability and the variance of the random shock that is implied by the theory. All coefficients on prediction ability are positive and significant, which is indicated by the theory. The table shows that prediction ability raises expected Q of the firm.

Table 6.2 reports the same regression over the period 1985-1994. It also reports the result of two stage OLS. For the most part, coefficients are still significant and positive, but a simple OLS with the simple correlation as the measure of prediction ability has a negative coefficient. However, two stage OLS recovers the positive significant relation. Notice that 2SLS increases the magnitude of the coefficients. This indicates that simple OLS underestimates the importance of prediction ability.

Prediction ability raises the expected growth rate: Table 7.1 discloses the effect of prediction ability on the expected growth rate. Again the first two columns show the effect of prediction ability and the last column shows the effect of the product of prediction ability and the variance of the random shock. All of the results of the simple OLS are positive and significant, which is expected. This indicates that prediction ability has a positive impact on the expected growth rate.

Table 7.1: The growth effect (1975-1994)

Dependent variable is an expected growth rate

intercept	0.061^{***}	0.053^{***}	0.062^{***}
	(0.004)	(0.003)	(0.004)
z^e	0.112^{***}	0.111^{***}	0.102^{***}
	(0.010)	(0.008)	0.010
h_Q	$2\text{x}10^{-4**}$		
	(6×10^{-5})		
ρ_{Qg}		0.029^{***}	
		(0.005)	
$\sigma_z^2 \times h_Q$			0.012^{***}
			(0.003)
$Adj - R^2$	0.148	0.177	0.160
obs	755	1058	755

* means significant at 5 % level.

** means significant at 0.5% level.

***means significant at 0.05% level.

standard error in parentheses

Table 7.2: The growth effect (1985-1994)

Dependent variable is an expected growth rate

OLS

intercept	0.022***	0.022***	0.023***
	(0.004)	(0.003)	(0.004)
z^e	0.126***	0.135***	0.128***
	(0.013)	(0.009)	(0.013)
h_Q	4x10^{-5}		
	(7×10^{-5})		
ρ_{Qg}		0.002	
		(0.005)	
$\sigma_z^2 \times h_Q$			-0.002
			(0.004)
$Adj-R^2$	0.234	0.236	0.233
obs	315	736	315

* means significant at 5 % level.

** means significant at 0.5% level.

***means significant at 0.05% level.

standard error in parentheses

2SLS

intercept	-0.054	-0.025	0.015
	(0.107)	(0.016)	(0.008)
z^e	0.178*	0.108**	0.083*
	(0.08)	(0.033)	(0.040)
h_Q	0.003		
	(0.005)		
ρ_{Qg}		0.246***	
		(0.069)	
$\sigma_z^2 \times h_Q$			0.075*
			(0.033)
$Adj-R^2$	0.012	0.033	0.071
obs	315	736	315

* means significant at 5 % level.

** means significant at 0.5% level.

***means significant at 0.05% level.

standard error in parentheses

Table 7.2 reports the results of the same regression and the two stage OLS over 1985-1994. Although all of coefficients on prediction ability using simple OLS lose significance and some are negative, 2SLS recovers a positive relation and shows significance. Moreover, two stage OLS again increases the magnitude of the coefficients. Again, OLS underestimates the importance of prediction ability.

Let me estimate the impact of prediction ability on the expected growth rate. 2SLS results in Table 7.2 show that the coefficient on $\sigma_z^2 * h^*$ is 0.075. Since the median of σ_z^2 over the period 1984-1993 is 0.003, the effect of h^* on the median firm is about 0.0002. The standard deviation of h^* is 45.17. This means that a one standard deviation change in prediction ability increases the expected growth rate by 1%. Hence, considering the median person and firm, a one standard deviation increase in h^* increases the expected growth rate from 4% to 5%.

9 Conclusion and Extensions

In this paper I constructed a micro foundation of one type of human capital: a manager's ability to predict the future profitability of a firm. My theory predicts that a manager who has high prediction ability goes into a risky industry, because risk increases the marginal productivity of prediction ability. I simulated my assignment model and found that the results fit B-school placement data quite well. I also employed a dynamic analysis, which shows that prediction ability increases a firm's expected Tobin's Q and allows a firm to attain a faster expected growth rate. The COMPUSTAT dataset confirms these points as well.

I am working on three different extensions: (1) generalization of the adjustment cost function, (2) the ability of a manager to learn and (3) consideration of an information collection cost. Each is discussed in turn.

One of problems in this paper is that the measure is model specific. Hence, it is difficult to answer several questions. How much does a tax policy affect the marginal productivity of prediction ability? Does capital market imperfection affect the marginal productivity of prediction ability? How does irreversible investment change the results? Despite this limit, the intuition behind the measure is quite robust. Takii (1999 a) extends the method to the case with more general adjustment costs. A key point is that the conditional expectation given a good signal must be constructed by a mean preserving spread of the conditional expectation given a bad signal. This approach can be applied to various analytical topics like search theory and technology adoption.

Another interesting question is the relationship between learning and prediction ability. In fact, Takii (1999, b) shows that knowing a good signal increases not only prediction ability but also the speed at which a manager learn an unknown parameter using Bayesian analysis, like in Jovanovic and Nyarko (1995), and Foster and Rosenzweig (1995). As long as you believe that education can increase our ability to determine a good signal, I can isolate two effects of education: an increase in prediction ability and an increase in learning speed. This shows that the effect on prediction ability is a persistent effect, although the effect on learning speed shrinks as time goes by. This may provide a testable framework in which to analyze the effect of education.

I also extend this model by adding a cost of collecting information. Takii (1999, c) shows that if Information Technology (IT) reduces costs, then IT and prediction ability are complementary. Complementarity is a crucial assumption in the skill-biased technological change argument. IT strengthens the value of prediction ability, because a good manager can access good information more often. This approach also shows that IT does not need to improve productivity, but it helps a manager to time investment better. This result helps to explain the productivity slowdown puzzle. Thus, my approach may reveal the economic value of IT.

10 Appendix 1

Proof of Theorem 3: I first construct a candidate equilibrium and then show that the candidate satisfies three equilibrium conditions.

First, construct $a(\cdot)$ so that $a(h) = \Psi^{-1}(\Gamma(h))$. Then a is a mapping from $\left[\underline{h},\ \bar{h}\right]$ into $\left[\underline{\sigma_z^2},\ \overline{\sigma_z^2}\right]$; it satisfies (9) and $a' > 0$ by construction given the assumptions that $\Gamma' > 0$ and $\Psi' > 0$.

Second, define h^* such that

$$\frac{\left(z^e - r\right)^2 + a\left(h^*\right)h^*}{2r^2A} = w_l.$$

There exists a unique $h^* \in \left[\underline{h},\ \bar{h}\right]$ since $a(h)$ is a strictly increasing continuous function of h and $\left\{\left[\left(z^e - r\right)^2 + \underline{\sigma_z^2 h}\right]/2r^2A\right\} \leq w_l \leq \left\{\left[\left(z^e - r\right)^2 + \overline{\sigma_z^2 h}\right]/2r^2A\right\}$ by assumption. Notice that $V\left(h^* : \eta\left(h^*\right)\right) = w_l$.

Third, construct $w(\cdot)$ so that

$$w(h) = \int_{h^*}^{h} \frac{a(\tau)}{2r^2A}d\tau + w_l.$$

The function exists since $a(h)$ is bounded. Notice that $w(h^*) = w_l$ and $w'(h) > 0$. Hence, if $h \geq (\leq)\, h^*$, a person becomes a manager (a worker).

If I show that the candidate equilibrium satisfies condition (10), the proof is complete. Since

$$V'\left(h : \sigma_z^2\right) - w'(h)\,|_{h=a^{-1}(\sigma_z^2)} = 0 \ for\ \forall\sigma_z^2$$

and

$$V''\left(h:\sigma_z^2\right)-w''\left(h\right)=-\frac{a'\left(h\right)}{2r^2A}<0\ for\ \forall\sigma_z^2,$$

the first order condition is always satisfied at $h=a^{-1}\left(\sigma_z^2\right)$, and a firm will attain the global maximum at this point. Hence the candidate satisfies condition (10). ■

Proof of Theorem 4: Suppose that $\Gamma\left(h\right)$ and $\Psi\left(\sigma_z^2\right)$ are log uniform. Then

$$\int_{\log\left(\underline{\sigma_z^2}\right)}^{\log a(h)}\frac{1}{\log\left(\overline{\sigma_z^2}\right)-\log\left(\underline{\sigma_z^2}\right)}dv=\int_{\log\left(\underline{h}\right)}^{\log h}\frac{1}{\log\left(\overline{h}\right)-\log\left(\underline{h}\right)}d\tau.$$

Hence

$$\begin{aligned}\log a\left(h\right) &= \frac{\log\left(\overline{\sigma_z^2}\right)-\log\left(\underline{\sigma_z^2}\right)\left(\log\left(h\right)-\log\left(\underline{h}\right)\right)}{\log\left(\overline{h}\right)-\log\left(\underline{h}\right)}+\log\left(\underline{\sigma_z^2}\right) \qquad (27)\\ &= \xi\log\left(h\right)+\Theta,\\ where\ \xi &= \frac{\log\left(\overline{\sigma_z^2}\right)-\log\left(\underline{\sigma_z^2}\right)}{\log\left(\overline{h}\right)-\log\left(\underline{h}\right)}\ and\\ \Theta &= \log\left(\underline{\sigma_z^2}\right)-\xi\log\left(\underline{h}\right).\end{aligned}$$

Therefore,

$$\begin{aligned}\log a\left(\delta\left(score\right)\right) &= \xi\log\left(score\right)+\Theta+\xi\log\left(\delta\right),\\ where\ \xi &= \frac{\log\left(\overline{\sigma_z^2}\right)-\log\left(\underline{\sigma_z^2}\right)}{\log\left(\overline{score}\right)-\log\left(\underline{score}\right)}\ and\\ \Theta &= \log\left(\underline{\sigma_z^2}\right)-\xi\log\left(\underline{score}\right)-\xi\log\left(\delta\right).\end{aligned}$$

Since $\log\left(\sigma_z^2\right)$ and $\log\left(score\right)$ are distributed uniformly,

$$\begin{aligned}\log\overline{\sigma_z^2} &= m_{\log\sigma}+\sqrt{3\sigma^2_{\log\sigma}},\\ \log\underline{\sigma_z^2} &= m_{\log\sigma}-\sqrt{3\sigma^2_{\log\sigma}},\\ \log\overline{score} &= m_{\log(score)}+\sqrt{3\sigma^2_{\log(score)}}\ and\\ \log\underline{score} &= m_{\log(score)}-\sqrt{3\sigma^2_{\log(score)}}.\end{aligned}$$

Hence

$$\xi = \sqrt{\frac{\sigma^2_{\log \sigma}}{\sigma^2_{\log(score)}}}$$

and

$$\begin{aligned}\Theta &= m_{\log \sigma} - \sqrt{3\sigma^2_{\log \sigma}} \\ &\quad - \sqrt{\frac{\sigma^2_{\log \sigma}}{\sigma^2_{\log(score)}}} \left(m_{\log(score)} - \sqrt{3\sigma^2_{\log(score)}} \right) - \xi \log(\delta) \\ &= m_{\log \sigma} - \xi m_{\log(score)} - \xi \log(\delta) \\ &= \theta - \xi \log \delta,\end{aligned} \tag{28}$$

where $\theta = m_{\log \sigma} - \xi m_{\log(score)}$. Finally, $\log a(\delta(score)) = \xi \log(score) + \theta$.

$$\log a(\delta(score)) = \xi \log(score) + \theta.$$

■

Proof of Theorem 5: By Theorem 3, equation (27) and equation (28),

$$\begin{aligned}w(\delta(score)) &= \int_{\delta(score)^*}^{\delta(score)} \frac{e^{\Theta}\tau^{\xi}}{2r^2 A} d\tau + w_l \\ &= \frac{(\delta)^{\xi+1} e^{(\theta - \xi \log \delta)} \left[(score)^{\xi+1} - (score^*)^{\xi+1} \right]}{2r^2 A(\xi+1)} + w_l \\ &= \frac{\delta e^{\theta} \left[(score)^{\xi+1} - (score^*)^{\xi+1} \right]}{2r^2 A(\xi+1)} + w_l.\end{aligned}$$

■

Proof of Theorem 7: Suppose that the space of u, U, is a compact Borel set and that $F(u'|u)$ has the Feller property. Suppose that $Q(\cdot) \in C$ where C is the space of bounded measurable continuous functions with the sup norm.

Since the reward function is continuous and the space of the state variable, U, and the strategy space, $[0,\ \alpha]$, are compact, the reward function is bounded. Since $\beta = [1+g]/[1+r] \leq [1+\alpha]/[1+r] < 1$, there exists a unique function $Q(u) = V^*(u,\ k)/k$ (See Harris [1987], and Stokey and Lucas [1989]). Since the reward function is strictly concave in g, the associated policy function $g(u)$ is continuous and unique.

Now I need to show that the solution is interior. If this is true, it is obvious that the unique solution is characterized by the first order condition (18) and the Bellman equation (19).

First, consider the upper bound. Denote $\bar{g} = \max_u g(u)$ and $\bar{u} = \arg\max_u g(u)$. They exist because $g(u)$ is continuous and the set U is compact. Consider the maximization of the Bellman equation without any boundary conditions. Then the first order condition and the Bellman equation imply that

$$\begin{aligned} 1 + A\bar{g} &= \frac{1}{1+r}\left[\begin{array}{c} E(z'|\bar{u}) + 1 + AE(g(u')|\bar{u}) \\ +\frac{A}{2}E\left(g(u')^2|\bar{u}\right) \end{array}\right] \\ &\le \frac{1}{1+r}\left[\bar{z} + 1 + A\bar{g} + \frac{A}{2}\bar{g}^2\right]. \end{aligned}$$

Hence,

$$\begin{aligned} \bar{g} &\ge r + \sqrt{r^2 + 2\left\{\frac{r-\bar{z}}{A}\right\}}, \ or \\ \bar{g} &\le r - \sqrt{r^2 + 2\left\{\frac{r-\bar{z}}{A}\right\}}. \end{aligned} \tag{29}$$

Condition (16) ensures the existence of $\bar{g} \in R$ and the existence of α such that

$$\bar{g} \le r - \sqrt{r^2 + 2\left\{\frac{r-\bar{z}}{A}\right\}}$$

does not attain the upper bound, α, since $g \le \alpha < r$. I find the maximizer of the unconstrained problem $\bar{g}$ at the interior. Hence this is also a maximizer with the boundary constraint. Since $g(u) \le \bar{g}$, for any u, $g(u)$ never attains the upper bound.

Similarly denote $\underline{g}= \min_u g(u)$ and $\underline{u}= \arg\min_u g(u)$. Then similarly it can be derived that

$$\underline{g}\ge r - \sqrt{r^2 + 2\left\{\frac{r-\underline{z}}{A}\right\}}.$$

Condition (16) ensures that $\underline{g}> 0$. Using the same logic as above, since for any u, $g(u) \ge\underline{g}$, $g(u)$ does not attain the lower bound. ■

Proof of Theorem 6: In the i.i.d. environment, z_t does not include any information with which to predict z_{t+1}. Hence the growth rate, g, only depends on the signal s_t. Hence, by (18) and (19),

$$E[Q(u')|s] = E(z'|s) + 1 + AE[g(s')] + \frac{A}{2}E\left[g^2(s')\right]$$

and

$$E\left[Q\left(u'\right)\right]=E\left(z'\right)+1+AE\left[g\left(s'\right)\right]+\frac{A}{2}E\left[g^{2}\left(s'\right)\right].$$

Hence

$$Var\left[E\left(Q\left(u'\right)|s\right)\right]=Var\left[E\left(z|s\right)\right].$$

The following lemma completes the proof.

Lemma 1 *The generalized measure can be rewritten as:*

$$h_{Q}=\frac{Var\left(E\left(Q\left(u_{t+1}|u_{t}\right)\right)\right)}{\sigma_{z}^{2}}. \tag{30}$$

Proof. Using equation (6), I can rewrite equation (25) to be

$$h_{Q}=\frac{Var\left(E\left(Q\left(u_{t+1}|u_{t}\right)\right)\right)}{\sigma_{Q^{*}}^{2}}.$$

To complete the proof of Lemma 1, consider Lemma 2.

Lemma 2 $\sigma_{Q^{*}}^{2}=\sigma_{z}^{2}$.

Proof. Taking expectations on both sides of (24)

$$\int Q^{*}\left(z\right)dF_{m}\left(z\right)=\int zdF_{m}\left(z\right)+1+Ag^{*}+\frac{A}{2}\left(g^{*}\right)^{2},$$

Combining (24) and this equation, I can calculate the unconditional variance of Q^{*}. The proof of Lemma 2 follows. ■
Using Lemma 2, the result of Lemma 1 is immediate. ■
By equation (30), the proof of Theorem 6 follows. ■

Proof of Theorem 10: First I show that prediction ability positively affects the variance of the growth rate. Using this result I will show all of the results in Theorem 10.

Lemma 3 *The variance of the growth rate is a strictly increasing function of prediction ability:*

$$\sigma_{g}^{2}=\frac{\sigma_{z}^{2}h_{Q}}{A^{2}\left(1+r\right)^{2}}.$$

Proof. Taking the expectation on both sides of (18),

$$1 + A\int g(u)\, F_u(u) = \frac{1}{1+r}\int Q(u)\, dF_u(u) \;\; and$$

$$A^2\sigma_g^2 = \frac{Var\left(E\left(Q\left(u'|u\right)\right)\right)}{\left(1+r\right)^2}.$$

By equation (30),

$$\sigma_g^2 = \frac{\sigma_z^2 h_Q}{A^2\left(1+r\right)^2}. \tag{31}$$

■

- First I prove the effect on the expected growth rate.

  By (18) and (19)

$$\begin{aligned} & 1 + Ag(u) \\ = & \frac{1}{1+r}\left\{\begin{array}{c} E\left[z'|u\right] + 1 + AE\left[g\left(u'\right)|u\right] \\ +\frac{A}{2}E\left[g^2\left(u'\right)|u\right] \end{array}\right\}, \end{aligned} \tag{32}$$

where $E[z|u] = \int zF(u'|u)$, $E[g(u')|u] = \int g(u')\, F(u'|u)$ and $E[g^2(u')|u] = \int g^2(u')\, F(u'|u)$. Taking expectations on both sides,

$$0 = g^{e2} - 2rg^e + 2(A)^{-1}[z^e - r] + \sigma_g^2.$$

Hence,

$$g^e = r \pm \sqrt{r^2 + 2\frac{r - z^e}{A} - \sigma_g^2}.$$

Since $g \leq \alpha - 1 < r$,

$$g^e = r - \sqrt{r^2 + 2\frac{r - z^e}{A} - \sigma_g^2}. \tag{33}$$

I know that for any u there exists a unique $g(u)$ such that $g(u) \in \left[\underline{g},\ \bar{g}\right]$; in turn the expectation exists.

- Next, I prove the effect on expected Tobin's Q.

By (18), (31) and (33)

$$\begin{aligned}
& \int Q(u)\, dF_u(u) \\
= \;& (1+r)\left(A \int g(u)\, dF_u(u) + 1\right) \\
= \;& (1+r)\left[Ar+1\right] \\
& - \sqrt{(1+r)^2 A\left[Ar^2 + 2(r - z^e)\right] - \sigma_z^2 h_Q}.
\end{aligned}$$

Moreover,

$$\begin{gathered}
E\left[V^*(u_0,\ k_0)\right] \\
= E\left[Q(u_0)\, k_0\right] \\
= Q^e k_0.
\end{gathered}$$

- Finally, I show how the prediction measure can be estimated.

By multiplying (18) by $g(u_t)$ and taking expectations on both sides,

$$g^e + A \int g^2(u_t)\, F_u(u_t) = \frac{1}{1+r} \int Q(u_{t+1})\, g(u_t)\, dF(u_{t+1}, u_t). \tag{34}$$

By taking expectations on both sides of (18) and multiplying the result by g^e,

$$g^e + A(g^e)^2 = \frac{1}{1+r} \int Q(u)\, dF_u(u)\, g^e. \tag{35}$$

Then subtracting (35) from (34) gives

$$Cov(g(u_t)\, Q(u_{t+1})) = (1+r)\, A\sigma_g^2.$$

Hence,

$$\begin{aligned}
\rho_{gQ} &= \frac{Cov(g(u_t)\, Q(u_{t+1}))}{\sqrt{\sigma_g^2 \sigma_Q^2}} \\
&= \frac{(1+r)\, A\sigma_g^2}{\sqrt{\sigma_g^2 \sigma_Q^2}} \\
&= \sqrt{\frac{\sigma_z^2 h_Q}{\sigma_Q^2}} \; and \\
h_Q &= \frac{\sigma_Q^2}{\sigma_z^2} (\rho_{gQ})^2.
\end{aligned}$$

■

11 Appendix 2: Heterogeneous Depreciation Rate

Since Tobin's Q is an endogenous variable, several factors could affect the value of the generalized measure. Typically a different firm has a different capital stock and a different rate of physical depreciation. The following theorem, however, says that I can compare managers' ability over firms with different economic depreciation rates.

Suppose that there are adjustment costs on net investment[20] such that

$$\begin{aligned} x_t &= \frac{(k_{t+1}-k_t)^2}{k_t},\ if\ k_{t+1}\geq k_t, \\ &= 0,\ if\ \ (1-\delta)\,k_t \leq k_{t+1}\leq k_t, \\ &= \infty,\ otherwise, \end{aligned} \tag{36}$$

where δ is the physical depreciation rate.

Theorem 11 *The generalized measure of prediction ability, h_Q, is depreciation rate invariant. That is, a change in δ does not affect the value of h_Q.*

Proof. Using the adjustment cost function (36), I can derive

$$Q\,(u) = z + 1 - \delta + Ag\,(u) + \frac{A}{2}g^2\,(u)\,,$$

instead of equation (19). The following lemma proves the theorem.

Lemma 4 *Suppose that $Q^*\,(u) = aQ\,(u) + b$ and $z^* = az + c$, then $h_{Q^*} = h_Q$.*

Proof. Using equation (30)

$$h_{Q^*} = \frac{Var\,(E(aQ\,(u_{t+1}) + b|u_t))}{\int\left((az+c) - \int(az+c)\,dF_m\,(z)\right)^2 dF_m\,(z)} = h_Q.$$

■

The proof of Theorem 11 follows. ■

---

[20]With depreciation, the capital accumulation function is replaced by

$$\begin{aligned} k_{t+1} &= I_t + (1-\delta)\,k_t, \\ where\ k_{t+1} &\in [(1-\delta)\,k_t,\ \alpha k_t]\,. \end{aligned}$$

Finally, the technical conditions are replaced by

$$\begin{aligned} \bar{z} &< \frac{Ar^2}{2} + r + \delta\ and \\ \underline{z} &> r + \delta. \end{aligned}$$

12 Appendix 3 : Data constructions

12.1 Some detail on the construction of the measure of risk

I did not consider data with any of the following features:

1. The absolute value of (investment / net capital stock) is greater than 1.
2. The number of observations are less than or equal to 5.
3. The mean of (operating income / net capital stock) is less than -30.

I need Condition 1 to avoid large mergers. I use Condition 2 to avoid an unreliable mean and correlation coefficient. Condition 3 avoids use of firms that are extremely poor performers.

12.2 Tobin's Q

Tobin's Q

$$Q_t = \frac{S_t + D_t - pinv_t}{p_t k_t},$$

where S_t is the market value of equity at the beginning of period t, D_t is the market value of firm debt at the beginning of period t, $pinv_t$ is the replacement cost of inventory at the beginning of period t and $p_t k_t$ is the replacement cost of capital at the beginning of period t.

Growth Rate: g

$$g_t = \frac{p_t I_t}{p_t k_t} - 1/L,$$

where $p_t I_t$ is capital expenditure, $p_t k_t$ is the replacement cost of capital at the beginning of period t and L is average lifetime of capital.

Random Shock: z

$$z_t = \frac{OI_t}{p_t k_t}$$

where OI_t is operating income before depreciation in period t and $p_t k_t$ is the replacement cost of capital at the beginning of period t.

Market Value of Equity: S

1. the market value of common stock (the common stock outstanding at the end of the previous year multiplied by the end of the previous year's common stock price)

2. the market value of preferred stock (the firm's preferred dividend payout at the end of the previous year divided by the previous year's average Standard and Poor's preferred dividend yield)

Debt: D

1. short term debt at the end of the previous year (the book value)
2. long term debt at the end of the previous year (the book value)

Capital Stock (Replacement Cost): pk

$$p_{t+1}k_{t+1} = \left(p_t k_t \left(1 - 1/L\right) + p_t I_t\right) \frac{p_{t+1}}{p_t},$$

$$L = \frac{\sum^T L_t^*}{T} \; and$$

$$L_t^* = \frac{GPPE_t}{DEP_t},$$

where L_t^* is the lifetime of capital, $GPPE_t$ is the book value of gross property, plant and equipment in year t, $p_t I_t$ is nominal investment and DEP_t is book depreciation in year t. The ratio p_{t+1}/p_t is the inflation rate of investment goods from the survey of current business.

Inventory (Replacement Cost): pinv
If a firm uses FIFO, the book value of the end of the previous period is the replacement cost. If a firm uses LIFO,

$$\begin{aligned} p_{t+1}inv_{t+1} &= p_t inv_t \frac{p_{t+1}}{p_t} + BINV_{t+1} - BINV_t, \; if \; BINV_{t+1} \geq BINV_t \\ &= \left(p_t inv_t + BINV_{t+1} - BINV_t\right) \frac{p_{t+1}}{p_t}, \; if \; BINV_{t+1} < BINV_t \end{aligned}$$

where $BINV_t$ is the book value of inventory at the end of period $t-1$. Again, p_{t+1}/p_t is the inflation rate of investment goods from the survey of current business.

Some details

1. For the process of estimating the replacement cost of the capital stock and inventory,

(a) I assume that the book value of the initial year is equal to the replacement value.

(b) if I encounter negative values or missing variables, I delete the observation.

2. I exclude extreme values by the following rules:

   (a) If the absolute value of the growth rate is greater than 1, I delete the observation.

   (b) If the absolute value of a random shock is greater than 10, I delete the observation.

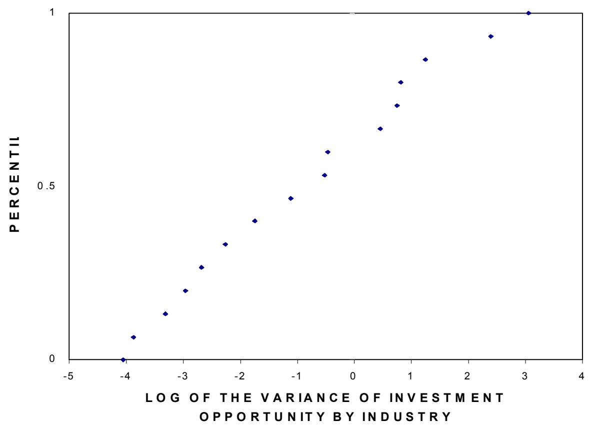
   (c) If the absolute value of Tobin's Q is greater than 10, I delete the observation.

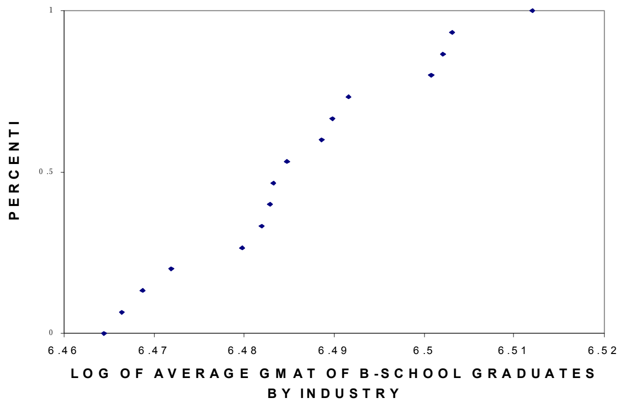
   Cummins, Hassett and Hubbard (1994) discussed that these extreme values mainly occur because of (1) large mergers, (2) extraordinary firm shocks or (3) COMPUSTAT coding errors.

3. I estimate the expected value, the variance and the correlation coefficient of the population using the mean, the variance and the correlation coefficient of the random sample over 1975-1994 for each firm. In order to calculate these, I delete samples which have fewer than 10 observations.

4. Finally I delete all observations, where the absolute value of the prediction ability measure is greater than 1000 to avoid extreme cases. This procedure eliminates only 2 or 3 observations.

13 Appendix 4





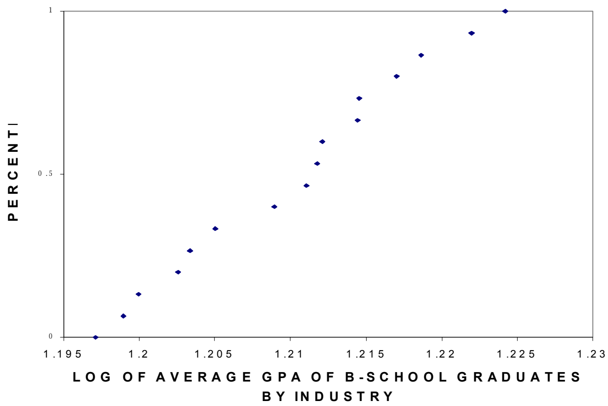


Figure : Estimated Cumulative Distributions of the log of Risk, log of GMAT and log of GPA

References

[1] Abel, Andrew B. (1983), “Optimal Investment Under Uncertainty,” American Economic Review, 73, 228-233.

[2] Athey, Susan and Levin, Jonathan (1998), “The Value of Information in Monotone Decision Problems,” unpublished

[3] Becker Gary S.(1973), “A Theory of Marriage: Part 1,” Journal of Political Economy, 81, 813-846.

[4] Blackwell, D. (1953), “Equivalent Comparisons of Experiments,” Annals of Mathematical Statistics, 24, 265-272.

[5] Cummins, Jason G., Hassett, Kevin A. and Hubbard, Glenn R. (1994), “A Reconsideration of Investment Behavior Using Tax Reforms as Natural Experiments,” Brookings Papers on Economic Activity, 2, 1-74.

[6] Demers, Michel (1991), “Investment under Uncertainty, Irreversibility and the Arrival of Information Over Time,” Review of Economic Studies, 58, 333-350.

[7] Fazzari, Steven, Hubbard, Glenn R. and Petersen, Bruce C. (1988), “Financing Constraints and Corporate Investment,” Brookings Papers on Economic Activity, 1, 141-205.

[8] Foster, Andrew D. and Rosenzweig, Mark R., (1995), “Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture,” Journal of Political Economy, 103, 1176-1209.

[9] Gaver, Jennifer J. and Gaver, Kenneth M. (1993), “Additional Evidence on the Association between the Investment Opportunity Set and Corporate Financing, Dividend, and Compensation Policies,” Journal of Accounting and Economics, 26, 125-160.

[10] Harris, Milton (1987), *Dynamic Economic Analysis*, New York, Oxford University Press.

[11] Hartman, Richard (1972), “The Effect of Price and Cost Uncertainty on Investment,”Journal of Economic Theory, 5, 256-266.

[12] Hayashi, Fumio (1982), ”Tobin’s Marginal and Average q: A Neoclassical Interpretation,”Econometrica, 50, 213-224.

[13] Hayashi, Fumio and Inoue, Tohru (1991), “The Relation between Firm growth and Q with Multiple Capital Goods: Theory and Evidence From Panel Data on Japanese Firms,” Econometrica, 59, 731-753.

[14] Holmes, Tomas J. and Schmitz, James A. (1990) "A Theory of Entrepreneurship and Its Application to the Study of Business Transfers," Journal of Political Economy, 98, 265-294.

[15] Jovanovic, Boyan and Nyarko, Yaw (1995), "A Bayesian Learning Model Fitted to a Variety of Empirical Learning Curves," Brookings Economics: Macroeconomics,247-305.

[16] Kihlstrom, Richard (1984), "A 'Bayesian' Exposition of Blackwell's Theorem on the Comparison of Experiment," *In Bayesian Models in Economic Theory*, edited by Marcel Boyer and Richard E. Kihlstrom, New York,Elsevier Science Publishers.

[17] Kihlstrom, Richard E. and Laffont, Jean-Jacques (1979), "A General Equilibrium Entrepreneurial Theory of Firm Formation Based on Risk Aversion," Journal of Political Economy, 87,719-748.

[18] Koopmans, Tjalling C. and Beckmann Martin (1957), "Assignment Problems and the Location of Economic Activities," Econometrica, 25, 53-76.

[19] Lucas, Robert E. (1978),"On the Size Distribution of Business Firms," Bell Journal of Economics, 9, 508-523.

[20] Lucas, Robert E., Jr. and Prescott, Edward C.(1971), "Investment and Uncertainty," Econometrica, 39, 659-682.

[21] Mansfield, Edwin (1962),"Entry, Gibrat's Law, Innovation, and the Growth of Firms," American Economic Review, 52, 1023-1051.

[22] Murphy, Kevin J., (1998), "Executive Compensation," unpublished.

[23] Nelson, Richard R. (1961), "Uncertainty, Prediction, and Competitive Equilibrium," Quarterly Journal of Economics, 75, 41-62.

[24] Prescott, Edward C and Visscher, Michael (1980), "Organizational Capital," Journal of Political Economics, 88, 446-461

[25] Prendergast, Canice and Stole, Lars (1996), "Impetuous Youngsters and Jaded Old-Timers: Acquiring a Reputation for learning," Journal of Political Economy, 104, 1105-1133.

[26] Scharfstein, David S. and Stein, Jeremy C. (1990), "Herd Behavior and Investment," American Economic Review, 80, 465-479.

[27] Salinger, Michael A. and Summers, Lawrence H. (1983), "Tax Reform and Corporate Investment: A Microeconometric Simulation Study," *In Behavioral Simulation Methods in Tax Policy Analysis*, edited by Martin Feldstein, Chicago, University of Chicago Press.

[28] Sattinger Michael (1993), "Assignment Models of the Distribution of Earnings," Journal of Economic Literature, 31, 831-880.

[29] Schultz, Theodore W. (1975), "The Value of the Ability to Deal with Disequilibria," Journal of Economic Literature, 13, 827-846.

[30] Schultz, Theodore W. (1980), "Investment in Entrepreneurial Ability," Scandinavian Journal of Economics, 82, 437-448.

[31] Smith, Clifford W. and Watts, Ross (1992), "The Investment Opportunity Set and Corporate Financing, Dividend, and Compensation Policies," Journal of Financial Economics, 32, 263-292.

[32] Stokey, Nancy L. and Lucas, Robert E. (1989), *Recursive Methods in Economic Dynamics*, Cambridge, Harvard University Press.

[33] Summers, Lawrence H. (1981), " Taxation and Corporate Investment: A q-Theory Approach," Brookings Papers on Economic Activity, 1, 67-139.

[34] Takii, Katsuya (1999 a), "The Value of Information, Investment and Adjustment Cost," unpublished

[35] Takii, Katsuya (1999 b), "Learning by Doing and Education," unpublished.

[36] Takii, Katsuya (1999, c), "Information Technology and Prediction Ability," unpublished

[37] Ueda, Kazuo and Yoshikawa, Hiroshi (1986), "Financial Volatility and the q Theory of Investment," Economica, 53, 11-27.

[38] Welch, F. (1971), "Education in Production," Journal of Political Economics, 78, 35-59.