RIETI Discussion Paper Series 15-E-027

University Prestige, Performance Evaluation, and Promotion: Estimating the employer learning model using personnel datasets

ARAKI Shota
RIETI

KAWAGUCHI Daiji
RIETI

ONOZUKA Yuki
Hitotsubashi University / University of Western Ontario

RIETI
Research Institute of Economy, Trade & Industry, IAA

The Research Institute of Economy, Trade and Industry
http://www.rieti.go.jp/en/

RIETI Discussion Paper Series 15-E-027
March 2015

University Prestige, Performance Evaluation, and Promotion: Estimating the employer learning model using personnel datasets*

ARAKI Shota
Hitotsubashi University and RIETI

KAWAGUCHI Daiji
Hitotsubashi University, RIETI and IZA

ONOZUKA Yuki
Hitotsubashi University and University of Western Ontario

Abstract

The employer learning model postulates that employers form employees' prior ability distribution from educational credentials and update its distribution by observing workers' performance on the job. This paper estimates the employer learning model for university-graduate white-collar workers using personnel datasets from two large manufacturers that contain rich information, including the name of the university from which the worker graduated, annual performance evaluations, and position in the promotion ladder. The estimates indicate that employers learn workers' ability relatively quickly through observing their performance on the job. The initial expectation errors on ability decline by a half in about three to four years in the two companies. Companies promote graduates of elite schools quickly mainly because they tend to perform better on the job.

Keywords: Bayesian learning, Promotion, Personnel data, Internal labor market.
JEL classification: J46

*This paper is conducted as a part of the project "Economic analysis of human resource allocation mechanisms within the firm: Insider econometrics using HR data," undertaken at Research Institute of Economy, Trade and Industry (RIETI). We thank two anonymous companies for providing their personnel records. We also appreciate comments from Takao Kato, Katsuyuki Kubo, Hideo Owan, Katsuya Takii, Fabian Lange, Edward Lazear, Alister Munro, Michael Waldman, Kengo Yasui, Shintaro Yamaguchi, and seminar participants at University of Tokyo, RIETI, Waseda University, Academia Sinica, Aahus University, Kansai Labor Workshop, and GRIPS.

1 Introduction

Hiring a worker from the labor market is a risky endeavor for an employer, because even careful examination of a resume and several interviews do not entirely reveal the qualifications of a worker that will determine her performance on a specific job. Consequently, employers are both routinely fortunate to have over-performers who surpass prior expectations and doomed by under-performers. Then, how long does it take until the employer is convinced whether he is fortunate or doomed by a specific employee? This is the question that we address in this paper. Analysis of personnel data sets from two large manufacturers of Japan reveals that it does not take long before employers are convinced that they are either fortunate or doomed.

The employer learning model postulates the process by which an employer learns about a specific employee's ability as a series of Bayesian updates of the prior distribution of the employee's ability; the employer first forms each worker's prior distribution of ability from educational background, prior labor-market experience, and impressions from interviews, and then the employer subsequently updates the ability distribution by observing her performance on the job. Estimating the speed of this employer's learning process is crucial to shed light on the two fundamental roles of education in the labor market: to promote human capital production and to credibly signal high-ability workers to employers. Disentangling these two roles from observations is difficult because two theories produce an observationally equivalent prediction: the higher a worker's educational attainment, the higher the worker's earnings. Distinguishing the two roles is critically important, however, because public funds are spent on education based on the presumption that education enhances workers' productivity. Moreover, disentangling the signaling role of education helps quantify the significance of information asymmetry in the labor market.

A recent literature empirically estimate the employer learning in which the employer determines the prior distribution of a worker's ability based on his schooling and updates the posterior distribution based on his performance on the job. Farber and Gibbons (1996) and

Altonji and Pierret (2001) test the model's prediction using the National Longitudinal Survey of Youth 1979 (NLSY 1979) by regressing earnings on education and the Armed Force Qualification Test (AFQT) score, allowing for variable slopes depending on the worker's labor-market experience. Both studies find that the return to schooling decreases, while the return to AFQT score increases as a worker accumulates labor-market experience. The findings are consistent with the employer learning model under the presumption that the worker's ability, approximated by the AFQT score, determines performance on the job and the employer sequentially updates the posterior distribution of the worker's ability observing his performance. Lange (2007) recovers the structural estimates of the speed of employer's learning from the reduced-form estimates and concludes that the initial expectation error on ability reduces by one half after three years of labor-market experience. This relatively speedy learning implies the relative importance of the human-capital role of education and the less significant information asymmetry on an employee's ability between an employer and an employee. Recent developments of the literature based on NLSY 1979 further ask whether the learned information about an employee's ability is the current employer's private information or public information that is shared by employers in the labor market to determine the significance of the labor-market friction caused by information asymmetry among employers (Schönberg (2007), Pinkston (2009), Kim and Usui (2012), and Kahn (2013)). Furthermore, Mansour (2012) and Light and McGee (2012) examine heterogeneity in the importance of learning across occupations.

Acknowledging the innovation that previous studies bring to the literature, existing studies have several limitations because of the usage of NLSY 1979, which only records workers' earnings, years of schooling, and AFQT score. Researchers need to make several critical assumptions to fill the gap between the employer learning model and observable variables. We point to two important limitations resulting from the assumptions.

First, absence of the employer's performance evaluation on the job requires that the AFQT score correlates with the performance that the employer observes. This is a reasonable assumption to maintain, but it significantly limits the scope of the signal that education can send to the labor market. Recent studies reveal that non-cognitive ability is a crucial determinant of both labor-market success and educational attainment. Therefore, education can potentially send credible signals of both cognitive and non-cognitive abilities to the labor market (Heckman et al. (2006)). Since AFQT is a proxy variable for cognitive ability, previous studies that capture the signaling role of education for cognitive ability but not for non-cognitive ability may have underestimated the signaling value of education.

Second, absence of the name of a worker's alma mater compels researchers to compare the wage paths of workers with different years of schooling assuming that all high-school graduates and college graduates, respectively, have identical amounts of human capital accumulation. The assumption on identical human capital accumulation paths is not innocuous, given the evidence for the heterogeneity of ability development suggested by personnel data regarding performance evaluations (Kahn and Lange (2013)). Moreover, it is widely known that individuals with higher ability tend to attend elite schools (Dale and Krueger (2002), Hoekstra (2009) and Dale and Krueger (2011)) and that employers are likely to form different expectations about workers' ability based on the prestige level of the university from which the workers graduated. Indeed, using a Chilean administrative data set, a recent study by Bordón and Braga (2013) reports that employers infer workers' ability from the prestige of the universities from which workers graduate at the early stage of workers' career, but its importance fades quickly as workers accumulate labor-market experience.[1]

Kahn and Lange (2013)'s contribution is the most closely related to our paper, in that they estimate an employer learning model with a personnel panel data set. Their goal departs from the original motivation of the employer learning model to decompose the increasing wage heterogeneity over careers into the effect of employer learning and the heterogeneous evolution of human capital. Examining the inter-temporal covariance structure of wage and

[1] Bordón and Braga (2013) use the test score and the admission cutoff to implement a regression discontinuity analysis. They do not structurally recover the parameters that govern the speed of employer learning.

performance evaluation, they find that current wage is correlated not only with past performance but also with future performance. They interpret the correlation of current wage and future performance evaluation as evidence for heterogeneous human capital accumulation, because workers with high learning ability tend to have high future human capital, resulting in high future performance evaluation, and it is correlated with current wage through current human capital. At the same time, they find that the covariance of current wage and past performance is discontinuously larger than the covariance between current wage and future performance. They interpret this as an evidence for employer learning, because the employer learns about its workers' ability only from past and present information and determines their wages. From their findings, they conclude that both employer learning and heterogeneous human capital accumulation are important elements to explain the heterogeneous wage process across workers. According to their results, employers and employees continuously update their beliefs about employers' abilities because they are hitting moving targets.

This paper contributes to the literature on the estimation of the employer learning model using personnel panel datasets, departing from Kahn and Lange (2013) in two important directions. The first contribution is to include the role of education into the estimation of the employer learning model, as education was the original motivation of the literature. We transform the name of the university from which each worker graduated into a standardized test score and examine how university prestige affects promotion decisions across career stages. Incorporating the role of education enables us to quantify the importance of the signaling role of education relative to the annual performance evaluation, as Lange (2007) did based on NLSY 1979. The second contribution is integrating the employer learning model into the model of the promotion decision by Gibbons and Waldman (1999) and Gibbons and Waldman (2006) and deriving the explicit formula for the promotion probability as a function of school prestige and past performance evaluation; the formula nests Lazear (2004)'s result as a special case. The closed-form formula allows us to directly estimate the structural parameters of the employer learning model. Relating employer learning with promotion, but not with wage, frees us from modeling the relationship between wage and expected productivity.

Our analysis assumes that each employee's ability is fixed and that employers gradually learn the fixed target, but this assumption does not conform with the Kahn and Lange (2013)'s finding that agents continuously learn workers' ability over careers. We argue that an employee's ability evolves over a career partly because of promotion and the reallocation of workers across jobs because workers' ability does not have single dimension. For example, a worker who works well as a staff member may not be a good manager. To allow for the evolution of ability over an employee's career because of promotions, we analyze the promotions to different job ranks separately; namely, we analyze the promotion from staff member to junior manager independently from the promotion from junior manager to senior manager. Focusing on promotions from a single job grade to the next grade arguably makes the fixed ability assumption more plausible.

We analyze personnel data from two large-scale manufacturers of Japan, which hire about 6,000 and 8,000 employees, respectively. The sample period is between 2005 and 2010 for the first company and between 2004 and 2010 for the second company. Annual panel data for each worker include the name of the worker's alma mater, current job rank, current performance evaluation, and current wage. The career tree analysis indicates that workers who graduated from prestigious universities are promoted more quickly than their peers who graduated from non-elite universities. Descriptive statistics indicate that workers with higher ranks have higher past performance evaluations than workers with lower rank, on average, while higher-ranked workers did not necessarily graduate from a higher-ranked school. The fact that those who are in higher-ranked jobs did not necessarily graduate from better schools implies that those elite-school graduates are promoted more quickly than others because they tend to receive higher performance evaluations on the job. The analysis of the determination of subjective evaluations indeed reveals that graduates from better schools tend to receive better subjective performance evaluations. A duration analysis of a promotion from a current rank to the next rank reveals that school prestige and past performance evaluations are equally important for promotion in the early career stage, whereas past performance evaluations become more important in the later career stage. Therefore, the quick promotion of elite-school graduates, particularly the promotion to higher job ranks, is largely due to these graduates' better performance on the job. All this descriptive evidence corroborates with the prediction of the employer learning model that employers shift weight from the name of the school to past performance on the job as they obtain more information.

To quantify the importance of performance evaluation relative to school prestige, we structurally estimate a promotion model with employer learning that recovers the model's primitive parameters. The prestige of the university from which a worker graduated contributes to forming the employer's prior distribution of the worker's ability. With the caveat that estimation results depend on functional-form assumptions, the estimates indicate that signaling and human capital roles of university prestige are equally important on the initial promotion decision but that the signaling role of education depreciates over time. Employers Bayesian update workers' ability relatively quickly through annual performance evaluations. The variance of the initial subjective ability distribution decreases by one half in about 3 to 4 years in both Companies A and B.

Overall, the empirical analysis shows that employers form a subjective distribution of workers' ability, combining information on school prestige and performance evaluations. Learning on the job occurs relatively quickly, and elite-school graduates are promoted quickly because they tend to perform better on the job.

2 Model of promotion with employer learning

We develop a model of the employer's promotion decision when the worker's ability is not perfectly observable by integrating the employer learning model by Farber and Gibbons (1996) and Altonji and Pierret (2001) into the promotion model developed by Gibbons and

Waldman (1999) and Gibbons and Waldman (2006).

2.1 Employer learning of worker's ability

Let the productivity of the employee i in the period t be y_{it}. The productivity of the employee with the educational credential s_i, the labor-market experience by the year before x_{it-1}, and unobserved ability a_i, is decomposed into a deterministic human capital part, unobserved ability, and random shock u_{it}, as

$$y_{it} = h(s_i, x_{it-1}) + a_i + u_{it}, \tag{1}$$

where $h(s_i, x_{it-1})$ is the human capital function whose functional form is known to the employer. We assume that the human capital production function $h(s_i, x_{it-1})$ is linearly separable in terms of educational credential s_i and the experience by the previous year x_{it-1} so that $\partial^2 h(s_i, x_{it-1})/\partial s_i \partial x_{it-1} = 0$. Human capital depends only on schooling and general labor-market experience, abstracting from firm-specific human capital. The ability a_i is not known by the employer, but the employer has its prior distribution, $a_i|s_i, x_{it-1} = 0 \sim N(\mu_0(s_i), \nu_0^2)$ when the employee enters labor market. We denote the subjective distribution function as $F_t(a_i)$ and its probability density function as $f_t(a_i)$.

The employer updates its subjective distribution $F_t(a_i)$ from the realization of y_{it} but never can pin down the value of a_i within a finite time because of the random shock to the productivity u_{it} that is identically independently distributed (i.i.d.) and independent from the distribution of a_i. The productivity shock follows the distribution $u_i|s_i, x_{it-1}, a_i \sim N(0, \sigma^2)$. The employer learns the employee's ability from the performance adjusted for the human capital function, $z_{it} = y_{it} - h(s_i, x_{it-1}) = a_i + u_{it}$. The performance adjusted for human capital function z_{it} corresponds to the employer's evaluation of the employee that discounts the expected performance based on the employee's educational credential and experience. We assume symmetric learning, meaning that the updated distribution of the current employers is shared by all employers in the labor market. Using the standard results of Bayesian learning, the subjective distribution of employee i's ability in the period t is expressed as:

$$\begin{aligned} a_i|s_i, x_{it}, z_{i1}, z_{i2}, ..., z_{it} &\sim N(\mu_t, \nu_t^2) \\ &= N\left(\frac{\mu_0(s_i) + x_{it}(\nu_0^2/\sigma^2)\bar{z}_{it}}{1 + x_{it}(\nu_0^2/\sigma^2)}, \frac{\nu_0^2}{1 + x_{it}(\nu_0^2/\sigma^2)}\right), \end{aligned} \tag{2}$$

where $\bar{z}_{it}$ is the past average of unexpected productivity that is $\bar{z}_{it} = (1/x_{it})\sum_{j=1}^{x_{it}} z_{ij}$. The mean expression shows that, as time goes by, the employer shifts weight from the educational credential $\mu_0(s_i)$ to the past average performance evaluation $\bar{z}_{it}$. The speed of the weight shift from school prestige to past performance is rapid when ν_0^2/σ^2 is large; when the employee's ability is heterogeneous (large ν_0^2); and when noise is small (small σ^2). As x_{it} goes to infinity, given $\nu_0^2/\sigma^2 > 0$, $\bar{z}$ converges to a_i because $(1/x_{it})\sum_{j=1}^{x_{it}} u_{ij} \to E(u_{it}) = 0$. The speed of convergence again depends on the relative variance of the employee's ability to the noise in performance, which is ν_0^2/σ^2.

2.2 Promotion rule when the employee's ability is uncertain

Our model of promotion is based on Gibbons and Waldman (1999) and Gibbons and Waldman (2006), with the modification of functional form assumptions on human capital formation. The employer can assign the worker to either an easy job or a difficult job. The output of the worker with the productivity y_{it} in the easy job is

$$\alpha + \beta y_{it} \tag{3}$$

and the productivity in the difficult job is

$$\gamma + \delta y_{it} \tag{4}$$

with $\alpha > \gamma$ and $\delta > \beta$. A worker with high productivity has a comparative advantage in the difficult job, and thus the employer can maximize production by assigning the worker to the difficult job if the worker satisfies $y_{it} > (\alpha - \gamma)/(\delta - \beta)$. Assigning an employee to a difficult job is called a promotion.

In each period, the employer assigns the employee to either job based on the expected output of each employee in either job. The expected productivities of the employee with $s_i, x_{it}, \bar{z}_{it}$ in easy job at $t+1$ is

$$\begin{aligned} & E(\alpha + \beta y_{it+1} \mid s_i, x_{it}, \bar{z}_{it}) \\ =& \alpha + \beta E(y_{it+1} \mid s_i, x_{it}, \bar{z}_{it}) \end{aligned} \tag{5}$$

and that in difficult job is

$$\begin{aligned} & E(\gamma + \delta y_{it+1} \mid s_i, x_{it}, \bar{z}_{it}) \\ =& \gamma + \delta E(y_{it+1} \mid s_i, x_{it}, \bar{z}_{it}). \end{aligned} \tag{6}$$

The employee is promoted if the expected output in the difficult job exceeds that in the easy job, which is

$$\begin{aligned} & E(y_{it+1} \mid s_i, x_{it}, \bar{z}_{it}) \\ =& h(s_i, x_{it}) + E(a_i \mid s_i, x_{it}, \bar{z}_{it}) \\ =& h(s_i, x_{it}) + \frac{\mu_0(s_i) + x_{it}(\nu_0^2/\sigma^2)\bar{z}_{it}}{1 + x_{it}(\nu_0^2/\sigma^2)} \\ >& \frac{\alpha - \gamma}{\delta - \beta}. \end{aligned} \tag{7}$$

Figure 1 illustrates this promotion rule. Workers with expected productivity exceeding the threshold are promoted, because the expected output in a difficult job is more sensitive to worker's expected productivity than in an easy job.

This solution includes the promotion rule proposed in Lazear (2004) as a special case. His model abstracts human capital accumulation so that $h(s_i, x_{it}) = 0$. Consequently, $z_{it} =$

$y_{it} - h(z_i, x_{it}) = y_{it}$. He also abstracts the signaling role of education, and thus $\mu_0(s_i) = \mu_0$. Lazear (2004) considers the employer's promotion decision after observing the employer's single-period performance y_{i1}. Substituting $t = 1$ and $\bar{z}_{it} = y_{i1}$ into the above inequality results in the promotion rule

$$y_{i1} > \frac{\alpha - \gamma}{\delta - \beta} + \frac{\sigma^2}{\nu_0^2}\left(\frac{\alpha - \gamma}{\delta - \beta} - \mu_0\right). \tag{8}$$

Although Lazear (2004) did not obtain the closed-form solution, our normality assumptions on a_i and u_{it} allow us to obtain it.[2] If there is no idiosyncrasy in productivity, the promotion threshold is set at $(\alpha - \gamma)/(\delta - \beta)$, whereas the employer either inflates or deflates the promotion threshold in the presence of idiosyncrasy in productivity. If an employee with average ability is ineligible for the promotion, which is the case close to reality, then $(\alpha - \gamma)/(\delta - \beta) - \mu_0 > 0$. Therefore, the employer inflates the promotion threshold, expecting the employee's decline after the promotion because of the mean reversion. The inflation of the threshold is larger when the employee's performance heavily depends on luck and lightly depends on ability, so that σ^2/ν_0^2 is large. This inflated promotion threshold corroborates with the fact that demotions are rarely observed in reality. Performance decline after a promotion typically does not result in demotion, because the employer promotes only a small fraction of workers who far exceed the break-even point. This observation asserts that we do not need an extra model to explain why we rarely observe demotion in our data set.

Going back to the promotion rule expressed in (7), we obtain several important comparative statics results. An employee who graduated from a reputable university is more likely to be promoted as

$$\frac{\partial E(y_{it+1} \mid s_i, x_{it}, \bar{z}_{it})}{\partial s_i} = \frac{\partial h(s_i, x_{it})}{\partial s_i} + \frac{1}{1 + x_{it}(\nu_0^2/\sigma^2)} \frac{\partial \mu_0(s_i)}{\partial s_i} > 0. \tag{9}$$

The two terms in the partial derivative correspond to two effects through which education

[2] Araki and Kawaguchi (2014) directly solve the Lazear (2004)'s model under the bivariate normality assumption of ability and idiosyncratic shock distributions to obtain an identical result.

enhances promotion probability. The first effect is to increase human capital, and the second effect is to credibly signal better ability to employers. The human capital effect is time constant because of the separability assumption $\partial^2 h(s_i, x_{it})/\partial s_i \partial x_{it} = 0$, while the signaling effect declines over time because the signaling value depreciates, as

$$\frac{\partial^2 E(y_{it+1} \mid s_i, x_{it}, \bar{z}_{it})}{\partial s_i \partial x_{it}} = -\frac{\nu_0^2/\sigma^2}{[1 + x_{it}(\nu_0^2/\sigma^2)]^2} \frac{\partial \mu_0(s_i)}{\partial s_i} < 0. \tag{10}$$

These distinct properties of the human capital and signaling effects of education on promotion enable us to empirically distinguish two roles of education on promotion. Therefore, the separability assumption on the human capital production function, $\partial^2 h(s_i, x_{it})/\partial s_i \partial x_{it} = 0$, is the crucial assumption for separating the two effects of education. In the discussion section, we will extensively discuss how the estimation results are affected by the violation of this assumption. Note that the decline in the signaling role would not be observed without holding past performance ($\bar{z}_{it}$) constant, because education and past performance are likely to be positively correlated.

Receiving a higher average evaluation in the past increases the promotion probability, and its importance increases as the employee's career progresses, as

$$\frac{\partial E(y_{it+1} \mid s_i, x_{it}, \bar{z}_{it})}{\partial \bar{z}_{it}} = \frac{x_{it}(\nu_0^2/\sigma^2)}{1 + x_{it}(\nu_0^2/\sigma^2)} > 0 \tag{11}$$

and

$$\frac{\partial^2 E(y_{it+1} \mid s_i, x_{it}, \bar{z}_{it})}{\partial \bar{z}_{it} \partial x_{it}} = \frac{\nu_0^2/\sigma^2}{[1 + x_{it}(\nu_0^2/\sigma^2)]^2} > 0. \tag{12}$$

The importance of evaluation increases because the average past evaluation becomes a reliable proxy for ability as evaluation points increase.

3 Data

We use personnel data sets of two manufacturing companies that hire about 8,000 and 6,000 employees, respectively, at domestic establishments in the sample period. Both firms operate internationally, each hiring about 20,000 employees worldwide, but information about employees working outside of Japan is not included in our data sets. Each firm uses a skill-based grading system (*Shokuno shikaku kyu*), in which each employee is ranked based on his skill and compensation heavily depends on the job rank. Job ranks are related to job titles, such as section manager, but ranks do not have one-to-one correspondence with job titles. Throughout this paper, we use job rank as a measure of promotion, because the job rank is clearly ordered, while there are many lateral job titles that cannot be rank ordered.

The monthly record of each employee includes ability rank, name of alma mater, annual performance evaluation, along with the usual information, such as the amount of monthly salary, the amount of biannual bonus payment, and hours worked per month. The sample periods of the first company, Company A, and the second company, Company B, are between 1991 and 2010, but performance evaluations are available between 2005 and 2010 for Company A and between 2004 and 2010 for Company B. We use the all past evaluations available in our empirical study but its length is limited up to 6 years for Company A and 7 years for Company B. We assume that employers make promotion decisions in year t based on the past evaluations up to year $t-1$. Therefore, the promotions are analyzed between 2006 and 2010 for Company A and between 2005 and 2010 for Company B.

We restrict the analysis sample to college-graduate white-collar workers, because white-collar workers and blue-collar workers are on different career tracks in the sample manufacturing firms. We also restrict the sample to male workers, because there are relatively few female workers in these companies and male and female workers are typically on different career tracks. We further drop doctoral-degree holders from the analysis sample, because they typically follow a different career track prepared for professionals. The name of the university from which a worker graduated is transformed into the standardized test score index published by Kawaijuku, a preparatory school that administers a series of large-scale mock examinations and collects follow-up surveys after university entrance examinations, asking to which universities mock-examination takers are admitted. Kawaijuku analyzes the relationship between performance on the mock examinations and performance on the entrance examinations to determine the competitiveness of admission to each school, summarized by the score called *hensachi*. The score index has a mean of 50 and a standard deviation of 10, and its distribution roughly follows a normal distribution. We list several examples to give a sense of the index: The economics department of the University of Tokyo has 70, Kyoto University 67.5, Hitotsubashi University 65, Osaka University 62.5, Waseda 67.5, and Keio 67.5. Since national universities impose more academic subjects on applicants in the entrance examination, the score is not directly comparable between national and private universities. To make the scores comparable, we add 5 to the score of national universities, because this attains good fit in the duration analysis for Companies A and B.

4 Descriptive analysis of promotion in two companies

4.1 Company A

Company A is a large manufacturing company hiring university graduates as engineers and other clerical staff members (sales, marketing, accounting, and legal, for example). All workers in Company A are classified into one of ten skill ranks, from J1 to G1. To convey an overview of the promotion ladder, career trees are presented in Figures 2 and 3.

Figure 2 illustrates the annual skill-rank transition of first-tier national university graduates in Company A.[3] Among freshly recruited workers, those who finish undergraduate and Master's degrees are hired as J1. All who are hired as J1 are promoted to J2 in the next year, and almost all of them are subsequently promoted to SA in the following year. One third of SAs repeat the same grade, but the rest are promoted to SB. So, the promotion up

[3] 1st tier national universities include Tokyo, Kyoto, Tokyo Institute of Technology, Hitotsubashi and Osaka.

to SB is almost automatic for those who graduated from elite schools, but the real challenges come at promotions to G6 and above.

Figure 3 presents the transition matrix for third-tier private university graduates.[4] Compared with the graduates from first-tier national universities, graduates from the third tier are significantly more likely to repeat J2, SA, and SB. Their transition probability from SB to G6 is significantly lower than that of graduates from first-tier national universities. Once promoted to G6, however, the transition probabilities to the ranks above are not significantly different from those of the graduates from first-tier national universities.

Company A implements an annual employee evaluation. This evaluation is used mainly for determining the bonus payment and promotion decision. Evaluators are generally immediate section chiefs in the case of non-managerial workers. In the case of managerial workers, upper-level managers generally serve as evaluators. Evaluations are given relative to the expected performance set for each worker. At the beginning of the evaluation period, an evaluator and an evaluatee meet in face to face to confirm the expected performance in the evaluation period. At the end of the evaluation period, the evaluator evaluates each worker's performance relative to the expected performance.

Table 1 tabulates the distribution of evaluations by skill rank. In the first stage of the career, J1, almost everyone receives A2 as her evaluation, but as workers' careers progress, some start receiving a better evaluation, S, while others receive a worse evaluation, A3. The divergence of the evaluation distribution continues until SB. After G6, a non-negligible fraction of workers receives the high evaluation, S, while the fraction receiving an average evaluation A1 increases again. The divergence of evaluations along the career progression shows that each worker's productivity becomes apparent at the upper level of the corporate hierarchy. This tendency is reasonable, considering that workers are given more discretion as they are promoted.

One may be concerned that those who graduated from elite universities do not necessarily

[4]third-tier private universities include private universities other than Keio, Waseda, Sophia, Meiji, Aoyama, St. Paul's, Chuo, Hosei, Kwanseigakuin, Kansai, Doshisha, Ritsumeikan, Gakushuin, and Nanzan.

have greater ability than those who graduated from non-elite universities, given that they are recruited by the same company after an intensive screening process. To address this reasonable concern, we regress the current evaluation on the university score of the worker's alma mater and the worker's tenure, along with rank dummy variables. Table 2 tabulates the estimation results. The estimated coefficients for the university score indicate that workers from better universities are more likely to receive higher evaluations, given tenure and job rank. Conditional on job rank, workers with longer tenure tend to receive lower evaluations. This relationship is perhaps because those who stay in the same rank for a longer period are those who are passed over for promotion and are less able workers, on average. Combined with the evidence that graduates from elite universities are promoted faster than graduates from non-elite universities, it is fair to claim that those who graduated from elite universities have higher ability than those who did not.

Table 3 tabulates the descriptive statistics of the analysis sample. The mean of university score does not vary significantly across job ability ranks. Years of job tenure generally increase as the rank increases. On average, those who are in entry-level managerial positions, G6, have ten years of job tenure. The evaluation distribution in this table tabulates the past average of dummy variables corresponding to each evaluation category. The mean of receiving high evaluation increases significantly as the ability rank goes up. The combination of stable mean university score and increase of high performance evaluation implies that a high performance evaluation becomes more important for upper-level promotions. It is also worth noting that the stable mean school score and early promotion of elite-school graduates evident in the career tree implies that elite-school graduates are promoted quickly because of high-performance evaluations.

We regress the speed of promotion from the current ability rank to the next ability rank on school name and the average past evaluation to quantify the relative importance of the two variables using a duration model. Let T be the time to be promoted with cdf $F(t) = P(T \leq t)$ and pdf $f(t)$. The hazard to be promoted at time t is denoted as $\lambda(t) = f(t)/(1 - F(t))$.

We allow for the hazard rate to depend on workers' characteristics by adopting the Cox proportional hazard model:

$$\lambda(t, x) = \lambda_0(t) \exp(\beta_0 + \beta_1 s_i + \beta_2 \bar{z}_{it}),$$

where t is the tenure in the company, s_i is the test score of the university from which the worker i graduated, and $\bar{z}_{it}$ is the average performance evaluation in the past years. Practically, we reclassify SS and S as high evaluation; A1 and A2 as medium evaluation; and A3, B, and C as low evaluation, and create dummy variables for high and low evaluation, and take the average of past years. The baseline hazard function is not parametrically specified.

The maximum-likelihood estimates are tabulated in Table 4. The estimated coefficients for university score are all positive and statistically significant, except for promotion to SA. The estimated coefficient ranges from 0.0104 to 0.0419, meaning that a 10-point (1 standard deviation) increase in university score increases the promotion probability by 10 to 40 percent at any given time period. The effect of graduating from a prestigious university is most pronounced at the promotion to G6 and declines afterward, though it is difficult to give a clear explanation for this non-monotonic relationship. The estimated coefficient for low evaluation is negative and statistically significant for all promotions except for the promotion to SA. The estimates monotonically decrease from -0.2671 in the promotion to SA to -2.7543 in the promotion to G4, implying that receiving a low evaluation has a detrimental effect on promotion and that its negative impact grows at the upper level of the career ladder. The coefficients for high evaluation are not statistically significant in lower-level promotions, that is, promotions to SA, SB, and G6. Receiving a high evaluation has statistically significant positive impacts on promotions to G5 and G4, however. Receiving a high evaluation consistently in the past increases the promotion probability by 37% in the promotion to G5 and 78% in the promotion to G4. Overall, as workers climb the career ladder, school prestige become less important, particularly after G6, and performance evaluation becomes more important. These multiple regression results corroborate with the prediction from the employer learning model that school prestige is important in the early career stage, while performance evaluation is important in the late career stage.

4.2 Company B

Company B is a large car parts manufacturing company that hires university graduates as engineers and other clerical staff (e.g., sales, marketing, accounting, and legal). All workers in Company B are classified into one of seven skill ranks, from 1st class to manager. The promotion ladders for prestigious national university graduates and non-elite private university graduates are illustrated in Figures 4 and 5. All newly hired workers are assigned to the 1st class. Among 1st class workers, about 45% of prestigious national university graduates are promoted to the 2nd rank, while only 22% of non-elite private university graduates are promoted. The significant difference in the promotion speed depending on workers' alma mater persists until the 3rd rank. There is no significant difference in the promotion probabilities to vice supervisor, supervisor, and vice manager, but the probability of promotion to manager again differs significantly. Overall, elite-school graduates are promoted quickly, particularly at the early career stage. It is worth noting that only a small fraction of workers quit during the entire promotion process until manager status.

Each worker receives an annual performance evaluation on a four-point scale from A to D. The section chief, who supervises 10 to 20 workers, typically is responsible for the initial evaluation, and the human-resource-management section curves the distribution. As in Company A, evaluations are given relative to the expected performance set for each worker. Table 5 tabulates the distribution of annual performance evaluations by skill rank. We restrict our analysis to the ranks up to the 5th skill rank, supervisor, because the number of observations above supervisor is limited. In the first stage of skill rank, about 70% of workers receive C, 28% of workers receive B, and only a few receive A and D. The distribution of the evaluations becomes slightly heterogeneous up to the second level but stays constant afterward. Among supervisors, about 4% receives A, 36% receives B, 60% receives C, and less than 1% receives D. Since only a few workers receive A as the evaluation, we bundle A and B together in the subsequent promotion analysis.

Table 6 examines if graduates from prestigious universities receive high performance evaluations, to assert whether university prestige works as a signal for high ability by regressing evaluation on university score, tenure, and rank dummy variables. Although the size of the effects is not very pronounced, workers who graduated from more reputable universities tend to receive higher evaluations. Similar to the results in Company A, conditional on rank, those who have longer tenure are less likely to receive high evaluations. To summarize, while many receive Cs as the performance, as shown in Table 5, graduates from prestigious universities are more likely to receive A and B and less likely to receive D. This finding conforms with the assumption that graduates from better schools have higher ability.

Table 7 reports the descriptive statistics of the analysis sample of company B. The average school score slightly decreases as the skill rank goes up. The average years of tenure help us understand at which career stage each skill rank stands. The average tenure of supervisors, for example, is about 15 years. Because this is a sample of college graduates, supervisors are slightly younger than 40 years old, on average. Finally, the fractions of high and low evaluations in the past years are almost equally distributed across ability ranks, except for 1st class. Combined with the decreasing importance of the university score as the rank goes up, this finding is consistent with the prediction from the employer learning model that performance evaluations become relatively more important as the career progresses.

Table 8 reports the regression estimates of the duration until promotion on school score and past evaluation dummy variables using the Cox proportional hazard model. As the ability rank goes up, the coefficient for the university score increases slightly, whereas the coefficient for low evaluation decreases and that for high evaluation increases. This descriptive result indicates that performance evaluations become more important relative to the alma mater's reputation as a worker's career progresses.

5 Estimation of the structural model of promotion

We discuss the structural estimation of the employer learning model in this section. The model assumes that the employer learns the employee's fixed ability from university prestige and performance evaluation. Learning on a fixed ability is the key assumption, and we argue that this assumption is realistic when we analyze a promotion from one job rank to the next job rank. We regard Kahn and Lange (2013)'s continuous evolution of the employee's ability on the job, and consequent continuous learning on the moving target, as very realistic, when we consider the long-term career developments of workers represented by wage evolution. We consider, however, that the continuous evolution of a worker's ability on the job career occurs because of promotion and relocation. In a textbook example, a worker who is eligible as a staff member is not necessarily a good manager, because the required skill sets for a staff member and a manager are presumably different (Lazear and Gibbs (2009)). Since a worker staying in a skill rank is likely to be involved in similar tasks, the fixed ability assumption is arguably innocuous when we analyze the employer's promotion decision from one level to the next. Of course, the problem of misalignment remains when the abilities required in the current job rank and those in the next job rank are different, but as far as two abilities are positively correlated, the best the employer can do is to promote workers whose expected productivity is high in the current job rank.[5]

The theoretical discussion derives the promotion rule as the equation (7) that is

$$E(y_{it+1}|s_i, x_{it}, \bar{z}_{it}) > \frac{\alpha - \gamma}{\delta - \beta}, \tag{13}$$

where y_{it+1} is the productivity of the worker i in year $t{+}1$, s_i is university score, x_{it} is tenure in the company, and $\bar{z}_{it}$ is the average past evaluation. This condition implies that workers who graduate from the same university s_i and share the same experience x_{it} and the same past evaluation $\bar{z}_{it}$ should share the same outcome: either promoted or not promoted. Within

[5] If different job ranks require heterogeneous skill sets, we need to consider multidimensional human capital to analyze the promotion. Such an analysis is beyond the scope of this paper and left for future research.

a group of workers who share the same observable characteristics in the analysis sample, however, there is heterogeneity regarding promotion. To allow for promotion determinants that are unobserved by econometricians but observed by the firm, we add a stochastic factor e_{it} to the equation for the promotion rule so that the promotion occurs.

$$E(y_{it+1}|s_i, x_{it}, \bar{z}_{it}) + e_{it} > (\alpha - \gamma)/(\delta - \beta). \tag{14}$$

The unobserved promotion determinant includes a sudden vacancy of a managerial position as an example (Ariga (2006)) and is assumed to be exogenous from school prestige or past performance evaluations. Imposing the normality assumption, the idiosyncratic determinant of promotion is distributed as $e_{it}|s_i, x_{it}, z_{i1}, z_{i2}, ..., z_{it} \sim N(0, 1)$.[6]

Using the relationship $E(y_{it+1}|s_i, x_{it}, \bar{z}_{it}) = h(s_i, x_{it}) + E(a_i|s_i, x_{it}, \bar{z}_{it})$, the probability of promotion is given as:

$$P\left[e_{it} < h(s_i, x_{it}) + E(a_i|s_i, x_{it}, \bar{z}_{it}) - \frac{\alpha - \gamma}{\delta - \beta}\Big| s_i, x_{it}, \bar{z}_{it}\right]. \tag{15}$$

We specify the human capital equation as $h(s_i, x_{it}; \theta) = \theta_0 + \theta_1 s_i + \theta_2 t_i$, where t_i is the tenure in the current ability rank of worker i. The prior distribution of a worker's ability is a linear function of university prestige, such as $\mu_0(s_i) = \pi_0 + \pi_1 s_i$. Given these functional-form assumptions and $E(a_i|s_i, x_{it}, \bar{z}_{it}) = \frac{\mu_0(s_i)+t(\nu_0^2/\sigma^2)\bar{z}_{it}}{1+t(\nu_0^2/\sigma^2)}$, the probability of promotion ($y_{it} = 1$) is specified as:

$$P(y_{it} = 1|t_i, s_i, z_{it}, ..., z_{i1}) = \Phi\left[(\theta_0 + \theta_1 s_i + \theta_2 x_{it-1}) + \frac{\pi_0 + \pi_1 s_i + x_{it-1}(\nu_0^2/\sigma^2)\bar{z}_{it-1}}{1 + x_{it-1}(\nu_0^2/\sigma^2)} - \frac{\alpha - \gamma}{\delta - \beta}\right]. \tag{16}$$

This probit equation includes estimable parameters: $\theta_1, \theta_2, \pi_0, \pi_1, (\nu_0^2/\sigma^2)$. Note that the human capital value of education, θ_1, and the signaling value of education, π_1, can be separately identified because the human capital role is time constant, while the information

[6]This assumption implies $e_{it}|s_i, x_{it}, \bar{z}_{it} \sim N(0, 1)$.

role diminishes as x_{it} increases. The coefficient for the constant term only estimates a mixture of $(\alpha-\gamma)/(\delta-\beta)$ and θ_0, but we are not interested in separately identifying each parameter. We assign the value 1 for high evaluation (S and SS for Company A; A and B for Company B) and -1 for low evaluation (A3, B and C for Company A; D for Company B) and take the average of past years to calculate $\bar{z}_{it}$.

Through the estimation of the structural model, we are most interested in the speed of employer learning, because speedier employer learning implies that the signaling value of education is less important and the human capital value of education is more important. The parameter (ν_0^2/σ^2) expresses the speed of employer learning. To illustrate this, remember that the variance of the subjective ability in year t is:

$$\frac{\nu_0^2}{1+x_{it}(\nu_0^2/\sigma^2)}.$$

From this expression, we know that the initial variance is ν_0^2, because $x_{it}=0$. This initial variance declines by one half after x_{it}^* years so that

$$\frac{\nu_0^2}{1+x_{it}^*(\nu_0^2/\sigma^2)}=\frac{\nu_0^2}{2}$$

holds. Therefore, the variance of the subjective ability distribution of each worker declines by one half after $x_{it}^*=(\sigma^2/\nu_0^2)$; we call this x_{it}^* the half-life period of the subjective variance of the ability distribution. Lange (2007) backs out this the half-life period from reduced-form estimates and concludes that it is about 3 years, using years of education as the educational credential and AFQT as the proxy variable for on-the-job performance, based on NLSY 1979.

To estimate the probit model, we separately estimate the promotion to SA, SB, G6, and G5 for Company A and the promotion to 2nd class, 3rd class, vice supervisor, and supervisor for Company B. Promotions to these ranks are considered to be the promotion to an entry-level managerial positions. Remember that average tenure is about 10 years among G6 workers in Company A and 9 years among 3rd class workers in Company B. We include all rank workers below G5 in Company A and below Supervisor in Company B in the analysis sample; for example, to analyze the promotion to G6 in Company A, workers in SB and below are assigned $y_{it} = 0$ and workers in G6 and above are assigned $y_{it} = 1$.

We treat the annual evaluation of each worker as the empirical counterpart of $z_{it} = y_{it} - h(s_i, x_{it-1})$. As explained before, in both companies A and B, evaluators are generally immediate section chiefs in the case of non-managerial workers. In case of managerial workers, upper-level managers generally serve as evaluators. In both companies, evaluations are given relative to the expected performance set for each worker. At the beginning the evaluation period, an evaluator and an evaluatee meet face to face to confirm the expected performance in the evaluation period. At the end of the evaluation period, the evaluator evaluates each worker's performance relative to the expected performance. We assume that the evaluator forms the expectation for the evaluatee based on the human capital of each worker, $h(s_i, x_{it-1})$, along with other factors such as job assignment.

Table 9 tabulates the maximum-likelihood estimates of the structural model for Company A. We report the estimates for the promotions to SA, SB, G6, and G5, because the estimations for G4 and above render unstable results and sometimes do not converge, most probably because of the small number of observations with $y_{it} = 1$. Graduating from a better university and accumulating job tenure contribute to promotions to the upper levels through the accumulation of human capital. Graduating from a better university also contributes to promotions through signaling better ability to employers. To demonstrate the relative importance of human capital and information values of education, we pick up the promotion to G6 in particular, because this is a promotion to entry-level management in the company that generally takes place in the mid-30s and the company seems to make cautious decisions. The coefficients for university scores in the human capital equation are larger than the coefficients in the signaling equation. The relationship of two estimated coefficients implies that the human capital role of university score is more important than the signaling role when a worker initiates his career (i.e., $x_{it} = 0$). The human capital role of university score dominates the signaling role over the whole career of the worker, because the signaling role depreciates over time, while the human capital role remains constant.

The implied half-life period of the subjective variance of ability distribution, σ^2/ν_0^2, is 3.46 (=1/0.2894) years for the promotion to SA, 4.48 years for SB, 3.06 for G6, and 17.33 years for G5. Except for the promotion to G5, the estimated half period is around 3 to 4 years, and these results imply that the signaling value of graduating from a prestigious university reduces by one half in 3 to 4 years. The signaling role of university prestige depreciates quickly, whereas the human capital value persists.

Table 10 tabulates the maximum-likelihood estimates of the structural model for Company B. Graduating from a better university and accumulation of job tenure both increase the probability of promotion. The sizes of coefficients for university score and tenure are comparable for the promotion to 2nd class or 3rd class. This implies that graduating from a better school by 1 standard deviation (10 points) and accumulating 10 years of job tenure have equivalent effects on early-stage promotions. The size of the coefficient for university score in the learning equation is about one half of the size of the university score in the human capital equation. This result suggests that the signaling role of education is relatively small in Company B from the initial period. The implied half-life period of the subjective variance of the ability distribution is 4.48 for the promotion to 2nd class, 3.06 for 3rd class, 1.73 for vice president, and 2.84 for supervisor. While there are variations of the estimated half-life depending on ranks, the estimates between 2 to 4 years imply that the signaling value of graduating from a prestigious university depreciate quickly and its value declines by one half in 2 to 4 years.

For both companies, the structural estimates indicate significant effects of educational credentials on promotion through human capital and signaling channels and a significant effect of job tenure on promotion. Comparing the estimates of the two companies reveals a reasonable similarity of estimates between them, but the estimated information role of education is larger in Company A than in Company B. As discussed in the theory section, the distinction between the two roles of education is based on the distinctive effects of education on promotion. The effect of education's signaling role decreases over a worker's career, whereas that of human capital role persists. Except for G5 of Company A, the estimates for ν^2/σ^2 that governs the speed of learning for the two companies are similar. Thus the information role of educational credentials is more significant in Company A than in Company B over workers' careers. The reason behind this difference is elusive, but the university score may be more strongly correlated with the ability required in Company A than that required in Company B.

The effects of tenure on human capital accumulation are significant in both companies. The size of the tenure coefficients in the human capital equation is comparable to the size of the university score coefficients, implying that graduating from a school with a 1 point (0.1 standard deviation) higher score is equivalent to accumulating one more year of tenure in terms of human capital accumulation.

The estimates imply that employers learn employees' ability relatively quickly. The implied half-life period of the subjective variance of the ability distribution, σ^2/ν_0^2, is around 3 to 4 years across specifications, except for a few cases. These estimates are quite similar to the preferred estimate obtained by Lange (2007) based on completely different methods and datasets. This quick learning of employees' ability by employers implies the importance of the human capital role of education relative to the signaling role. Graduates from elite universities quickly move up career ladders in both companies because they actually perform well on the job, as indicated by the ordered probit estimates reported in Tables 2 and 6.

6 Discussion

6.1 Human capital vs. signaling roles of education

We add a caveat in interpreting the human capital role of education. The key identifying assumption to separate the human capital role of education from the signaling role in our analysis is that the human capital role does not depreciate over time while the signaling role does. If prestigious universities endow its graduates with high human capital and the graduates perform well on the job, then the effect of graduating from prestigious universities is classified as the human capital role, as we intend. If those with high innate ability to persistently perform well on the job tend to have lower cost to prepare for competitive entrance examinations and attend elite universities for the purpose of signaling, however, the effect of graduating from elite universities is attributed to the human capital role, because the effect of graduating from prestigious universities on promotion does not depreciate over time. In this case, the signaling role of education is identified when an individual with low ability is admitted to an elite school by chance and performs poorly on the job, or an individual with high ability is not admitted to an elite school and performs well on the job.

Previous studies on employer learning share the same difficulty in disentangling human capital and signaling roles of schooling. Typical previous studies based on the NLSY79 regress log hourly wage on schooling and test scores with their interaction terms with experience:

$$\ln(wage) = \beta_0 + \beta_1 s + \beta_2 s \times exp + \beta_3 iq + \beta_4 iq \times exp + f(exp) + u,$$

where $wage$ is hourly wage, s is years of schooling, exp is years of labor-market experience, iq is test score such as AFQT score, and $f(exp)$ is a flexible function approximating the experience-wage profile. Consistent with the predictions of the employer learning model, previous studies found $\beta_2 < 0$ and $\beta_4 > 4$. If s and iq are linearly related without error, however, the model cannot be identified because of multicollinearity. Thus, the model is identified off the individuals with high IQ and short schooling or low IQ and long schooling.

We should keep in mind that what is estimated as the human capital role of education includes the effect of ability that is related to better educational attainment and persistent productivity on the job.

6.2 Violation of linear separability

The linear separability of the human capital production function with respect to school quality and experience is the key assumption to separately identify the human capital and signaling roles of education, as discussed in the model section. The observed diminishing effect of school quality on promotion corresponds to the theoretical prediction that workers' true ability is revealed to employers as workers accumulate experiences and the signaling role of education diminishes. If school quality affects the speed of skill acquisition on the job and makes the human capital production function not linearly separable with respect to school quality and experience, the human capital role of education also produces an empirical prediction that the effect of school quality on promotion changes as workers accumulate experiences. Thus, the violation of linear separability of the human capital production function disables the identification strategy to distinguish signaling and human capital roles of school quality from the observed diminishing effect of school quality on promotion.

The violation of linear separability of the human capital production function biases the estimate of the signaling role of education. Although the differential speed of skill acquisition across workers with different educational backgrounds creates increasing or decreasing effects of school quality on promotion, the model attributes all observations to the signaling role of education, and this model misspecification causes the biased estimate. What then is the probable direction of the bias?

The differential speed of skill acquisition among workers with different educational backgrounds is likely to result in the underestimation of the signaling role of education. Previous studies of the internal labor market, such as Gibbons and Waldman (2006), assume that workers with stronger educational backgrounds acquire skills on the job more quickly than workers with weaker educational backgrounds and claim that this assumption is consistent with observations found by Baker et al. (1994). The notion that better-educated workers acquire skills more quickly is intuitively appealing, given the copious empirical research reporting that the nature of human capital production is "skill begets skill" (Heckman (2000)). If workers graduating from better schools acquire skills more quickly, the effect of school quality becomes a more important determinant for promotion as workers accumulate experiences. The increasing influence of education on promotion counteracts the decreasing effect of education predicted from the signaling role of education. Consequently, the speed of the diminishing effect of education on promotion attributed to the signaling role of education becomes attenuated. Therefore, the speed of employer learning estimated in the previous section should be interpreted as the lower-bound estimate of the actual speed of employer learning under a probable scenario in which those workers with better educational backgrounds acquire skills more quickly on the job.

7 Conclusion

We set up the model for how employers learn each employee's productivity from the employee's alma mater and past performance and use its knowledge on promotion decisions. We then estimate the model's parameters governing the speed of learning using personnel records of white-collar workers in two large manufacturing companies in Japan. We are the first to estimate the employer learning model with the role of education based on personnel data sets. The information on each employee's alma mater and job rank in the data sets bring the employer learning model to a realistic testing ground.

The empirical results show that employers are more likely to promote workers with better educational backgrounds and better performance evaluations. The weight given to past performance relative to education increases as the worker's career progresses. This finding corroborates with the employer learning model, a model incorporating employers' Bayesian learning of employees' productivity. Furthermore, the estimates of structural parameters imply that employers learn each employee's productivity quickly, as the subjective variance of each employee's productivity decreases by one half in around 3 to 4 years.

The relatively quick learning of each employee's productivity cannot be generalized to the whole labor market, because our results are based on only two firms in a specific labor market. The evidence for high speed learning, however, is consistent with the evidence provided by Lange (2007) based on NLSY 1979, even though his study is based on survey data from the US. We need to accumulate more studies on the employer learning model based on different data setting and from different economies to derive a definitive conclusion on the speed of employer learning.

Our work does not tell whether employers in the labor market share the information of each employee's productivity or the current employer privately holds the information. Distinguishing whether the learned information is public or private is important, because private information about each worker's productivity is the fundamental source of information asymmetry in the labor market resulting in labor-market friction (Gibbons and Katz (1991) and Hu and Taber (2011)). The comparison of the career progressions of workers with and without prior labor-market experience might enable us to distinguish between private and public learning. This is an area for potential future research.

Our model treats each employee's productivity as a scaler, but an employee's ability could be multidimensional; a good employee in the engineering department may not be a good employee in the sales department. In this multidimensional skill setting, the employer probably rotates its employees to several sections to learn about each employee's comparative advantage. Extending our model to the multidimensional setting and fitting the model to the personnel data sets utilizing work-section information is left for future research.

References

Altonji, J. G. and Pierret, C. R. (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics*, 116(1):313–350.

Araki, S. and Kawaguchi, D. (2014). The promotion rule under imperfect observability of the employees ability. *Theoretical Economics Letters*, 4(8):662–665.

Ariga, K. (2006). Horizontal transfer, vertical promotion, and evolution of firm organization. *Journal of the Japanese and International Economies*, 20(1):20–49.

Baker, G., Gibbs, M., and Holmstrom, B. (1994). The Wage Policy of a Firm. *The Quarterly Journal of Economics*, 109(4):921–55.

Bordón, P. and Braga, B. (2013). Employer learning, statistical discrimination and university prestige. University of Wisconsin and University of Michigan Working Paper.

Dale, S. and Krueger, A. B. (2011). Estimating the return to college selectivity over the career using administrative earning data. IZA Discussion Papers 5533, Institute for the Study of Labor (IZA).

Dale, S. B. and Krueger, A. B. (2002). Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables. *The Quarterly Journal of Economics*, 117(4):1491–1527.

Farber, H. S. and Gibbons, R. (1996). Learning and wage dynamics. *The Quarterly Journal of Economics*, 111(4):1007–47.

Gibbons, R. and Katz, L. F. (1991). Layoffs and Lemons. *Journal of Labor Economics*, 9(4):351–80.

Gibbons, R. and Waldman, M. (1999). A theory of wage and promotion dynamics inside firms. *The Quarterly Journal of Economics*, 114(4):1321–1358.

Gibbons, R. and Waldman, M. (2006). Enriching a theory of wage and promotion dynamics inside firms. *Journal of Labor Economics*, 24(1):pp. 59–107.

Heckman, J. J. (2000). Policies to foster human capital. *Research in Economics*, 54(1):3 – 56.

Heckman, J. J., Stixrud, J., and Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24(3):411–482.

Hoekstra, M. (2009). The effect of attending the flagship state university on earnings: A discontinuity-based approach. *The Review of Economics and Statistics*, 91(4):717–724.

Hu, L. and Taber, C. (2011). Displacement, Asymmetric Information, and HeterogeneousHuman Capital. *Journal of Labor Economics*, 29(1):113–152.

Kahn, L. B. (2013). Asymmetric Information between Employers. *American Economic Journal: Applied Economics*, 5(4):165–205.

Kahn, L. B. and Lange, F. (2013). Employer learning, productivity and the earnings distribution: Evidence from performance measures. Yale University Working Paper.

Kim, S. and Usui, E. (2012). Employer learning, job changes, and wage dynamics. Working Papers UWEC-2012-01, University of Washington, Department of Economics.

Lange, F. (2007). The speed of employer learning. *Journal of Labor Economics*, 25:1–35.

Lazear, E. P. (2004). The peter principle: A theory of decline. *Journal of Political Economy*, 112(S1):pp. S141–S163.

Lazear, E. P. and Gibbs, M. (2009). *Personnel Economics in Practice, 2nd Edition*. Wiley.

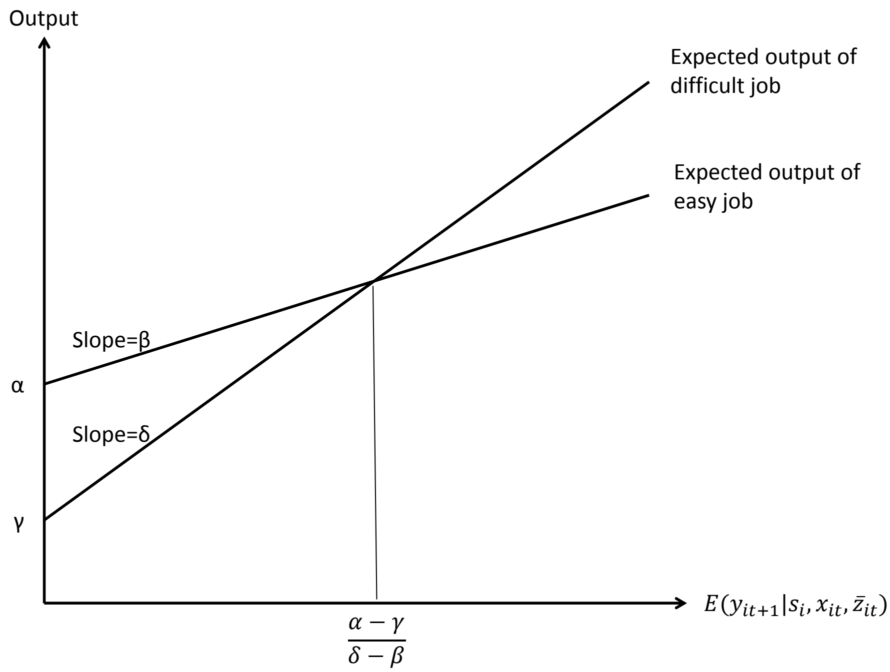
Light, A. and McGee, A. (2012). Employer learning and the importance of skills. IZA Discussion Papers 6623, Institute for the Study of Labor (IZA).

Mansour, H. (2012). Does employer learning vary by occupation? *Journal of Labor Economics*, 30(2):415 – 444.

Pinkston, J. C. (2009). A model of asymmetric employer learning with testable implications. *Review of Economic Studies*, 76(1):367–394.

Schönberg, U. (2007). Testing for asymmetric employer learning. *Journal of Labor Economics*, 25(4):pp. 651–691.

Figure 1: Expected productivity and expected output in easy and difficult jobs

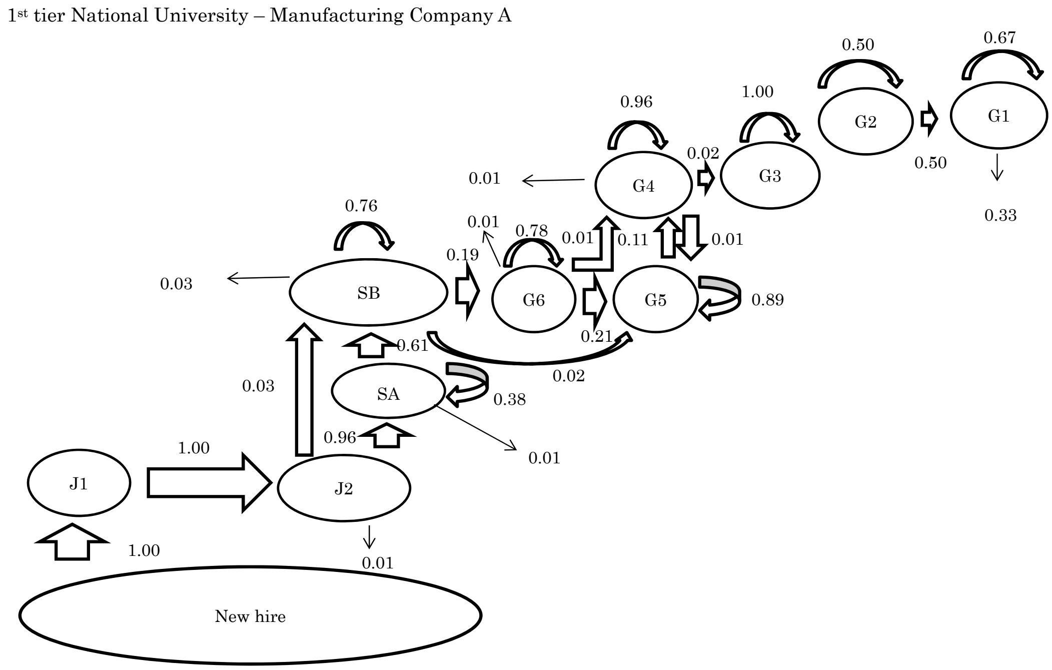


Note: $E(y_{it+1}|s_i, x_{it}, \bar{z}_{it})$ is the expected productivity conditional on schooling (s_i), experience (x_{it}), and average past performance evaluation ($\bar{z}_{it}$).

Table 1: Evaluation by ability rank in Company A, 2005-2010, %

Rank Evaluation	J1	J2	SA	SB	G6	G5	G4	G3
SS	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S	0.00	1.86	3.13	4.50	8.33	9.21	12.71	0.00
A1	3.56	16.15	21.93	27.14	15.64	20.66	11.86	0.00
A2	95.17	71.74	64.10	59.22	63.30	64.52	67.80	90.00
A3	1.27	7.45	7.47	6.60	7.75	4.19	4.45	10.00
B	0.00	2.48	1.69	1.95	4.39	1.35	2.97	0.00
C	0.00	0.31	1.69	0.60	0.58	0.07	0.21	0.00
N	393	322	415	667	684	1336	472	10
N of individual	390	305	299	280	270	414	181	4

Figure 2: 1st tier national university graduates' annual transition of Company A, 2005-2010



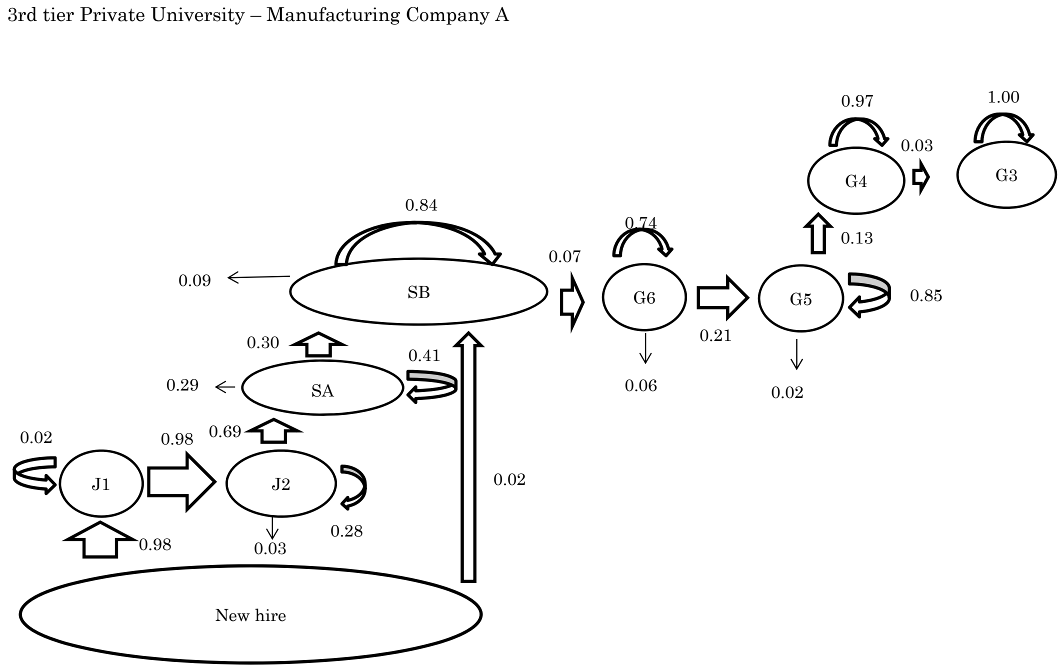
Note: 1st tier national universities include Tokyo, Kyoto, TIT, Hitotsubashi, and Osaka.

Table 2: The determination of evaluation in Company A, Marginal effects of ordered probit model

	C	B	A3	A2	A1	S
University score	-0.0001	-0.0004	-0.0008	-0.0015	0.0017	0.0011
	(0.0000)	(0.0001)	(0.0002)	(0.0004)	(0.0005)	(0.0003)
Tenure	0.0006	0.0028	0.0056	0.0110	-0.0121	-0.0078
	(0.0002)	(0.0005)	(0.0008)	(0.0017)	(0.0016)	(0.0011)
N	4291					
LL	-4291.4664					

Note: Marginal effects evaluated at the means are reported. Standard errors are in parentheses. Eight rank dummy variables are included in the regression model, but the estimated coefficients are not reported.

Figure 3: 3rd tier private university graduates' annual transition of Company A, 2005-2010



Note: 3rd tier private universities include private universities other than Keio, Waseda, Sophia, Meiji, Aoyama, St. Paul's, Chuo, Hosei, Kwanseigakuin, Kansai, Doshisha, Ritsumeikan, Gakushuin, and Nanzan.

Table 3: Descriptive statistics of Company A sample

	J1	J2	SA	SB	G6	G5	G4	G3
University score	61.37	61.96	61.14	62.4	62.7	61.54	62.69	65.25
	(7.62)	(7.42)	(8.22)	(7.15)	(7.58)	(7.33)	(6.60)	(4.92)
Tenure	0.11	1.49	2.83	5.48	10.55	13.67	15.49	15.9
	(1.22)	(2.11)	(1.60)	(2.20)	(2.66)	(2.33)	(1.66)	(1.79)
Average past evaluation								
Low(A3, B, C)	-	0.04	0.09	0.13	0.17	0.08	0.04	0.14
		(0.18)	(0.21)	(0.28)	(0.31)	(0.20)	(0.10)	(0.19)
Medium(A1, A2)	-	0.91	0.73	0.55	0.47	0.45	0.45	0.44
		(0.27)	(0.35)	(0.40)	(0.38)	(0.37)	(0.30)	(0.23)
High(S, SS)	-	0.05	0.18	0.32	0.37	0.47	0.5	0.43
		(0.21)	(0.30)	(0.39)	(0.39)	(0.38)	(0.31)	(0.28)
Obs.	393	322	415	665	681	1332	472	10

Note: Standard deviations are in parentheses. Average past evaluation is generally not available for J1 workers, because most J1 workers are promoted to J2 in the following year.

Figure 4: 1st tier national university graduates' annual transition of Company B, 2004-2010

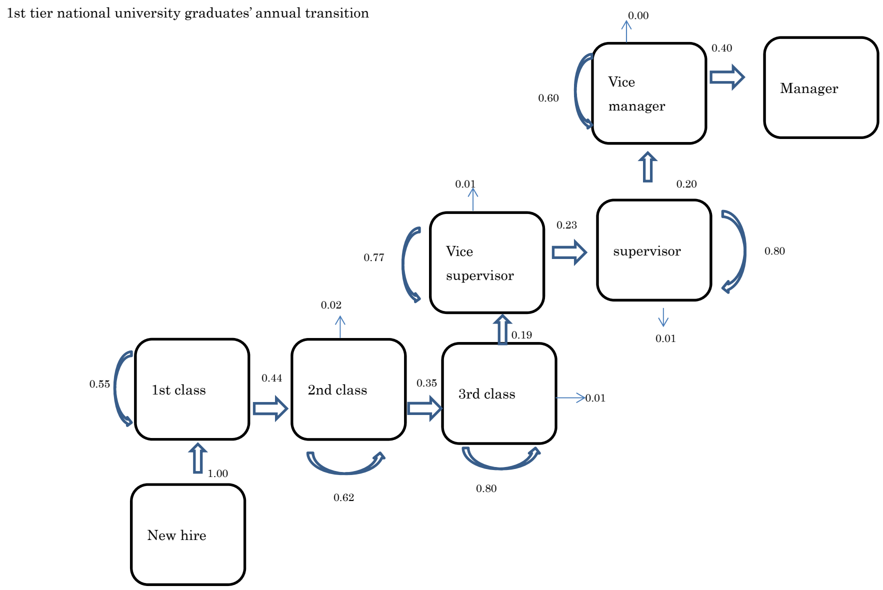


Table 4: The determinants of promotion in Company A, Cox proportional hazard model

	SA	SB	G6	G5	G4
University score	0.0104	0.0278	0.0660	0.0419	0.0221
	(0.0089)	(0.0096)	(0.0135)	(0.0100)	(0.0100)
Low eval.	-0.2671	-0.7750	-1.1483	-1.3710	-2.7543
(A3, B, C)	(0.3714)	(0.3479)	(0.5606)	(0.3426)	(0.7038)
High eval.	0.0190	-0.0869	0.3509	0.3719	0.7845
(S, SS)	(0.2512)	(0.2203)	(0.2519)	(0.2040)	(0.2250)
N	753	1272	2074	2937	4524
N of subjects	429	531	661	843	1101
N of failures	298	219	121	174	183
LL	-1686.4671	-1108.1541	-544.5146	-821.0675	-973.3288

Figure 5: 3rd tier private university graduates' annual transition of Company B 2004-2010

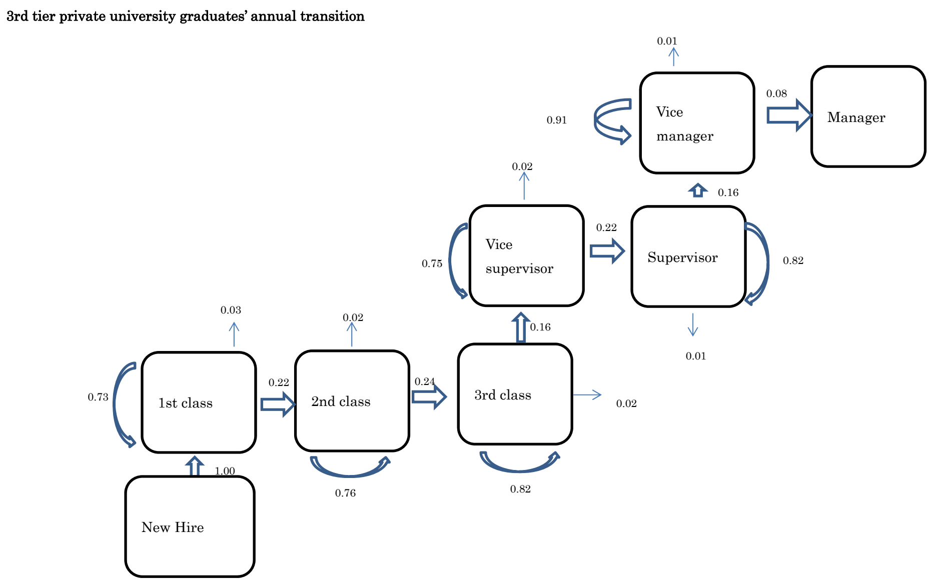


Table 5: Evaluation in Company B, 2004-2010

Ability rank Evaluation	1st class	2nd class	3rd class	Vice supervisor	Supervisor
A	0.70	2.26	1.08	4.12	3.65
B	27.96	41.65	36.67	34.08	35.55
C	69.72	53.98	59.17	59.55	60.13
D	1.62	2.11	3.08	2.25	0.66
N	862	665	649	534	301
Individual N	436	312	246	201	122

Table 6: The determination of evaluation in Company B, Marginal effects of ordered probit model

	D	C	B	A
University score	-0.0002	-0.0016	0.0016	0.0002
	(0.0001)	(0.0010)	(0.0010)	(0.0001)
Tenure	0.0033	0.0233	-0.0234	-0.0031
	(0.0006)	(0.0039)	(0.0039)	(0.0006)
N	4291			
Log LL	-4291.4664			

Note: Marginal effects evaluated at the means are reported. Standard errors are in parentheses. Eight rank dummy variables are included in the regression model, but the estimated coefficients are not reported.

Table 7: Descriptive statistics of Company B sample

	1st class	2nd class	3rd class	Vice supervisor	Supervisor
University score	56.56	57.05	56.00	54.41	53.85
	(6.63)	(6.16)	(6.28)	(7.32)	(8.13)
Tenure	1.85	4.73	9.4	13.56	15.05
	(0.92)	(1.52)	(2.70)	(2.18)	(1.58)
Average Past Evaluation					
Low (D)	0.02	0.03	0.03	0.03	0.00
	(0.11)	(0.12)	(0.15)	(0.12)	(0.02)
Medium (C)	0.88	0.62	0.56	0.67	0.58
	(0.28)	(0.35)	(0.37)	(0.34)	(0.34)
High (AB)	0.11	0.36	0.41	0.3	0.42
	(0.25)	(0.35)	(0.37)	(0.33)	(0.34)
N	862	665	649	534	301
Individual N	436	312	246	201	122

Note: Standard deviations are in parentheses.

Table 8: The determinants of promotion in Company B, Cox proportional hazard model

	2nd class	3rd class	Vice supervisor	Supervisor
University score	0.0055	-0.0039	0.0236	0.0208
	(0.0098)	(0.0124)	(0.0135)	(0.0129)
Low eval.	-0.6861	-2.5308	-2.4581	-6.1169
(D)	(0.9234)	(1.1873)	(0.9818)	(3.0596)
High eval.	0.3696	0.1553	0.3431	1.1497
(AB)	(0.1869)	(0.1960)	(0.2136)	(0.2675)
N	1154	1823	2488	3043
N of subjects	475	591	701	790
N of failures	251	177	142	114
LL	-1305.6914	-841.4124	-629.7786	-528.5440

Table 9: Structural parameters of the promotion model in Company A

	SA	SB	G6	G5
Human capital				
University score	1.4343	1.9185	1.5344	0.9524
	(0.2251)	(0.0491)	(0.2463)	(0.0718)
Tenure	2.9959	2.1615	1.2413	0.7222
	(0.3886)	(0.2031)	(0.0593)	(0.1095)
Constant	-19.2695	-12.1346	-11.0417	-17.7172
	(4.4478)	(5.7569)	(2.3897)	(10.0089)
Learning				
University score	1.5415	1.3126	1.1705	0.8253
	(0.2996)	(0.1481)	(0.1847)	(0.1363)
Constant	19.6469	26.4541	28.2082	65.9221
	(4.4618)	(14.3720)	(11.4456)	(11.6313)
ν^2/σ^2	0.2894	0.2231	0.3263	0.0577
	(0.1534)	(0.0984)	(0.1052)	(0.0139)
N	4524	4524	4524	4524
LL	-257.19	-340.23	-1048.13	-1036.21

Table 10: Structural parameters of the promotion model in Company B

	2nd class	3rd class	Vice supervisor	Supervisor
Human capital				
University score	1.9180	1.5347	0.9465	1.1873
	(0.4905)	(0.2466)	(0.1838)	(0.2009)
Tenure	2.1612	1.4333	1.5222	0.5405
	(0.2031)	(0.0593)	(0.1095)	(0.0536)
Constant	-8.6024	-11.0416	-8.7171	-16.5392
	(2.5241)	(2.3897)	(1.0008)	(1.9772)
Learning				
University score	0.9564	0.7015	0.5305	0.5090
	(0.1481)	(0.1847)	(0.1363)	(0.1465)
Constant	2.1346	5.7716	6.5922	6.4792
	(0.2875)	(2.3168)	(1.1631)	(2.0357)
ν^2/σ^2	0.2231	0.3263	0.5765	0.3517
	(0.0983)	(0.1052)	(0.1394)	(0.1130)
N	4322	4322	4322	4322
LL	-200.52	-240.99	-646.25	-779.88