Acquired Skills and Learned Abilities

Wage Dynamics of Blue-collar Workers in Internal Labor Markets*

ISS Discussion Paper Series

F-153

January 2015

Masaki Nakabayashi†

Institute of Social Science

The University of Tokyo

Abstract

Workers' abilities are hidden information. Thus, when hiring, firms first use education as a proxy for abilities, and then learn about workers' abilities by tracking products. If this learning is asymmetric inside and outside major firms' internal labor markets, the market expects work experience and schooling to be complements for experience before workers gain long-term employment, which hides the learning effect. Once workers gain long-term employment, the learning effect becomes evident. Furthermore, the employer learns more quickly in the early stages of internal career, and this privately learned information could improve the efficiency of in-house training programs.

Key words: asymmetric employer learning; cross-sectional skill distribution; longitudinal employer learning; internal labor markets.

JEL: J31; L22; M52.

*This research has been funded by a JSPS Grant-in-Aid (22243022).

†The author is motivated and encouraged by comments from Hideshi Itoh, Michael Waldman, Susumu Cato, Takashi Shimizu, Daiji Kawaguchi, Yuji Genda, Rosa Ferrer, Fumitoshi Moriya, Hideo Owan, Fumio Ohtake, Masaru Sasaki, Kengo Yasui, Carol Scotese, Ryuichi Tanaka, Konosuke Odaka, Takashi Kurosaki, Norihito Abe, Shingo Ishiguro, Chiaki Moriguchi, Naofumi Nakamura, Junichiro Ishida, Kazufumi Yugami, Shinpei Sano, Keisuke Kawata, Keiko Yoshida, and John M. Abowd. The author also thanks the participants at the Annual International Industrial Organization Conference 2014, the North American Winter Meeting of the Econometric Society 2014, Royal Economic Society Annual Conference 2013, Society of Labor Economists Meetings 2012, Osaka Workshop on Economics of Institutions and Organizations, Contract Theory Workshop, Kansai Labor Workshop, Western Economic Association International Conference 2011, Southern Economic Association Meeting 2011, and at seminars held in Hitotsubashi University, Keio University, National Graduate Institute for Policy Studies, the University of Tokyo, University of California, Irvine, Free University of Berlin, London School of Economics and Political Science, and University of Tübingen. Corresponding address: Institute of Social Science, The University of Tokyo, Hongo 7-3-1, Tokyo 103-0033, Japan. E-mail: mn@iss.u-tokyo.ac.jp

1 Introduction: Warp and weft

Workers' innate abilities and acquired skills determine their current productivity. Thus, it is efficient to correlate the wage distribution with the cross-sectional distribution of innate abilities and acquired skills. Abilities also determine the speed and depth of future skill acquisition. Therefore, when recruiting, employers are interested in workers' innate abilities as well as skills they have already acquired. However, workers' innate abilities are generally private information when they enter the labor market and, hence, employers use a proxy to predict these abilities. If innate abilities affect skill acquisition at school as well as productivity and skill acquisition in the workplace, educational background can be used as a proxy for abilities. Thus, employers often use education as a signal of abilities, and not just as certificates of acquired skills (Spence (1973); Arrow (1973); Riley (1979); Hungerford and Solon (1987); and Jaeger and Page (1996)). Then, after workers join the labor market, employers gradually learn about workers' true on-the-job abilities from their products in the longitudinal dimension (Farber and Gibbons (1996)).

The cross-sectional distribution of acquired skills affects differences in current productivity. At the same time, employers learn about workers' abilities in the longitudinal dimension. As a result, the two dimensions can provide a mixed picture in terms of empirical results, which have focused attention on the interaction term between years of schooling and work experience. If skill acquisition at both school and the workplace are affected by abilities, these acquired skills should be complements in the cross-sectional distribution, as assumed by Gibbons and Waldman (2006). Then, the interaction term between schooling and work experience in a wage regression should have a positive coefficient. This assumption is supported by some empirical research, such as Rubinstein and Weiss (2006) and Habermalz (2006).

However, Mincer (1974) found that this interaction term has a negative coefficient,[1] which other studies have confirmed. Following the literature on the signaling role of schooling, Farber and Gibbons (1996) give a clear-cut prediction. Employers first use education as a signal of ability when workers join the market. Then, they learn more about these abilities based on workers' experience after joining the labor market. Wages increase as workers' productivity increases owing to skill acquisition from work experience, but the signaling role of schooling declines as employers learn about workers' abilities. Thus, the relative impact of schooling on wage growth declines. This employer learning effect provides a non-positive coefficient for the interaction term between schooling and work experience in a wage equation in the antilogarithmic term, and a negative coefficient in a wage equation in the logarithmic term, as in Mincerian equations. Based on US data sets, this prediction is supported by studies such as Altonji and Pierret (2001), Pinkston (2006), Lange (2007), and Schönberg (2007).

In summary, as we show in section 2, if the cross-sectional complementarity between schooling and work experience dominates, then the interaction term should have a positive coefficient. However, if the employers' longitudinal learning effect dominates, then the coefficient should be negative. This simple point has been overlooked in the existing literature on this topic. Of course, the reality lies somewhere between these two extremes. In the United

[1] See Mincer (1974), pp. 92–93.

States, the complementarity is observed more strongly in the case of young workers.[2] In Germany, the school system is closely linked to the apprentice system and, hence, schooling and work experience are more complementary than in the United States (Pischke and von Wachter (2008)). Thus, not surprisingly, the employer learning effect is only weakly observed in Germany (Bauer and Haisken-DeNew (2001) and Lluis (2005)). These real-real world situations have been recognized as mixed, without considering why they occur, though it is essential in order to understand the diversity of firm organizations in different economies (Waldman (2013)).

Another factor that causes observations to be mixed is that learning by current and potential employers in the market is asymmetric. Current employers learn about their employees' abilities more quickly than do outside potential employers (Schönberg (2007) and Pinkston (2009)). This implies that longer-term employment makes better employer learning possible (Mansour (2012)).

This asymmetry relates to a reality in the labor market. Employers in an industrial economy are far from uniform. Larger and more productive employers tend to manage longer-term employment and predominantly promote workers from within. Accordingly, they have a longer time to learn about their workers. Such personnel practices are called internal labor markets. We focus on these internal labor markets, using a newly built long-term panel data set of blue-collar workers in a major manufacturing firm. Here, we examine the warps of learning and the wefts of ability and skill distribution, enabling us to decompose the cross-sectional skill distribution of workers and the longitudinal learning by employers.

Section 2 presents the underlining theoretical framework and shows how it is transmitted to empirical contexts. A theoretical key point is that wage determination is distorted particularly for workers upgraded to upper notches of wage ladders. This, in turn, distorts the market expectation about workers' skill elements. Section 2 predicts that if employer learning is asymmetric between current and potential employers, and if skills acquired at workplaces are sufficiently firm-specific, then, in the mid-career market, employers expect work experience and schooling to be complements for workers who have not been promoted in previous employment and substitutes for workers who have been promoted. In other words, the market expectation of workers' skills in the cross-sectional dimension differs before and after workers are promoted. Therefore, assuming that job seekers in the mid-career recruiting market are dropouts of internal labor markets, our estimation framework separates labor market experience into before and after gaining long-term employment with a major firm.

Section 3 describes the data set, which needs to be large and to contain detailed intra-firm data to test our prediction. Thus, we build a new data set from first-hand wage records of blue-collar workers in a Japanese ironworks. This source of data provides us with two advantages. One is that major Japanese firms upgrade the basic wages for both blue-collar and white-collar regular workers every year, depending on observed and predicted performance. This means that every worker faces a test of fine-tuned promotion every year and the outcome is recorded. Basic wages do not include bonuses, overtime compensation, or other allowances, which depend on current performance or conditions. This means they do not change during a fiscal year after being set at the beginning of the year. Therefore, they capture the outcome

[2]See Rubinstein and Weiss (2006), pp. 11–16, and Habermalz (2006), p. 130–133.

of the employer's learning in the previous year, providing us with rich information about the annual promotion decisions in their wage records. The second advantage is that we can examine the wage dynamics of blue-collar workers under lighter institutional restrictions. Major Western firms, either American or European, operate under a collective bargaining framework with trade unions to determine the wages of blue-collar workers. In contrast, in Japan, unions do not intervene in determining individual wages and, hence, the wages of blue-collar workers are determined solely by management's evaluation as those of white-collar workers are.

Section 4 presents our empirical results. Our findings show that schooling and short-term work experience at younger ages are expected to be complements in the market and that the employer learning effect is obscured. However, the employer learning effect is strongly evident once workers gain long-term employment at the case firm. Furthermore, once a worker gains employment with the case firm, learning is faster in the earlier internal career stages. The information learned in these earlier stages is used to screen employees with potential and, thus, to become trainees of the in-house training program.

2 Theoretical framework

2.1 Skill acquisition and asymmetric employer learning

To describe the skill acquisition process under asymmetric employer learning, our approach is based on DeVaro and Waldman (2012). The skeleton of their model was provided by Gibbons and Waldman (1999, 2006). They captured both skill acquisition and symmetric employer learning within a model, which had been requested by empirical works such as Ariga, Ohkusa and Brunello (1999). Then DeVaro and Waldman (2012) introduced an asymmetric learning environment, assuming a competitive labor market. Here, realized performance was only observable by current employers, while potential employers could only observe whether workers had been promoted, which was the essence of Waldman (1984).

Following the model of DeVaro and Waldman (2012), let $\phi_i \in (\phi_L, \phi_H)$ denote the innate ability of worker i, $i = 1, 2, \ldots, n$, which is a random draw from a probability density function $g(\phi)$, with $g(\phi) > 0$ for $\phi \in (\phi_L, \phi_H)$, and $g(\phi) = 0$ otherwise; let S_i denote worker i's years of schooling; and let $M_{i,t}$ denote worker i's employment experience until period t. Then, assume that the "on-the-job" skill of worker i, who has S_i years of schooling and has M_i years of work experience in period t is $\eta_{i,t} = (\phi_i + bS_i)f(M_{i,t}) \equiv \theta_i f(M_{i,t})$, where $b > 0$, f is increasing in M, and $f(0) > 0$. All firms have homogenous production functions and each firm consists of job 1 and job 2. The product of worker i assigned to job j in period t is given by

$$y_{i,j,t} = (1 + k_{i,t})(d_j + c_j \eta_{i,t}) + G(S_i), \tag{1}$$

where $0 < d_2 < d_1$, $0 < c_1 < c_2$, G is increasing in S, and $k_{i,t} > 0$ if worker i was employed at the same firm in period $t-1$. Defining $\eta_{i,t} = (\phi_i + bS_i)f(M)$, it is assumed that schooling S and experience M are complements in production. This assumption is justified if ability positively affects skill acquisition both at school and at the workplace. While $M_{i,t}$,

S_i, $f(\cdot,\cdot)$, $G(\cdot)$, b, d_j, c_j, and $k_{i,t}$ are public information, $y_{i,j,t}$ is privately observed by the current employer, and ϕ_i is privately known to worker i. We do not specify anything about the possible correlation between ϕ and S, which allows S to potentially depend on ϕ and to function as a signal of ϕ. At the end of worker i's first period, the current employer privately observes $y_{i,j,t}$ and, thus, learns about θ_i. Hereafter, we assume that an increase in productivity owing to a promotion ($c_2 - c_1$) and/or return on firm-specific skill k is sufficiently large such that $k > c_1/(c_2 - c_1)$.

Consider $\eta' = (d_1 - d_2)/(c_2 - c_1)$ that solves $d_1 + c_1\eta' = d_2 + c_2\eta'$ and assume that $(E[\phi \mid S] + bS)\, f(0) \equiv \theta^E(S)f(0) < \eta'$ for any S, which implies that any worker is efficiently assigned to job 1 in the first period. Further, assume that $(\phi_L + bS)\, f(1) < \eta' < (\phi_H + bS)\, f(1)$, which implies that some workers in their second period are efficiently assigned to job 1, while the others are assigned to job 2.

The structure of the game is as follows. At the beginning of workers' second period, each firm offers each existing worker employed in the previous period a job assignment, or fires the worker. This decision is publicly observed. Then wages are determined before each period by spot-market contracting. Observing worker i's job assignment or discharge, employers other than the worker's first-period employer offer a wage, and the worker's first-period employer offers a wage weakly greater than the wage offered by others. For simplicity, we assume no transaction costs and a common discount factor.

Further, letting $w_{i,1,t}$ denote the wage paid to worker i assigned to job 1, and $w_{i,2,t}$ denote the wage paid to job 2, consider an ability level $\eta^+(S)$ in worker i's second period such that $y_{i,1,t} - w_{i,1,t} = y_{i,2,t} - w_{i,2,t}$ if $\eta_{i,t} = \eta^+(S)$; that is, profit is indifferent, regardless of whether worker i is promoted to job 2 in his/her i's second period. Under this setting, DeVaro and Waldman (2012) established that there is a perfect Bayesian equilibrium such that, in the second period of worker i who was employed by firm A in the first period, if $\eta_{i,t} \geq \eta^+(S_i)$, then worker i remains at firm A, is assigned to job 2, and is paid $w_{i,2,t}(S_i, \eta_{i,t}) = d_2 + c_2\eta^+(S_i) + G(S_i)$; however, if $\eta_{i,t} < \eta^+(S_i)$, then worker i remains at firm A, is assigned to job 1, and is paid $w_{i,1,t}(S_i, \eta_{i,t}) = d_1 + c_1\,(\phi_L + bS_i)\, f(1) + G(S_i)$.

On this equilibrium strategy, outside employers offer wages equal to the possible least on-the-job skill given the publicly available information about job assignments at the current employer. Then, the current employer counteroffers with a wage only weakly greater than that offered by the others. Offering the expected productivity as a wage, given the publicly available information about the job assignment outcome, which is equal to or greater than the lowest possible productivity, cannot be an equilibrium strategy. This is because workers whose productivity is strictly lower than such an offer would take it, and employers would predict this adverse selection outcome.

The definition of η^+, $y_{i,1,t} - w_{i,1,t} = (1+k)\,(d_1 + c_1\eta^+(S_i)) - [d_1 + c_1(\phi_L + bS_i f(1))] = (1+k)\,(d_2 + c_2\eta^+(S_i)) - (d_2 + c_2\eta^+(S_i)) = y_{i,2,t} - w_{i,2,t}$, is rearranged to

$$\eta^+(S_i) = \frac{k(d_1 - d_2) - c_1\,(\phi_L f(1) + bS_i f(1))}{k(c_2 - c_1) - c_1}. \tag{2}$$

This is a tradeoff of employer's benefits between promotion and non-promotion. By curbing promotion, wage payment is constrained. However, an increase in $bSf(1)$ pushes up the wage by the rate c_1, even for job 1. Thus, an increase in $bSf(1)$ implies a decrease in the promotion threshold η^+.

Then, consider a modified two-period setting. Now, when worker i joins the firm, worker $i+1$ is already there and has not been promoted to job 2. In the second period of worker i, which is the third period for worker $i+1$, both workers i and $i+1$ are promoted to job 2 if $\eta_{i,t} \geq \eta^+$ and $\eta_{i+1,t} \geq \eta^+$, respectively, and retained at job 1 otherwise. Then, we can derive a lemma for the market expectation on the cross-sectional skill distribution for $i = 1, 2, \ldots, n$.

Lemma 1. *Among the on-the-job skill components, work experience is expected in the market to be complementary to schooling for workers who have not been promoted, and substitutive for schooling for workers who have been promoted.*

Proof See Appendix I.

Lemma 1 states that, under asymmetric employer learning and a sufficiently high return on firm-specific skill, work experience and schooling are expected in the market to be complements before workers are promoted, but substitutes once workers have been promoted. This switch comes from the characteristics of the promotion threshold η^+. The production function equation (1) assumes that schooling and experience are complements. Thus, the promotion threshold η^+ decreases in the product of schooling and experience, $Sf(M)$. At the same time, the market can only observe the outcome of the promotion and offers a possible least productivity as a wage. This implies that wages offered for promoted workers increase in η^+. It turns out that schooling and experience are valued as substitutes when determining wages for promoted workers. Such a distortion does not occur for non-promoted workers.

Threshold $\eta^{'}$ is the value under symmetric learning. In other words, the difference $D \equiv \eta^+ - \eta^{'}$ captures the distortion due to the asymmetry of information between the current employer and other employers about workers' on-the-job skills. Under assumption $k > c_1/(c_2 - c_1)$, $D > 0$; that is, a less than optimal number of workers are promoted under asymmetric learning.

2.2 Skill complementarity and learning in panel estimations

Next, let us discuss how **Lemma 1** can be placed within empirical contexts. Hereafter, we consider a setting where $t = M$ for $i = 1, 2, \ldots, n$, and for simplicity, assume that $b = 1$ and $f(M) = M = t$. Again, let $y_{i,t}$ denote the output of worker i in period t ($t = 1, \ldots, T$), and let $\eta_{i,t}$ denote the ith worker's on-the-job skill in period t, which is not observable by employers. Then, suppose that $\eta_{i,t} = (\phi_i + S_i)t = \theta_i t$, where $\theta_i = \phi_i + S_i$ denotes the ith worker's ability, which is a time-invariant multiplier of skill acquired at the workplace. Schooling S_i and experience t are observable to employers. However, worker i's innate ability ϕ_i is not observable, and thus $\eta_{i,t}$ is not observable to employers when the worker joins the labor market, but is later learned by the employers. Further, let $\boldsymbol{x}_i$ denote a vector of the time-invariant characteristics of worker i other than years of schooling, which are observable to employers.

Following Farber and Gibbons (1996), we assume that the conditional distribution $F_1(y_{i,t} \mid \theta_i, S_i, \boldsymbol{x}_i)$ and the joint distribution $F_2(\theta_i, S_i, \boldsymbol{x}_i)$ can be arbitrary and that outputs $y_{i,t}$ are independently drawn from $F_1(y_{i,t} \mid \theta_i, S_i, \boldsymbol{x}_i)$. We assume that all employers know $F_2(\theta_i, S_i, \boldsymbol{x}_i)$ and $F_1(y_{i,t} \mid \theta_i, S_i, \boldsymbol{x}_i)$ and can observe $y_{i,1}, \ldots, y_{i,t}$ for each worker $i = 1, \ldots, n$. Thus, both the current and potential employers in the market symmetrically learn about the ith employee's ability in the market. Furthermore, we assume that owing to competition between employers, the wage paid to the ith worker in period t is equal to the expected output given all available information in period t about the ith worker:

$$w_{i,t} = E\big(y_{i,t} \mid S_i, \boldsymbol{x}_i, y_{i,1}, \ldots, y_{i,t-1}\big). \tag{3}$$

We further assume that the conditional expectation $E\big(y_{i,t} \mid S_i, \boldsymbol{x}_i, y_{i,1}, \ldots, y_{i,t-1}\big)$ is a linear combination of S_i, $\boldsymbol{x}_i$, and $y_{i,1}, \ldots, y_{i,t-1}$.

Then, the linear projection of an n-dimensional vector $\boldsymbol{w}$ whose ith element is w_i, denoted by $E^*(\boldsymbol{w} \mid \cdot)$, yields $E^*(\boldsymbol{w} \mid \boldsymbol{X}) = \boldsymbol{X}\hat{\boldsymbol{\alpha}}$.[3] Normal equations give

$$\hat{\boldsymbol{\alpha}} = [\boldsymbol{X}'\boldsymbol{X}]^{-1}\boldsymbol{X}'\boldsymbol{w}, \tag{4}$$

where the hth element of $\hat{\boldsymbol{\alpha}}$, $\hat{\alpha}_h$, is increasing in $\sum_{t=1}^{T}\sum_{i=1}^{n} x_{h,i}w_{i,t} - TnE(x_j)E(w) = \operatorname{Cov}(x_{i,h}, w_{i,t})$.

Consider an example of panel estimation of worker i's wage at time t, $w_{i,t}$,

$$w_{i,t} = \alpha_0 + \alpha_1 S_i + \alpha_2 t + \alpha_3 S_i t + \alpha_4 x_{4,i} + \cdots + \alpha_j x_{j,i} + \cdots + \alpha_m x_{m,i} + \theta_i^E(S_i) + \epsilon_{i,t}, \tag{5}$$

where S_i, t, and $\boldsymbol{x}_i$ are observable and $\theta_i^E \equiv E(\phi_i|S_i) + S_i$. Then, we obtain

$$\varDelta_t w_{i,t} = \alpha_2 + \alpha_3 S_i + \varDelta_t \theta_i^E(S) + \varDelta_t \epsilon_{i,t} \equiv \alpha_2 + \alpha_3 S_i + \varphi_{i,t}, \tag{6}$$

where $\varDelta_t \epsilon_{i,t}$ is independent of other independent variables and is also serially independent.

Here, $\hat{\alpha}_3$ in equation (5) is increasing in $\operatorname{Cov}(S_i t, w_{i,t}) = \sum_{\tau=2}^{T} \operatorname{Cov}(S_i\tau, \varphi_{i,\tau})$. In addition, $\operatorname{Cov}(S_i t, w_{i,t})$ contains a two-dimensional effect composed of the cross-sectional effect over workers $i = 1, \ldots, n$ and the longitudinal effect over periods $t = 1, \ldots, T$. In the cross-sectional dimension, for each τ ($\tau = 2, \ldots, T$), $\operatorname{Cov}(S\tau, \varphi_\tau)$ is increasing in the degree of complementarity between years of schooling (S) and years of work experience (τ). Thus for each period t, the covariance between φ_τ and $S\tau$ should be positive in the cross-sectional dimension of workers $i = 1, \ldots, n$ if schooling (S) and experience (τ) are complements for the productivity difference ($\varDelta_\tau \epsilon$) among workers $i = 1, \ldots, n$ and non-positive otherwise.

In the longitudinal dimension, let us assume that the employers have learned about the employees' time-invariant abilities that were hidden when the employees were recruited, given as ϕ_i. This is included in θ_i, such that $\varDelta_\tau \theta_i^E(S) = \varDelta_\tau E(\theta_i \mid S_i, \tau - 1)$ is decreasing in τ and $\lim_{\tau\to\infty} \varDelta_\tau E(\theta_i \mid S_i, \tau - 1) = 0$ as θ_i^E approaches a stationary state, which is worker i's true ability. Then, for each i, $\operatorname{Cov}(S_i\tau, \varphi_i)$ is decreasing in τ and $\lim_{\tau\to\infty} \operatorname{Cov}(S_i\tau, \varphi_i) = 0$.

[3]Note that $E^*(y \mid S, \boldsymbol{x}) = E(y \mid S, \boldsymbol{x})$ because E is assumed to be linear.

Thus, $\hat{\alpha}_3$ depends on the relative impact of the effect of the complementarity between schooling and work experience in the cross-sectional dimension and the effect of employer learning in the longitudinal dimension. Then, $\hat{\alpha}_3$ is increasing in the relative impact of the complementarity effect over the employer learning effect and, fixing the complementarity effect, $\hat{\alpha_3}$ decreases to 0 as the employer learning effect increases. In addition, suppose that wages increase with experience t owing to skill acquisition. Then, $\hat{\alpha_3}$ also depends on the relative impact of $S_i t$ on the wage growth compared to other independent variables. Taking the logarithmic terms, if the complementarity effect dominates the employer learning effect, then $\hat{\alpha}_3 > 0$. However, if the employer learning effect dominates, then $\hat{\alpha_3} \leq 0$.

This reasoning differs from that of Farber and Gibbons (1996) in two regards. First, we deal directly with the interaction term between schooling and work experience. Farber and Gibbons (1996) presented a model highly tailored for the US National Longitudinal Survey of Youth, which includes Armed Forces Qualification Test (AFQT) scores that are thought to be correlated with innate abilities of respondents but are unknown to employers of the respondents. Their estimation strategy is not currently applicable to other countries that do not provide data equivalent to AFQT scores. Then, a negative coefficient of the interaction term between schooling and work experience is a convenient indicator of the employer learning effect. However, Farber and Gibbons (1996) only mention that it does not contradict with their model, but do not directly inquire how it is related to employer learning. Second, we explicitly differentiate the cross-sectional effects and longitudinal effects in the interaction term between schooling and work experience. Farber and Gibbons (1996) emphasized that "[schooling and other observable variables] play a declining role in the market's inference process but have a constant estimated effect."[4] Instead, we decompose the coefficient of the interaction term between schooling and experience into the cross-sectional dimension and the longitudinal dimension. In the former case, our framework predicts that the estimated effect of schooling does not change, as in Farber and Gibbons (1996). In the latter case, our framework predicts that the estimated effect of schooling declines as work experience is acquired. We infer that this longitudinal effect generates a non-positive coefficient for the interaction term between schooling and work experience.

2.3 Semi-public estimation framework of employer learning

Next, we consider the internal labor market of the major firm discussed in subsection 2-1: 1) the current employer learns about workers' abilities better than the other employers; 2) the other employers can less correctly guess workers' abilities from publicly available information about job assignments; 3) the return on firm-specific skills is positive, and, therefore the current employer produces more by hiring current workers in the next period than other employers do; and 4) the current employer faces a competitive market composed of other employers.

Then, in the next period, the firm that currently employs a worker offers wages weakly greater than those offered by other employers, and the current employee does not leave on the equilibrium path. Here, we can assume that the current employer of worker i knows $F_1(y_{i,t} \mid$

[4] Farber and Gibbons (1996), p. 1014.

$\theta_i, S_i, \boldsymbol{x}_i)$ and $F_2(\theta_i, S_i, \boldsymbol{x}_i)$ for $i = 1, \ldots, n$ and $t = 1, \ldots, T$, and that the firm observes $y_{i,1}, \ldots, y_{i,t}$. That is, wage growth depends on the current employer's learning with arbitrage with the outside market, where worker i's ability is signaled by her/his job assignment. The competitive environment guarantees that $w_{i,t} = lE(y_{i,t} \mid S_i, \boldsymbol{x}_i, y_{i,1}, \ldots, y_{i,t-1})$, where $l \leq 1$, which captures the efficiency loss due to asymmetric learning in internal labor markets. While employees' abilities are learned within the internal labor market, $F_1(y_{i,t} \mid \theta_i, S_i, \boldsymbol{x}_i)$ and $F_2(\theta_i, S_i, \boldsymbol{x}_i)$ for the current employees, for $i = 1, \ldots, n$, remain only imperfectly known to outside employers by job assignment. We refer to these properties as semi-public, which is public in the sense that the current employers face a competitive market and wages are determined by spot-contracting, but is "semi" in the sense that the wages are affected by the asymmetry of employer learning.

Lemma 1 argues that the mid-career recruiting market expects workers who have not been promoted by previous employers to have acquired work experience complementary to schooling. On the equilibrium path, workers do not leave their first-period employers. If a worker leaves his/her first-period employer, this is when the worker's belief and his/her employer's belief may not be consistent, in which case the worker is not satisfied by the job assignment outcome. That is, this occurs when a worker believe that he/she should be promoted, but is not promoted. Thus, workers in the mid-career recruiting market are likely to be those who have not been promoted. In other words, schooling and work experience are expected to be complements for workers seeking employment in the mid-career recruiting market. On the other hand, the market expects workers who have been promoted to have acquired work experience that is substitutive for schooling. This belief, shared by employers, turns out to reduce workers' opportunity cost of time to acquire firm-specific skills, which are likely to be less complementary to schooling.

To capture this effect of internal labor markets, we separate the ith employee's experience into two components, such that $t \equiv t^e$ and $t^T = t^p + t^e$. Here, t^T is total labor market experience, t^p is labor market experience prior to joining the case firm, and t^e denotes tenure after being employed by this firm. Then, assuming that $\boldsymbol{\beta} = l\boldsymbol{\alpha}$, wage equation (4) is reformulated as

$$(7) \quad w_{i,t} = \beta_0 + \beta_1 S_i + \beta_2 t_i^p + \beta_3 t_{i,t}^e + \beta_4 S_i t_i^p + \beta_5 S_i t_{i,t}^e + \boldsymbol{\gamma}' \boldsymbol{x_i} + \boldsymbol{\delta}' \boldsymbol{x_i} t_{i,t}^e + \theta_i + \epsilon_{i,t}.$$

Lemma 1 concerns an implication for the cross-sectional distribution. Schooling and work experience are expected to be complements for non-promoted workers who have left short-term employment for the mid-career market, which is presumed to be captured by t_i^p. Standardizing such that t takes 0 when $t_{i,t}^e = 0$, St_i^p captures a summary of the employer learning process until worker i gains long-term employment at $t_{i,t}^e = 0$ after he/she entered the labor market if he/she did not gain long-term employment immediately after graduation and left former employers. At the same time, for worker i whose $t_{i,t}^e > 0$ after gaining long-term employment, he/she has started to step up promotion ladder in a internal. Then, his/her schooling and work experience are expected to be substitutes as those for already promoted workers, which is presumed to be captured by $S_i t_{i,t}^e$. Then, combining **Lemma 1** with the characteristics of the panel estimation discussed above, a prediction about the market's expectation formed by employer learning, as well as about skill distribution is as follows.

Prediction 1. *If employer learning is asymmetric between current and potential employers and the return on firm-specific skill is sufficiently large, then the coefficient of the interaction term between years of schooling and previous labor market experience before gaining employment with a firm that commits to long-term employment* (St^p) *is greater than that of the interaction term between years of schooling and tenure after gaining employment with the firm* (St^e)*; thus,* $\hat{\beta}_4 > \hat{\beta}_5$*.*

3 Case firm and data

3.1 Case plant

The case plant is one of the oldest modern ironworks in Japan. As a part of a company-wide investment plan from the 1950s to the 1960s, the company that operated the ironworks decided to build a new state-of-the-art plant in a city far from the original plant. The plan was to decrease the capacity of the original ironworks, and to relocate skilled workers of the case ironworks and other old ironworks to the new plants. The selection for relocation was handled in cooperation with the union, and in principle, anyone willing to move was allowed to be relocated.

3.2 Data

This research uses preserved personnel documents for 1,558 blue-collar employees who were relocated from the ironworks, tracking them from the late 1920s or later, depending on the year when the employee joined the ironworks, to the 1960s, when they left the ironworks. Owing to the relocation process where anyone who wanted was allowed to be relocated, a possible sample selection bias can be assumed to be small. However, an attrition bias is inevitable, since the sample includes only employees who worked for the ironworks in the late 1960s, at which time the sample period ends. Thus, the sample does not include employees who left the ironworks before the late 1960s, the time of relocation. It means that the sample does not include dropouts. Since the firm does not preserve wage records of such dropouts, we cannot statistically correct the bias. However, this bias is not necessarily serious to our specific context. Our framework separates previous work experience and tenure at the ironworks, and, for tenure, we only deal with employees who continue to serve on the equilibrium path. All sample employees are Japanese males. The documents contain all important employee information from the time of recruiting, such as physiological characteristics when hired, educational background, and basic wage upgrades every year. Recorded wages are basic wages, which were upgraded annually and did not change through the year. Thus, they captures the notch at which the employee was placed in a finely designed promotion ladder in each year. Basic wages do not include bonuses, overtime compensation or other allowances, which depend on current performance or effort. Instead, they capture predictions about the employee's ability, which was upgraded every year based on the previous year's learning by the employer. Definitions and descriptive statistics of the variables used are in **Appendix III**. The total number of observations is 23,120.

There are three noteworthy characteristics of the data set. First, different from Western manufacturing firms, individual wage determination is not affected by negotiations with the union. In Japanese manufacturing, enterprise unions rather than trade unions are dominant. In addition, in collective bargaining between management and the enterprise union, they discuss only the average wage of the firm such that a productivity increase be distributed to the labor as well, but only on average. Unions do not intervene in individual performance evaluations. Thus, we can extract less noisy information from blue-collar workers' wages that we could from white-collar workers' wages. Second, as a custom of Japanese firms, all regular employees are eligible for a possible upgrade to their basic wages every year. Thus, focusing on basic wages in the data set means that we can track records of finely tuned promotions. Third, the data set is not dominated by those who were employed immediately after graduation. The mean of previous labor market experience (years after graduating from school and before employment with the firm, $t^p \equiv \mathrm{PreExperience}$) did not decrease through the sample period. Workers had on average three to eight years of previous labor market experience, often at smaller workplaces, through the sample period.[5] During the early twentieth century, when heavy manufacturing was introduced from the Western world, the typical career pattern for male skilled workers involved gaining experience at several workplaces to acquire relevant skills and then either gaining employment with a large firm on a long-term basis or starting one's own workshop.

Compulsory education was extended from six years to nine years in 1947. Therefore, the difference in educational background across employees who graduated before 1947 is distributed mainly between those with six years of schooling, who attended mandatory elementary schools, and those with eight years of schooling, who attended an additional two-year high elementary school. Here, high elementary school graduates are the majority. The difference for graduates after 1947 is distributed mainly between those who spent nine mandatory years attending a six-year elementary school and a three-year junior high school, and those who spent twelve years attending an additional three-year high school, with junior high school graduates as the majority.

3.3 Learning within an internal labor market

The existence of an internal labor market, which "shields" wage determination from the outside market by asymmetric employer learning, is to be empirically established. Persistent cohort effects are thought to be an indicator of the shielding effect of internal labor markets (Baker, Gibbs and Holmstrom (1994)). The market environment would be fully reflected only at entry. Following that, internal wage dynamics would be shielded from the market if there exists an internal labor market in the manner discussed in **Lemma 1**. **Table 1** regresses real wages as a logarithmic expression ($\log(w_{i,t})$) on the interaction terms of the two-year-joined dummy variables ($\mathrm{Yearjoined}_i^{19\mathrm{XX-YY}}$), which takes 1 if worker i joined the ironworks in 19XX or 19YY, and the first-lagged terms ($\mathrm{Yearjoined}_i^{1928-29} \cdot \log(w_{i,t-1})$, $\mathrm{Yearjoined}_i^{1930-31} \cdot \log(w_{i,t-1})$, etc.), controlling for years of schooling ($S \equiv \mathrm{School}_i$), years of total experience in the labor market ($t^T \equiv \mathrm{TotExperience}_{i,t}$), tenure at the ironworks

[5] See Nakabayashi (2013), **Table 1**.

($t^e \equiv \mathrm{Tenure}_{i,t}$), their squared terms, and growth in the real gross national product as a representative variable of exogenous shocks. Then, for all cohorts, cohort effects are significant. Furthermore, significant non-parallel wage curves are observed, even between adjacent cohorts. This indicates that the cohort effects should be considered when examining **Prediction 1**.

INSERT **Table 1** HERE

4 Empirical results

4.1 Standard test of employer learning

Before estimating equation (7), we describe the benchmark results for equation (5). **Table 2** gives the results of the random effect estimation regressing the logarithm of the real wages ($\log(w_{i,t})$) on relative height when employed by the firm (Height_i);[6] years of schooling (School_i); total experience in the labor market ($t^T_{i,t} = \mathrm{TotExperience}_{i,t}$); tenure at the firm ($t^e_{i,t} = \mathrm{Tenure}_{i,t}$), their squared terms; the interaction term between relative height and total labor market experience ($\mathrm{Height}_i \cdot \mathrm{TotExperience}_{i,t}$), the interaction between relative height and tenure ($\mathrm{Height}_i \cdot \mathrm{Tenure}_{i,t}$); the interaction between years of schooling and total labor market experience ($\mathrm{School}_i \cdot \mathrm{TotExperience}_{i,t}$); the interaction between years of schooling and tenure ($\mathrm{School}_i \cdot \mathrm{Tenure}_{i,t}$); the dummy variables for completing in-house training programs, namely the Development Center for Youth ($\mathrm{Training}^{1927-35}_{i,t}$, operated in 1927–35), School of Youth ($\mathrm{Training}^{1935-48}_{i,t}$, operated in 1935–48), Development Center for Technicians ($\mathrm{Training}^{1935-46}_{i,t}$, operated in 1939–46), and Development Center ($\mathrm{Training}^{1946-73}_{i,t}$, operated in 1946–73), which takes 1 when and after worker i completes a program; and the interaction of these dummy variables with tenure ($\mathrm{Training}^{1927-35}_{i,t} \cdot \mathrm{Tenure}$, $\mathrm{Training}^{1935-48}_{i,t} \cdot \mathrm{Tenure}_{i,t}$, $\mathrm{Training}^{1935-46}_{i,t} \cdot \mathrm{Tenure}_{i,t}$, $\mathrm{Training}^{1946-73}_{i,t} \cdot \mathrm{Tenure}_{i,t}$).[7] We control for the potential impact of extended compulsory schooling using the postwar-education generation dummy variable ($\mathrm{Postwar}_i$), which takes 1 if worker 1 is twelve years or older in 1947, when the US-led education reform was implemented.

INSERT **Table 2** HERE

In **Table 2**, the coefficient of tenure ($\mathrm{Tenure}_{i,t}$), controlling for total labor market experience ($\mathrm{TotExperience}_{i,t}$), implies that the return on firm-specific skill is considerable. Then, the interaction of years of schooling with total labor market experience after graduation ($\mathrm{School}_i \cdot \mathrm{TotExperience}_{i,t}$) has significant negative coefficient in specifications 2-1 and 2-3, as does the interaction of years of schooling with tenure ($\mathrm{School}_i \cdot \mathrm{Tenure}_{i,t}$) in specifications 2-2 and 2-4. The employer learning effect is clearly observed.

[6] To control for improved nutrition throughout the period, we use relative as compared to the national average height, provided by the Ministry of Education's statistics for estimation. Thus, (employee i's observed height)/(national average height at employee i's age in the year of i's joining) is used as i's "height (Height_i)."

[7] The information on height, weight, and lung capacity is not included in the wage records of the employees who joined the firm before 1939.

Along with years of schooling, proxies for abilities observable to the employer are physiological characteristics, such as height. Physical strength was important for blue-collar workers, and height is a good proxy of such physical strength. Indeed, with regard to height, the employer learning effect is observed. The interaction terms of relative height with both total labor market experience and with tenure ($\mathrm{Height}_i \cdot \mathrm{TotExperience}_{i,t}$, $\mathrm{Height}_i \cdot \mathrm{Tenure}_{i,t}$, respectively) have negative coefficients in specifications 2-3 and 2-4.

4.2 Learned abilities and acquired skills in the internal labor market

Next, we examine equation (7) and **Prediction 1**. A straightforward specification without control for the cohort effect by the random effect estimation is presented in specifications 3-1 and 3-2 in **Table 3**. After controlling for the changes in return on schooling using the interaction term between the year dummy variables and years of schooling ($\mathrm{Year}^{19\mathrm{XX}} \cdot \mathrm{School}_i$), specification 3-1 regresses the logarithm of real wages ($\log(w_{i,t})$) on years of schooling (School_i), labor market experience after graduation and before employment with the firm ($t_i^p = \mathrm{PreExperience}_i$), tenure after employment with the firm ($\mathrm{Tenure}_{i,t}$), their squared terms, the interaction between years of schooling and previous labor market experience ($\mathrm{School}_i \cdot \mathrm{PreExperience}_i$), and the interaction between years of schooling and tenure ($\mathrm{School}_i \cdot \mathrm{Tenure}_i$). Then, the interaction between years of schooling and previous labor market experience ($\mathrm{School}_i \cdot \mathrm{PreExperience}_i$) has a strictly greater coefficient than that between years of schooling and tenure ($\mathrm{School}_i \cdot \mathrm{Tenure}_{i,t}$) has, implying that **Prediction 1** holds: $\hat{\beta}_4 > \hat{\beta}_5$.

Specification 3-2, in addition to specification 3-1, controls for marginal decreases in the returns on previous labor market experience and tenure in their interactions with years of schooling ($\mathrm{School}_i \cdot \mathrm{PreExperience}_i^2$, $\mathrm{School} \cdot \mathrm{Tenure}^2$). Then, the interaction between years of schooling and previous labor market experience ($\mathrm{School}_i \cdot \mathrm{PreExperience}_i$) has a positive coefficient and the interaction between years of schooling and tenure ($\mathrm{School}_i \cdot \mathrm{Tenure}_{i,t}$) has a negative coefficient. Hence, **Prediction 1** holds: $\hat{\beta}_4 > \hat{\beta}_5$.

Specifications 3-3 and 3-4, in addition to specifications 3-1 and 3-2 respectively, control for cohort effects by inserting interactions between the two-year-joined dummy variables and years of schooling ($\mathrm{Yearjoined}_i^{1930-31} \cdot \mathrm{School}_i$, $\mathrm{Yearjoined}_i^{1932-33} \cdot \mathrm{School}_i$, etc.) as regressors. In specification 3-3, as in specification 3-1, both the interactions of years of schooling with previous labor market experience and with tenure ($\mathrm{School}_i \cdot \mathrm{PreExperience}_i$, $\mathrm{School}_i \cdot \mathrm{Tenure}_{i,t}$, respectively) have negative coefficients. However, the former is strictly greater than the latter. In specification 3-4, as in specification 3-2, the former has a significant positive coefficient and the latter has a significant negative coefficient. Thus, **Prediction 1** holds: $\hat{\beta}_4 > \hat{\beta}_5$.

INSERT **Table 3** HERE

Non-parallel wage curves in **Table 1** suggest the necessity to control for different effects of learning in different cohorts when checking the robustness of the results in **Table 3**. Thus, **Table 4** regresses the logarithm of real wages ($\log(w_{i,t})$) on years of schooling (School_i); previous labor market experience ($\mathrm{PreExperience}_i$); tenure ($\mathrm{Tenure}_{i,t}$); their squared terms, and, motivated by **Table 1**, the interaction terms of the two-year-joined dummy variables, years of schooling, and previous labor market experience ($\mathrm{Yearjoined}_i^{1936-37} \cdot \mathrm{School}_i \cdot \mathrm{PreExperience}_i, \mathrm{Yeearjoined}_i^{19}$

$\mathrm{School}_i \cdot \mathrm{PreExperience}_i$, etc.), and the interaction of the two-year-joined dummy variables, years of schooling, and tenure ($\mathrm{Yearjoined}_i^{1936-37} \cdot \mathrm{School}_i \cdot \mathrm{Tenure}_i$, $\mathrm{Yearjoined}_i^{1938-39} \cdot \mathrm{School}_i \cdot \mathrm{Tenure}_i$, etc.) to control for the cohort effects on learning. Possible marginal decreases in the interaction terms of previous labor market experience and tenure with years of schooling are controlled for by the interaction of their squared terms with years of schooling ($\mathrm{School}_i \cdot \mathrm{PreExperience}_i^2$, $\mathrm{School}_i \cdot \mathrm{Tenure}_{i,t}^2$) in specification 4-1 and by the interaction of their squared terms with squared years of schooling ($\mathrm{School}_i^2 \cdot \mathrm{PreExperience}_i^2$, $\mathrm{School}_i^2 \cdot \mathrm{Tenure}_{i,t}^2$) in specification 4-2. In addition, changes in the return on schooling over time are controlled for using the interaction between the year dummy variable and years of schooling ($\mathrm{Year}^{19XX} \cdot \mathrm{School}_i$). Since cohorts that contain two or more employees who had acquired positive previous labor market experience include only those who joined in 1934 or later, the sample cohorts are restricted to this period, and the control two-year-joined dummy variable is $\mathrm{Yearjoined}_i^{1934-35}$. Then with the exception of the only nonsignificant cohort 1936–37 in specification 4-1 and of those 1936–37, 1938-39, and 1940-1941 in specification 4-2, the interactions years of schooling with previous labor market experience in all significant cohorts have positive positive coefficients and the interactions with tenure in all cohorts have negative coefficients. Therefore, we still have $\hat{\beta}_4 > \hat{\beta}_5$ and **Prediction 1** holds, even after controlling for different learning curves in different cohorts.

INSERT **Table 4** HERE

While the regression of wages on the interaction term between years of schooling and total labor market experience ($\mathrm{School}_i \cdot \mathrm{TotExperience}_{i,t}$) in specifications 2-1 and 2-3 in **Table 2** suggests that the employer learning hypothesis holds, the results in **Table 3** indicate that the effect should be divided into before and after gaining employment with the firm ($\mathrm{School}_i \cdot \mathrm{PreExperience}_i$, $\mathrm{School}_i \cdot \mathrm{Tenure}_{i,t}$). In the mid-career recruiting market, workers were expected to have experience complementary to their schooling before gaining employment with the firm. After gaining employment with the firm, the complementarity was less valued and the employer learning process was more strongly observed. While the learning process differs in different cohorts, as shown in **Table 4**, the same tendency is still obvious after controlling for the difference due to cohort effects.

Table 4 also shows that the absolute value of the negative coefficient of the interaction between years of schooling and tenure ($\mathrm{School}_i \cdot \mathrm{Tenure}_{i,t}$) increases as the cohort nears the end of the covered period. Since the marginally decreasing return in a better match in the labor market or that in investment in firm-specific skill is captured by the interaction term between years of schooling and tenure squared ($\mathrm{School}_i \cdot \mathrm{Tenure}_{i,t}^2$) in specification 4-1 and by that between squared years of schooling and tenure squared ($\mathrm{School}_i^2 \cdot \mathrm{Tenure}_{i,t}^2$), the result is thought to come from the learning process. The greater absolute value of the negative coefficients for cohorts closer to the end suggest that the employer learning effect had a larger impact in the earlier tenure in the internal labor markets. Faster learning in the earlier career stages have been reported for the United States and Germany as well (Lluis (2005); Gibbons, Katz, Lemieux and Parent (2005) and Lange (2007)).

4.3 Employer learning for better trainee selection

Employers statistically discriminate when hiring. Then having hired, they begin to privately learn about employees' abilities and assign them to notches in the basic wage ladder based on their own private learning and employees' skill acquisition in the previous period. This is the primary channel of employers' learning and employees' skill acquisition. Both employers' learning and employees' skill acquisition are factored in when upgrading basic wages, but do not directly interact with each other.

There is another channel, in which learning might directly affect skill acquisition: In-house training programs in which decisions on who acquires what skills is made by the employer, not the employees. If, when making a decision on trainee selection, an employers uses employee information that he/she has learned privately since hiring, the in-house training program might improve efficiency by more than if selection was based only on information available in the market. If the employer does not use information he/she learned privately, internal training programs supplied by the employer would not outperform human resource allocations attained by the market. Thus, we next focus on in-house training programs provided by the case firm.

In the sample period, the firm operated four consecutive programs: The Development Center for Youth operated from 1927 to 1935 ($\mathrm{Training}^{1927-35}$);the School for Youth operated from 1935 to 1948 ($\mathrm{Training}^{1935-48}$); the Development Center for Technicians operated from 1939 to 1946 ($\mathrm{Training}^{1939-46}$); and the Development Center operated from 1946 to 1973 ($\mathrm{Training}^{1946-73}$).

Regulations behind the programs differed before and after the education reform in 1947. From 1926 to 1935, the government requested that major firms have a Development Center for Youth ($\mathrm{Training}^{1927-35}$). In 1935, the program was extended to a School for Youth ($\mathrm{Training}^{1935-48}$). From 1939, major firms were required to have a School for Youth ($\mathrm{Training}^{1935-48}$). Before 1947, secondary schooling was not mandatory, and the Development Center for Youth ($\mathrm{Training}^{1927-35}$) and School for Youth ($\mathrm{Training}^{1935-48}$t) were designed to complement shorter schooling. The Development Center for Technicians ($\mathrm{Training}^{1935-48}$) was not required by law and was operated according to the firm's own training plan.

The regulatory requirement was abandoned when junior high school became compulsory in 1947 and the School for Youth ($\mathrm{Training}^{1935-48}$) was abolished. The firm's original Development Center for Technicians ($\mathrm{Training}^{1939-46}$) program was reorganized as the Development Center ($\mathrm{Training}^{1946-73}$).

The Development Center for Youth ($\mathrm{Training}^{1927-35}$) and School for Youth ($\mathrm{Training}^{1935-48}$) offered a program of three days a week for four years, 800 hours in total. The Development Center for Technicians ($\mathrm{Training}^{1939-46}$) offered a full-time program for three years, 6,453 hours in total. The Development Center ($\mathrm{Training}^{194673}$) offered a program of three days a week before 1950, and six days a week from 1950, for two years. From 1963, only high school graduates were admitted as trainees.

Table 5 presents the estimated probabilities of acceptance to the in-house training programs, namely the School for Youth ($\mathrm{Training}^{1935-48}$), the Development Center for Technicians ($\mathrm{Training}^{1939-46}$), and the Development Center ($\mathrm{Training}^{1946-73}$), given age ($\mathrm{Age}$), years of schooling ($\mathrm{School}$), previous labor market experience ($\mathrm{PreExperience}$), and their squared terms. We assume that cohorts who joined the firm two years before introduction of a new program or later and until the year when the program was amended or abandoned were eligible to apply for the new program. Then, the sample period is defined as the one during which the relevant program was operated and the sample cohorts are defined as those who joined the firm from two years before the introduction of the relevant program to the year when the program ended. For the Development Center ($\mathrm{Training}^{194673}_{\mathrm{dc}}$), since high school graduation became an explicit condition to be accepted as a trainee from 1963, the sample period and cohorts are defined as being until 1962. For the earliest program, the Development Center for Youth from 1927 to 1935, the data set does not include sufficient cross-sections.

INSERT **Table 5** HERE

For schooling, the results are opposite for the School for Youth ($\mathrm{Training}^{1935-48}$) and for the Development Center ($\mathrm{Training}^{1946-73}$). In the former case, as required by regulations, employees with less schooling were more likely to be accepted. In the latter case, better educated employees were more likely to be accepted.

Next, we examine whether information learned privately by the employer was used for trainee selection in the post-1947 program, which was not motivated by the government regulations. In **Table 6**, specification 6-1, as a benchmark regresses the logarithm of the real wage ($\log\left(w_{i,t}\right)$) on tenure ($\mathrm{Tenure}_{i,t}$), its squared term, the interaction term of years of schooling, previous labor market experience, the Development Center dummy variable ($\mathrm{School}_i \cdot \mathrm{PreExperience}_i \cdot \mathrm{Training}^{1946-73}_{i,t}$), and the interaction term of years of schooling, tenure, and the Development Center dummy variable ($\mathrm{School}_i \cdot \mathrm{Tenure}_{i,t} \cdot \mathrm{Training}^{1946-73}_{i,t}$). Then, specification 6-2, as an inquiry on learning in trainee selection, regresses the logarithm of the real wage ($\log\left(w_{i,t}\right)$) on tenure ($\mathrm{Tenure}_{i,t}$), its squared term, the interaction term of years of schooling, previous labor market experience, and the probability of being accepted by Development Center estimated by specification 5-3 in **Table 5** ($\mathrm{School}_i \cdot \mathrm{PreExperience}_i \cdot \mathrm{E}[\mathrm{Training}^{1946-73}]$), and the interaction term of years of schooling, previous labor market experience, and the estimated probability of being accepted by Development Center ($\mathrm{School}_i \cdot \mathrm{Tenure}_{i,t} \cdot \mathrm{E}[\mathrm{Training}^{1946-73}_i]$). We also control for the Development Center dummy variable ($\mathrm{Training}^{1946-73}_{i,t}$), its interaction with tenure ($\mathrm{Training}^{1946-73}_{i,t} \cdot \mathrm{Tenure}_{i,t}$) in specification 6-1, and the estimated probability of being accepted by the Development Center ($\mathrm{E}[\mathrm{Training}^{1946-73}]$), its interaction with tenure ($\mathrm{E}[\mathrm{Training}^{1946-73}] \cdot \mathrm{Tenure}_{i,t}$) in specification 6-2. We also control for the interactions of years of schooling with the year dummy variable ($\mathrm{School}_i \cdot \mathrm{Year}^{19\mathrm{XX}}$) in both specifications.

INSERT **Table 6** HERE

First in specification 6-1, the benchmark specification, the coefficient of $\mathrm{School}_i \cdot \mathrm{PreExperience}_i \cdot \mathrm{Training}^{1946-73}_{i,t}$ is positive, while that of $\mathrm{School}_i \cdot \mathrm{Tenure}_{i,t} \cdot \mathrm{Training}^{1946-73}_{i,t}$ is negative, which is consistent with **Prediction 1**, $\hat{\beta}_4 > \hat{\beta}_5$.

Next, in specification 6-2, after inserting the estimated $\mathrm{E}[\mathrm{Training}^{1946-73}]$ for the observed $\mathrm{Training}^{1946-73}_{i,t}$, the coefficient of $\mathrm{School}_i \cdot \mathrm{PreExperience}_i \cdot \mathrm{E}[\mathrm{Training}^{1946-73}]$ is negative, while that of $\mathrm{School}_i \cdot \mathrm{Tenure}_{i,t} \cdot \mathrm{E}[\mathrm{Training}^{1946-73}]$ is positive, which contradicts the theoretical prediction. Furthermore, in specification 6-4, after controlling for marginal decreases in the interactions of tenure and previous work experience with years of schooling by their squared terms, the coefficients of both $\mathrm{School}_i \cdot \mathrm{PreExperience}_i \cdot \mathrm{E}[\mathrm{Training}^{1946-73}]$ and $\mathrm{School}_i \cdot \mathrm{Tenure}_{i,t} \cdot \mathrm{E}[\mathrm{Training}^{1946-73}]$ are negative. That is, the employer learning effect is mostly absorbed by the expected probability of being accepted as trainees.

This result indicates that the firm utilized information learned privately after hiring when selecting trainees and, hence, on average, the trainee selection process absorbs the employer learning effect, which would have appeared otherwise. Thus, overall, we can see that the in-house training program capitalized on information learned privately inside the internal labor market and, in this sense, likely improved efficiency. The information about abilities learned faster in earlier career stages was exploited efficiently for better training.

5 Discussion: A source of the mixed picture

We conjectured theoretically that, in the market, young workers are expected to have work experience complementary to schooling in the cross-sectional dimension. This complementarity of acquired skills in the cross-sectional distribution could hide the employer learning effect in the longitudinal dimension.

Then, we have empirically shown that the employer learning effect in the longitudinal dimension is dominated by the cross-sectional complementarity expectation and is hidden for previous labor market experience before workers gained long-term employment with the case firm. We then showed that the employer learning effect is clearly observed once they gained long-term employment. At the same time, the employer learning effect is more weakly observed in the latter stages of workers' internal careers. Furthermore, the case firm drew on information learned privately in the earlier internal career stages to screen potentially more competent employees as trainees for their in-house training program.

While this research addresses a Japanese experience, we believe that our results have relevance to labor markets in other developed economies. For example, the complementarity effect between schooling and work experience is also reported for young US workers in a similar to our results (Habermalz (2006)). Then, recent empirical research found that internal labor markets are still a widely used incentive device in the United States (Ben-Ner, Kong and Lluis (2012)). Given these related findings, internal labor markets that affect both the direction of workers' skill acquisition and the speed of employers' learning still seem to be prevalent in developed economies. In addition, and the diversity of their forms might explain the puzzling heterogeneity of workers' skill acquisition and employers' learning among developed economies. For instance, the schooling and training system in Germany (Pischke and von Wachter (2008)) might explain a puzzling result of employer learning observed in intra-firm wage dynamics in the country (Lluis (2005)). Our study suggest a viewpoint for further comparative studies on such diversity of firm organizations.

References

Altonji, Joseph G. and Charles R. Pierret, "Learning and statistical discrimination," *The Quarterly Journal of Economics*, Feb 2001, *116* (1), 313–350.

Ariga, Kenn, Yasushi Ohkusa, and Giorgio Brunello, "Fast track: is it in the genes?: The promotion theory of a large Japanese firm," *Journal of Economic Behavior & Organization*, Apr 1999, *38* (4), 385–402.

Arrow, Kenneth, "Higher education as a filter," *Journal of Public Economics*, Jul 1973, *2* (3), 193–216.

Baker, George, Michael Gibbs, and Bengt Holmstrom, "The wage policy of a firm," *The Quarterly Journal of Economics*, Nov 1994, *109* (4), 921–955.

Bauer, Thomas K. and John P. Haisken-DeNew, "Employer learning and the returns to schooling," *Labour Economics*, May 2001, *8* (2), 161–180.

Ben-Ner, Avner, Fanmin Kong, and Stéphanie Lluis, "Uncertainty, task environment, and organization design: An empirical investigation," *Journal of Economic Behavior & Organization*, Apr 2012, *82* (1), 281–313.

DeVaro, Jed and Michael Waldman, "The signaling role of promotions: Further theory and empirical evidence," *Journal of Labor Economics*, Jan 2012, *30* (1), 91–147.

Farber, Henry S. and Robert Gibbons, "Learning and wage dynamics," *The Quarterly Journal of Economics*, Nov 1996, *111* (4), 1007–1047.

Gibbons, Robert and Michael Waldman, "A theory of wage and promotion dynamics inside firms," *The Quarterly Journal of Economics*, Nov 1999, *114* (4), 1321–1358.

____ and ____ , "Enriching a theory of wage and promotion dynamics inside firms," *Journal of Labor Economics*, Jan 2006, *24* (1), 59–107.

____ , Lawrence F. Katz, Thomas Lemieux, and Daniel Parent, "Comparative advantage, learning, and sectoral wage determination," *Journal of Labor Economics*, Oct 2005, *23* (4), 681–724.

Habermalz, Steffen, "More detail on the pattern of returns to educational signals," *Southern Economic Journal*, Jul 2006, (1), 125–135.

Hungerford, Thomas and Gary Solon, "Sheepskin effects in the returns to education," *The Review of Economics and Statistics*, Feb 1987, *69* (1), 175–177.

Jaeger, David A. and Marianne E. Page, "Degrees matter: New evidence on sheepskin effects in the returns to education," *The Review of Economics and Statistics*, Nov 1996, *78* (4), 733–740.

Lange, Fabian, “The speed of employer learning,” *Journal of Labor Economics*, Jan 2007, *25* (1), 1–35.

Lluis, Stéphanie, “The role of comparative advantage and learning in wage dynamics and intrafirm mobility: Evidence from Germany,” *Journal of Labor Economics*, Oct 2005, *23* (4), 725–767.

Mansour, Hani, “Does employer learning vary by occupation?,” *Journal of Labor Economics*, Apr 2012, *30* (2), 415–444.

Mincer, Jacob, *Schooling, Experience, and Earnings*, New York: National Bureau of Economic Research, 1974.

Nakabayashi, Masaki, “Career experiences replaced: Emergence of Japanese internal labor markets,” April 2013. ISS Discussion Paper Series, The University of Tokyo, F-157, http://ideas.repec.org/p/itk/issdps/f157.html.

Nippon Tokei Kyokai (Japan Statistical Association), ed., *Nippon Choki Tokei Soran (Historical Statistics of Japan)*, Vol. 4, Tokyo: Nippon Tokei Kyokai (Japan Statistical Association), 1988.

Ohkawa, Kazushi, Nobukiyo Takamatsu, and Yuzo Yamamoto, *Choki Keizai Tokei 1 Kokumin Shotoku (Estimates of Long-term Economic Statistics of Japan since 1968, volume 7, National Income)*, Tokyo: Toyo Keizai Shinposha, 1974.

——, **Tsutomu Noda, Nobukiyo Takamatsu, Saburo Yamada, Minoru Kumazaki, Yuichi Shinomiya, and Ryoshin Minami**, *Choki Keizai Tokei 8 Bukka (Estimates of Long-term Economic Statistics of Japan since 1968, volume 8, Price Index)*, Tokyo: Toyo Keizai Shinposha, 1967.

Pinkston, Joshua C., “A test of screening discrimination with employer learning,” *Industrial and Labor Relations Review*, Jan 2006, *59* (2), 267–284.

——, “A model of asymmetric employer learning with testable implications,” *The Review of Economic Studies*, Jan 2009, *76* (1), 367–394.

Pischke, Jörn-Steffen and Till von Wachter, “Zero returns to compulsory schooling in Germany: Evidence and interpretation,” *The Review of Economics and Statistics*, Aug 2008, *90* (3), 592–598.

Riley, John G., “Testing the educational screening hypothesis,” *The Journal of Political Economy*, Oct 1979, *87* (5-2), s227–s252.

Rubinstein, Yona and Yoram Weiss, “Post schooling wage growth: Investment, search and learning,” in Eric A. Hanushek and Finis Welch, eds., *Handbook of the Economics of Education, volume 1*, Elsevier Amsterdam 2006, pp. 1–67.

Schönberg, Uta, "Testing for asymmetric employer learning," *Journal of Labor Economics*, Oct 2007, *25* (4), 651–691.

Spence, Michael, "Job market signaling," *The Quarterly Journal of Economics*, Aug 1973, *87* (3), 355–374.

Waldman, Michael, "Job assignments, signalling, and efficiency," *Rand Journal of Economics*, Summer 1984, *15* (2), 255–267.

____, "Theory and evidence in internal labor markets," in Robert Gibbons and John Roberts, eds., *The Handbook of Organizational Economics*, Princeton University Press Princeton, NJ 2013, pp. 520–569.

Appendix I: Proof

Proof of Lemma 1. For worker $i+1$ who is promoted in his/her third period, we have

$$w_{i+1,2,t}(S_{i+1},\eta_{i+1,t}) = d_2 + c_2\frac{k(d_1-d_2)-c_1(\phi_L f(2)+bS_{i+1}f(2))}{k(c_2-c_1)-c_1} + G(S_{i+1}),$$

$$\begin{aligned}&w_{i+1,2,t}(S_{i+1},\eta_{i+1,t}) - w_{i,2,t}(S_i,\eta_{i,t})\\ =&-\frac{c_1c_2(bS_{i+1}f(2)-bS_if(1))}{k(c_2-c_1)-c_1} - \frac{c_1c_2\phi_L(f(2)-f(1))}{k(c_2-c_1)-c_1} + G(S_{i+1}) - G(S_i).\end{aligned}$$

For worker $i+1$ who is not promoted in his/her second or third period, we have

$$w_{i+1,1,t}(S_{i+1},\eta_{i+1,t}) = d_1 + c_1(\phi_L + bS_{i+1})f(2) + G(S_{i+1}),$$

$$\begin{aligned}&w_{i+1,1,t}(S_{i+1},\eta_{i+1,t}) - w_{i,1,t}(S_i,\eta_{i,t})\\ =&c_1(bS_{i+1}f(2)-bS_if(1)) + c_1\phi_L(f(2)-f(1)) + G(S_{i+1}) - G(S_i).\end{aligned}$$

Suppose that $S_{i+1} > S_i$. The term $(bS_{i+1}f(2) - bS_if(1))$ has a positive multiplier if work experience and schooling are complements and a negative multiplier if they are substitutes in the cross-sectional dimension $i = 1, 2, \ldots, n$. Since $k > c_1/(c_2 - c_1)$, $c_1 > 0 > -c_1c_2/\left[k(c_2-c_1)-c_1\right]$, which implies that work experience and schooling are expected to be complements for non-promoted workers and substitutes for promoted workers. $\square$

Appendix II: Data sources

Wages and workers' characteristics Original wage records of the case firm in Japan.

Series of national data Consumer prices (to deflate nominal wages): Nippon Tokei Kyokai (Japan Statistical Association), ed (1988), p.362. National average height: the School Health Statistics surveyed by the Ministry of Education, Science, Sports and Culture (http://www.e-stat.go.jp/). Real gross national product: Ohkawa, Takamatsu and Yamamoto (1974), pp. 232 (1885-1929) – 233 (1930–70); to connect series before and after 1955, when governmental statistics are not continuous, a deflator from Ohkawa, Noda, Takamatsu, Yamada, Kumazaki, Shinomiya and Minami (1967), p. 134, is used.

Appendix III Definition and descriptive statistics of variables.

Variable	Definition	Mean	Median	Maximum	Minimum	Standard deviation	Skewness	Kurtosis	Number of observations
w	Real daily wage: yen per day.	3.5782	3.3700	72.0600	0.3400	1.9650	2.4475	66.7437	23,120
Height	Relative height when employed by the firm: (observed height)/(national average height at his age in the year).	0.9957	1.0000	1.1000	0.8000	0.0406	-0.4750	6.6180	16,637
Age	Age.	30.5638	30.0000	55.0000	13.0000	8.1126	0.3644	2.5497	23,120
School	Years of schooling.	8.7093	8.0000	15.0000	5.0000	1.6194	1.1881	4.3356	23,120
Postwar	Postwar education generation dummy variable: =1 if 12 years old or younger in 1947, and 0 otherwise.	0.1805	0.0000	1.0000	0.0000	0.3846	1.6615	3.7606	23,120
TotExperience	Years of total labor market experience: Age–(6+School).	15.8309	15.0000	42.0000	0.0000	8.5340	0.3159	2.5205	23,120
PreExperience	Years of previous labor market experience prior joining the firm: Age–(6+School+Tenure). Every sample employee had worked at the firm until the last year of his record.	6.3631	6.0000	35.0000	0.0000	5.1436	0.7689	3.4393	23,120
Tenure	Tenure: (years after employed by the firm).	10.0591	9.0000	37.7500	0.0000	6.9391	0.6156	2.7515	23,120
$Training^{1927-35}$	=1 if completed Development Center for Youth (operated from 1927 to 1935), and 0 otherwise.	0.0010	0.0000	1.0000	0.0000	0.0308	32.3714	1,048.9100	23,120
$Training^{1935-48}$	=1 if completed School for Youth (operated from 1935 to 1948), and 0 otherwise.	0.0419	0.0000	1.0000	0.0000	0.2004	4.5720	21.9034	23,120
$Training^{1939-46}$	=1 if completed Development Center for Technician (operated from 1939 to 1946).	0.0513	0.0000	1.0000	0.0000	0.2205	4.0700	17.5646	23,120
$Training^{1946-73}$	=1 if completed Development Center (operated from 1946 to 1973), and 0 otherwise.	0.1257	0.0000	1.0000	0.0000	0.3316	2.2577	6.0970	23,120
$Yearjoined^{19XX-Y}$	2-year-joined dummy variable: =1 if joined the firm from 19XX to 19YY(=19XX+1), and 0 otherwise.								
$Year^{19XX}$	Year dummy variable: =1 if the year is 19XX, and 0 otherwise.								
GNP	Real gross national product.								

Sources : See **Appendix II**.

Table 1 Cohort effect on wage curves.

	1 – 1		
Estimation method	panel extended generalized least squares		
Dependent variable	$\log(w_{i,t})$		
Cross-section	random effect		
Period (year)	pooled (no year dummies inserted)		
Independent variables	coefficient	t statistic	
Constant	-0.0781	-2.2423	**
School_i	0.0358	5.0124	***
$\text{School}_i^{\,2}$	-0.0015	-4.0319	***
$\text{TotExperience}_{i,t}$	0.0112	15.7277	***
$\text{TotExperience}_{i,t}^{\,2}$	-0.0001	-7.8801	***
$\text{Tenure}_{i,t}$	0.0125	13.2765	***
$\text{Tenure}_{i,t}^{\,2}$	0.0004	11.3815	***
$\text{Yearjoined}^{1930-31}{}_i \cdot \log(w_{i,t-1})$	0.4359	15.0252	***
$\text{Yearjoined}^{1932-33}{}_i \cdot \log(w_{i,t-1})$	0.4755	32.3269	***
$\text{Yearjoined}^{1934-35}{}_i \cdot \log(w_{i,t-1})$	0.5053	42.7458	***
$\text{Yearjoined}^{1936-37}{}_i \cdot \log(w_{i,t-1})$	0.5218	57.3005	***
$\text{Yearjoined}^{1938-39}{}_i \cdot \log(w_{i,t-1})$	0.5650	92.1537	***
$\text{Yearjoined}^{1940-41}{}_i \cdot \log(w_{i,t-1})$	0.5815	102.6212	***
$\text{Yearjoined}^{1942-43}{}_i \cdot \log(w_{i,t-1})$	0.6017	97.8593	***
$\text{Yearjoined}^{1944-45}{}_i \cdot \log(w_{i,t-1})$	0.6253	90.2239	***
$\text{Yearjoined}^{1946-47}{}_i \cdot \log(w_{i,t-1})$	0.6642	82.7922	***
$\text{Yearjoined}^{1948-49}{}_i \cdot \log(w_{i,t-1})$	0.6853	168.9371	***
$\text{Yearjoined}^{1950-51}{}_i \cdot \log(w_{i,t-1})$	0.7132	134.0633	***
$\text{Yearjoined}^{1952-53}{}_i \cdot \log(w_{i,t-1})$	0.7488	68.2213	***
$\text{Yearjoined}^{1954-55}{}_i \cdot \log(w_{i,t-1})$	0.7746	85.3240	***
$\text{Yearjoined}^{1956-57}{}_i \cdot \log(w_{i,t-1})$	0.7629	152.6817	***
$\text{Yearjoined}^{1958-59}{}_i \cdot \log(w_{i,t-1})$	0.7967	107.5379	***
$\text{Yearjoined}^{1960-61}{}_i \cdot \log(w_{i,t-1})$	0.8036	77.0475	***
$\text{Yearjoined}^{1962-63}{}_i \cdot \log(w_{i,t-1})$	0.8219	66.1291	***
$\text{Yearjoined}^{1964-65}{}_i \cdot \log(w_{i,t-1})$	0.8532	57.0154	***
$\text{Yearjoined}^{1966-67}{}_i \cdot \log(w_{i,t-1})$	0.8847	28.1190	***
ΔGNP	Yes		
cross-sections included		1,555	
periods included (years)	40 (1930 – 69)		
included observations		21,562	
adjusted R^2		0.8936	
F statistic		6,962.4719	***

Notes : The control cohort dummy variable is $\text{Yearjoined}^{1928-1929}$. *** and ** respectively denote significance at the 1- and 5-percent levels.

Table 2 Skill elements and employer learning effect in wage determination.

	2−1		2−2		2−3		2−4	
Estimation method	panel extended generalized least squares							
Dependent variable	$\log(w_{i,t})$							
Cross-section	random effect							
Period (year)	pooled (no year dummies inserted)							
Independent variables	coefficient	*t* statistic	coefficient	*t* statistic	coefficient	*t* statistic	coefficient	*t* statistic
Constant	-1.3797	-17.6256 ***	-1.0374	-13.7473 ***	-8.2014	-11.2220 ***	-7.9041	-10.3975 ***
Height_i					12.3992	8.5047 ***	13.2281	8.7190 ***
$\text{Heigt}_i^{\,2}$					-5.4457	-7.4297 ***	-6.1167	-8.0237 ***
School_i	0.1904	12.3740 ***	0.1395	9.1231 ***	0.1617	8.9997 ***	0.0868	4.8363 ***
$\text{School}_i^{\,2}$	-0.0064	-8.5673 ***	-0.0050	-6.5236 ***	-0.0051	-5.8399 ***	-0.0024	-2.7609 ***
Postwar_i	0.4703	57.5718 ***	0.4937	60.3669 ***	0.4932	62.4583 ***	0.5166	64.0350 ***
$\text{TotExperience}_{i,t}$	0.0587	30.7057 ***	0.0412	35.8644 ***	0.1274	20.6904 ***	0.0296	25.5080 ***
$\text{TotExperience}_{i,t}^{\,2}$	-0.0003	-8.7299 ***	-0.0002	-6.7225 ***	-0.0002	-6.9956 ***	-0.0002	-5.6679 ***
$\text{Tenure}_{i,t}$	0.0870	77.4515 ***	0.0922	46.6436 ***	0.1356	103.5312 ***	0.2114	28.8327 ***
$\text{Tenure}_{i,t}^{\,2}$	-0.0015	-37.4222 ***	-0.0015	-36.9969 ***	-0.0030	-54.7404 ***	-0.0030	-54.5433 ***
$\text{Height}_i \cdot \text{TotExperience}_{i,t}$	-0.0018	-11.5768 ***			-0.0759	-12.3038 ***		
$\text{Height}_i \cdot \text{Tenure}_{i,t}$							-0.0616	-8.6816 ***
$\text{School}_i \cdot \text{TotExperience}_{i,t}$					-0.0024	-16.2488 ***		
$\text{School}_i \cdot \text{Tenure}_{i,t}$			-0.0006	-3.1841 ***			-0.0016	-9.2743 ***
$\text{Training}^{1927-35}_{i,t}$	-0.8494	-5.0896 ***	-0.8338	-4.9526 ***				
$\text{Training}^{1927-35}_{i,t} \cdot \text{Tenure}_{i,t}$	0.0221	2.4901 **	0.0212	2.3904 **				
$\text{Training}^{1935-48}_{i,t}$	-0.1644	-7.6893 ***	-0.1707	-7.8796 ***				
$\text{Training}^{1935-48}_{i,t} \cdot \text{Tenure}_{i,t}$	0.0075	5.6422 ***	0.0076	5.7226 ***				
$\text{Training}^{1939-46}_{i,t}$	-0.2251	-11.2331 ***	-0.2305	-11.3440 ***				
$\text{Training}^{1939-46}_{i,t} \cdot \text{Tenure}_{i,t}$	0.0107	8.6292 ***	0.0109	8.7899 ***				
$\text{Training}^{1946-73}_{i,t}$	0.1539	13.0095 ***	0.1372	11.5120 ***				
$\text{Training}^{1946-73}_{i,t} \cdot \text{Tenure}_{i,t}$	-0.0067	-6.1725 ***	-0.0056	-5.1606 ***				
cross-sections included		1,558		1,558		1,246		1,246
periods included (years)	41 (1929−1969)		41 (1929−69)		31 (1939−69)		31 (1939−69)	
included observations		23,210		23,210		16,637		16,637
adjusted R^2		0.7730		0.7745		0.8640		0.8639
F statistic		4,922.0817 ***		4,964.6115 ***		9,609.0120 ***		9,599.4685 ***

Notes : *** and ** respectively denote significance at the 1- and 5-percent levels. The information about physiological characteristics is not included in the wage records of the employees who joined the firm before 1939.

Table 3 Detection of employer learning effects in previous experience and tenure.

	3−1		3−2		3−3		3−4	
Estimation method	panel extended generalized least squares							
Dependent variable	$\log(w_{i,t})$							
Cross-section	random effect							
Period (year)	pooled (no year dummies inserted)							
Independent variables	coefficient	t statistic	coefficient	t statistic	coefficient	t statistic	coefficient	t statistic
Constant	-0.6740	-15.2121 ***	-0.4701	-18.9872 ***	-0.6386	-14.8029 ***	-0.6996	-15.7167 ***
School_i	0.0957	11.1396 ***	0.0538	11.7441 ***	0.1212	12.1921 ***	0.1204	11.9630 ***
$\text{School}_i^{\,2}$	-0.0059	-14.4278 ***	-0.0038	-17.7145 ***	-0.0056	-13.9784 ***	-0.0054	-13.6782 ***
PreExperience_i	0.0306	18.4545 ***	0.0198	9.2105 ***	0.0295	17.9188 ***	0.0134	3.6080 ***
$\text{PreExperience}_i^{\,2}$	-0.0004	-11.0913 ***	0.0004	3.1714 ***	-0.0002	-5.2303 ***	0.0006	3.2492 ***
$\text{Tenure}_{i,t}$	0.1093	157.6259 ***	0.1246	63.1873 ***	0.1092	157.2888 ***	0.1427	84.6012 ***
$\text{Tenure}_{i,t}^{\,2}$	-0.0007	-52.9482 ***	-0.0019	-23.6766 ***	-0.0006	-47.8240 ***	-0.0021	-29.9396 ***
$\text{School}_i \cdot \text{PreExperience}_i$	-0.0004	-2.6272 ***	0.0007	2.9480 ***	-0.0010	-6.2116 ***	0.0009	2.1320 **
$\text{School}_i \cdot \text{Tenure}_{i,t}$	-0.0054	-78.3730 ***	-0.0071	-31.2024 ***	-0.0075	-34.1524 ***	-0.0114	-40.1782 ***
$\text{School}_i \cdot \text{PreExperience}_i^{\,2}$			-0.0001	-6.8403 ***			-0.0001	-4.2313 ***
$\text{School}_i \cdot \text{Tenure}_{i,t}^{\,2}$			0.0001	14.0995 ***			0.0002	21.6793 ***
$\text{Year}^{19XX} \cdot \text{School}_i$	Yes		Yes		Yes		Yes	
$\text{Yearjoined}^{19XX-YY}_i \cdot$	No		No		Yes		Yes	
cross-sections included		1,558		1,558		1,558		1,558
periods included (years)	41 (1929−69)		41 (1929−69)		41 (1929−69)		41 (1929−69)	
included observations		23,120		23,120		23,120		23,120
adjusted R^2		0.9766		0.9687		0.9771		0.9776
F statistic		20,076.5892 ***		14,300.3232 ***		14,741.6266 ***		14,624.3724 ***

Notes: *** and ** respectively denote significance at the 1- and 5-percent levels. Control year dummy variable for $\text{Year}^{19XX} \cdot$ School is Year^{1929}. Control 2-year-joined dummy variable for $\text{Yearjoined}^{19XX\text{-}YY} \cdot$ School is $\text{Yearjoined}^{1928\text{-}1929}$.

Table 4 Detection of employer learning effects in previous experience and tenure with controlling for cohort effects.

		4−1			4−2		
Estimation method		panel extended generalized least squares					
Dependent variable		$\log(w_{i,t})$					
Cross-section		random effect					
Period (year)		pooled (no year dummies inserted)					
Independent variables		coefficient	t statistic		coefficient	t statistic	
	Constant	-0.6047	-14.6739	***	-0.3901	-9.2345	***
	School_i	0.0866	10.7921	***	0.0420	5.1172	***
	$\text{School}_i^{\ 2}$	-0.0063	-16.1554	***	-0.0040	-10.1654	***
	PreExperience_i	0.0115	4.3058	***	0.0143	5.3403	***
	$\text{PreExperience}_i^{\ 2}$	0.0013	8.7910	***	0.0000	-0.2235	
	$\text{Tenure}_{i,t}$	0.0867	87.6503	***	0.0852	86.2228	***
	$\text{Tenure}_{i,t}^{\ 2}$	0.0001	3.0339	***	-0.0005	-15.3233	***
previous experience	$\text{Yearjoined}^{1936-37}_i \cdot \text{School}_i \cdot \text{PreExperience}_i$	-0.0001	-0.3144		-0.0004	-0.9107	
	$\text{Yearjoined}^{1938-39}_i \cdot \text{School}_i \cdot \text{PreExperience}_i$	0.0007	2.0291	**	0.0004	1.1647	
	$\text{Yearjoined}^{1940-41}_i \cdot \text{School}_i \cdot \text{PreExperience}_i$	0.0008	2.3674	**	0.0004	1.2302	
	$\text{Yearjoined}^{1942-43}_i \cdot \text{School}_i \cdot \text{PreExperience}_i$	0.0016	4.6001	***	0.0013	3.7370	***
	$\text{Yearjoined}^{1944-45}_i \cdot \text{School}_i \cdot \text{PreExperience}_i$	0.0014	3.4760	***	0.0011	2.6563	***
	$\text{Yearjoined}^{1946-47}_i \cdot \text{School}_i \cdot \text{PreExperience}_i$	0.0036	9.3756	***	0.0032	8.3964	***
	$\text{Yearjoined}^{1948-49}_i \cdot \text{School}_i \cdot \text{PreExperience}_i$	0.0034	11.2480	***	0.0030	10.0146	***
	$\text{Yearjoined}^{1950-51}_i \cdot \text{School}_i \cdot \text{PreExperience}_i$	0.0032	10.2964	***	0.0028	9.1290	***
	$\text{Yearjoined}^{1952-53}_i \cdot \text{School}_i \cdot \text{PreExperience}_i$	0.0046	8.7995	***	0.0042	8.1084	***
	$\text{Yearjoined}^{1954-55}_i \cdot \text{School}_i \cdot \text{PreExperience}_i$	0.0049	8.4089	***	0.0047	7.9865	***
	$\text{Yearjoined}^{1956-57}_i \cdot \text{School}_i \cdot \text{PreExperience}_i$	0.0029	9.6702	***	0.0025	8.4724	***
	$\text{Yearjoined}^{1958-59}_i \cdot \text{School}_i \cdot \text{PreExperience}_i$	0.0019	5.8496	***	0.0015	4.7576	***
	$\text{Yearjoined}^{1960-61}_i \cdot \text{School}_i \cdot \text{PreExperience}_i$	0.0023	7.3898	***	0.0019	6.2218	***
	$\text{Yearjoined}^{1962-63}_i \cdot \text{School}_i \cdot \text{PreExperience}_i$	0.0028	9.0709	***	0.0024	7.8368	***
	$\text{Yearjoined}^{1964-65}_i \cdot \text{School}_i \cdot \text{PreExperience}_i$	0.0043	14.0531	***	0.0039	12.8821	***
	$\text{Yearjoined}^{1966-67}_i \cdot \text{School}_i \cdot \text{PreExperience}_i$	0.0038	7.5879	***	0.0035	6.9848	***
tenure	$\text{Yearjoined}^{1936-37}_i \cdot \text{School}_i \cdot \text{Tenure}_{i,t}$	-0.0011	-12.0123	***	-0.0011	-12.2029	***
	$\text{Yearjoined}^{1938-39}_i \cdot \text{School}_i \cdot \text{Tenure}_{i,t}$	-0.0011	-14.1849	***	-0.0011	-13.9519	***
	$\text{Yearjoined}^{1940-41}_i \cdot \text{School}_i \cdot \text{Tenure}_{i,t}$	-0.0011	-14.2744	***	-0.0011	-13.3863	***
	$\text{Yearjoined}^{1942-43}_i \cdot \text{School}_i \cdot \text{Tenure}_{i,t}$	-0.0012	-14.1783	***	-0.0012	-13.3786	***
	$\text{Yearjoined}^{1944-45}_i \cdot \text{School}_i \cdot \text{Tenure}_{i,t}$	-0.0012	-12.4599	***	-0.0011	-11.5020	***
	$\text{Yearjoined}^{1946-47}_i \cdot \text{School}_i \cdot \text{Tenure}_{i,t}$	-0.0008	-7.5235	***	-0.0008	-6.9188	***
	$\text{Yearjoined}^{1948-49}_i \cdot \text{School}_i \cdot \text{Tenure}_{i,t}$	-0.0023	-23.9710	***	-0.0022	-22.7975	***
	$\text{Yearjoined}^{1950-51}_i \cdot \text{School}_i \cdot \text{Tenure}_{i,t}$	-0.0022	-19.5376	***	-0.0021	-18.6083	***
	$\text{Yearjoined}^{1952-53}_i \cdot \text{School}_i \cdot \text{Tenure}_{i,t}$	-0.0022	-11.6990	***	-0.0021	-11.0768	***
	$\text{Yearjoined}^{1954-55}_i \cdot \text{School}_i \cdot \text{Tenure}_{i,t}$	-0.0021	-13.2253	***	-0.0020	-12.4992	***
	$\text{Yearjoined}^{1956-57}_i \cdot \text{School}_i \cdot \text{Tenure}_{i,t}$	-0.0027	-17.9852	***	-0.0026	-16.9971	***
	$\text{Yearjoined}^{1958-59}_i \cdot \text{School}_i \cdot \text{Tenure}_{i,t}$	-0.0026	-14.1252	***	-0.0025	-13.3082	***
	$\text{Yearjoined}^{1960-61}_i \cdot \text{School}_i \cdot \text{Tenure}_{i,t}$	-0.0031	-12.6286	***	-0.0029	-12.0812	***
	$\text{Yearjoined}^{1962-63}_i \cdot \text{School}_i \cdot \text{Tenure}_{i,t}$	-0.0039	-11.4366	***	-0.0038	-10.9617	***
	$\text{Yearjoined}^{1964-65}_i \cdot \text{School}_i \cdot \text{Tenure}_{i,t}$	-0.0052	-10.1331	***	-0.0051	-9.7979	***
	$\text{Yearjoined}^{1966-67}_i \cdot \text{School}_i \cdot \text{Tenure}_{i,t}$	-0.0081	-8.8014	***	-0.0080	-8.7226	***
	$\text{School}_i \cdot \text{PreExperience}_i^{\ 2}$	-0.0003	-15.4276	***			
	$\text{School}_i^{\ 2} \cdot \text{PreExperience}_i^{\ 2}$				0.0000	-13.9982	***
	$\text{School}_i^{\ 2} \cdot \text{Tenure}_{i,t}^{\ 2}$	-0.0001	-36.8799	***			
	$\text{School}_i \cdot \text{PreExperience}_i^{\ 2}$				0.0000	-38.1034	***
	$\text{Year}^{19XX} \cdot \text{School}_i$	Yes			Yes		
	cross-sections included		1,551			1,551	
	periods included (years)	36 (1934-69)			36 (1934-69)		
	cohorts included	joined in 1934-67			joined in 1934-67		
	included observations		22,928			22,928	
	adjusted R^2		0.9767			0.9767	
	F statistic		12,793.6875	***		12,799.9763	***

***Notes*:** Control 2-year-joined dummy variable is $\text{Yearjoined}^{1934-1935}$ and control year dummy variable for interaction with schooling is Year^{1934}. *** and ** respectively denote significance at the 1-percent and 5-percent levels.

Table 5 Probability of acceptance as a trainee for in-house training programs.

	5−1				5−2				5−3			
Estimation method	binary probit				binary probit				binary probit			
Dependent variable	$\text{Training}^{1935-48}$				$\text{Training}^{1939-46}$				$\text{Training}^{1946-73}$			
Independent variables	coefficient	*z* statistic		marginal effect	coefficient	*z* statistic		marginal effect	coefficient	*z* statistic		marginal effect
Constant	-5.5027	-2.9138	***		-34.4557	-5.3953	***		-21.0175	-20.4493	***	
Age	0.7384	5.1269	***	0.0009	3.2124	5.8500	***	0.0000	0.0968	3.0832	***	0.0034
Age^2	-0.0159	-4.8545	***		-0.0821	-5.5880	***		-0.0029	-4.5522	***	
School	-0.8331	-3.2905	***	-0.0010	0.5146	0.6187		0.0000	4.5331	22.8440	***	0.1592
School^2	0.0414	3.1259	***		-0.0181	-0.4422			-0.2445	-23.7262	***	
PreExperience	-0.1676	-4.4759	***	-0.0002	-0.8942	-7.0762	***	0.0000	-0.4530	-37.0305	***	-0.0159
PreExperience^2	0.0077	2.5979	***		0.0721	6.5262	***		0.0203	24.1173	***	
periods included (years)	13 (1936−48)				8 (1939−46)				17 (1946−62)			
cohorts included	joined in 1933−48				joined in 1937−46				joined in 1944−62			
included observations				1,822				844				12,741
Log likelihood				-342.6259				-137.5664				-2,769.0706
McFadden R^2				0.0957				0.4121				0.4700
LR statistic				72.4878 ***				192.8855 ***				4,911.9308 ***

***Notes*:** Marginal effects are calculated by mean values of independent variables. *** and ** respectively denote significance at the 1- and 5-percent levels.

Table 6 Employer learning for trainee selection.

	6−1			6−2			6−3			6−4		
Estimation method	panel extended generalized least squares											
Dependent variable	$\log(w_{i,t})$											
Cross-section	random effect											
Period (year)	pooled (no year dummies inserted)											
Independent variables	coefficient	*t* statistic		coefficient	*t* statistic		coefficient	*t* statistic		coefficient	*t* statistic	
Constant	-0.0368	-8.0315	***	-0.1534	-31.4632	***	-0.1269	-28.3160	***	-0.1713	-36.9536	***
$\text{Tenure}_{i,t}$	0.0926	125.9735	***	0.0809	111.8764	***	0.0843	121.1516	***	0.0811	117.1807	***
$\text{Tenure}_{i,t}^2$	-0.0008	-18.9280	***	-0.0009	-22.1007	***	0.0008	16.3695	***	-0.0006	-10.2334	***
$\text{School}_i \cdot \text{PreExperience}_i \cdot \text{Training}^{1946-73}_{i,t}$	0.0018	5.8888	***				0.0004	1.3816				
$\text{School}_i \cdot \text{Tenure}_i \cdot \text{Training}^{1946-73}_{i,t}$	-0.0034	-4.6671	***				-0.0018	-2.6263	***			
$\text{School}_i \cdot \text{PreExperience}_i \cdot \text{E}[\text{Training}^{1946-73}]$				-0.0002	-13.5066	***				0.0003	9.9936	***
$\text{School}_i \cdot \text{Tenure}_i \cdot \text{E}[\text{Training}^{1946-73}]$				0.0015	76.6054	***				0.0012	29.3802	***
$\text{School}_i^{\,2} \cdot \text{PreExperience}_i^{\,2}$							0.0000	32.5343	***	0.0000	23.4869	***
$\text{School}_i^{\,2} \cdot \text{Tenure}_{i,t}^{\,2}$							0.0000	-77.0932	***	0.0000	-7.4097	***
$\text{Training}^{1946-73}_{i,t}$	0.0556	8.0619	***				0.0483	7.6216	***			
$\text{Training}^{1946-73}_{i,t} \cdot \text{Tenure}_{i,t}$	0.0338	5.3983	***				0.0176	2.9762	***			
$\text{E}[\text{Training}^{1946-73}]$				-0.0242	-10.6473	***				-0.0419	-19.6426	***
$\text{E}[\text{Training}^{1946-73}] \cdot \text{Tenure}_{i,t}$				-0.0114	-52.2662	***				-0.0099	-26.4809	***
$\text{School}_i \cdot \text{Year}^{19XX}$	Yes			Yes			Yes			Yes		
cross-sections included		1,229			1,229			1,229			1,229	
periods included (years)	17 (1946 – 62)			17 (1946 – 62)			17 (1946 – 62)			17 (1946 – 62)		
cohorts included	1944 – 62			1944 – 62			1944 – 62			1944 – 62		
included observations		12,741			12,741			12,741			12,741	
adjusted R^2		0.9731			0.9786			0.9782			0.9784	
F statistic		20,962.8151	***		26,423.3359	***		23,774.2416	***		24,055.6607	***

Notes : $\text{E}[\text{Training}_{dc}^{1946-73}]$ is estimated by specification 5 – 3 in **Table 5**. ***, ** and * respectively denote significance at the 1-, 5-, and 10-percent levels. Control year dummy variable for $\text{School} \cdot \text{Year}^{19XX}$ is Year^{1946}.